

REVUE SEMESTRIELLE EN SCIENCES DE L'ÉDUCATION



Contextes
et
Didactiques

« Contextes et évaluations : quelles interactions ? »

NUMÉRO 9

de la Revue « *Contextes et Didactiques* »

Coordonné par

Elisabeth ISSAIEVA MOUBARAK NAHRA

et Lucie MOTTIER LOPEZ

Juin 2017



POLITIQUE EDITORIALE

La revue « *Contextes et Didactiques* » est une revue numérique semestrielle internationale en libre accès. La diversité, la coexistence, voire la confrontation de contextes de différentes natures jouent un rôle fondamental dans de nombreuses situations d'enseignement. Depuis les contextes historiques, géographiques, écologiques, sociolinguistiques, et culturels notamment, et leurs interactions en contextes didactiques, la revue « *Contextes et Didactiques* » publie des articles qui rendent compte de cette dimension encore peu explorée sous toutes ses facettes (interculturalité, multilinguisme, identité, symbolique, institution) en proposant des recherches empiriques, des comptes-rendus de pratiques innovantes, des analyses de dispositifs didactiques, des revues de littérature et des réflexions théoriques liés aux interactions entre une pluralité de contextes et de situation d'éducation, de formation et d'apprentissage. La revue accorde une attention particulière aux travaux portant sur les questions de contextualisation didactique dans des situations post-coloniales, des territoires ultrapériphériques français et plus généralement dans les pays de la Francophonie, mais s'intéresse à toutes situations spécifiques mettant en évidence des décalages entre les prescriptions institutionnelles, les contextes d'enseignement et les conceptions des acteurs, ou « effets de contextes », et contribuant à leur identification et leur prise en considération dans les processus d'enseignement, d'apprentissage et de formation. Ainsi, la revue est ouverte à toute contribution traitant des questions de didactique en prise avec différents éléments contextuels.

Les propositions de contributions sont à envoyer au secrétariat de la revue (contextes.didactiques@espe-guadeloupe.fr), ainsi qu'au(x) responsable(s) de chaque numéro thématique dans les délais impartis (cf. Appels à contributions). Les appels à contribution sont lancés deux fois par an, en juin et en décembre. La revue « *Contextes et Didactiques* » publie aussi occasionnellement des textes hors thématiques dans une rubrique *varia*. La revue est dotée d'un comité de rédaction et d'un comité scientifique constitués d'experts garants de la qualité des articles publiés. Les articles reçus font l'objet d'une double évaluation en aveugle par le comité scientifique. Ce dernier a la charge de lire les manuscrits soumis, de les expertiser et de proposer les modifications qu'il estime nécessaire en vue de leur publication. La revue n'est pas responsable des manuscrits qui lui sont envoyés. L'envoi des documents implique l'accord de l'auteur pour leur publication. Les textes et les images publiés dans la revue engagent la responsabilité de leurs seuls auteurs.

Historique de la revue

L'ancienne revue intitulée « *Recherches et Ressources en Éducation et Formation* » créée en 2007 par des enseignants-chercheurs en Didactique de l'IUFM (Institut Universitaire de Formation des Maîtres) de Guadeloupe, et publiée annuellement sous forme papier par le CRDP (Centre Régional de Documentation Pédagogique) de Guadeloupe a changé en 2014, de titre et de format éditorial pour devenir une revue semestrielle en sciences de l'éducation et en ligne, intitulée : « *Contextes et Didactiques* ». Cette revue est soutenue par l'ESPE (École Supérieure du Professorat et de l'Éducation) de l'académie de Guadeloupe de l'université des Antilles et par le Centre de Recherches et de Ressources en Éducation et Formation (CRREF EA 4538). La consultation et le téléchargement des numéros précédents et des articles sont gratuits et s'effectuent en ligne sur le site de la revue (<http://web.espe-guadeloupe.fr/la-recherche/revue-contextes-et-didactiques/>).

Les trois premiers numéros sont également en vente et disponibles en version papier au CRDP de Guadeloupe (Route de la Documentation, Lotissement Petit Acajou, BP 385, 97183 Abymes Cedex) et sur sa boutique en ligne : <http://www.cndp.fr/crdp-guadeloupe/>.

Types de contributions

La revue « contextes et didactiques » publie :

- des travaux de recherche empirique et des résultats originaux ;
- des revues de littérature et des approches théoriques ;
- des comptes-rendus critiques d'innovations pédagogiques ;
- des comptes-rendus d'ouvrages scientifiques ;
- des présentations de thèses, de colloques et de manifestations scientifiques.

Public concerné

La revue « *Contextes et Didactiques* » s'adresse à un large public : chercheurs, formateurs, enseignants, étudiants, professionnels de l'éducation.

Partenaires

- ESPE de l'Académie de Guadeloupe
- Université des Antilles
- Le Rectorat de l'académie de la Guadeloupe
- Centre de Recherches et de Ressources en Education et Formation (CRREF, EA-4538)

ORGANIGRAMME DE LA REVUE

Directeur de publication

- Antoine DELCROIX, Université des Antilles

Secrétariat de rédaction

- Frédéric ANCIAUX, Université des Antilles
- Béatrice JEANNOT-FOURCAUD, Université des Antilles

Comité de rédaction

- Jacqueline BOURDEAU, Télé université, Québec (TELUQ), Canada
- Thomas FORISSIER, Université des Antilles
- Laurent GAJO, Université de Genève, Suisse
- Muriel MOLINIÉ, Université Sorbonne Nouvelle, Paris 3
- Marie-Paule POGGI, Université des Antilles

Comité scientifique

- Michel ACACIA, Université d'État d'Haïti, Haïti
- Sophie ALBY, Université de la Guyane
- Christian ALIN, Université de Lyon 1
- Philippe BILAS, Université des Antilles
- Régis BLACHE, Université des Antilles
- Christos CLAIRIS, Université de Paris 5
- Denis COSTAOUEC, Université de Paris 5
- Jean-Baptiste COYOS, Université de Bordeaux 3 et de Pau
- Robert DAMOISEAU, Université des Antilles
- Jacques DUMONT, Université des Antilles
- Colette FEUILLARD, Université de Paris 5
- Diane GERIN-LAJOIE, Université de Toronto, Canada
- Françoise GUERIN, Université de Paris 4
- Marie-Christine HAZAËL-MASSIEUX, Université d'Aix-en-Provence
- Christine HELOT, Université de Strasbourg
- Élisabeth ISSAIEVA, Université des Antilles
- Gabriel LANGOUËT, Université de Paris 5
- Joël LEBEAUME, Université de Paris 5
- Mylène LEBON-EYQUEM, Université de la Réunion
- Malory LECLERE, Université Sorbonne Nouvelle, Paris 3
- Danièle MANESSE, Université de Paris 3
- Jean-Louis MARTINAND, École Normale Supérieure
- Yves MAZABRAUD, Université des Antilles
- Alain MERCIER, Institut Français de l'Éducation
- Sébastien RUFFIÉ, Université des Antilles
- Jean-Pierre SAINTON, Université des Antilles
- Jean-Claude SALLABERRY, Université Montesquieu-Bordeaux 4
- Laurence SIMONNEAUX, École Nationale de Formation Agronomique
- Jocelyne TROUILLOT-LÉVY, Université Caraïbe, Haïti
- Frédéric TUPIN, Université de la Réunion

RECOMMANDATIONS AUX AUTEURS

La revue « *Contextes et Didactiques* » est publiée en français, mais des contributions d'intérêt majeur peuvent être acceptées en anglais, espagnol et créole. Les soumissions se font par voie électronique à l'adresse du coordonateur du numéro (cf. numéros en préparation) et à celle de la revue : contextes.didactiques@espe-guadeloupe.fr

Les manuscrits proposés doivent être originaux et ne pas être soumis simultanément à une autre revue. Une feuille de style est disponible et téléchargeable sur le site de la revue (<http://web.espe-guadeloupe.fr/la-recherche/revue-contextes-et-didactiques/>) afin de mettre les contributions aux normes rédactionnelles. Le texte ne doit pas dépasser 30 000 signes (espaces compris) et comporter les éléments suivants :

- un titre,
- la liste des auteurs avec leurs noms et prénoms en entier,
- leur institution de rattachement, l'adresse de correspondance, et leur adresse électronique,
- un résumé de 250 mots environ en français et en anglais indiquant les éléments essentiels du travail,
- une liste de 5 mots clés en français et en anglais.

La forme de présentation préférée est un fichier sous format .doc, .docx, .txt, .rtf, .odt envoyé en tant que pièce jointe dans un courriel. Dès leur réception, les articles sont confiés pour évaluation anonyme à deux experts du comité scientifique de la revue ou choisi par le comité de rédaction en fonction de ses compétences. Les deux expertises sont renvoyées à l'auteur dans les deux mois suivant la réception de l'article. Dans le cas de deux évaluations contradictoires, le comité de rédaction se réserve le droit de soumettre le manuscrit à un troisième expert. Une fois l'article accepté pour publication, un formulaire de transfert de droits est envoyé à l'auteur. Ce formulaire est à signer et à renvoyer au comité de rédaction.

La correspondance s'effectue avec l'auteur cité en premier. Chaque présentation doit être accompagnée d'une lettre de l'auteur mentionnant que :

- le manuscrit n'est pas et ne sera pas soumis à une autre publication ;
- que tous les cosignataires de l'article ont lu et approuvé l'article présenté.

La liste des références bibliographiques est organisée par ordre alphabétique en fin d'article. Toutes les références citées dans le corps du texte doivent nécessairement y figurer. La présentation des références bibliographiques doit suivre les normes APA (version française).

Des notes de bas de page sont acceptées, ainsi que des figures, des tableaux et des images.

Les auteurs qui soumettent leurs articles à la revue « *Contextes et Didactiques* » acceptent de publier gratuitement leurs articles. Ils cèdent, à titre gratuit et non exclusif, les droits de diffusion papier et électronique qui permettent à la revue d'assurer la diffusion libre et gratuite de leurs articles, de même que la vente sur tous les supports et formats, en conformité avec la licence CC BY-NC (<https://creativecommons.org/licenses/by-nc/4.0/deed.fr>).

Les auteurs conservent le droit de diffuser, selon les modalités de leur choix, les articles publiés dans la revue à la seule condition, comme l'exige la licence Creative Commons concédée, qu'apparaissent les références bibliographiques complètes, mentionnant la revue. Les auteurs peuvent notamment déposer la version publiée, en texte intégral, sur l'entrepôt universitaire de leur établissement.

SOMMAIRE

Éditorial p. 6
Elisabeth ISSAIEVA MOUBARAK NAHRA et Lucie MOTTIER LOPEZ

**D'un référentiel d'évaluation fixe à une co-constitution référentielle dynamique,
ce que nous apprend le jugement situé de l'enseignant** p. 12
Lucie MOTTIER LOPEZ et Lionel DECHAMBOUX

**« Formative, on est vraiment dans cette perspective-là » : interventions et intentions
d'une formatrice universitaire dans l'évaluation de stages** p. 30
Kristine BALSLEV, Anne PERREARD VITE et Edyta TOMINSKA

**Faire vivre une expérience d'évaluation mutuelle en formation initiale :
quelle transposition vers les pratiques d'enseignement projetées ?** p. 46
Christelle GOFFIN et Annick FAGNANT

**Les raisons du faible usage des résultats d'évaluation externe par les enseignants.
Étude croisée dans trois contextes éducatifs** p. 60
Gonzague YERLY

**Analyse sociodidactique des évaluations internationales inspirées du Cadre
européen commun de référence pour les langues** p. 72
M'hand AMMOUDEN

**Pierre angulaire de l'enseignement des langues étrangères dans un
environnement d'apprentissage virtuel : l'évaluation** p. 84
Ana Maria de JESUS FERREIRA NOBRE

Rubrique VARIA

**Contextualisations didactiques et didactologiques. Suite du débat. Réponse à l'article
de Philippe Blanchet dans le n°7 de la revue « Contextes et Didactiques »** p. 96
Véronique CASTELLOTTI, Marc DEBONO et Isabelle PIEROZAK

Éditorial

Elisabeth ISSAIEVA MOUBARAK NAHRA¹ et Lucie MOTTIER LOPEZ²

¹ CRREF (EA 4538), Université des Antilles

² Université de Genève, Faculté de psychologie et des sciences de l'éducation

Dans le champ de la recherche en évaluation, il est largement admis que toute pratique évaluative est située dans un contexte donné et qu'il est essentiel d'en tenir compte lorsque l'on envisage l'étude de l'évaluation comme objet de recherche, outil méthodologique ou pratique pédagogique (De Ketele, 1993, 2006 ; Figari et Achouche, 2001 ; Mottier Lopez, 2009 ; Stufflebeam et Shinkfliend, 1985). Si la prise en compte du contexte est reconnue comme importante par la communauté scientifique de la recherche en évaluation, la définition proposée du contexte n'est pas toujours claire et sa conception est loin d'être unique et partagée. De manière générale, on peut différencier au moins deux conceptions, plus exactement deux modèles du contexte dans le domaine de l'évaluation : le modèle hiérarchique et le modèle interactionniste.

Le modèle hiérarchique, proposé par De Ketele (2006), est compris comme une pluralité de contextes pouvant être située sur trois niveaux :

- un niveau micro-contextuel comprenant les caractéristiques de l'enseignement et de l'apprentissage en classe ou en groupe de formation ;
- un niveau méso-contextuel se situant au plan institutionnel (école, établissement, institut de formation) ;
- un niveau macro-contextuel renvoyant au système éducatif ou de formation.

Dans ce modèle hiérarchique, les niveaux contextuels sont envisagés comme s'emboîtant les uns dans les autres. Les relations entre eux sont souvent peu conceptualisées et opérationnalisées ou alors dans un mouvement plutôt descendant, le macro-contextuel sur le méso ou micro-contextuel. Par exemple, il est examiné dans quelle mesure les résultats des épreuves externes standardisées infléchissent les pratiques évaluatives en classe (par exemple, Mons et Pons, 2013 ; Yerly, 2014) ou encore quelles sont les perceptions des enseignants à la lumière des contextes culturels et des politiques évaluatives (par exemple, Brown, Lake et Matters, 2009 ; Issaieva, Yerly, Petkova, Marbaise et Crahay, 2015). Peu de travaux existent par contre sur l'influence inverse, à savoir des pratiques évaluatives en classe sur les décisions prises au niveau des établissements (niveau méso) et/ou sur la régulation du système éducatif ou de formation (niveau macro), peut-être parce que la conception dominante reste un pilotage par le haut.

Le modèle interactionniste vise à dépasser cette conception hiérarchique entre les niveaux de contexte. Gallego, Cole et *the Laboratory of comparative human cognition* (2001) incitent, par exemple, à concevoir les différents contextes comme constitutifs d'un processus dynamique d'influences réciproques. Comme le précise Mottier Lopez (2008 : 286), « *une microculture de classe se constituerait en interaction avec d'autres cultures plus larges et contribuerait, elle aussi, à construire et à assurer la pérennité de ces cultures* ». Pour ce qui concerne

L'évaluation des apprentissages des élèves, des travaux empiriques adoptant cette perspective interactionniste sont encore isolés, mais semblent prometteurs. Ainsi, dans une analyse de l'élaboration d'un dispositif d'évaluation par les formateurs, dans un institut de formation, Mottier Lopez et Van Nieuwenhoven (2009) mettent en évidence que le contexte institutionnel (niveau méso) n'est pas imposé d'en haut (mouvement *top-down*). Celui-ci s'élabore en interaction et réajustement avec le niveau micro-contextuel, au travers notamment, des modalités de concertation entre formateurs, et ceci sur la base de leurs expériences en situation avec les étudiants. Dans une autre recherche, Mottier Lopez (2013) examine la façon dont des enseignants négocient des cultures d'évaluation dans des dispositifs de modération sociale, prenant en considération à la fois des référents institués et standardisés et des référents informels ancrés dans des savoirs d'expérience.

L'objectif scientifique de ce numéro 9 de la revue « *Contextes et Didactiques* » est de poursuivre la réflexion et l'analyse de l'évaluation à la lumière des contextes et de leurs interactions, dans divers ordres d'enseignement et de formation. Ce numéro propose de questionner également l'impact des contextes sur l'évaluation quand celle-ci recueille des informations à partir des apprentissages (au sens large, incluant des compétences) et acquis des élèves, ou d'étudiants et stagiaires en formation universitaire ou professionnelle.

Ce numéro comporte des études empiriques mobilisant des dispositifs méthodologiques complexes divers (entretiens, observations, analyses de productions), mais aussi des analyses théoriques et empiriques ciblant la conceptualisation de la notion de contextes en évaluation selon différentes définitions possibles. Au regard des questionnements posés et des problématiques traitées, les différents articles qui composent ce numéro apportent des développements intéressants et proposent des pistes de réflexion qui visent à s'inscrire dans un ou plusieurs des axes suivants :

- la définition des contextes en jeu et leur relation avec la question évaluative (comprise comme une pratique évaluative, réelle ou déclarée) ;
- la modélisation des relations et des interactions conçues entre les contextes concernés du point de vue de l'évaluation des apprentissages et acquis des élèves, des étudiants et des stagiaires ;
- la problématisation et l'interrogation des éventuelles tensions et transformations liées à la relation ou interaction entre contextes ;
- l'identification de l'impact des différents contextes au niveau macro (politiques éducatives, pilotage d'un système), méso (établissements, équipes disciplinaires, par exemple) et micro (classe, discipline, contenu, acteurs) sur l'évaluation et inversement ;
- la compréhension des mécanismes pouvant expliquer l'impact réciproque entre contextes au regard d'enjeux évaluatifs.

L'article de Lucie Mottier Lopez et Lionel Dechamboux intitulé « D'un référentiel d'évaluation fixe à une co-constitution référentielle dynamique, ce que nous apprend le jugement situé de l'enseignant » propose une modélisation théorique de la référentialisation dans des pratiques d'évaluation certificative réelles d'enseignants de l'école primaire genevoise, en français et en mathématiques. À partir de l'hypothèse d'une double relation dialectique entre le couple *réfèrent-référé* et les *situations*, les auteurs tentent, non seulement, d'expliquer le processus sous-jacent des interactions des contextes évaluatifs pluriels

convoqués (en termes de dynamiques co-constituantes médiatisées par l'activité interprétative de l'enseignant), mais aussi d'identifier les éléments qui les composent et qui sont réellement mobilisés dans l'agir évaluatif. Ce faisant, l'enjeu scientifique vise aussi à mieux appréhender les conditions des interactions des contextes co-constitutifs de la pratique évaluative.

L'article de Kristine Balslev, Anne Perréard Vité et Edyta Tominska, intitulé « "Formative, on est vraiment dans cette perspective-là" : interventions et intentions d'une formatrice universitaire dans l'évaluation de stages » s'interroge sur la mise en œuvre de l'évaluation formative lors d'un entretien tripartite de stage en responsabilité dans la formation initiale à l'enseignement primaire à Genève, en Suisse. En partant du postulat que les pratiques réelles prennent forme dans un mouvement double, entre des contextes prédéfinis et émergents dans l'interaction, les auteurs interrogent l'évaluation formative du stage en termes de réseaux et d'influences réciproques. Les résultats révèlent l'importance des rôles endossés par les trois interactants (enseignante stagiaire, tutrice de terrain et formatrice universitaire), ainsi que des dimensions relationnelles informelles engagées pendant l'entretien.

L'évaluation formative en contexte constitue aussi l'objet de l'article proposé par Christelle Goffin et Annick Fagnant et intitulé « Faire vivre une expérience d'évaluation mutuelle en formation initiale : quelle transposition vers les pratiques d'enseignement projetées ? ». En partant du constat d'un écart entre le prescrit en matière d'évaluation formative et les pratiques réellement mises en classe, en Belgique francophone, les auteurs considèrent qu'un dispositif d'expérience d'évaluation mutuelle, en formation initiale, permettra aux futurs enseignants de comprendre ses enjeux et d'envisager des transferts possibles vers le terrain. En adoptant cette perspective, les chercheurs distinguent quatre niveaux de contextes et étudient plus particulièrement l'interaction entre les contextes de la formation et des futures pratiques professionnelles.

Si dans les articles précédents, des interactions entre des contextes évaluatifs sont constatées, c'est moins le cas dans la recherche réalisée par Gonzague Yerly et intitulée « Les raisons du faible usage des résultats d'évaluation externe par les enseignants. Étude croisée dans trois contextes éducatifs ». Ainsi, l'auteur analyse le regard de 41 enseignants quant à leur dispositif et à leur politique d'évaluation externe au Québec, en Ontario et en Suisse. Malgré les différences importantes entre les dispositifs et les politiques d'évaluation externe dans les trois contextes, les enseignants accordent peu de validité à ces données externes et ceci pour des raisons proches. L'auteur discute ces résultats en termes de dissonances qui sépareraient les approches d'évaluation externe et d'évaluation interne, tant sur le plan de leurs méthodes, de leurs usages que de leurs finalités. Afin de trouver un dialogue possible entre le contexte évaluatif macro et micro, l'auteur suggère de cibler le contexte méso (l'établissement) en tant qu'espace potentiel de rencontre et de rapprochement.

L'article de M'hand Ammouden intitulé « Analyse sociodidactique des évaluations internationales inspirées du Cadre européen commun de référence pour les langues » suggère de prendre en considération la diversité des contextes et des cultures éducatives, et par extension des cultures d'évaluation, en Algérie. Les expérimentations menées examinent les niveaux de compétences proposés pour les langues, et leurs descripteurs supposés prendre en considération le contexte immédiat de la pratique langagière. Les auteurs soulignent les

difficultés de leur application dans des contextes non européens. Les auteurs concluent, d'une part, que ces difficultés découlent essentiellement de la nature des consignes et des thématiques, et d'autre part, que les contraintes qui en résultent désavantagent les publics qui ne connaissent pas la France et sa culture. La fiabilité et l'équité des évaluations sont alors mises en question.

L'article d'Ana Maria de Jesus Ferreira Nobre, intitulé « Pierre angulaire de l'enseignement des langues étrangères dans un environnement d'apprentissage virtuel : l'évaluation », envisage une redéfinition du concept d'évaluation en langues étrangères dans l'enseignement supérieur à distance au Portugal. Les auteurs proposent de passer d'une approche d'évaluation des connaissances à une approche qui vise le développement des compétences, en établissant la distinction bien connue aujourd'hui dans la littérature anglophone entre une évaluation des apprentissages (*assessment of learning*) et une évaluation pour les apprentissages (*assessment for learning*). Sans être directement théorisée par les auteurs, la relation entre évaluation et contexte est ici abordée du point de vue des outils pédagogiques et des ressources notamment sous forme d'informations et fichiers multimédias.

Un dernier article de synthèse sera inséré au dossier au mois de septembre 2017. Il visera à discuter les contributions de ce numéro au regard de la littérature de recherche.

En interrogeant la notion du contexte en évaluation dans une perspective plurielle, ce numéro 9 de la revue « *Contextes et Didactiques* » a pour objectif de mieux définir et comprendre les contextes au regard des enjeux évaluatifs et de modéliser plus explicitement certaines pratiques évaluatives situées. Ceci permettrait sans doute de prolonger les investigations afin de comprendre à quelles conditions certaines interactions contextuelles se produisent et pas d'autres.

Références bibliographiques

- Brown, G. T. L., Lake, R. et Matters, G. (2009). Assessment policy and practice effects on New Zealand and Queensland teachers' conceptions of teaching. *Journal of Education for Teaching*, 35(1), 61-75.
- De Ketele, J.-M. (1993). L'évaluation conjugquée en paradigmes. *Revue Française de Pédagogie*, 103, 59-80.
- De Ketele, J.-M. (2006). La recherche en évaluation : propos synthétiques et prospectifs. *Mesure et évaluation en éducation*, 29(1), 99-118.
- Figari, G. et Achouche, M. (dir.) (2001). *L'activité évaluative réinterrogée : regards scolaires et socioprofessionnels*. Bruxelles : De Boeck.
- Gallego, M. A., Cole, M. et the Laboratory of comparative human cognition. (2001). Classroom cultures and cultures in the classroom. Dans V. Richardson (dir.), *Handbook of research on teaching. Fourth edition* (p. 951-997). Washington, DC : American Educational Research Association.
- Issaieva, E., Yerly, G., Petkova, I., Marbaise, C. et Crahay, M. (2015). Conceptions et prises de positions des enseignants face à l'évaluation scolaire dans quatre systèmes éducatifs : quel est le reflet des cultures et politiques évaluatives ? Dans L. Belair et P.-F. Coen

- (dir.), *Évaluation et auto-évaluation : quels espaces de formation ?* (p. 73-98). Bruxelles : De Boeck.
- Mons, N. et Pons, X. (2013). Pourquoi n'y a-t-il pas eu de « choc PISA » en France ? Sociologie de la réception d'une enquête internationale (2001-2008). *Revue Française de Pédagogie*, 182, 9-18.
- Mottier Lopez, L. (2008). *Apprentissage situé : la microculture de classe en mathématiques*. Berne : Peter Lang.
- Mottier Lopez, L. (2009). L'évaluation en éducation : des tensions aux controverses. Dans L. Mottier Lopez et M. Crahay (dir.), *Évaluations en tension : entre la régulation des apprentissages et le pilotage des systèmes* (p. 7-25). Bruxelles : De Boeck.
- Mottier Lopez, L. (2013). Pour une culture partagée de l'évaluation. Du retour à des pratiques d'évaluation traditionnelles vers la construction d'un nouveau rapport entre la recherche, le terrain et le politique. *Revue Éducation Canada*. Disponible en ligne sur : <http://www.cea-ace.ca/education-canada/article/pour-une-culture-partagée-de-l'évaluation>
- Mottier Lopez et Van Nieuwenhoven, C. (2009). Co-élaboration d'une pratique d'évaluation entre formateurs: quelles relations de régulation au sein de la communauté de pratique ? Dans L. Mottier Lopez et M. Crahay (dir.), *Évaluations en tension : entre la régulation des apprentissages et le pilotage des systèmes* (p. 181-199). Bruxelles : De Boeck.
- Stufflebeam, D. L. et Shinkfliend, A. J. (1985). *Systematic evaluation: A self-instructional guide to theory and practice*. Boston : Kluwer-Nijhoff Publishing.
- Yerly, G. (2014). *Les effets de l'évaluation externe des acquis des élèves sur les pratiques des enseignants. Analyse du regard des enseignants du primaire*. Thèse de doctorat en Sciences de l'éducation, Université de Fribourg, Fribourg.

D'un référentiel d'évaluation fixe à une co-constitution référentielle dynamique, ce que nous apprend le jugement situé de l'enseignant

Lucie MOTTIER LOPEZ¹ et Lionel DECHAMBOUX²

¹ Université de Genève, Faculté de psychologie et des sciences de l'éducation

² Haute École Pédagogique du canton de Vaud et Université de Genève, Suisse

Résumé

Il est aujourd'hui largement admis que l'évaluation des apprentissages des élèves par les enseignants doit s'appuyer sur un référentiel d'évaluation exprimé sous forme d'objectifs et/ou de critères d'évaluation à partir de plans d'études institués. Peu de travaux ont examiné la façon dont ce référentiel est susceptible d'être transformé au fil de l'activité évaluative de l'enseignant. Quelle relation observe-t-on entre les éléments préexistants formalisés dans le référentiel d'évaluation et d'autres éléments informels qui interviennent également dans le jugement de l'enseignant ? Quel rôle les contextes jouent-ils dans la définition de ces éléments et leur convocation parfois émergente dans les processus évaluatifs ? L'objectif de cet article est de présenter quelques éléments de modélisation autour de ces questions à partir d'études empiriques préalables et de propositions de la littérature. Les études concernent des pratiques d'évaluation certificative d'enseignants du deuxième cycle (élèves de 8 à 12 ans) de l'école primaire genevoise, en français (production écrite) et en mathématiques (résolution de problèmes). À partir de l'exposition de résultats, l'article interroge le modèle théorique de l'évaluation autour de deux hypothèses de travail présentées successivement, celle d'une relation dialectique entre référent et référé dans le jugement évaluatif en acte de l'enseignant, et celle d'une part prise par des contextes pluriels ayant des fonctions particulières dans les processus d'interprétation et de prise de décision évaluatives. L'enjeu scientifique, en proposant des éléments de modélisation, est de mieux comprendre comment référents et référés agissent dans les pratiques d'évaluation situées des enseignants.

Mots-clés

Évaluation certificative, modélisation théorique, référentialisation, référentiel d'évaluation, référé, référent.

Abstract

It is now widely accepted that the evaluation of student learning by teachers should be based on an evaluation framework expressed in terms of objectives and/or evaluation criteria based on the established curricula. Little research has examined how this framework is likely to be affected in the course of teachers' evaluative activity, in particular during the grading process. What is the relation between preexisting elements, as formalized in the evaluation framework, and other informal elements that also operate during the grading process? What role do contexts play in defining these elements and in their convocation by teachers during the grading process? The objective of this article is to present some elements of modeling about these issues, based on empirical studies and proposals from the literature. The field studies concerned classroom assessment practices of teachers of primary school in the canton of Geneva (pupils aged 8 to 12) in French (writing) and in mathematics (problem solving). From the empirical results, the article questions the theoretical model of evaluation around two working hypotheses: first, a dialectical relationship between referents and referees in teachers' evaluative judgement in action; secondly, the role played by multiple contexts in the interpretation and decision-making processes underlying teachers' evaluations. The scientific challenge is to elaborate elements of a theoretical modeling in order to better understand how referents and referees act in situated evaluation practices of teachers.

Keywords

Evaluation framework, referee, referent, referentialization, summative assessment, theoretical modeling.

1. Introduction

Il semble aujourd'hui largement admis que l'évaluation des apprentissages des élèves par les enseignants doit s'appuyer sur un référentiel d'évaluation prescrit, exprimé sous forme d'objectifs et/ou de critères d'évaluation à partir de plans d'études institués. Le référentiel désigne « un construit social qui clarifie les normes d'une activité ou un sens donné à des systèmes sociaux. Il est ce par rapport à quoi un jugement ou un sens est donné à un objet ou une action » (Cros et Raisky, 2010 : 107). Toutefois, même si de nombreux travaux ont porté sur la question du référentiel de l'évaluation des apprentissages des élèves, nos analyses systématiques de la littérature montrent que peu examinent la façon dont ce référentiel est susceptible d'être spécifié et transformé au fil de l'activité évaluative tout en infléchissant en retour le jugement produit au cours de cette activité (Allal et Mottier Lopez, 2005 ; Mottier Lopez et Laveault, 2008 ; Mottier Lopez, 2015). Peu de travaux également problématisent le rapport entre les éléments préexistants formalisés dans le référentiel d'évaluation et d'autres éléments informels qui interviennent dans le jugement de l'enseignant. Il en est de même pour ce qui concerne le rôle des contextes dans la définition de ces éléments et leur convocation parfois émergente dans les processus évaluatifs. En lien avec la thématique générale de ce numéro de la revue *Contextes et Didactiques*, l'objectif de cet article est de présenter quelques éléments de modélisation autour de ces questions à partir d'études empiriques préalables et de propositions de la littérature. Rappelons que la modélisation offre au travail scientifique une « construction abstraite pour comprendre le réel » (Figari, 1994 : 50), dont les unités d'analyse et les relations entre elles sont susceptibles d'être re-questionnées tout au long du processus de recherche en vue de leur confirmation, de leur ajustement ou de leur infirmation (Mottier Lopez et Figari, 2012).

Contrairement à un article de recherche classique qui commence par présenter le cadre théorique de l'étude, nous choisissons de renverser ici la logique, en faveur d'une approche *abductive* de la recherche (Anadón et Guillemette, 2007 ; Dechamboux et Mottier Lopez, 2015). À partir de données empiriques analysées dans une perspective exploratoire, nous réinterrogeons le modèle théorique pour mieux problématiser et comprendre comment les référents agissent dans les pratiques d'évaluation des enseignants et les contextes qu'ils convoquent. Des exemples concrets seront donnés, fondés sur des analyses dont les bases seront explicitées. Nous donnerons ainsi à voir les conceptions initiales à l'origine de ces analyses pour, ensuite, à la lumière des résultats obtenus, les remettre en partie en question. Ce faisant, cet article expose une démarche dans la recherche scientifique qui est souvent passée sous silence, à savoir la re-conceptualisation théorique du phénomène étudié à partir de données empiriques analysées.

Cet article commence par présenter notre recherche et son cadre d'analyse. Puis, à partir de l'exposition de constats principaux, nous réinterrogeons le modèle théorique initial autour de deux hypothèses de travail présentées successivement. Premièrement, nous posons l'hypothèse d'une relation dialectique entre référent et référé dans le jugement évaluatif en acte de l'enseignant¹. Rappelons que le référent est « ce par rapport à quoi un jugement de valeur est porté [...] le référent appartient à l'ordre des représentations d'objectifs » (Barbier, 1985 : 294). L'ensemble des référents forme le référentiel de l'évaluation (Figari, 1994). Quant au référé, il désigne « ce qui est constaté ou appréhendé de façon immédiate » (Lesne, 1984 : 132) par

¹« Jugement en acte » ou « activité évaluative » de l'enseignant. Dans cet article, les deux termes sont utilisés dans une relation de synonymie.

l'évaluateur ; il « désigne la partie de la réalité choisie » par ce dernier (Figari, 1994 : 44). À la suite des travaux scientifiques qui ont montré que des référents autres que ceux qui sont strictement spécifiques à l'objet évalué interviennent également dans les jugements évaluatifs des enseignants (par exemple, Allal et Mottier Lopez, 2008 ; Laveault, 2008), nous faisons l'hypothèse, deuxièmement, que la relation dialectique référent-référé dans le jugement de l'enseignant convoque des contextes pluriels dans les processus d'interprétation et de prise de décision évaluatives. Nous terminons en présentant la schématisation de notre modélisation renouvelée. Puis notre conclusion revient sur les points principaux de re-conceptualisation suggérée par notre modélisation et formule quelques axes de questionnement adressés à des travaux futurs autour de l'enjeu scientifique d'étudier la façon dont les différents éléments explicites et implicites du référentiel se combinent dans le jugement évaluatif de l'enseignant.

2. Une recherche sur les jugements évaluatifs des enseignants pendant qu'ils corrigent des travaux en français et en mathématiques

La recherche que nous avons entreprise² étudie, d'une part, les jugements évaluatifs des enseignants quand ils attribuent des notes certificatives aux travaux de leurs élèves en français (production écrite) et en mathématiques (résolution de problèmes) et, d'autre part, des pratiques dites de « modération sociale » entre ces mêmes enseignants visant à construire une culture d'évaluation commune. Dans cet article, nous ne nous intéresserons qu'au premier pan de la recherche, à savoir les jugements évaluatifs en acte des enseignants. Le détail du dispositif de recherche est présenté dans Mottier Lopez, Tessaro, Dechamboux et Morales Villabona (2012) ; nous renvoyons le lecteur intéressé par les pratiques de modération sociale à l'article de Tessaro, Dechamboux, Morales Villabona et Mottier Lopez (2016).

2.1. Participants et observation de leurs pratiques de correction

La recherche a impliqué deux cohortes d'enseignants de l'école primaire genevoise : 18 enseignants des degrés 5-6 HarmoS³ (élèves de 8 à 10 ans), et 13 enseignants des degrés 7-8 HarmoS (élèves de 10 à 12 ans). Conformément à ce qui avait été défini dans le contrat de recherche, chaque enseignant nous a contactés à quatre reprises pour que nous puissions assister à la correction de contrôles écrits de son cru, qu'il avait fait passer dans sa propre classe. Les travaux des élèves ne devaient pas être préalablement lus ou corrigés. Le règlement de l'enseignement primaire demande aux enseignants d'attribuer une note sur une échelle de un à six⁴. Sur chaque épreuve, les prescriptions institutionnelles exigent que l'enseignant insère un « entête » (ou champ de l'épreuve) pour indiquer les objectifs et/ou critères d'évaluation et le barème utilisé. Les épreuves ont été élaborées par les enseignants selon leur manière habituelle de procéder portant sur des objectifs d'apprentissage communs à chaque cohorte. Pour la recherche,

² Cette recherche est soutenue par le Fonds National Suisse de la recherche (n°100013-143453/1). Nous le remercions, ainsi que les enseignants qui ont participé et les personnes qui ont contribué à la recherche en plus des deux auteurs de cet article, Fernando Morales Villabona, Sophie Serry et Walther Tessaro.

³ Ce terme fait référence au concordat suisse sur l'harmonisation de la scolarité obligatoire entre les différents cantons suisses et accompagne la nouvelle désignation des degrés de la scolarité dans les cantons ayant approuvé ce texte. Disponible en ligne sur : <http://www.edk.ch/dyn/11737.php>.

⁴ Depuis la rentrée scolaire de 2014, les enseignants du cycle moyen de l'école genevoise doivent attribuer des notes de 1,5 à 6,0 par tranche de demi-notes et non plus par notes entières comme c'était le cas au moment de notre recherche.

une centration a été faite sur le processus évaluatif de chacun de ces enseignants tant au niveau des instruments utilisés (contrôles écrits proposant des tâches complexes) que de leur activité évaluative pendant la correction de ces mêmes épreuves.

Chaque observation⁵ s'est déroulée en trois temps successifs :

1. Un entretien semi-dirigé *ante*, fondé sur un canevas portant sur la description des objectifs visés, des choix évaluatifs, de l'inscription de l'évaluation dans le temps d'enseignement, ainsi que la description de chaque élève concerné.
2. Une verbalisation de l'enseignant pendant qu'il corrigeait quatre épreuves écrites d'élèves successivement, à partir de la méthode du *think-aloud* dans une approche qualitative (Charters, 2003). Les élèves choisis par l'enseignant devaient être d'un niveau scolaire moyen à faible dans la discipline concernée. Lafortune et Allal (2008) ont montré en effet que les enseignants tendent à davantage verbaliser avec ce profil d'élèves, l'évaluation étant moins « évidente ».
3. Un entretien semi-dirigé de clôture, sollicitant des remarques de l'enseignant à propos de la correction effectuée et des résultats obtenus par les élèves.

Les trois temps de l'observation ont été intégralement transcrits. Comme indiqué plus haut, ces transcriptions, mises en regard des traces écrites produites par l'enseignant pour son évaluation (par exemple, le corrigé de l'épreuve) et par l'élève (les épreuves elles-mêmes), ont fait l'objet de premières études exploratoires, notamment en examinant les critères d'évaluation formulés par l'enseignant dans l'entête de son épreuve comparativement aux propositions des moyens didactiques utilisés (plus particulièrement en français, Mottier Lopez, Serry et Sales Cordeiro, 2016) et la façon dont ces critères, ou d'autres non formalisés, sont mobilisés dans le jugement en acte de l'enseignant (également en français, Mottier Lopez, 2017 ; Mottier Lopez, Tessaro, Dechamboux, Morales Villabona et Serry, 2015). Les études se sont poursuivies en mathématiques également (Mottier Lopez et Dechamboux, 2016 ; Mottier Lopez et Morales Villabona, 2016). Nous présentons donc dans cet article un prolongement de ces études en nous préoccupant plus particulièrement de l'activité évaluative des enseignants au cours de la correction de ces contrôles écrits. Notre intention est de montrer comment les données analysées nous amènent à actualiser notre première modélisation.

2.2. Analyse du jugement en acte de l'enseignant et premiers constats

On pourrait penser aujourd'hui qu'il n'est peut-être plus nécessaire de démontrer que l'évaluation des apprentissages des élèves par un enseignant s'éloigne radicalement d'une opération de mesure. Comme le développe Hadji (1997) « l'évaluation est un acte de confrontation entre une situation réelle et des attentes concernant cette situation » (Hadji, 1997 : 38) ; elle « consiste à mettre en relation un référé et un référent » (Hadji, 1997 : 43). C'est à partir de cette conception, de plus en plus partagée par la communauté scientifique, que nous avons analysé nos données, en établissant une distinction, à la suite de Figari et Remaud (2014), entre :

- Le *référentiel explicite préexistant*, sous forme d'objectifs d'apprentissage, critères d'évaluation, indicateurs⁶ « déjà là ». Dans ce cas, les attentes sont formelles, communicables, pouvant être communes et partagées.

⁵ A quatre reprises, deux fois en français et deux fois en mathématiques, voir Mottier Lopez, Tessaro, Dechamboux et Morales Villabona (2012).

- Le *référentiel implicite*, qui émerge lors de la verbalisation de l'enseignant pendant la correction. Par exemple, l'enseignant admet que l'élève n'explique pas toutes les phases de son raisonnement mathématique dans une résolution de problèmes si celui-ci est correct. Dans ce cas, les attentes sont informelles, susceptibles d'être différentes entre les classes, voire entre les élèves d'une même classe.

Fortes de ces catégories, les premières analyses effectuées sur notre corpus visaient à mieux comprendre la relation entre les attentes explicites et implicites du référentiel, et à explorer la manière dont elles se combinent dans le jugement évaluatif de l'enseignant. Ce faisant, il s'agissait d'appréhender le processus de référentialisation qui implique, selon Figari (1994), une « recherche de références pertinentes (c'est-à-dire à la fois universelles et particulières) » (Figari, 1994 : 48) effectuée de manière consciente par l'évaluateur, susceptible de remettre en question des critères préexistants, de définir de nouveaux critères et, toujours, de spécifier les attentes. Car comme le soulignait déjà Figari dans ses travaux de 1994, le référentiel de l'évaluateur n'est pas fixé une fois pour toutes mais il est un *construit* susceptible de se reconstruire et de se transformer au fil de l'évaluation. Pour notre part, nous ajoutons « en acte » car nous cherchons à observer plus spécialement ce processus *au cours de l'activité évaluative de l'enseignant*. Nous élargissons donc la référentialisation au jugement évaluatif « en train de se faire », au-delà de la phase consciente et méthodique qui précéderait le jugement évaluatif.

Pour ce faire, le verbatim de la verbalisation de l'enseignant en cours de correction a été catégorisé à partir des unités conceptuelles suivantes :

- Le *référent* de l'évaluation, c'est-à-dire le critère d'évaluation ou tout autre « repère » verbalisé par l'enseignant ou inféré par le chercheur à partir des verbalisations, en observant dans quelle mesure ce référent était préexistant ou non.
- Le *référé* de l'évaluation, à savoir les « signes parlants » (Hadji, 2012) sélectionnés par l'enseignant dans le texte de l'élève et énoncés par l'enseignant en cours de correction ; cette deuxième unité permet d'identifier très précisément les éléments de la « réponse » de l'élève retenus par l'enseignant pour l'évaluation.
- Les *interprétations et raisonnements* verbalisés par l'enseignant donnant à voir en partie la confrontation entre référent et référé et la façon dont l'enseignant construit du sens et de la valeur à la réponse de l'élève.
- Les *décisions évaluatives* prises par l'enseignant, en termes de points attribués à chaque critère, puis concernant la note finale.

C'est à partir de cette analyse que nous présentons, dans la suite de cet article, quelques éléments de modélisation de la « référentialisation en acte » opérant dans le jugement évaluatif de l'enseignant.

3. Eléments d'une modélisation de la « référentialisation en acte » dans le jugement évaluatif de l'enseignant, comme processus de co-constitution référentielle située

Cette partie commence par présenter quelques constats issus des analyses, à partir desquels nous exposons ensuite des éléments de notre modélisation fondée sur deux hypothèses de travail qui amènent à définir la *référentialisation en acte* comme étant non plus seulement une modification

⁶ Ces trois notions sont évidemment à distinguer entre elles.

et mise à jour du référentiel pendant l'évaluation, mais comme étant aussi un processus dynamique de co-constitution :

- entre référents et référés ;
- foncièrement située, c'est-à-dire par rapport, et avec, des contextes significatifs agissant.

La partie se termine par l'exposition schématique de la modélisation à laquelle nous aboutissons et qui, comme toute modélisation scientifique, a pour vocation d'évoluer et d'être questionnée par des nouveaux travaux scientifiques.

3.1. Constats autour de la distinction entre le référentiel de l'évaluation préexistant et les référents-référés agissant dans le jugement évaluatif de l'enseignant

Les constats issus de nos études⁷ donnent à voir la complexité du jugement évaluatif de l'enseignant, y compris quand celui-ci s'appuie sur des critères explicités dans l'entête de l'épreuve certificative, que ce soit en termes :

- De *critères de réalisation*, désignant les qualités des processus ou démarches à l'œuvre pour réaliser la production attendue (Bonniol et Genthon, 1989 ; Veslin et Velsin, 2001). Dans notre étude, cela concerne l'écriture de différents genres textuels en français en respectant diverses composantes propres à chaque genre et la résolution de problèmes additifs et multiplicatifs en suivant un certain nombre d'étapes, dans des épreuves certificatives réunissant plusieurs « exercices » ou tâches évaluatives.
- De *critères de réussite* portant sur le résultat (Veslin et Veslin, 2001), qui se traduisent dans les épreuves certificatives de nos enseignants par des points, des seuils de réussite par objectif ou par critère, par la fixation d'un seuil de réussite global matérialisant le niveau d'exigence attendu.

Dans leurs entêtes, les enseignants utilisent des formulations variées entre ce qui relève littéralement d'un objectif d'apprentissage tel qu'il est proposé dans le plan d'études officiel (*résoudre des problèmes additifs et multiplicatifs...*), d'objectifs qui sont reformulés par eux, parfois à la première personne (du point de vue de l'élève), souvent proches de la formulation d'un critère d'évaluation (*je fournis au minimum trois arguments*), parfois donnant à voir des indicateurs au sein même de la formulation (*j'organise mon texte au moyen d'organiseurs tels que : tout d'abord, par ailleurs, ensuite, pour terminer*). Les annexes 1 et 2 présentent respectivement un exemple d'entête en français, pour l'écriture d'un texte argumentatif (élèves de 9-10 ans), et un exemple d'entête en mathématiques (élèves de 9-10 ans), typiques des pratiques des enseignants de notre recherche. Pour des questions d'écriture, par la suite, nous parlerons des « critères d'évaluation » de l'entête, tout en attirant l'attention sur le fait que les formulations des enseignants combinent fréquemment des objectifs, critères et indicateurs en français, et privilégient plutôt la formulation d'objectifs en mathématiques avec des critères implicites qui apparaissent pendant la correction (par exemple : *raisonnement mathématique pertinent, résultat final correct*).

⁷ Le format de cet article ne permet pas d'exposer dans le détail l'ensemble des résultats des études que nous avons citées. Nous ne repreneons ici que les constats principaux.

Nos analyses montrent différents cas de figure dans la façon dont les enseignants s'appuient sur les critères formulés dans leurs entêtes pour évaluer les productions des élèves⁸. Ces cas montrent une grande diversité dans leur rapport avec le référentiel explicite préexistant. Nous avons tenté d'ordonner ci-dessous les principaux types de cas en fonction d'un rapport de moins en moins étroit avec les éléments préexistants du référentiel formalisé dans les entêtes :

- L'enseignant effectue une mise en relation rapide entre le critère d'évaluation et le référent retenu dans la production de l'élève ; le ou les points sont rapidement attribués sans verbalisations particulières.
- L'enseignant se pose des questions par rapport à ses attentes liées à un critère au regard d'un élément de la production de l'élève (un référent) inattendu, singulier, partiellement correct, difficile à comprendre. L'enseignant reformule alors le critère, précise le sens de celui-ci par rapport à la production particulière de l'élève, parfois aussi par rapport aux autres productions d'élèves déjà corrigées, ou aux prochaines à venir.
- Deux critères distincts sont formulés (soit dans l'entête, soit dans la verbalisation) mais, dans la confrontation à la production réelle de l'élève, l'enseignant peine à les distinguer ; les critères apparaissent finalement liés les uns aux autres dans certaines productions d'élèves. À nouveau, l'enseignant (se) précise (à lui-même) ses attentes propres à chaque critère et ce qu'il interprète de la production de l'élève notamment quand celle-ci garde une certaine opacité (par exemple, un raisonnement mathématique partiellement correct ; une structuration textuelle qui influence le contenu thématique du texte rédigé ou inversement).
- L'enseignant a formulé dans le même critère deux dimensions et envisageait préalablement d'attribuer un point par dimension. Par exemple : *mon histoire est divertissante, il y a de la fantaisie*. Au cours de son jugement en acte, l'enseignant n'évalue finalement qu'une seule des dimensions énoncées ou choisit de les regrouper. Il est dans l'impossibilité de les différencier avec des indicateurs précis ou de lever un implicite dans la production de l'élève qui lui permettrait de le faire.
- Pendant la correction, d'autres référents se mêlent aux critères explicités dans l'entête et prennent finalement le dessus, tels *le progrès de l'élève entre le brouillon initial du texte à rédiger et le texte final* ou *l'effort de l'élève*, pour attribuer les points à un critère formel portant, par exemple, sur le *respect de l'orthographe*.
- Au regard d'une production originale ou non attendue d'un élève ou de plusieurs élèves, l'enseignant convoque un référent non anticipé dans son entête, par exemple en mathématiques afin de faire une différence entre une *petite erreur de calcul* et un *raisonnement partiellement correct* ou encore quand *les calculs sont corrects dans un raisonnement incorrect*.
- Enfin, citons le cas de figure dans lequel c'est l'atteinte du seuil de réussite qui apparaît comme déterminante⁹ et qui amène l'enseignant à revoir certaines de ses décisions intermédiaires à propos de critères pour lesquels il avait hésité à attribuer ou non des points. Le référent pour l'enseignant est ici de considérer que, si ses hésitations ont une conséquence négative sur la note finale, l'hésitation doit alors être *au bénéfice de l'élève*, notamment quand il considère que l'élève *mérite* d'avoir la note minimale. Le mérite, en tant que référent émergent, oriente alors la décision évaluative de l'enseignant.

⁸ Dans cet article, nous n'aurons pas la place de discuter les différences entre le français et les mathématiques.

⁹ Note finale de l'élève en-dessous ou non de la note de réussite minimale de 4. Voir annexes.

Parmi ces différents cas de figure, non exhaustifs, nous retiendrons que la dynamique référentielle est complexe entre :

- des critères préexistants qui se spécifient, voire se modifient, au contact des productions d'élèves à évaluer ;
- des référents informels préexistants qui se combinent à des critères formalisés (par exemple, la prise en compte des efforts d'un élève, ou de son « mérite ») ;
- des référents qui émergent, inattendus, afin de pouvoir faire face à la singularité d'une ou plusieurs productions d'élève ;
- des critères qui apparaissent tout particulièrement déterminants (notamment l'atteinte du seuil de réussite) et qui influencent l'appréciation liée à d'autres critères d'évaluation.

Par rapport à notre projet de modélisation conceptuelle, notamment relativement aux référents dits préexistants, ces constats nous amènent désormais à ne parler de référents que lorsque ceux-ci sont *effectivement* mobilisés dans le jugement évaluatif en acte de l'enseignant, c'est-à-dire quand ils deviennent *ce par rapport à quoi* la production de l'élève est appréciée et qui orientent la prise de décision évaluative (Mottier Lopez, 2017). Ce choix vise à différencier ce qui relève d'attentes et de repères préexistants (sous forme d'objectifs d'apprentissage, critères d'évaluation, indicateurs) mais qui, dans le jugement évaluatif, sont susceptibles de ne pas être réellement convoqués, ou dont le sens peut être transformé, partiellement ignoré, combiné, etc.

Nos analyses montrent également que le sens des référents, qui font *système* entre eux comme on l'a vu plus haut, se construit au contact des productions réelles des élèves, plus spécifiquement par rapport aux indices ou « signes parlants » (Hadj, 2012) que l'enseignant sélectionne dans la production de l'élève à des fins d'interprétation et de prise de décision évaluatives. Les pratiques de modération sociale que nous avons observées (Tessaro, Dechamboux, Morales Villabona et Mottier Lopez, 2016), qui consistent à demander aux enseignants de confronter leurs évaluations pour construire des consensus sur la note à attribuer à l'élève (Wyatt-Smith, Klenowski et Gunn, 2010), montrent que les enseignants¹⁰ ne donnent pas sens aux mêmes éléments de la production de l'élève, notamment quand celle-ci est complexe. Autrement dit, le référé relève aussi d'une construction dans le jugement interprétatif de l'enseignant. Ceci nous amène à réserver également le concept de référé à l'unité que l'enseignant retient et interprète effectivement de la production de l'élève qu'il est train d'évaluer. Nous utilisons le concept d'indicateur pour désigner les éléments formulés et anticipés par l'enseignant avant sa correction, en termes de repères concrets à observer dans les productions des élèves.

La figure 1 synthétise ces premiers choix conceptuels, puis nous poursuivons sur nos hypothèses d'une relation dialectique entre référés et référents en lien avec des contextes agissant.

¹⁰ Il s'agit des mêmes enseignants que ceux de notre étude sur les jugements évaluatifs en actes, à propos des mêmes entêtes (telles celles présentées dans les annexes) et épreuves certificatives.

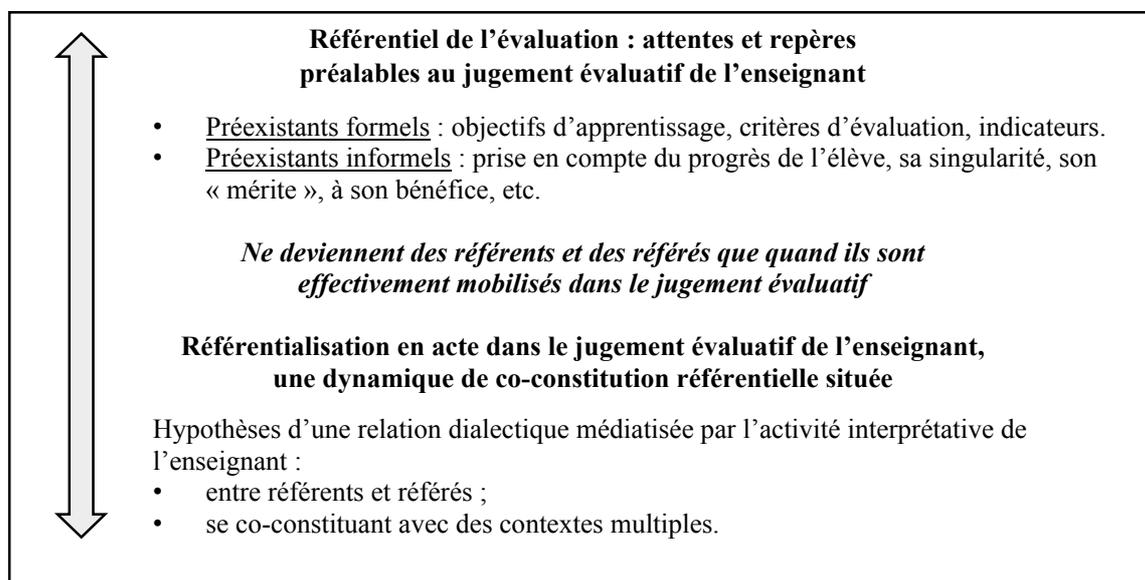


Figure 1 : Premiers éléments conceptuels fondant notre proposition de modélisation

3.2. Conceptualisation d'une relation dialectique entre référent et référé...

Dans la littérature scientifique, la relation entre référent et référé a été différemment conceptualisée, notamment en termes de « mesure d'écart » dans une approche d'évaluation-mesure, ou en termes de « comparaison », voire de « confrontation » dans une approche qualitative d'évaluation consistant à déterminer une valeur¹¹ (voir par exemple Goasdoué et Vantourout, 2017). Quelle que soit la conceptualisation, chaque terme est ici appréhendé comme une unité distincte de l'autre, dont le sens lui reste propre.

Notre hypothèse de travail est plutôt celle d'une *relation dialectique* (au sens de Lave, 1988) entre référent et référé, c'est-à-dire vus comme étant co-constitutifs l'un de l'autre¹². Concrètement cela signifie que le sens des référents et référés dans le jugement en acte de l'enseignant s'établit dans leur mise en rapport et constitutions réciproques. Il s'agit d'une relation *indirecte*, comme montré dans les travaux de Lave et successeurs (Cobb et Bowers, 1999 ; Greeno, 1997 ; Lave, 1988 ; Lave et Wenger, 1991 ; Mottier Lopez, 2008) : la relation dialectique suppose en effet des *processus demédiation* qui, dans notre étude, relèvent de l'activité interprétative et de prise de décision de l'enseignant.

Cette conceptualisation amène à considérer que l'évaluation de productions complexes s'éloigne de l'idée de rechercher des « correspondances » entre des attendus et des réponses d'élèves, comme montré également par d'autres travaux (par exemple, Vial, 2012). Dans notre recherche, les attentes préexistantes du référentiel de l'évaluation (objectifs, critères, indicateurs formulés dans les entêtes des épreuves) fournissent des catégories importantes qui structurent le jugement

¹¹ Détermination d'une valeur qui n'exclut pas l'usage de nombres (des points, des notes chiffrées) mais qui se fonde sur un raisonnement qualitatif et interprétatif qui n'est pas référencé à un modèle quantitatif de la mesure.

¹² Mais sans pour autant verser dans un syncrétisme réducteur qui amènerait à ne plus les différencier.

évaluatif de l'enseignant. Ces catégories conservent cependant une forme de variation ou d'instabilité irréductible. Contrairement à la conception d'une évaluation-mesure, cette instabilité n'est pas *de facto* considérée comme un biais ou une limite du jugement évaluatif de l'enseignant. Dans une approche située de l'évaluation, telle celle que nous développons dans nos travaux, cette instabilité est vue comme intrinsèque à la relation dialectique qui se construit entre référents et référés, et si besoin se déconstruit / se reconstruit, s'adapte à la singularité des productions des élèves. C'est dans ce processus dynamique d'interprétation et de construction de sens et de valeurs propre à l'évaluation (Vial, 2012) que l'enseignant tente de stabiliser ses référents et référés, fondés sur des éléments contextuels préexistants (*l'arena*, dans la terminologie de Lave, 1988, qui relèverait par exemple des prescriptions des plans d'études, des formulations dans les entêtes des épreuves) et émergents (*le setting*, au sens de Lave). Cette recherche de stabilisation, cette mise en relation, qui peut même s'avérer parfois créatrice, peut alors amener l'enseignant à vouloir modifier son référentiel d'évaluation dont certains éléments lui semblent inadaptés, soit de façon explicite (nouvelle formulation d'un critère ou modification du nombre de points dans l'entête de l'épreuve ou pour la prochaine épreuve ; décidés pour tous les élèves ou pour certains, comme ceux à besoins éducatifs particuliers par exemple), soit de façon implicite (c'est-à-dire en spécifiant ou transformant le sens d'un critère mais sans forcément le rédiger et le communiquer ; soit pour tous les élèves ou seulement pour certains élèves ayant un profil particulier).

3.3. ... au cœur d'un jugement évaluatif foncièrement situé

La dernière hypothèse examinée dans cet article découle directement de la conception de la *relation dialectique* développée dans une perspective de l'apprentissage situé (par exemple, Cobb et Bowers, 1999 ; Greeno, 1997 ; Lave, 1988 ; Lave et Wenger, 1991 ; Mottier Lopez, 2008). Cette perspective postule en effet un caractère foncièrement incarné aux activités sociales et individuelles (le jugement évaluatif de l'enseignant, dans notre recherche), situées dans des contextes socioculturels, dans *le hic et nunc* de la situation expérimentée (ici, la situation *immédiate* de correction des productions des élèves et ses dimensions contextuelles) et également par la convocation mentale de contextes *non immédiats* mais agissant néanmoins sur l'activité humaine (par exemple, le contexte de la classe, celui de l'établissement scolaire, le contexte familial d'un élève, etc.). L'hypothèse, qui fonde notre proposition de modélisation, est que les significations produites par le moyen du jugement évaluatif de l'enseignant sont également le produit des situations expérimentées et de leurs contextes. Réciproquement, ces situations et leurs contextes sont elles-mêmes constituées et susceptibles d'être transformées en partie par l'activité évaluative motivée par des buts, des intentions, des pratiques sociales et culturelles (Mottier Lopez, 2016).

Nos analyses fines des verbalisations des enseignants, croisées aux traces écrites produites par les élèves et par eux-mêmes, montrent que la construction de significations dans la relation dialectique référent-référent est résolument liée à des contextes qui peuvent être pluriels et qui participent à part entière à la dynamique de co-constitution référentielle. Prenons un exemple emblématique de cette co-constitution : en mathématiques, une enseignante accepte une réponse d'élève non canonique (sans formulation mathématique, mais dont le raisonnement est correct et qui débouche sur une réponse correcte). L'enseignante justifie : *donc moi je compte juste, parce que je leur ai dit qu'ils faisaient comme ils voulaient* (consignes orales transmises aux élèves dans le contexte de passation de l'épreuve et/ou contrat didactique plus général, la procédure experte n'a pas encore été institutionnalisée dans le contexte de la classe). Pour le même

problème, un autre élève de la même classe a tenté, quant à lui, une résolution mathématique en posant des algorithmes pertinents mais non maîtrisés. La réponse annoncée est incorrecte. L'élève *perd un point*. L'implicite, lié au contrat didactique et normes de la microculture de classe (Mottier Lopez, 2008), est ici qu'il vaut mieux pour les élèves proposer un raisonnement que l'on maîtrise, quitte à ne pas mobiliser certains outils mathématiques, plutôt que de proposer des procédures plus expertes mais non maîtrisées – tout au moins à ce moment précis de la progression didactique de la classe.

Prenons un autre exemple significatif en français, la co-constitution référentielle s'observant dans les deux disciplines concernées par notre recherche¹³. Au cours de sa verbalisation, une autre enseignante convoque le critère-référent *pour éviter les répétitions et les ambiguïtés on désigne les personnages de manière claire* et elle pointe plusieurs référés dans le texte de l'élève : *L'enfant de 15 ans [...] le gamin*. Dans le texte de l'élève, les reprises anaphoriques sont peu nombreuses et l'anaphore « le gamin » apparaît peu adéquate aux yeux de l'enseignante quant aux normes langagières scolaires. Cependant, l'enseignante interprète sa mise en relation référent-référent au regard du contexte de la classe et du contrat didactique négocié avec ses élèves : *on avait insisté sur le fait qu'on ne voulait pas tout le temps 'un enfant de 15 ans' [...] le but c'était qu'ils changent*. Cette interprétation se fait dans une relation de couplage avec l'activité manifestée par l'élève en rédigeant son texte : *il a fait un effort même si ce n'est pas top[...] il a fait un effort de changer*. Cet exemple montre le caractère situé de la dynamique référentielle qui sous-tend le jugement évaluatif de l'enseignant, ici pour ce qui concerne plus spécialement l'interprétation de la production de l'élève sur le plan de son usage des reprises anaphoriques :

- *pour éviter les répétitions et les ambiguïtés on désigne les personnages de manière claire* (dans l'entête du contrôle écrit, critère préexistant lié au modèle de l'objet de savoir en jeu, ici le genre de texte) ;
- *on ne voulait pas tout le temps un enfant de 15 ans* (contrat didactique négocié entre l'enseignant et ses élèves dans le contexte de la classe) ;
- *effort de l'élève* (implicite valorisé par l'enseignant relatif à l'implication active de l'élève).

Dans son entête, l'enseignante avait décidé d'attribuer un seul point au critère d'évaluation concerné. Dans son jugement en acte, elle souhaiterait disposer de deux points, ce qui lui permettrait de valoriser « l'effort » de l'élève en lui octroyant un point sur deux. Si l'attente relative à ce critère n'est pas totalement remplie par l'élève, sa production ne « mérite » cependant pas de n'obtenir aucun point. Mais cette décision évaluative exige, selon l'enseignante, une concertation avec la collègue avec qui elle élabore les épreuves certificatives : *j'appelle ma collègue après pour voir si elle me suit là-dedans ou pas*. Cette décision semble acceptable car *la demande de la directrice* (de l'établissement scolaire), *c'est implicite, mais c'est de pas couler les élèves*. Ainsi, la prise de décision dans le jugement évaluatif de l'enseignant est, elle également, liée à des contextes, ici de collaboration professionnelle entre enseignants et de culture d'établissement scolaire.

¹³ Pour les lecteurs intéressés, ce cas est présenté dans son intégralité dans Mottier Lopez (2017).

En bref, nous retiendrons que les processus d'interprétation et de prise de décision qui sous-tendent le jugement en acte de l'enseignant s'inscrivent dans des contextes pluriels qui participent pleinement à la dynamique de co-constitution référentielle qui nous intéresse dans cet article.

3.4. Essai de modélisation de la référentialisation en acte, comme processus de co-constitution référentielle située

Au regard des développements proposés dans cet article, la figure 2 synthétise les principaux éléments de modélisation de la référentialisation en acte que nous proposons, en tant que processus dynamique de co-constitution référentielle située. Elle demandera, par la suite, à s'intégrer à un modèle plus général de l'activité évaluative située.

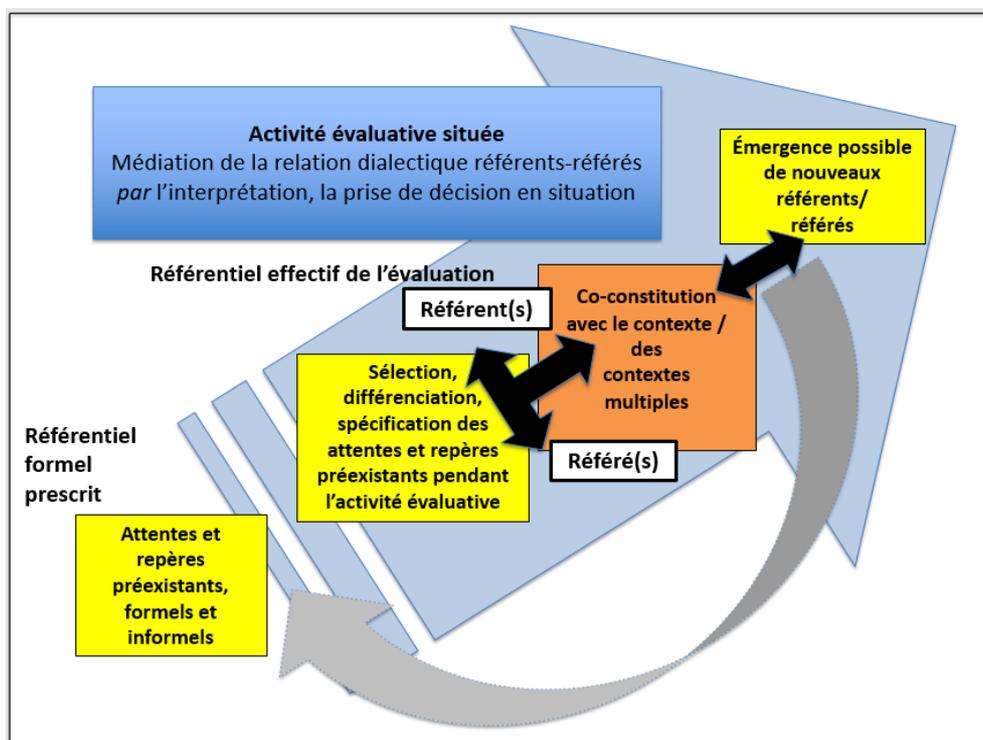


Figure 2 : Modélisation de la référentialisation au cours de l'activité évaluative

Notre modélisation conceptualise, en termes de système, une double relation dialectique (flèches noires dans la figure) entre référents et référés au cœur du jugement évaluatif, et avec les contextes possiblement pluriels de l'activité évaluative. Ce faisant, notre modélisation met radicalement en question une conception positiviste, encore courante dans la littérature contemporaine en évaluation bien que parfois diffuse, d'un rapport référent-référé qu'il suffirait de « comparer » pour, notamment, assurer l'objectivité et l'équité de l'évaluation des apprentissages des élèves. Elle argumente au contraire la nature co-construite des référents-référés et de leur mise en relation dans le jugement interprétatif de l'enseignant. Ce jugement, en tant qu'activité foncièrement située, intègre des éléments de sens propres aux contextes signifiants pour les interprétations et décisions évaluatives à prendre. Ce peut être le contexte immédiat à l'activité évaluative et sa situation, mais aussi des contextes distaux convoqués par

l'activité interprétative de l'évaluateur, liés à ses expériences et vécus en tant qu'enseignant, évaluateur, élève, etc., liés à une culture d'établissement, à un contexte familial de l'élève, etc. C'est en cela que nous proposons de parler de *multiréférentialité* du jugement évaluatif de l'enseignant. Dans des précédents travaux, nous avons déjà convoqué ce concept, notamment dans le cadre du jugement professionnel de l'enseignant en évaluation (par exemple, Mottier Lopez et Allal, 2008), mais sans l'observer aussi finement dans les processus de jugement en acte et dans le rapport constitutif avec les référents-référés et contextes convoqués.

Notre modélisation pointe également la dimension processuelle de cette mise en relation, en prenant en considération la façon dont les attentes préexistantes formelles (référentiel prescrit de l'évaluation) et informelles sont effectivement prises en compte dans le jugement évaluatif de l'enseignant (sélectionnées, parfois écartées, différenciées en fonction des productions des élèves et/ou de leurs situations possiblement singulières) et toujours spécifiées comme l'argumentent Figari et Remaud (2014). La thèse de la relation d'indissociabilité entre activités et contextes amène ici à considérer que tous référents, y compris ceux qui sont préexistants et formels, relèvent « toujours en partie d'une interprétation située qui les spécifie tout en devant garantir un niveau de généralisation qui puisse concerner l'ensemble des élèves (d'une même classe, et au-delà quand les enseignants se coordonnent entre eux) » (Mottier Lopez, 2017 : 187). Ce faisant, « il est légitime de parler aussi de référents émergents à leur propos » (Mottier Lopez, 2017 : 187). Un des éléments de notre modélisation est aussi de considérer l'émergence de nouveaux référents-référés au cours du jugement évaluatif et la façon dont ils opèrent pendant l'activité en train de se faire et dans des situations d'évaluation ultérieures – en tant qu'attentes pouvant devenir formelles ou restant dans un informel conscient par exemple – comme la prochaine production d'élève à corriger ou pour la prochaine situation d'évaluation comparable.

4. Discussion conclusive

Depuis les travaux de Barbier (1985), Figari (1994), Hadji (1997) et Lesne (1984) notamment, les concepts de référent et de référé sont régulièrement invoqués dans les travaux scientifiques en évaluation, mais sans être pleinement stabilisés par rapport aux concepts d'attente, objectif, critère, indicateur, y compris entre ce qui relève du prescrit, préexistant, « représenté », et ce qui relève de l'activité évaluative en train de se faire, du concret, du singulier, de l'agissant. A partir de cette littérature existante et de nos études exploratoires, nous avons choisi, dans cet article, de soumettre à la communauté scientifique une modélisation qui explicite la relation que nous établissons entre ces différentes unités conceptuelles. Cette modélisation demandera à se développer, s'affiner, se préciser, et évidemment à être mise en question – le propre du travail scientifique – avec des nouvelles recherches et débats qu'elle suscitera certainement.

L'analyse de nos données empiriques nous a amenés à observer que non seulement certaines des attentes explicitées dans le référentiel d'évaluation ne sont pas toujours « ce par rapport à quoi » le jugement de l'enseignant opère de manière effective, mais que d'autres attentes plus ou moins implicites peuvent également émerger au cours du jugement en acte au regard de référés inattendus par exemple. C'est pourquoi, notre modélisation propose de parler de référents et référés que lorsque ceux-ci sont *effectivement* mobilisés dans l'activité évaluative de l'enseignant. Cette distinction nous paraît particulièrement importante pour mieux appréhender les phénomènes de référentialisation qui se jouent dans la complexité de l'activité évaluative. Elle amène à poser certaines questions à propos, entre autres :

- Des attentes affichées dans un référentiel d'évaluation mais qui s'avèrent non opérantes dans le jugement évaluatif de l'enseignant. À quoi servent-elles ? Pourquoi le sont-elles ? Servent-elles, par exemple, juste à répondre à des prescriptions institutionnelles ? Ou est-ce plutôt lié à un manque de connaissances liées à l'objet de savoir en jeu dans l'évaluation et des critères qui peuvent être réellement évalués dans une épreuve ? Etc.
- Des attentes implicites mais néanmoins préexistantes et celles affichées dans l'entête. Pourquoi certaines ne peuvent-elles pas être formalisées dans le référentiel de l'évaluation, sachant aussi que tout ne peut pas être explicité (Hameline, 1979) ? Dans quelle mesure est-ce fonctionnel pour l'enseignant de garder une certaine souplesse, marge de manœuvre ? Dans quelle mesure certaines de ces attentes sont-elles indicibles, en quoi, pourquoi ? Etc.
- Des référents émergents au cours du jugement de l'enseignante. Qu'est-ce qui les fait émerger ? Dans quelle mesure l'enseignant choisit-il, ou non, de les « sortir » de l'implicite pour les faire figurer dans son référentiel d'évaluation par exemple ? À quelles conditions ? Pourquoi certains référents émergents ne peuvent-ils pas, aux yeux de l'enseignant, se formaliser dans le référentiel de l'évaluation ? Etc.

Un risque majeur d'une lecture rapide de notre conceptualisation serait de conclure que l'évaluation serait toujours relative, instable, mouvante, subjective. Nos analyses montrent que la recherche de stabilité fait aussi pleinement partie de la dynamique de co-constitution entre référents-référés et contextes agissants dont nous avons vu qu'ils étaient multiples entre interprétations et prises de décision évaluatives. Dans cette perspective, l'objectivité et l'équité de l'évaluation demandent à être également considérées comme des construits – ou des produits – de la dynamique de co-constitution référentielle. À nos yeux, un enjeu scientifique est alors de mieux appréhender les conditions de cette construction, nécessitant de toujours mieux comprendre en situation d'évaluation les systèmes référents-référés mobilisés et les contextes pluriels convoqués par elle.

Enfin, pour conclure, il est important de préciser que certains des phénomènes que nous avons pointés sont déjà bien documentés dans la littérature scientifique, tels que la prise en compte de l'effort de l'élève et de son mérite, le choix de l'enseignant de ne pas attribuer une note globale en dessous du seuil de réussite quand celle-ci ne correspond pas à sa représentation de l'élève, etc. Malgré leur pertinence, nous avons cependant choisi de ne pas citer ces travaux, souvent d'orientations psychologique ou sociologique. Dans le cadre restreint de cet article, notre objectif était en effet de rester au plus près des catégories conceptuelles propres à une *théorie de l'évaluation* (Figari et Remaud, 2014 ; Mottier Lopez et Figari, 2012), plus spécialement du point de vue des phénomènes de référentialisation en acte. Des mises en perspective entre les travaux seront évidemment à envisager pour une appréhension plus large des questions que nous avons soulevées.

Références bibliographiques

- Allal, L. et Mottier Lopez, L. (2005). Formative Assessment of Learning : A Review of Publications in French. Dans *Formative Assessment – Improving Learning in Secondary Classrooms* (p. 241-264). Paris : OECD Publication.
- Anadón, M. et Guillemette, F. (2007). La recherche qualitative est-elle nécessairement inductive ? *Recherches Qualitatives*, 5, 26-37.

- Bain, D. et Schneuwly, B. (1993). Pour une évaluation formative intégrée dans la pédagogie du français : de la nécessité et de l'utilité de modèles de référence. Dans L. Allal, D. Bain et P. Perrenoud (dir.), *Évaluation formative et didactique du français* (p. 51-79). Neuchâtel : Delachaux et Niestlé.
- Barbier, J.-M. (1985). *L'évaluation en formation*. Paris : Presses Universitaires de France.
- Bonniol, J.-J. et Genthon, M. (1989). L'évaluation et ses critères : les critères de réalisation. *Repères pour la rénovation de l'enseignement du français*, 79, 106-115.
- Charters, E. (2003). The use of think-aloud methods in qualitative research. An introduction to think-aloud methods. *Brock Education*, 12, 68-82.
- Cobb, P. et Bowers, J. (1999). Cognitive and situated learning perspectives in theory and practice. *Educational Researcher*, 3, 4-15.
- Cros, F. et Raisky, C. (2010). Référentiel. *Recherche et formation*. Mis en ligne le 01 mai 2010, consulté le 13 février 2017. URL : <http://rechercheformation.revues.org/215>
- Dechamboux, L. et Mottier Lopez, L. (2015). L'enquête de l'enseignant-évaluateur et du chercheur : l'abduction au cœur du processus et quelques raisons de les comparer. *Recherches Qualitatives*, 34 (1), 173-195.
- Figari, G. (1994). *Evaluer : quel référentiel ?* Bruxelles : De Boeck.
- Figari, G. et Remaud, D. (2014). *Méthodologie d'évaluation en éducation et formation*. Bruxelles : De Boeck.
- Gagnebin, A., Guignard, N. et Jaquet, F. (1997). *Apprentissage et enseignement des mathématiques, commentaires didactiques sur les moyens d'enseignement pour les degrés 1 à 4 de l'école primaire*. Neuchâtel : Institut de recherche et de documentation pédagogique.
- Gouasdoué, R. et Vantourout, M. (2016). Evaluations scolaires et étude du jugement des enseignants : pour une didactologie cognitive. Dans P. Detroz, M. Crahay et A. Fagnant (dir.), *L'évaluation à la lumière des contextes et des disciplines*. (p. 141-168). Bruxelles : De Boeck.
- Greeno, J. G. (1997). On claims that answer the wrong questions. *Educational Researcher*, 26(1), 5-17.
- Hadji, C. (1997). *L'évaluation démystifiée : mettre l'évaluation scolaire au service des apprentissages*. Paris : ESF.
- Hadji, C. (2012). *Faut-il avoir peur de l'évaluation ?* Bruxelles : De Boeck.
- Hameline, D. (1979). *Les objectifs pédagogiques en formation initiale et continue*. Paris : ESF éditions.
- Lafortune, L. et Allal, L. (dir.). (2008). *Jugement professionnel en évaluation : pratiques enseignantes au Québec et à Genève*. Québec : Presses de l'Université du Québec.
- Lave, J. (1988). *Cognition in practice : Mind, mathematics and culture in everyday life*. Cambridge, UK : Cambridge University Press.
- Lave, J. et Wenger, E. (1991). *Situated learning : Legitimate peripheral participation*. Cambridge, MA : Cambridge University Press.
- Laveault, D. (2008). Le jugement professionnel : foyer de tensions et de synergies nouvelles en évaluation scolaire. *Revue suisse des sciences de l'éducation*, 30 (3), 483-500.
- Lesne, M. (1984). *Lire les pratiques de formation d'adultes – essai de construction théorique à l'usage des formateurs*. Paris : EDILIG.
- Mottier Lopez, L. (2008). *Apprentissage situé : la microculture de classe en mathématiques*. Berne : Peter Lang.
- Mottier Lopez, L. (2015). *Evaluations formative et certificative des apprentissages. Enjeux pour l'enseignement*. Bruxelles : De Boeck.

- Mottier Lopez, L. (2016). La compétence à l'école pensée à partir de la perspective située de l'apprentissage. *Education et francophonie*, 44(2), 152-171. Disponible en ligne sur : http://www.acelf.ca/c/revue/pdf/EF-44-2-152_MOTTIER_LOPEZ.pdf
- Mottier Lopez, L. (2017). *Une modélisation pour appréhender la référentialisation dans l'évaluation des apprentissages des élèves*. Dans M. Crahay, A. Fagnant et P. Detroz (dir.), *L'évaluation à la lumière des contextes et des disciplines* (p. 169-192). Bruxelles : De Boeck.
- Mottier Lopez, L. et Allal, L. (2008). Le jugement professionnel en évaluation : un acte cognitif et une pratique sociale située. *Revue suisse des sciences de l'éducation*, 3, 465-482.
- Mottier Lopez L. et Dechamboux, L. (2016). *La place et le rôle de la référence dans les jugements évaluatifs des enseignants*. Communication au congrès annuel 2016 de la SSRE, Université de Lausanne, Suisse.
- Mottier Lopez, L. et Figari, G. (2012). Modèles et modélisations face à une crise du sens de l'évaluation en éducation. Dans L. Mottier Lopez et G. Figari (dir.), *Modélisations de l'évaluation en éducation : questionnements épistémologiques* (p. 7-23). Bruxelles : De Boeck.
- Mottier Lopez, L. et Laveault, D. (2008). L'évaluation des apprentissages en contexte scolaire : développements, enjeux et controverses. *Mesure et évaluation en éducation*, 3, 5-34.
- Mottier Lopez, L. et Morales Villabona, F. (2016). *Étude de la pratique évaluative de cinq enseignants quand ils corrigent des productions d'élèves en français et en mathématiques : entre hésitations, difficultés et jugement professionnel*. Communication au congrès annuel 2016 de la SSRE, Université de Lausanne, Suisse.
- Mottier Lopez, L., Serry, S. et Sales Cordeiro, G. (2016). Evaluation certificative de productions écrites d'élèves : référence au modèle didactique du genre textuel dans l'épreuve et lors de la correction de « faits divers ». *Recherche en didactique du français*, 8, 27-61.
- Mottier Lopez, L., Tessaro, W., Dechamboux, L. et Morales Villabona, F. (2012). La modération sociale : un dispositif soutenant l'émergence de savoirs négociés sur l'évaluation certificative des apprentissages des élèves. *Questions vives*, 6(18). Disponible en ligne sur : <http://questionsvives.revues.org/1235>.
- Mottier Lopez, L., Tessaro, W., Dechamboux, L., Morales Villabona, F. et Serry, S. (2015, août). *Situated judgment in summative assessment practices about mathematics and writing in primary school*. Communication in the symposium organized by L. Mottier Lopez « From teachers' assessment practices to shared professional judgment cultures », EARLI, Chypre.
- Tessaro, W., Dechamboux, L., Morales Villabona, F. et Mottier Lopez, L. (2016). La modération sociale : pour se former à l'évaluation des apprentissages. *Education permanente*, 208(3), 65-76.
- Velsin, O. et Veslin, J. (2001). Evaluation formatrice et critères de réalisation. Dans G. Figari et M. Achouche (dir.), *L'activité évaluative réinterrogée : regards scolaires et socioprofessionnels* (p. 89-101). Bruxelles : De Boeck.
- Vial, M. (2012). *Se repérer dans les modèles de l'évaluation. Méthodes, dispositifs, outils*. Bruxelles : De Boeck.
- Wyatt-Smith, C., Klenowski, V. et Gunn, S. (2010). The centrality of teachers' judgement practice in assessment: a study of standards in moderation. *Assessment in Education : Principles, Policy et Practice*, 17(1), 59-75.

Annexe 1 : Exemple d'entête, écriture d'un texte argumentatif (élèves de 9-10 ans)

Points obtenus :	23	Seuil de réussite :	20/30	Note obtenue :	5
GUIDE DE PRODUCTION ET D'EVALUATION					
	Objectifs	Elève : contrôlé	Enseignant : contrôlé		
Français 1 : texte argumentatif	1) J'adresse mon texte à un lecteur – je lui réponds en exposant le problème dès le début du texte.	<input checked="" type="checkbox"/>	2/2		
	2) J'exprime clairement mon opinion pour commencer .	<input checked="" type="checkbox"/>	2/2		
	3) Je fournis au minimum 3 arguments .	<input checked="" type="checkbox"/>	6/6		
	4) Dans mon texte apparaissent 2 avantages et 1 inconvénient .	<input checked="" type="checkbox"/>	2/2		
	5) Je rédige une conclusion pour clore mon texte afin de convaincre mon lecteur du choix à faire.	<input checked="" type="checkbox"/>	3/4		
	6) Je m'assure que mon texte soit signé .	<input checked="" type="checkbox"/>	1/1		
	7) Pas de longue phrase .	<input checked="" type="checkbox"/>	0/1		
	8) J'organise mon texte au moyen des organisateur s comme par exemple : tout d'abord, par ailleurs, ensuite, pour terminer.	<input checked="" type="checkbox"/>	4/6		
Français 2 : structuration	9) Je souligne en gris les mots pour lesquels j'ai des doutes. Je vérifie l' orthographe dans un ouvrage de référence.	<input checked="" type="checkbox"/>	1/2		
	10) Je souligne les formes verbales en rouge puis les vérifie au moyen de mon tableau de conjugaison.	<input checked="" type="checkbox"/>	1/2		
	11) Je soigne la propreté, l'écriture et la ponctuation dans mon travail.	<input checked="" type="checkbox"/>	1/2		
Barème					
6	5	4	3	2	1
30-27	26-23	22-19	18-13	12-7	6-1
Commentaires :					

Annexe 2 : Exemple d'entête en mathématiques (élèves de 9-10 ans)

Mes objectifs :		J'auto-évalue mon travail en coloriant la case : (vert : je suis sûr de moi orange : je doute)			
Points Attribués	Objectifs				
/5	• J'effectue des additions et soustractions avec retenue,	😊 😐			
/2	• Je rédige correctement un nombre qui m'est dicté oralement,	😊 😐			
/3	• Je complète une bande numérique,	😊 😐			
/12	• Je résous des problèmes en utilisant correctement les additions et les soustractions,	😊 😐			
/5	• Je me repère dans un plan.	😊 😐			
Barème :					
6	5	4	3	2	1
27-25	24-22	21-18	17-15	14-8	7-1

**« *Formative, on est vraiment dans cette perspective-là* » :
interventions et intentions d'une formatrice universitaire
dans l'évaluation de stages**

Kristine BALSLEV, Anne PERRÉARD VITÉ et Edyta TOMINSKA
Groupe de recherche TALEs – Faculté de Psychologie et des Sciences de
l'Éducation, Université de Genève, Suisse

Résumé

Cette contribution porte sur l'évaluation formative de stages en responsabilité dans la formation initiale à l'enseignement primaire. À partir des notions de contexte, de réseau (Latour, 1994) et de dispositions d'agir (Lahire, 2012), les interventions verbales d'une formatrice universitaire dans un entretien tripartite visant l'évaluation formative d'une étudiante-stagiaire seront analysées puis complétées par les propres intentions et interprétations de cette formatrice. Les analyses mettent en évidence que les interventions de la formatrice visent avant tout à actualiser le contexte préalable tout en lui donnant une coloration propre, en s'adaptant aux « débordements » du contexte et préservant les dimensions relationnelles.

Mots-clés

Formation des enseignants, évaluation formative, interactions verbales, développement professionnel, dispositions à agir.

Abstract

In this paper the role of university supervisors in triadic mentoring dialogues aiming at the formative assessment of practicums is considered as nodal. We begin by describing the institutional context of practicum assessment in initial primary teacher education in Geneva (Switzerland) and then analyse the supervisor's role in one specific triadic mentoring dialogue. Our aim is to study how she gives life to the institutional context, first by analysing her verbal interventions during the mentoring dialogue, then by giving place to her own analysis and interpretations of what is being said. Our study shows that the supervisor's interventions aim at respecting the institutional context while giving a personal touch to it, adapting to « overflows » and preserving relational aspects.

Keywords

Teacher education, formative assessment, verbal interactions, professional development, triadic mentoring dialogues.

1. Introduction

Cette contribution s'intéresse à l'évaluation des stages dans la formation à l'enseignement. Le texte de cadrage de ce numéro affirme que « toute pratique évaluative est située dans un contexte donné et il est essentiel d'en tenir compte lorsque l'on envisage l'étude de l'évaluation comme objet de recherche, outil pédagogique ou pratique pédagogique ». Le contexte dont il est question ici est celui de la formation initiale à l'enseignement primaire du canton de Genève (Suisse) et les pratiques sont celles d'une formatrice universitaire expérimentée. Affirmé ainsi, les choses paraissent simples : l'étude de deux objets distincts, d'un côté les pratiques évaluatives, de l'autre côté le contexte d'évaluation. Or, il n'est pas aisé de séparer les deux : les pratiques – en partie visibles dans des interactions – sont constituées par le contexte tout en le constituant (Grossen, 2001 ; Lahire, 2012). C'est ce double mouvement, entre prédéterminé et émergent dans l'interaction, qui nous intéresse. Un entretien de stage sera analysé sous cet angle avec comme point de mire la question : entre le prédéterminé et l'émergence dans l'interaction, quelles formes prend l'évaluation formative ? La conception du contexte de Latour (1994) et de Lahire (2012) couplée avec le terme « dispositions à agir » (Lahire, 2011, 2012) serviront de ressources pour y répondre.

L'entretien dont il est question ici, à visée formative et se déroulant en fin de formation, implique une formatrice universitaire, une formatrice de terrain et une étudiante en stage. Adoptant le vocabulaire utilisé dans le domaine de la formation des enseignants (Chaliès, Cartaut, Escalié et Durand, 2009 ; Moussay, Méard et Etienne, 2009), nous parlerons de « superviseure universitaire » (S) pour désigner la formatrice universitaire, de « tutrice » (T) pour désigner la formatrice de terrain et d'« enseignante en formation » (EF) pour désigner l'étudiante-stagiaire.

Le rôle joué par la S et les gestes professionnels qu'elle déploie nous intéressent particulièrement pour étudier à la fois le contexte créé pendant l'entretien de stage, et l'effet du contexte préexistant sur ses gestes professionnels. Fondée sur la grille d'Analyse des Discours d'Apprentissage Professionnel (ADAP – Vanhulle, 2013, 2015), l'analyse mettra en évidence une série d'interventions verbales de la part de la S, qui contribuent à renforcer ou modifier le contexte préalable. Pour mieux comprendre ses interventions, un accès à ses intentions s'impose. À ces fins, l'entretien analysé est complété par les commentaires de la S confrontée à l'enregistrement audio et à l'analyse de l'entretien effectué par les chercheuses. Ainsi cet article est à trois voix : deux voix de chercheuses et une voix de superviseure (partie 7). D'une voix à l'autre, les points de vue diffèrent : pour les chercheuses, un point de vue externe, détaché des acteurs et pour une des actrices principales de la situation, un point de vue impliqué. Du dialogue entre ces voix différemment positionnées émerge une compréhension plus complète des contraintes et des possibles sous-jacents à une situation d'évaluation formative.

Comme toute formation en alternance, le contexte genevois de la formation initiale à l'enseignement primaire est complexe (Perréard Vité et Leutenegger, 2007). Il implique plusieurs acteurs et s'ancre dans deux lieux principaux : les écoles primaires du Canton et l'institut de formation. Deux formateurs émanant de chacun de ces lieux jouent un rôle clé dans l'évaluation des EF : les S et les T. Cette contribution commence par modéliser sommairement le contexte de l'évaluation formative, visant essentiellement le développement professionnel de l'EF et non le contrôle (Mottier Lopez, 2009) des connaissances et des compétences de ce dernier. Précisons que nous considérons l'évaluation formative de manière praxéologique, en tant que pratique sociale et culturelle (Mottier Lopez, 2009) dans une situation ayant explicitement cette fonction.

2. Le contexte d'évaluation des stages dans une formation en alternance : cadre ou réseau ?

Pour situer le contexte où prennent place les interactions, plusieurs niveaux (Lahire, 2012) peuvent être considérés : celui de l'espace national et culturel ; celui de l'organisation ou de l'institution ; celui des interactions antérieures entre les partenaires de l'interaction ou des savoirs partagés par ces mêmes partenaires, ou encore, le contexte immédiat de l'interaction. Dans notre cas, le contexte institutionnel est marqué par une alternance intégrative (Vanhulle, Ronveaux et Mehran, 2007). Par conséquent, l'évaluation des stages implique la création et la mise en place de dispositifs incitant l'étudiant à articuler les apports des différents pans de sa formation et à s'adresser à des formateurs provenant du terrain professionnel et de l'université. Comment caractériser le contexte institutionnel d'évaluation formative des stages ? Qu'est-ce qui participe et agit sur ce contexte ? Les réflexions de Latour (1994) sont éclairantes pour répondre à ces questions, notamment :

On dit, sans y regarder de trop près, que nous interagissons face à face. Certes, mais l'habit que nous portons vient d'ailleurs et fut fabriqué il y a longtemps ; les mots que nous employons n'ont pas été formés pour la situation : les murs sur lesquels nous nous appuyons furent dessinés par un architecte pour un client et construits par des ouvriers, toutes personnes aujourd'hui absentes bien que leur action continue à se faire sentir. La personne même à laquelle nous nous adressons provient d'une histoire qui déborde de beaucoup le cadre de notre relation. De plus, elle n'est pas forcément présente dans l'interaction, non seulement parce qu'elle peut avoir « l'air absent », mais parce qu'il peut s'agir du masque d'une fonction définie ailleurs par d'autres. Si l'on voulait dessiner la carte spatio-temporelle de ce qui se présente dans une interaction, et si l'on voulait dresser la liste de tous ceux qui sous une forme ou une autre y participent, on ne discernerait pas un cadre bien délimité, mais un réseau très échevelé multipliant des personnes, des dates et des lieux fort divers (Latour, 1994 : 40).

Dans le dispositif genevois de la formation initiale à l'enseignement du primaire, les étudiants de dernière année font trois stages en responsabilité (d'une durée allant de 3 à 8 semaines). Ces stages ont comme but la prise de responsabilité de la mise en œuvre des enseignements et de la gestion de la classe par les EF, donc « le passage à l'action », sous l'œil avisé de leurs T. Des objectifs spécifiques à chaque stage guident le travail enseignant des EF et permettent de centrer l'accompagnement des formateurs sur un aspect choisi : didactique, organisationnel ou relationnel. En plus de constituer le cadre prescriptif de chacun des stages, précisant leur déroulement, leurs objectifs, un document d'accompagnement propose des instruments d'évaluation formative et certificative¹ facilitant la compréhension des trois acteurs quant à l'aspect crucial de la réussite ou non de la pratique (*Stages en responsabilité. Document d'accompagnement*). Ce même document précise les tâches et les rôles de chaque acteur de l'échange lors des entretiens de stage.

Les formateurs ont des rôles différents : le T est en classe avec l'EF, l'observe en action avec les élèves, le conseille, guide, co-prépare, aide, reconnaît, encourage, fait verbaliser, expliciter et réguler, au sens de « corriger » son action enseignante face aux élèves. Le rôle de S est plus distancé, il se tient à disposition des deux autres partenaires, et intervient seulement lors des rencontres tripartites, formative (en cours de stage) et certificative (après la fin du stage). Ces

¹ L'entretien certificatif ne sera pas traité ici, mais mentionnons qu'il fait office du bilan de stage, il se déroule après la fin de celui-ci. Il possède une autre structure que l'entretien formatif : il commence par un bilan auto-évaluatif de la part de l'EF. Ensuite, les deux formateurs interviennent à la base de cette présentation, en donnant leurs évaluations respectives.

rencontres se déroulent dans la classe de stage, hors des heures scolaires selon un horaire convenu entre les trois partenaires.

Pour ce qui est de la rencontre formative (objet de cet article), elle a lieu au milieu du parcours de chaque stage et dure entre 60 et 90 minutes – comme d’ailleurs la réunion certificative. En plus de la brochure, deux documents « agissent » directement sur la situation : 1) la « feuille de route » (contrat de stage) élaborée conjointement par l’EF et son T selon des consignes rédigées par les formateurs responsables des stages ; 2) l’instrument d’évaluation formative qui peut être complété tant par l’EF (auto-évaluation) que le T (évaluation) en vue du travail de co-évaluation qui est attendu d’eux. Quant au S, son rôle est surtout celui de facilitateur : d’un côté, il accompagne le déroulement partagé entre les interactants, et de l’autre, au cours de la tripartite formative essentiellement, il décèle les points à améliorer dans la pratique de l’EF en classe, en se basant sur le discours conjoint de ses deux partenaires (EF et T).

La question est alors : qui – ou qu’est-ce qui – agit sur l’évaluation formative d’un stage ? S’agit-il des éléments propres au dispositif des stages, élaborés préalablement par d’autres acteurs que ceux en présence et posant – potentiellement artificiellement – des attentes « académiques » propres aux stages ? Ou, au delà du dispositif seul, s’agit-il pour les acteurs en présence de « réinventer » le dispositif, à la recherche d’un cheminement propre aux acteurs « physiquement » en présence, porteurs d’une recherche commune : le développement d’un EF particulier ? Dans ce cas, comme le dit Latour (1994 : 43), nous serions « obligés, pour suivre une interaction, de dessiner un écheveau assez fantasque qui mêle des temps, des lieux et des acteurs hétérogènes, ce qui nous force à barbouiller sans cesse le cadre fixé ».

Ainsi, le contexte de l’entretien d’évaluation peut être qualifié de *réseau* (Latour, 1994), qui, contrairement à l’utilisation métaphorique de *cadre*, met en évidence non ce qui limite les interactions mais ce qui les influence.

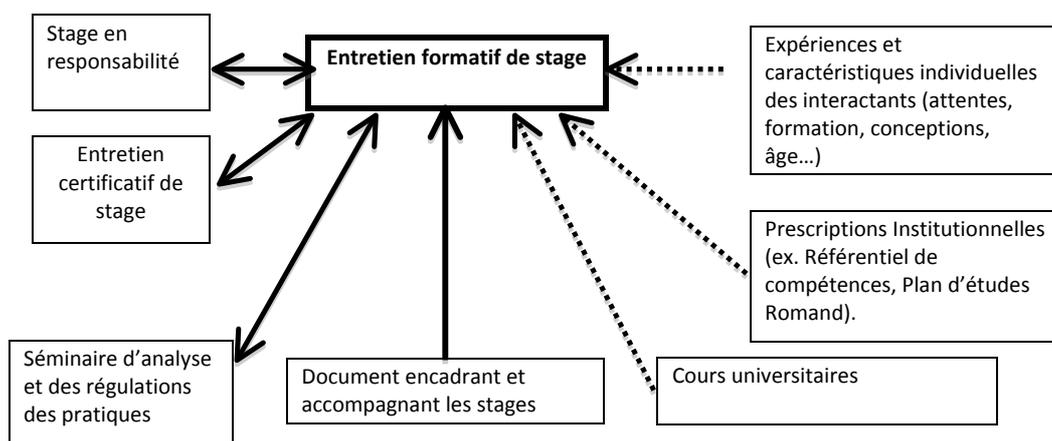


Figure 1 : Le réseau de l’évaluation formative des stages en responsabilité

À partir de notre connaissance du dispositif de stage, nous avons tenté de modéliser sous la forme d’un réseau (cf. Figure 1) le contexte genevois de l’évaluation formative des stages en responsabilité, qui met en évidence à la fois les activités de la formation, les documents et les expériences antérieures agissant sur le déroulement et le contenu de l’entretien de stage (mis

au centre du réseau), constituant à la fois des contraintes et des ressources. Ce qui influence directement et de manière interactive cet entretien est représenté par des flèches pleines et à deux sens :

- *Le stage en responsabilité* : thématique centrale de l'entretien formatif, le stage a des conséquences inévitablement sur le contenu de l'entretien et, comme l'entretien a pour vocation de réguler et d'ajuster l'enseignement de l'EF, il a un effet sur la suite du stage. Les principaux acteurs du stage sont l'EF, le T et les élèves de la classe.
- *L'entretien certificatif* fait suite à l'entretien formatif : il se déroule juste après le stage et réunit les mêmes interlocuteurs. Les principaux éléments discutés pendant l'entretien formatif réapparaissent habituellement sous une forme plus analytique dans le discours de l'EF.
- *Les séminaires d'analyse et de régulation des pratiques* regroupent une vingtaine d'enseignants en formation, et sont animés par des formateurs universitaires : ils se déroulent à l'université, en parallèle aux stages en responsabilité, et ont lieu à une fréquence moyenne de deux séances par mois. Ces séminaires « agissent » sur les entretiens de stage par leurs liens étroits avec les stages et par le fait que, dans certains cas, les S sont également responsables de ces séminaires. Ces séminaires permettent d'analyser certaines situations marquantes, de discuter des concepts ou prescriptions en jeu et d'y réfléchir en groupe.

Les influences que nous considérons comme unilatérales sont indiquées par des flèches à sens unique, avec des pointillés lorsque l'influence est, d'après nous, faible ou plus implicite :

- *Les documents encadrant et accompagnant les stages* (dont certains ont déjà été brièvement mentionnés) sont élaborés par les responsables universitaires de la formation. D'une part, une brochure nommée « Stages en responsabilité – Document d'accompagnement » destinée aux EF, S et T explicite le déroulement et les objectifs de chaque stage en responsabilité mis en discussion lors de l'entretien formatif. D'autre part, plusieurs instruments permettent de documenter le stage. Notamment, l'« instrument d'évaluation formative » sert de support pendant les entretiens. Cet instrument est divisé en quatre grandes parties : gestion de la classe et organisation de la vie quotidienne ; mise en œuvre de l'enseignement dans la classe ; développement des relations avec l'ensemble des partenaires ; implications et attitudes personnelles, sens des responsabilités. À ces parties s'ajoutent les domaines relevant des objectifs spécifiques de chaque stage.
- *Les cours universitaires* : lors de chaque entretien certificatif, il est demandé aux EF d'analyser leur pratique. Dans ce cadre, la « mobilisation d'éléments théoriques en lien avec la problématique choisie » constitue un critère d'évaluation de cette analyse. Les EF puisent généralement dans les apports de leur cursus universitaire pour mobiliser des éléments théoriques, ce qui permet d'inférer que les cours ont un impact sur les contenus abordés pendant les entretiens.
- Le métier d'enseignant est régi par des *prescriptions institutionnelles* émanant du Département de l'Instruction Publique du Canton de Genève, comme un référentiel de compétences. Le Plan d'études Romand, les moyens d'enseignement et autres prescrits institutionnels ont aussi une fonction de cadre de référence pour la profession enseignante, comme d'ailleurs le Cahier des charges de l'enseignant primaire genevois.
- Enfin, il n'est pas possible d'ignorer l'impact des *expériences et caractéristiques* inhérentes aux personnes en présence. Selon leurs expériences, leurs convictions ou leurs parcours, les interactions qui ont lieu sont d'office soumises à des influences diverses (pour les EF : propres par exemple à leur parcours scolaire, ou à une

expérience professionnelle préalable dans l'enseignement ou ailleurs ; pour les T ou les S : propres par exemple à leur parcours de formation initiale ou à des expériences professionnelles préalables dans l'enseignement ou ailleurs).

La Figure 1 permet donc de constater que les interactions en face à face sont largement influencées par tout un réseau d'éléments servant de ressources et de contraintes et relevant d'autres temps, lieux et acteurs. Par conséquent, nous ne pouvons sciemment adopter une approche uniquement interactionnelle qui considérerait les interactions en face à face comme seule échelle d'observation (Lahire, 2012). Pourtant, c'est bien à partir de ces interactions que nous étudions les pratiques en situation d'évaluation formative d'une S ; mais il s'agit davantage d'un *interactionnisme méthodologique* (Lahire, 2012) visant à identifier comment les acteurs – particulièrement l'un d'entre eux – agit sur le contexte :

Les *pratiques* considérées (...) ne se comprennent donc que si l'on étudie, d'une part, les *contraintes contextuelles* qui pèsent sur l'action (ce que le contexte exige ou sollicite de la part des acteurs) et, d'autre part, les *dispositions* socialement constituées à partir desquelles les acteurs perçoivent et se représentent la situation, et sur la base desquelles ils agissent dans cette situation (Lahire, 2012 : 24).

3. Dispositions à intervenir dans un entretien de stage

Nous avons évoqué les contraintes et ressources contextuelles de l'entretien de stage. Cet entretien est donc une interaction cadrée, contenant toutefois aussi une part informelle et ouverte. Pour sa part, Vanhulle (2016) qualifie le genre réflexif attendu dans ces entretiens comme un genre « appelé à naviguer entre informel et formel » (Vanhulle, 2016 : 7). La part informelle exige de la part des interactants une adaptation aux éléments émergents de l'interaction tout en faisant preuve d'une capacité à « maintenir le cap » pour ne pas perdre de vue l'objectif d'évaluation formative et éviter de glisser vers une conversation ordinaire. Cette part informelle offre la possibilité aux acteurs d'agir sur la situation, d'actualiser des dispositions au sens de Lahire (2012). En effet, selon cet auteur :

(...) les interactants en présence sont porteurs de dispositions et de compétences qu'ils engagent dans l'interaction, autant pour la « définir » que pour agir, des schèmes de perception, d'appréciation, d'interprétation, des compétences rhétoriques ou gestuelles, des manières de voir, de sentir et de réagir qu'ils ont constitués au cours de leurs expériences sociales (et en partie interactionnelles) passées (Lahire, 2012 : 49).

Les interactions dont il est question ici sont langagières, et essentiellement verbales, et le langage est un moyen de contextualisation par excellence. En effet, « le langage n'est pas seulement inscrit dans les situations, mais en grande partie les détermine » (Béguin et Clot, 2004 : 38). C'est pourquoi nous nous intéresserons aux interventions verbales de la S, identifiables dans ses énoncés. Ainsi, nous considérons que les dispositions se révèlent à travers des interactions verbales, que les fonctions, expériences passées et statuts des interactants ne leur confèrent pas les mêmes possibilités d'actualiser les dispositions et, enfin, que ces dispositions peuvent renforcer ou, au contraire, dénaturer la mission formative de l'entretien.

4. Les interactantes et leurs possibilités d'agir sur le contexte

Nos analyses précédentes – basées sur plus de 60 entretiens (Balslev, Dobrowolska, Mosquera et Tominska., 2015 ; Dobrowolska, Balslev, Tominska, Mosquera Roa et Vanhulle, 2016) – montrent que dans ceux-ci, le rôle de « garant du cadre » est quasiment toujours dévolu au S. Celui-ci ouvre généralement la discussion en rappelant les objectifs de

l'entretien, gère les tours de paroles, etc. Ajoutons à cela que la liberté prise quant au cadre imposé augmente souvent avec le degré d'expertise du professionnel (voir Béguin et Clot, 2004). C'est pourquoi l'analyse présentée dans cet article se focalise sur les interventions d'une S expérimentée qui nous semblent emblématiques d'une capacité à « coller » à la situation de développement professionnel en cours, combinant un respect des contraintes, une exploitation des ressources et la recherche d'un cheminement individuel, propre à l'EF concerné.

La S (Louise) est expérimentée dans trois domaines-clé : l'enseignement (c'est une ancienne enseignante de l'école primaire), la formation universitaire (elle est enseignante universitaire depuis 20 ans), l'accompagnement et l'évaluation de stages (une partie importante de son travail universitaire consiste à accompagner et évaluer des enseignants en formation). La situation présente peu de risque et on suppose que son pouvoir d'agir sur le contexte est important. Elle endosse ici le rôle de formatrice.

L'EF (Anna) est jeune. Elle se trouve dans une situation où se joue en partie son avenir professionnel, puisque l'évaluation certificative qui suivra l'évaluation formative participera à l'obtention du certificat d'enseignant, ce qui peut limiter son pouvoir d'agir sur le contexte. Elle endosse les rôles d'évaluée et de formée.

La T (Clara) est une enseignante expérimentée, mais une formatrice d'enseignants relativement novice dans le domaine des stages en responsabilité. Elle est dans une situation moins risquée que sa stagiaire, son avenir professionnel ne se joue pas dans les entretiens de stage. Par contre, elle pourrait sentir que ses pratiques de formatrice sont jugées par la S. On peut néanmoins considérer que son pouvoir d'agir est plus important que celui de son EF.

Si l'entretien formatif se caractérise par une part informelle, laissant la place à de l'imprévu, à un ajustement sur le moment, on constate que les trois interactantes n'ont pas des dispositions égales à agir sur la situation. Autrement dit, nous posons comme hypothèse que la superviseure universitaire, ayant la disposition à agir la plus forte, va marquer l'entretien davantage que les autres interactantes. Ainsi, le contexte d'évaluation est en partie prédéterminé par le dispositif, en partie créé par les (inégaux) dispositions à agir des interactantes. Suivant cette hypothèse, c'est la S, davantage que les autres interactantes, qui va donner forme à l'évaluation et faire en sorte que ce soit un réel moment de régulation et une occasion de favoriser le développement professionnel. Ainsi, notre questionnement se décline comme suit :

1. Comment la S actualise-t-elle ou modifie-t-elle le contexte préalable d'évaluation formative ?
2. Quels rôles la S prend-elle et attribue-t-elle aux autres interactantes ?
3. En regard de sa disposition à agir, en quoi les propos de la S aident-ils à comprendre ce qui se passe (non dits et/ou implicites) entre les trois interactantes et quelles sont ses intentions formatives ?

5. Outil et démarche d'analyse

Les discours oraux ou écrits de formation, tels les entretiens de stage, constituent des outils de développement professionnel. Telle est du moins l'hypothèse principale à la base de la création de l'outil ADAP (Analyse de Discours d'Apprentissage Professionnel – Vanhulle, 2013, 2015). Celui-ci permet d'identifier et de caractériser les éléments discursifs et interactionnels en lien avec le développement professionnel, en identifiant les contenus et

références dans les échanges ainsi que les manières dont ces éléments sont mis en discours. La dimension interactive et sociale de la situation de production des discours est également considérée, notamment les rapports de place (Vion, 2000) et la dynamique des interactions.

Pour caractériser l'influence réciproque entre la S et le contexte préalable, nous choisissons trois catégories d'analyse. Les *thématiques* tout d'abord, concernent les principaux sujets abordés pendant l'entretien. Ensuite, les *rôles* font référence à des rôles tels que garant du cadre, enseignant universitaire, enseignant primaire, enseignant en formation, étudiant. Nous identifions d'où parlent les interactantes, les positions sociales qu'elles endossent en repérant les marqueurs explicites du type « en tant que... », ainsi que les pronoms utilisés et les acteurs auxquels ils renvoient. Nous relevons également l'effet de l'endossement d'un rôle sur les autres participants. Les *interventions verbales* (Balslev *et al.*, 2015) sont étroitement liées à ce que disent les interlocuteurs et ont notamment les fonctions suivantes : conseiller, comparer, questionner, demander de préciser, formuler, valider, etc. Ayant un effet plus ou moins formatif sur l'EF, ces interventions permettent, dans le cas qui nous occupe ici, de cerner finement la forme que prend l'évaluation formative.

L'entretien choisi dure 80 minutes. Il est constitué de 2 141 tours de paroles, dont 619 sont émis par Anna (EF), 816 par Louise (S), et 706 par Clara (T). Dans le cadre de l'analyse qui suit, les chercheuses ont sélectionné quelques extraits de l'entretien pour pointer les thématiques abordées, les types d'interventions verbales utilisées par Louise et les rôles qu'elle joue. Nous avons choisi des extraits illustrant la manière dont la S s'y prend pour mener et réguler l'entretien autour de la préoccupation de l'EF.

L'analyse, qui suit le fil de l'entretien tripartite, ne distingue pas systématiquement les trois axes d'analyse, mais montre à quel point ils s'entremêlent les uns avec les autres. Enfin, les chercheuses résument brièvement les tendances repérées tout au long de l'entretien.

De son côté, Louise (S) a réécouté l'entretien en prenant des notes, tentant de reconstruire l'histoire de cette situation et de retrouver ses intentions du moment. Elle a ensuite lu l'analyse et, en partie 7, elle y réagit en vue de compléter les analyses des chercheuses. Ceci permet tant une acceptation interne (Mucchielli, 2002) de nos analyses, qu'une triangulation des points de vue.

6. Rôles et interventions de la superviseure lors d'une évaluation formative

Avant d'entrer dans le vif de l'analyse, précisons que Clara (T) et Louise (S) ne se connaissent pas préalablement. Elles ont donc à établir une nouvelle relation pendant l'entretien. Louise (S) et Anna (EF), par contre, se connaissent car Anna (EF) participe au séminaire d'analyse et de régulations des pratiques dispensé par Louise (S).

Durant l'entretien, les thématiques abordées concernent des questions organisationnelles et didactiques. Le stage se déroule durant une période très chargée puisque l'école prépare une fête qui implique de nombreuses répétitions et activités préparatoires, rendant délicate l'insertion d'Anna (EF) dans la classe au sens où, par moments, elle n'a que peu d'autonomie ou de « marge de manœuvre ».

L'entretien est imprégné par une préoccupation, portée par Anna (EF), et reprise par Louise (S), concernant la place d'Anna (EF) en tant que stagiaire dans la classe de Clara (T). Cette préoccupation n'est pas seulement présente durant l'entretien mais aussi pendant le stage (d'après les dires d'Anna). Elle se manifeste différemment chez Anna (EF) et Louise (S).

Pour Anna (EF), il s'agit d'une problématique rencontrée au cours du stage, à savoir qu'elle perçoit n'avoir qu'une très faible implication dans l'organisation des activités d'enseignement dans la classe. Pour Louise (S), cette préoccupation concerne d'abord la place d'Anna (EF) pendant cet entretien formatif. En effet, elle doit constamment s'assurer que c'est bien Anna (EF) qui est au centre de l'entretien, et non Clara (T), sa tutrice, qui tend à prendre souvent et longuement la parole. Dès le début de l'entretien, Louise (S) explicite le contrat de communication :

TP16 S : (...) dans un premier temps on fait l'exercice LE plus difficile pour vous et pour moi c'est « on se tait » (rires)

TP18 S : donc on va se donner beaucoup de peine de se taire (rire)

Le principe d'écoute attentive, permettant à Anna (EF) d'explicitier le déroulement de son stage, souhaité par Louise (S), implique la nécessité de « se taire » de la part des deux formatrices. Cependant, tout au long de la première partie d'entretien, Clara (T) intervient beaucoup et principalement au sujet de ses propres préoccupations de tutrice. La place d'Anna (EF) et de sa préoccupation sera longuement négociée.

Louise (S) commence par préciser sa position externe qui a besoin d'entendre évoquer des éléments de la situation de classe pour se rendre compte du fonctionnement d'Anna (EF) en stage et avec sa formatrice. Elle annonce les thématiques qui devront être abordées par Anna (EF), précise son besoin de situer ses actions dans le stage par rapport à ce qui a été annoncé dans *la feuille de route*, annonce clairement les objectifs de la rencontre et son déroulement, et donne d'abord la parole à Anna (EF). Louise (S) se positionne donc comme garante du cadre et intervient en vue d'actualiser le contexte institutionnel comme on peut le voir ci-après :

TP1 S : (...) tripartite formative [intermédiaire (...)]

TP2 EF: [hmm / hmm]

TP3 S : FORMATIVE donc on est vraiment dans cette perspective-là / (...) moi c'que je vous propose bon on a jamais eu l'occasion d'fonctionner ensemble [mais voilà / c'est une]

TP4 T: [non / non mais voilà]

TP5 S : c'est une première / pis c'est bien sûr une première avec EF aussi mais voilà / (...) c'que j'vous propose c'est d'donner la parole à EF d'abord / pour que vous puissiez nous dire / deux choses alors une chose qui va / qui est plus pour moi que pour votre formatrice parce qu'elle elle sait déjà c'que vous [avez fait]

TP6 T: [hmm]

TP7 S : jusqu'à maintenant

TP8 EF: (...)

TP9 S : mais voir / juste situer qu'est-ce que vous avez pu commencer qu'est-ce que vous avez pas pu commencer etc / enfin juste que j'aie une p'tite idée par rapport à la feuille de route

TP10 EF: d'accord

TP11 T: hmm

TP12 S : et puis ensuite le gros de de la réunion c'est vraiment comment ça va ? qu'est-ce qui va bien ? / pourquoi ça va bien ? qu'est-ce qui va moins bien ? pourquoi ça va moins bien ? / et puis qu'on essaie de de d'affiner un peu ça pour aboutir sur quelques pistes de régulation

Louise (S) propose d'emblée de parler de « qu'est-ce qui va bien », « pourquoi ça va bien » mais aussi de « qu'est-ce qui va moins bien » et « pourquoi (...) ? » en se positionnant en tant qu'évaluatrice des éléments jugés positifs ou à améliorer par Anna (EF). Ceci permet de réfléchir ensemble sur les éléments énoncés, les discuter, juger de leur importance et peut-être les déplacer dans une autre catégorie qu'initialement prévue. Elle annonce clairement son intention « d'affiner un peu ça » afin de proposer à Anna (EF) et Clara (T) des « pistes de régulation ».

Malgré l'injonction de se taire, au début de l'entretien, Clara (T) parle beaucoup de ses propres préoccupations et c'est une intervention de Louise (S) qui ramène l'entretien sur Anna (EF), en l'incitant à s'autoévaluer :

TP297S: alors maintenant / le gros du gros (EF rit) / vos points forts et vos points / moins forts ? / [qu'est-ce que vous nous diriez ?]

Suite à cette réorientation et centration sur les préoccupations énoncées par Anna (EF), Louise (S) crée un espace de discussion autour du sentiment d'Anna (EF) de ne plus savoir s'organiser dans la classe. Pour ce faire, Louise (S) reformule la préoccupation entendue, relativise la situation initiale évoquée et valorise ce qu'Anna (EF) est en train d'apprendre. Elle provoque en même temps la réaction de Clara (T) qui confirme les capacités d'Anna (EF) qu'elle a elle-même observées en classe. L'ouverture du contexte discursif est donc double : s'adressant à Anna (EF), Louise (S) vise la mise en visibilité des apprentissages professionnels en cours, en renforçant chez elle une analyse de sa pratique axée sur les conditions du stage ; s'adressant à Clara (T), Louise (S) vérifie si la situation de stage mise en discussion porte des perspectives de réussite. Louise (S) encourage Clara (T) à s'inscrire dans la discussion comme une partenaire de l'analyse de la pratique de sa stagiaire, en y amenant son point de vue, ses conseils et ses pratiques.

Louise (S) puise ensuite dans d'autres types d'interventions, en questionnant Anna (EF) de façon à ce qu'elle voie les choses différemment et qu'elle puisse comprendre son sentiment de « ne pas savoir gérer » :

TP357S: donc / quelles sont les choses qui sont des choses que vous pouvez que vous pouvez initier et d'autres choses qui sont / qui roulent / donc le plan d'semaine ça roule ça [c'est normal]

L'utilisation du verbe *pouvoir* (« les choses que vous **pouvez** initier ») permet ici à Anna (EF) de parler de ses propres initiatives en termes de « possibilité » et lui offre une opportunité de dire que, si elle n'a pas initié certaines « choses », ce n'est pas par manque de volonté ou de capacité de gestion, mais par manque de possibilité contextuelle.

Dans la suite de l'entretien, Anna (EF) parle plus, et Louise (S) continue à jouer le rôle de garante du cadre, en montrant une écoute active (« hm hm »), en validant les propos d'Anna (EF), en lui posant quelques questions de compréhension. Après avoir écouté Anna (EF), elle valorise et reformule l'un des problèmes rencontrés :

TP360S: ouais / moi j'crois que ça c'est très important c'que vous êtes en train d'dire / c'est que quand on est dans une classe où y'a pas une / enfin ça veut pas dire que ça fonctionne mal du [tout hein / mais y'a]

TP361EF: [non pas du tout]

TP362S: pas une structure très très arrêtée / on a l'sentiment d'avoir une plus grande marge de liberté

TP363EF: [voilà] / c'est peut-être ça

TP364S: et donc on s'dit "ben j'vais faire ça comme ça comme ça comme ça" / bon / quand on arrive dans la classe de quelqu'un qui est beaucoup plus structuré ou [qui a]

TP365EF: [hmm]

TP366S: des habitudes ou des rituels ou des choses comme ça / on se repose beaucoup plus facilement sur ça / puisque ça existe [déjà]

TP367EF: [hmm]

TP368S: et en même temps / c'est un piège

TP369EF: voilà / c'est ça / [j'pense que c'est ça]

TP370S: [parce que] / pour les enfants

TP371EF: hmm
TP372S: quand c'est T qui fait
TP373EF: hmm
TP374S: ça roule
TP375EF: hmm
TP376S: mais quand c'est plus T / c'est plus T
TP377EF: hmm c'est ça
TP378S: et donc c'est automatique avec T n'est [pas forcément automatique]
TP379T: [n'est plus / ouais ne l'est plus]
TP380S: avec [EF]
TP381EF: [mais] c'est c'qu'on a [constaté]
TP382T: [ouais]
TP383EF: là avec le plan d'travail
TP384S: et j'pense que y'a / donc / tout en gardant / le fonctionnement parce que
c'est probablement intéressant [d'entrer]
TP385EF: [hmm]
TP386S: dans un [fonctionnement]
TP387EF: [hmm]
TP388S: et pis de [découvrir]
TP389EF: [ah ben oui]
TP390S: comment il [marche etc]
TP391EF: [ça c'est sûr]
TP392S: donc / ben / y'a pas d'innovation mais / trouvez votre EMPREINTE
personnelle pour [que]

Louise (S) reformule en élargissant le propos d'Anna (EF) à un problème non pas spécifique à ce stage mais plutôt comme un problème général qu'« on » peut rencontrer dans de nombreuses situations de stage. En utilisant le « on » elle change légèrement de rôle puisqu'elle se positionne comme égale à Anna (EF) – voire Clara (T) – et réduit l'asymétrie des rôles. Elle semble parler en se mettant à la place d'Anna (EF), en tant qu'enseignante débutante. À la fin de l'extrait pourtant, elle reprend son rôle de formatrice universitaire, et donne un conseil à Anna (EF) : « trouvez *votre* empreinte personnelle ». Par ces mots, elle l'incite, même dans une situation a priori contextuellement « bloquée », à l'investir à sa manière.

Plus loin, Louise (S) reformule de manière plus explicite sa compréhension des raisons du problème d'Anna (EF), en formulant le problème sans porter atteinte à l'image de Clara (T) :

TP1208 S: donc c'que j'suis simplement en train de dire à EF / c'est en fait / votre tâche n'est / je PERCOIS JE PRESSENS que votre tâche est pas facile
TP1209 EF: ouais
TP1210 S: PAS parce que / y'a quelqu'un qui veut pas vous donner la place
TP1211 EF: non
TP1212 S: mais c'est parce que c'est quelqu'un qui a une grande
TP1213 EF: [hmm]
TP1214 T: [place]
TP1215 S: place/ et du coup la place est plus dure à prendre à côté d'elle

Dans la suite de l'entretien, on observe que Louise (S) navigue entre plusieurs rôles : celui de garante du cadre, de formatrice, d'ancienne enseignante et d'ancienne T. Ses expériences passées et ses fonctions présentes lui donnent accès à un répertoire de rôles variés. Elle parvient à créer un contexte dans lequel les interactantes se rapprochent les unes des autres. Les points de vue adoptés par Louise (S) et les contenus de ses énoncés amènent les interactantes à parler de ce qu'elles partagent et à estomper les rapports de force que peut engendrer ce genre d'entretien (Bullough et Draper, 2004).

À la fin de l'entretien, Louise (S) revient une fois encore sur la préoccupation d'Anna (EF) et met l'accent sur la suite du stage et sur les éléments d'apprentissage à prioriser :

TP2082 S: donc / moi j'pense que la dernière chose que vous pouvez continuer / mais ça vous avez déjà commencé donc vous pouvez continuer / c'est cette question de trouver votre PLACE

TP2083 EF: hmm hmm

TP2084 S: dans la classe / de trouver votre autonomie / à l'intérieur d'un système de contraintes contextuelles qui sont finalement dues à une fête un machin un truc etc donc vous prenez l'char en route / donc comment est-ce que vous pouvez colorer votre char aussi avec votre [couleur] ?

Par ces mots, en présence de Clara (T), Louise (S) met en valeur la préoccupation d'Anna (EF), l'encourage à oser investir sa classe de stage à sa manière, tout en respectant et honorant le cadre contextuel complexe dans lequel elle se trouve. Indirectement, ces mots enjoignent aussi Clara (T) à laisser un peu plus de place à sa stagiaire et à ne pas se sentir culpabilisée de l'accueillir dans un contexte chargé d'une actualité difficilement compatible avec l'offre d'une « carte blanche » à sa stagiaire.

En résumé, on peut considérer que, durant cet entretien tripartite formatif, Louise (S) actualise bien un contexte institutionnel prédéfini, mais en lui donnant une coloration propre consistant à mettre l'accent sur le positif, estomper l'asymétrie des rôles et la concurrence des préoccupations. Elle se sert des éléments partagés par les interactantes pour avancer dans la discussion, trouver un moyen d'accorder une place à Anna (EF), tout en préservant la face de Clara (T), en la rassurant et la valorisant. Alors que l'évaluation – même formative – peut être perçue comme menaçante pour Anna (EF) et Clara (T), ici Louise (S) fait en sorte que celle-ci soit un moment de régulations des rapports de place entre la tutrice et sa stagiaire, tout en préservant la face, au sens de Goffman (1974), de la tutrice. On voit là que la situation d'évaluation du stage d'Anna (EF) dépasse son mandat premier puisqu'il s'agit aussi de faire en sorte que Clara (T) se sente sécurisée.

Voyons maintenant comment Louise (S), de son côté, perçoit cet entretien après l'avoir revisité au travers de cette analyse et de son propre retour sur la situation. Nous passons donc à une écriture en « je », prise en charge uniquement par Louise (S), en vue de laisser de la place à un point de vue impliqué, permettant un accès à ses intentions formatives.

7. L'entretien de stage commenté par Louise (S)

« Lorsque mes collègues chercheuses m'ont invitée à faire entendre ma voix de superviseure suite à leur analyse de l'entretien formatif vécu avec Anna (EF) et Clara (T), j'ai été envahie par un sentiment d'inconfort. J'aurais voulu pouvoir oublier cet épisode dont je n'ai pas gardé un bon souvenir. À l'époque, j'étais ressortie de cette rencontre en ayant le sentiment d'avoir “tout raté”. J'avais été “noyée” dans le flot de paroles de Clara (T), la tutrice d'Anna (EF) qui, en parlant de tout et de rien [du moins je l'avais ressenti ainsi], m'avait empêchée d'être réellement à l'écoute d'Anna (EF), de faire de cette tripartite un moment formateur.

L'analyse de mes deux collègues et mon écoute exhaustive de l'enregistrement audio conservé ont été éclairantes. Elles permettent, je crois, d'accéder à un autre niveau de compréhension de ce qui s'est – parfois implicitement – joué durant ces 2 141 tours de parole.

Dans leur analyse, mes collègues relèvent la prégnance d'une double préoccupation autour de la place réservée à Anna (EF) tant dans sa classe de stage que dans l'entretien lui-même. Je réalise aujourd'hui que durant toute la première partie de l'entretien, je n'arrivais pas à mettre un nom sur cette préoccupation. D'une part, Anna (EF) décrivait l'avancement de son travail de stagiaire avec beaucoup de professionnalisme, s'inquiétant d'un sentiment de "recul" par rapport aux capacités de planification qu'elle avait démontrées dans son stage précédent. D'autre part, je sentais Clara (T) sur la défensive, cherchant "au passage" à rassurer Anna (EF), mais racontant surtout mille anecdotes disparates, allant jusqu'à interroger sa capacité d'être tutrice. Pour moi, il y avait là quelque chose de très déstabilisant et je n'arrivais pas à saisir ce qui se jouait sous mes yeux.

Face à ce malaise, ne connaissant Clara (T) ni comme enseignante ni comme tutrice et n'ayant perçu d'Anna (EF) que la figure d'une étudiante sérieuse et discrète en cours, je réalise aujourd'hui que, pendant toute une première partie de l'entretien, j'ai surtout cherché à entrer en contact avec mes deux partenaires, à la recherche d'un terrain "fructueusement formatif". Avec Clara (T), j'ai essayé d'établir une forme de complicité, notamment en faisant de nombreuses allusions à ma propre pratique de tutrice et d'enseignante et en l'invitant par exemple, comme le relève l'analyse, à me "rejoindre dans un duo de formatrices", en nous enjoignant de nous taire pour écouter Anna (EF) et lui faire une place centrale. Avec Anna (EF), j'ai effectué une sorte de "lecture diagnostique positive" de tout ce qu'elle me disait – ou ne me disait pas – en la relançant et en allant à la pêche aux indices. Avec le recul, je réalise que j'ai été prise dans une sorte de jonglage relationnel, tentant de décrypter simultanément le message de mes deux partenaires qui, sous couvert d'un semblant de triologue, cachait deux dialogues parallèles dont j'étais chaque fois la destinataire.

Je me retrouve donc parfaitement dans l'analyse de mes collègues qui, au travers des tours de parole, détectent chez moi un discours doublement adressé. Au travers des mêmes mots, je cherche d'une part à rassurer Anna (EF) et à vérifier avec Clara (T) que tout se passe bien, malgré la préoccupation affirmée de sa stagiaire.

La réponse positive à cette quête est pour moi le déclencheur de toute la fin – enfin constructive ou formatrice – de l'entretien. Dès ce moment, j'ai un objectif clair en vue : "redonner la main à une stagiaire qui ne sait / ne peut pas trouver sa place dans la classe". L'analyse de mes collègues, qui décodent ma volonté d'encourager Anna (EF) à mettre sa propre couleur dans ce qu'elle entreprend dans la classe, devient un leitmotiv. Non seulement ça, mais en prenant Clara (T) à témoin et, en lisant à haute voix la réalité complexe du stage comme une richesse (fête d'école et rituels de classe à l'image du plan de semaine) et non comme un obstacle, je les engage explicitement à profiter de cette situation unique tout en exploitant au maximum les plages d'enseignement dévolues à Anna (EF), les rassurant chacune quant à leurs compétences respectives, d'enseignante en formation et de tutrice.

A posteriori, je réalise aujourd'hui que, si Anna (EF) et moi avions en effet une préoccupation commune – sa place dans le stage –, Clara (T), elle aussi, cherchait sa place. Dans un stage dont le contexte était "chamboulé", peut-être avait-elle perdu le sens de son travail de tutrice ? Se voyant responsable de faire respecter des attentes portées par moi-même comme superviseure et représentante de l'université, elle ne s'en sentait pas ou plus capable. Insidieusement, une nouvelle dimension s'est apparemment invitée dans cet entretien : implicitement Clara (T) attendait aussi que je valide son rôle de tutrice. Et me voilà piégée dans un double contrat de formation : de la stagiaire *et* de sa tutrice. Pas étonnant que j'aie ressenti tant de confusion ».

8. Conclusion à trois voix

Nous avons modélisé le contexte de l'évaluation formative du stage sous forme de réseau, montrant que la rencontre face à face est imprégnée par de nombreux éléments provenant d'autres lieux, d'autres acteurs. L'analyse des interventions de la S montre que ce sont davantage ses propres « expériences et caractéristiques » qui lui servent de ressources pour actualiser un contexte prédéfini visant entre autres à ce que l'entretien, ainsi que le stage en responsabilité, accorde la place principale à l'EF. On constate aussi qu'un élément imprévu par le contexte institutionnel est engendré par les interventions de la T qui, préoccupée par ses propres compétences tutorales, semble voir dans cet entretien l'occasion d'en parler avec une S. Par conséquent, et contrairement à ce que nous avons supposé, dans ce cas-là, c'est la T qui, par moments, transforme le contexte préalable et non pas la S qui intervient davantage pour actualiser et rappeler le contexte institutionnel, cherchant à remettre l'EF au centre des réflexions tripartites. La S actualise le contexte institutionnel en mettant l'accent sur les aspects positifs, en endossant un registre de rôles variés possibles grâce à ses expériences professionnelles diverses. Par contre, l'EF agit peu sur le contexte, subissant les préoccupations de sa T et cherchant plutôt à « limiter les dégâts » ou justifier ce qui pourrait lui être reproché.

Heureusement dans ce cas, l'évaluation formative telle que prévue dans le dispositif étudié, laisse une part importante à l'informel et par conséquent à la liberté des formateurs. Cette liberté se décline au travers de l'expression du style propre des formateurs et rend possible un dépassement du cadre prévu. Dans l'entretien analysé, nous pouvons voir que la S met ses forces à maintenir un contexte de formation et d'évaluation pour l'EF et à éviter que le contexte se modifie trop et devienne un contexte d'évaluation des compétences de la T. L'entretien formatif repose donc sur des savoirs, des compétences, voire un jugement professionnel (Mottier Lopez et Allal, 2010) de formateurs, pouvant être développés dans le cadre d'une formation de formateurs d'enseignants.

On voit aussi que ce qui dépasse du cadre prévu est souvent de l'ordre du relationnel, du réseau humain et qu'on peut décider ou non de lui laisser une place. En effet, l'entretien analysé aurait pu se passer autrement. Il n'y avait *a priori* pas de problème majeur et, une fois que la T avait confirmé que le stage se déroulait suffisamment bien, les deux formatrices auraient pu laisser de côté tout autre objectif de formation. Or, c'est justement le caractère informel et le jugement professionnel de la S qui ont permis aux trois interactantes de rester centrées sur le problème – en partie relationnel – de la prise de place ou de la place accordée à l'EF. Convaincue, malgré les éléments objectifs de réussite du stage, sur la base d'une forme d'inconfort communicationnel – ou relationnel –, la S s'obstine à « rétablir le courant » entre ses deux partenaires, à la recherche d'une action formatrice qui puisse – notamment au travers de son « diagnostic positif » – accompagner l'EF dans le développement d'une posture professionnelle confiante.

Lorsqu'on ne se contente pas d'analyser les faits tels qu'ils apparaissent au travers des discours, mais que l'on choisit de convoquer aussi l'informel, de donner une place à l'histoire des personnes qui déborde le cadre (Latour, 1994), de nouveaux horizons s'ouvrent et montrent l'indicible et l'unicité de chaque situation. Ces horizons comportent le risque d'une dérive où l'évaluation formative pourrait disparaître au profit d'une simple « discussion de bistrot » si ce n'est par le professionnalisme et la connaissance profonde des objectifs de formation portés par la S qui permettent un retour sur l'orientation de développement professionnel poursuivie. Dans un contexte politique où la formation des tuteurs est réclamée

tant par les instances politico-institutionnelles que par les tuteurs eux-mêmes, qu'en est-il du professionnalisme des superviseurs ? En quoi sera-t-il possible, à l'avenir, de renforcer leur professionnalisme et... quel professionnalisme ? En effet, on ne peut attendre de tous les formateurs les mêmes dispositions à agir en vue de mener une réelle évaluation formative, surtout si ces dispositions ont été acquises au travers d'autres expériences professionnelles. L'essentiel nous semble être de s'approprier le principe d'une évaluation formative au service du développement professionnel de l'EF, puis de forger des dispositions à agir adaptées aux situations et aux interlocuteurs.

Références bibliographiques

- Balslev, K., Dobrowolska, D., Mosquera, S. et Tominska, E. (2015). Interventions discursives des tuteurs dans l'évolution des préoccupations des enseignants en formations. Dans K. Balslev, L. Fillietaz, S. Cartaut-Ciavaldini et I. Vinatier (dir.), *La part du langage : pratiques professionnelles en formation* (p. 281-313). Paris : L'Harmattan, collection Action et savoirs.
- Béguin, P. et Clot, Y. (2004). L'action située dans le développement de l'activité. *@ctivités*, 1(2), 27-49.
- Bullough J. R. V. et Draper, R. J. (2004). Making Sense of a failed Triad. Mentors, university supervisors, and positioning theory. *Journal of Teacher Education*, 55(5), 407-420.
- Chaliès, S., Cartaut, S., Escalié, G. et Durand, M. (2009). L'utilité du tutorat pour de jeunes enseignants : la preuve par 20 ans d'expérience. *Recherche et Formation*, 61, 85-129.
- Dobrowolska, D., Balslev, K., Tominska, E., Mosquera Roa, S. et Vanhulle, S. (2016). Discours académiques et discours professionnels : positionnement et savoirs des enseignants en formation. *Phronesis*, 5(3-4), 28-41
- Grossen, M. (2001). La notion de contexte: quelle définition pour quelle psychologie? Un essai de mise au point. Dans J.-P. Bernié (dir.), *Apprentissage, développement et significations* (p. 59-76). Bordeaux : Presses Universitaires de Bordeaux.
- Lahire, B. (2011). *L'homme pluriel. Les ressorts de l'action*. Paris : Fayard.
- Lahire, B. (2012). *Monde pluriel. Penser l'unité des sciences sociales*. Paris : Éditions du Seuil.
- Latour, B. (1994). Une sociologie sans objet ? Remarques sur l'interobjectivité. *Sociologie du travail*, 4, 587-607.
- Moussay, S., Etienne, R. et Méard, J. (2009). Le tutorat en formation initiale des enseignants: orientations récentes et perspectives méthodologiques. *Revue française de pédagogie*, 166, 59-69.
- Mottier Lopez, L. (2009). L'évaluation en éducation : des tensions aux controverses. Dans L. Mottier Lopez et M. Crahay (dir.), *Évaluations en tension : entre la régulation des apprentissages et le pilotage des systèmes* (p. 7-25). Bruxelles : De Boeck.
- Mottier Lopez, L. et Allal, L. (2010). Le jugement professionnel en évaluation: quelles triangulations méthodologiques et théoriques ? Dans L. Paquay, C. Van Nieuwenhoven, et P. Wouters (dir.), *L'évaluation, levier du développement professionnel ? Tensions, dispositifs, perspectives* (p. 237-249). Bruxelles : De Boeck.
- Mucchielli, A. (1996/2002). *Dictionnaire des méthodes qualitatives en sciences humaines et sociales*. Paris : Armand Colin.
- Perréard Vité, A. (2011) Accompagner, former, évaluer : une « triple mission » délicate pour les formateurs de stages, de terrain et universitaires. Dans F. Regnard et L. Mottier Lopez (dir.), *Former à l'enseignement musical. Pratiques et problématiques évaluatives* (p. 79-94). Paris : L'Harmattan.

- Stages en responsabilité. Document d'accompagnement.* 2011-2012. Genève : Université de Genève.
- Vanhulle, S., Merhan, F. et Ronveaux, C. (2007). Du principe d'alternance aux alternances en formation des enseignants et des adultes. Dans F. Merhan, C. Ronveaux et S. Vanhulle (dir.), *Alternances en formation* (p. 7-45). Bruxelles : De Boeck.
- Vanhulle, S. (2013). Dialoguer avec les discours d'enseignants en devenir. *Ikastaria*, 19, 37-67.
- Vanhulle, S. (2015). L'alternance : des mondes socio-discursifs en déséquilibre. Dans K. Balslev, L. Fillettaz, S. Cartaut-Ciavaldini et I. Vinatier (dir.), *Pratiques professionnelles en formation : la part du langage* (p. 249-279). Paris : L'Harmattan, collection Action et savoirs.
- Vanhulle, S. (2016). Dire et écrire l'expérience pour réguler son agir professionnel : réflexions à propos d'un « genre réflexif ». *Pratiques*, 171-172.
- Vion, R. (2000). *La communication verbale. Analyse des interactions*. Paris : Hachette.

Faire vivre une expérience d'évaluation mutuelle en formation initiale : quelle transposition vers les pratiques d'enseignement projetées ?

Christelle GOFFIN et Annick FAGNANT

Université de Liège, Belgique

Résumé

Dans une étude préalable menée en Belgique francophone (Fagnant et Goffin, à paraître), nous avons notamment montré que les futurs enseignants du secondaire supérieur (grades 10-12) témoignaient d'expériences limitées en matière d'évaluation formative : leur vécu en tant qu'élève faisait davantage état de pratiques d'auto-notation que de réelles expériences d'autoévaluation. Voulant prendre en compte le décalage entre le cadre légal qui prescrit la mise en œuvre d'évaluations formatives et la réalité des classes où ce type d'évaluation semble être peu pratiqué, nous avons tenté d'agir dès la formation initiale. Ainsi, nous avons fait vivre à ces futurs enseignants une expérience d'évaluation mutuelle en tant qu'étudiant universitaire dans un cours pluridisciplinaire tout en les faisant réfléchir à la transposition de ces acquis vers leurs pratiques de stage en tant qu'enseignant d'une discipline donnée. Cet article présente le dispositif mis en place, tout en y portant un regard critique documenté par deux sources d'informations. La première concerne les commentaires rédigés par les étudiants sur ce qu'ils ont retenu du dispositif expérimenté. La seconde porte sur les pratiques d'évaluation que les futurs enseignants ont développées ou pourraient développer en classe avec leurs élèves, en lien avec l'expérience vécue en formation. Les analyses mettent en évidence que cette expérience a pu, dans une certaine mesure, nourrir leurs processus d'apprentissage et leur développement professionnel.

Mots-clés

Évaluation mutuelle, formation d'enseignants, pratiques d'évaluation, apprentissage de l'évaluation, isomorphisme.

Abstract

In a preliminary study conducted in the French-speaking community of Belgium (Fagnant & Goffin, forthcoming), we showed that future upper secondary school teachers (grades 10-12) reported limited experiences in formative evaluation during their schooling: they mentioned more self-grading practices than actual experiences of self-assessment. In order to deal with the discrepancy between the institutional context that prescribes the implementation of formative evaluation and the reality of the classes, where this type of practice is still uncommon, we chose to act at the level of initial training. We thus proposed to the future teachers an experience of peer-to-peer assessment as academic students in a multidisciplinary course, while making them reflect on possible transpositions of this knowledge into their practice as teachers in a given discipline. This article presents the implemented device and its critical assessment, documented by two sources of information. The first concerns the comments written by the students on this experience. The second relates to the assessment practices they developed or could develop as teachers, in relation to this experience. To some extent, the study shows positive results on their learning processes and their professional development.

Keywords

Peer to peer assessment, teacher training, assessment practices, learning of assessment, isomorphism.

1. Introduction

Comme dans de nombreux pays (Ocdé, 2005), la pratique de l'évaluation formative est vivement recommandée en Belgique francophone depuis de nombreuses années, et ceci tant dans les textes officiels décrivant les attentes du système scolaire (Décret « Missions », 1997) qu'au niveau de la formation initiale d'enseignants (Paquay, 1994). Des études menées notamment en France, en Suisse romande et en Belgique francophone (Issaieva et Crahay, 2010 ; Issaieva, Yerly, Petkova, Marbaise, et Crahay, 2015 ; Luisoni et Monnard, 2015 ; Soussi, Ducrey, Ferrez, Nidegger et Viry, 2006) ont montré que les enseignants semblaient reconnaître l'intérêt des évaluations formatives pour soutenir les apprentissages des élèves, mais qu'ils étaient toutefois relativement peu enclins à mettre en œuvre ce type d'approche dans leurs classes (Soussi *et al.*, 2010 ; Younès et Gaime, 2012). D'après l'enquête Thalys 2013 (Ocdé, 2014), si l'autoévaluation est très fréquente dans certains pays comme l'Angleterre ou le Canada, elle est très peu pratiquée dans d'autres, comme en France notamment où seuls 16,5% des enseignants du premier cycle du secondaire déclarent utiliser cette modalité impliquant l'élève dans le processus d'évaluation (Cnesco, 2014). La Belgique francophone n'a pas participé à cette enquête, mais, nous basant sur une étude que nous avons menée en 2015-2016 auprès de futurs enseignants du secondaire supérieur (grades 10-12) (Fagnant et Goffin, à paraître), nous pouvons penser que peu d'autoévaluations sont mises en place dans les classes belges francophones.

Dans cette étude, nous avons, entre autres, demandé à ces futurs enseignants de se positionner face à des affirmations visant à évaluer l'intérêt qu'ils portaient aux évaluations formatives, mais aussi de décrire des expériences auxquelles ils avaient été confrontés durant leur parcours scolaire. S'ils reconnaissaient volontiers l'intérêt de pratiquer des évaluations formatives, leur vécu en tant qu'élève concernant ces évaluations (et plus particulièrement ici, l'autoévaluation et l'évaluation mutuelle) s'est révélé particulièrement limité. En effet, ils faisaient principalement état de pratiques d'auto-notation (Rey et Feyfant, 2014), relatant davantage des procédures de correction (par exemple, noter leur propre copie en s'appuyant sur la grille de critères fournie par l'enseignant) que de réelles pratiques d'autoévaluation potentiellement porteuses d'apprentissage et invitant à un véritable recul métacognitif (Campanale, 2007).

Partant de l'hypothèse selon laquelle le vécu en tant qu'élève (socialisation secondaire – Campanale, 2007 ; Crahay, Wanlin, Issaieva et Laduron, 2010 ; Vause, 2010) influence les conceptions des futurs enseignants, on peut craindre que les accords de principes formulés face aux affirmations qui leur étaient proposées traduisent une forme de « désirabilité sociale » (Butori et Parguel, 2010), mais que leurs conceptions profondes soient en réalité moins enthousiastes à l'égard de l'évaluation formative. Suite à leur vécu en tant qu'élève, on peut en effet s'attendre à ce que les futurs enseignants aient en réalité une vision de l'évaluation formative assez « classique » ou peu « élargie » (Allal et Mottier Lopez, 2005 ; Morissette, 2010) et qu'ils imaginent finalement peu comment intégrer ces pratiques dans les classes.

Face à ces divers constats et nous appuyant sur un principe d'isomorphisme ou d'homologie (Campanale, 2007 ; Vanhule, 2002), nous avons décidé de faire vivre une expérience d'évaluation mutuelle (Bourgeois et Laveault, 2015) à ces futurs enseignants, tout en les faisant réfléchir à la transposition de ces acquis vers leurs pratiques. Il s'agissait ainsi d'envisager « l'évaluation comme objet d'apprentissage et comme outil de développement professionnel » (Fagnant, Etienne, Mottier Lopez et Hindryckx, soumis).

L'objet du présent article est de faire part de cette expérience et de quelques résultats tirés de l'évaluation de celle-ci en termes de transpositions entre l'expérience vécue et les pratiques d'enseignement projetées. En accord avec ce numéro thématique, une précision des contextes dans lesquels l'étude s'inscrit s'impose en préambule.

2. Les différents contextes en interaction

De façon certes simplifiée, nous allons situer la présente étude dans quatre contextes en étroite interaction : le système scolaire en Belgique francophone (contexte 1), la formation des enseignants, constituée de cours (contexte 2) et de stages (contexte 3) et, enfin, les écoles et les classes où exerceront les enseignants lorsqu'ils seront diplômés (contexte 4).

Le système scolaire (contexte 1) est notamment régi par le décret « Missions » (1997) « définissant les missions prioritaires de l'enseignement fondamental et de l'enseignement secondaire et organisant les structures propres à les atteindre ». Celui-ci prône la pratique de l'évaluation formative (article 15) définie comme une « évaluation effectuée en cours d'activité et visant à apprécier le progrès accompli par l'élève et à comprendre la nature des difficultés qu'il rencontre lors d'un apprentissage ; elle a pour but d'améliorer, de corriger ou de réajuster le cheminement de l'élève ; elle se fonde en partie sur l'auto-évaluation » (p.4). Certaines tensions sont sans doute perceptibles entre cette vision d'une évaluation au service des apprentissages et le contexte historique d'un système scolaire fortement marqué par le redoublement et les relents d'une logique méritocratique conduisant au classement et à la sélection des élèves (Crahay, 2013). On imagine aisément que ces tensions peuvent constituer un frein à la mise en œuvre de réelles pratiques d'évaluation formative (Issaieva *et al.*, 2015).

En Belgique francophone, le titre pédagogique requis pour enseigner dans le secondaire supérieur (grades 10-12, élèves âgés de 15 à 18 ans environ) est l'agrégation de l'enseignement secondaire supérieur ou le master à finalité didactique. Cette formation universitaire est constituée de 15 ECTS de cours généraux et de 15 ECTS de stages et de cours spécifiques à la didactique de la discipline qui sera enseignée. Les cours généraux sont fréquentés par plus de 200 étudiants issus de 18 disciplines différentes (mathématiques, histoire, langues modernes, éducation physique, etc.). L'expérience d'évaluation mutuelle dont il est question ici a été mise en place en avril 2016. Elle s'est déroulée dans le contexte d'un cours général, et plus particulièrement lors d'une des séances de pratiques réflexives. Ces séances sont menées en petits groupes composés chacun d'environ 25 étudiants. Elles ont entre autres pour objectif de créer des connexions entre les cours théoriques, notamment pluridisciplinaires (contexte 2), et les stages (contexte 3) où les étudiants doivent œuvrer en tant qu'enseignant d'une discipline donnée.

Enfin, dans le cadre de notre étude, nous discuterons des transferts envisagés vers les pratiques professionnelles futures que les étudiants, une fois diplômés, seraient susceptibles de mettre en œuvre dans les classes dans lesquelles ils exerceront en tant qu'enseignant (contexte 4).

3. Le dispositif d'évaluation mutuelle mis en place

Si l'évaluation mutuelle consiste à ce que deux ou plusieurs apprenants s'impliquent activement dans l'évaluation de leurs productions respectives, l'objectif est *in fine* de développer l'autoévaluation (Mottier Lopez, 2015). L'autoévaluation, qui s'inscrit comme

une modalité d'évaluation formative au service de l'apprentissage (Morissette, 2010), implique « d'encourager les élèves à réfléchir sur ce qu'ils ont appris, à chercher les moyens d'améliorer leur apprentissage, et à planifier ce qui leur permet de progresser en tant qu'apprenant et d'atteindre leurs objectifs » (Rey et Feyfant, 2014 : 28).

Par ailleurs, c'est notamment en référence au principe d'homologie ou d'isomorphisme que nous avons voulu que les futurs enseignants vivent réellement une évaluation mutuelle. Appliqué à l'éducation, ce principe postule que, en faisant vivre aux futurs enseignants une activité proche de celle qu'ils auront à faire vivre aux élèves, ils transféreront plus facilement cette activité dans leurs pratiques. Un tel dispositif s'appuie ainsi sur l'hypothèse selon laquelle les *modalités* de formation seraient plus puissantes que les *contenus* eux-mêmes (Campanale, 2007). Soulignons encore qu'il ne suffit pas non plus de vivre une activité pour que la transposition dans les pratiques de classe soit automatique. Il est en effet essentiel de consacrer du temps pour analyser l'expérience vécue (Vanhulle, 2002) : les enjeux de l'activité, ce qui a été fait et comment cela a été mis en place, les difficultés éventuellement rencontrées, et ce que cela implique si des activités similaires étaient construites pour leurs futurs élèves, avec la possibilité d'imager des variantes...

3.1. Description du dispositif

Avant la séance consacrée à l'évaluation mutuelle, nous avons demandé aux étudiants d'analyser, à l'aide de l'outil CompAS (Parmentier et Paquay, 2002)¹, une séquence de cours qu'ils avaient mise en œuvre en stage. Ils devaient rédiger un texte présentant cette analyse de façon à disposer d'un matériau qui servirait de base à un dialogue évaluatif constructif. Nous les avons informés que ce travail préparatoire serait lu et évalué par un autre étudiant lors de ladite séance et qu'il pourrait ensuite être retravaillé en vue de figurer dans le travail réflexif final, qui fait partie de l'évaluation certificative du cours.

Concrètement, une partie de ce travail réflexif final porte sur l'analyse d'une séquence d'enseignement/apprentissage que les étudiants ont mise en œuvre en stage. Ils doivent rédiger une présentation de la séquence choisie en la contextualisant et en décrivant les objectifs, le timing et les grandes étapes méthodologiques de celle-ci. Pour analyser cette séquence, ils sont amenés à compléter l'outil CompAS et à choisir deux facettes dites satisfaisantes (c'est-à-dire rencontrées dans leur séquence) et deux facettes dites interpellantes (c'est-à-dire peu rencontrées alors qu'elles leur semblent importantes). Dans le premier cas, ils sont invités à expliquer ce qu'ils ont mis concrètement en œuvre en classe pour rencontrer ces facettes ; dans le second cas, ils sont tenus d'expliquer ce qu'ils pourraient mettre concrètement en œuvre en classe afin de les rencontrer davantage. Parmi les quatre facettes, les étudiants sont contraints d'exploiter la facette relative à l'évaluation. Quant au travail préparatoire, il devait être réalisé pour deux des quatre facettes : une satisfaisante et une interpellante.

¹ L'outil CompAS a été conçu par des chercheurs de l'Université catholique de Louvain, en collaboration avec des formateurs d'enseignants et des enseignants du secondaire. Cet outil a été construit en vue d'apprécier comment les dispositifs d'enseignement favorisent le développement de compétences par les apprenants. Les auteurs ont élaboré un modèle pour la construction de compétences qui intègre et articule les différentes facettes des activités d'enseignement/apprentissage. Parmi ces facettes, l'une concerne l'évaluation et s'intéresse notamment à l'implication des élèves dans le processus.

Lors de la séance d'évaluation mutuelle, nous leur avons demandé, comme annoncé, de lire le texte rédigé par un de leurs condisciples et d'évaluer ce travail préparatoire sur la base d'une grille d'évaluation constituée de critères et d'indicateurs. Cette grille avait préalablement été construite et validée par les six formatrices impliquées dans le cours et serait ensuite utilisée pour l'évaluation certificative de cette partie du travail réflexif final.

Avant de réaliser cette évaluation mutuelle en duo pluridisciplinaire, nous avons présenté la grille d'évaluation aux étudiants. Comme le recommandent Bourgeois et Laveault (2015), nous avons veillé à la compréhension des objectifs attendus, des critères et des indicateurs. Pour évaluer le travail de leur condisciple, les étudiants devaient non seulement identifier les forces et les faiblesses en attribuant une appréciation (-/±/+) face à chaque indicateur de la grille, mais ils devaient également préciser sur quels éléments concrets ils fondaient leurs appréciations. Ils devaient aussi formuler des pistes pour aider l'étudiant concerné à améliorer son travail. Nous avons insisté sur la nécessité de justifier les appréciations données et d'indiquer des pistes à mettre en œuvre pour rencontrer davantage les critères et les indicateurs, afin que les étudiants ne se contentent pas d'attribuer une simple appréciation. En effet, si tel était le cas, l'évaluation mutuelle tomberait sans doute dans les mêmes travers que l'auto-notation : elle risquerait de ne pas conduire à un véritable recul métacognitif et à de réelles régulations. *A contrario*, les justifications des forces et faiblesses et les pistes d'amélioration devraient susciter les interactions verbales favorisant les régulations (Campanale, 2007). Après avoir pris connaissance de l'évaluation réalisée par le pair, les duos ont eu l'occasion d'en discuter et de retravailler leur travail préparatoire.

Par la suite, avec l'ensemble du petit groupe pluridisciplinaire, nous sommes alors revenues sur la méthodologie vécue afin de faire réfléchir les étudiants sur les éléments auxquels ils devaient être attentifs pour transférer cette expérience vers leurs futures pratiques d'enseignement (Bourgeois et Laveault, 2015 ; Vanhulle, 2002). Pour terminer la séance, nous avons amené les étudiants à poursuivre le travail initialement demandé en réfléchissant à la façon dont ils pourraient intégrer une évaluation dans la séquence de cours analysée ou, s'ils en avaient déjà prévue une, d'explicitier ce qu'ils avaient mis en œuvre.

3.2. Objectifs visés par le dispositif et liens entre les contextes

Le dispositif visait un double objectif, assorti de trois boucles possibles de rétroaction dans différents contextes.

Premièrement, il s'agissait d'amener les futurs enseignants à participer, en tant qu'étudiants universitaires, à une évaluation formative en vue de préparer l'évaluation certificative. Dans ce cas-ci, on peut situer l'objectif comme une boucle de rétroaction qui agit au sein même du contexte 2 (le cours lui-même).

Deuxièmement, en référence au principe d'homologie ou d'isomorphisme (Campanale, 2007 ; Vanhulle, 2002), participer à cette expérience d'évaluation mutuelle (mais aussi prendre du recul et analyser ce dispositif) visait à initier de nouvelles pratiques. Deux boucles de rétroaction sont ici envisagées : d'une part, une possible intégration d'expériences d'évaluation formative au sein même de la séquence de cours qu'ils ont analysée (liens entre le contexte 2 – cours général pluridisciplinaire – et le contexte 3 – stage d'enseignement disciplinaire –, explicitement demandés dans les consignes du travail réflexif final) et, d'autre part, une réflexion quant à de possibles transpositions vers leurs pratiques futures, lorsqu'ils auront terminé leur formation et qu'ils prendront leur fonction en tant qu'enseignant dans une

école (liens escomptés entre la formation dans son ensemble – contextes 2 et 3 – et les futures pratiques professionnelles – contexte 4).

4. L'évaluation du dispositif mis en place

Cette évaluation prend la forme d'un regard critique documenté par deux sources d'informations. La première, qui est centrale, concerne les commentaires rédigés par les étudiants suite à l'expérience d'évaluation mutuelle vécue. La seconde, intervenant à titre illustratif, porte sur les pratiques d'évaluation que les futurs enseignants ont développées ou pourraient développer en classe, pratiques décrites dans les travaux réflexifs de fin d'année.

4.1. Que nous racontent les étudiants au sujet de cette expérience ?

Nous avons demandé aux étudiants d'expliquer en quelques lignes ce qu'ils avaient retenu de l'expérience d'évaluation mutuelle vécue. Cette consigne était volontairement vague et non orientée de façon à permettre aux étudiants de choisir les enjeux qui leur semblaient importants ou non de souligner.

En nous inspirant de l'approche inductive proposée par Blais et Martineau (2006)², nous avons analysé les 181 commentaires recueillis en vue de faire ressortir les enjeux mentionnés spontanément par les étudiants. Notons que, si certains étudiants n'ont indiqué qu'un enjeu, d'autres ont fait part de plusieurs. Nous avons réparti les catégories ainsi créées selon le contexte auquel elles faisaient référence : le fait de vivre une évaluation mutuelle, soit en tant qu'étudiant universitaire devant préparer l'évaluation certificative d'un cours donné (rétroaction au sein du contexte 2), soit en tant que futur enseignant cherchant à transposer l'expérience vécue vers ses pratiques (transposition du contexte 2 – cours général pluridisciplinaire – vers le contexte 3 – stage d'enseignement disciplinaire – ou 4 – pratiques professionnelles futures).

4.1.1. Participer à une évaluation formative en tant qu'étudiant universitaire (boucle de rétroaction au sein du contexte 2)

Les trois premières catégories se centrent sur les qualités du « produit attendu » pour l'évaluation certificative. Les commentaires font référence au caractère formatif de l'expérience vécue, soit en se cantonnant au versant « informatif » (des consignes du travail à réaliser ou des qualités de la première ébauche de ce travail), soit en mentionnant explicitement les versants « information » et « régulation » associés à ce type de pratique :

- Une série de déclarations portent sur l'idée selon laquelle l'évaluation mutuelle a engendré une meilleure compréhension du travail attendu : les étudiants sont au nombre de 32 à déclarer avoir été mieux **informés des attentes** (par exemple, *J'ai apprécié le fait de me retrouver à mon tour « dans la peau de l'évaluateur », car cela m'a aidé à mieux cerner et comprendre ce qu'il était attendu de ce travail* – Étudiant en langues modernes).
- Pour 30 étudiants, l'évaluation mutuelle a permis d'être **informés des points forts et/ou à améliorer** de leur travail préparatoire. Ces étudiants ne disent toutefois pas explicitement ce qu'ils font ou feront de ce feedback formatif (par

² Plus concrètement, nous avons lu chaque commentaire plusieurs fois jusqu'à avoir une vue d'ensemble des enjeux mentionnés. Nous avons ensuite identifié les segments de commentaires qui constituaient des unités de sens et, sur cette base, nous avons créé des catégories. Chaque segment de commentaires a alors été intégré dans une catégorie déjà créée ou, le cas échéant, a conduit à créer une nouvelle catégorie.

exemple, *Le regard de l'autre est toujours positif car il permet de mettre l'accent sur les faiblesses et les points forts de notre production* – Étudiant en histoire de l'art).

- Cent étudiants rapportent avoir été **informés et avoir pu réguler leur travail au niveau du « contenu »**. Il est à noter que, si cette régulation se réalise, pour la grande majorité de ces étudiants, grâce aux commentaires reçus (par exemple, *Je me suis référée aux commentaires de mon partenaire afin de corriger les deux facettes déjà écrites et de poursuivre le travail grâce à ses conseils* – Étudiant en langues romanes), certains d'entre eux mentionnent qu'ils se sont régulés grâce à la lecture de l'autre travail (par exemple, *Cela nous a aussi permis de voir la production d'autres étudiants et donc de voir des pistes de réalisation pour les parties qui nous semblaient plus obscures* – Étudiant en chimie).

Les deux dernières catégories se réfèrent moins directement aux qualités du « produit attendu », mais elles n'en demeurent pas moins liées à leur position d'étudiant ayant vécu une expérience singulière dans un cours (*la rétroaction reste au sein du contexte 2*) :

- Six étudiants rapportent avoir été **informés au niveau des méthodes de travail** cette fois (par exemple, *Grâce à cette expérience, j'ai également pu voir comment les autres avaient procédé pour réaliser leur travail et confronter leur méthode à la mienne* – Étudiant en langues classiques). L'un d'entre eux évoque explicitement avoir pu réguler sa façon de travailler.
- Une douzaine d'étudiants déclarent que l'évaluation mutuelle les a **rassurés** (par exemple, *Mon collègue m'a conforté dans certains choix que j'avais faits* – Étudiant en histoire de l'art ; *Il m'a donné confiance dans ma démarche* – Étudiant en histoire de l'art).

4.1.2. *Vivre une expérience d'évaluation formative en tant que futur enseignant (transposition du contexte 2 vers les contextes 3 et 4)*

Les deux premières catégories évoquent explicitement l'objet d'apprentissage visé au travers de l'expérience vécue : l'évaluation formative dans une vision élargie et, singulièrement ici, l'évaluation mutuelle comme une pratique permettant d'impliquer l'élève dans le processus d'évaluation :

- Trois étudiants déclarent spontanément que l'expérience vécue leur a permis d'**élargir leurs conceptions** en matière d'évaluation : ils ont expliqué avoir mieux compris ce qu'était une évaluation mutuelle et/ou concevoir l'évaluation formative dans sa perspective élargie (par exemple, *L'évaluation mutuelle m'a permis de comprendre en quoi consistait une évaluation mutuelle* – Étudiant en biologie ; *Cela m'a permis de me rappeler que l'évaluation formative peut revêtir des formes différentes que celle de l'évaluation classique* – Étudiant en chimie).
- Les propos de 40 étudiants montrent que l'expérience vécue les a conduits à **réfléchir à une possible transposition de dispositifs d'évaluation mutuelle** dans leurs futures classes. Si la moitié d'entre eux a plutôt précisé l'intérêt pour les élèves (par exemple, *La méthode est intéressante : réaliser cet exercice en classe avec les élèves pourrait être pertinent. En effet, en réalisant cet exercice, l'élève apprend à évaluer, à vérifier que les consignes sont respectées. Cet apprentissage pourra lui être utile pour s'autoévaluer* – Étudiant en sciences sociales ; *Je pense que c'est une bonne idée à appliquer*

*en classe, cela pousse l'élève à s'investir puisqu'il doit aider son camarade en l'évaluant au mieux et en plus, la discussion qui vient après peut apporter des éclaircissements sur certains points de la matière moins bien/pas compris – Étudiant en chimie), l'autre moitié s'est davantage placée du point de vue de l'enseignant (par exemple, *Avoir vécu cette évaluation mutuelle m'a paru très utile car je saurai à l'avenir comment l'organiser au sein de mes futures classes – Étudiant en langues modernes ; Jusqu'à ce que je ne prenne moi-même part à cette expérience, je ne me rendais pas vraiment compte de la manière dont je pouvais m'y prendre pour appliquer ces méthodes dans mes futures classes – Étudiant en sciences économiques).**

Si les deux dernières catégories laissent elles aussi entrevoir une possible régulation des pratiques d'enseignement, ce n'est pas au niveau de l'évaluation que la transposition se situe ici. La première catégorie reprend des propos qui restent spécifiques à la séquence de cours analysée et envisagent des modifications de celle-ci (« si c'était à refaire... ») ; la deuxième imagine de transposer des éléments découverts dans une discipline donnée dans des cours portant sur une autre discipline. Même s'ils ne relèvent pas du type de transposition explicitement visée dans les objectifs du dispositif, ils sont intéressants dans la mesure où ils amorcent également des liens entre le contexte 2 (cours pluridisciplinaire) et les contextes 3 (stages) et 4 (futurs pratiques professionnelles) :

- Grâce aux commentaires reçus, l'évaluation mutuelle a permis à 4 étudiants d'être **informés et d'envisager de réguler la séquence de cours elle-même** (par exemple, *Ces quelques faiblesses pourront dès lors faire l'objet d'un remaniement dans le cadre de la réutilisation de cette séquence – Étudiant en histoire).*
- Grâce à la lecture de l'autre travail, l'évaluation mutuelle a permis à 15 étudiants de **découvrir de nouvelles méthodes d'enseignement**. Quatre d'entre eux envisagent explicitement de réguler leurs pratiques (par exemple, *Par la lecture de la séquence de mon condisciple, j'ai pu avoir un aperçu de ses propres méthodologies. Celles-ci s'appliquaient à la physique, un cours scientifique. J'ai pu en retirer des éléments pour ma pratique dans le cadre du cours de biologie (situation-problème, expérimentation et modélisation – Étudiant en bio).*

4.2. Que nous révèlent les travaux réflexifs ?

Quelques étudiants décrivent une expérience d'évaluation formative qu'ils ont mise en œuvre en stage dans le cadre de la séquence de cours analysée. C'est le cas de l'étudiante suivante qui avait développé une évaluation mutuelle au départ d'une activité de production écrite :

- (...) *J'ai ensuite redistribué les textes au hasard parmi la classe. Chacun était alors chargé de faire un commentaire sous forme de conseils sur la rédaction qu'il venait de recevoir en se basant sur des critères que j'avais inscrits au tableau (les péripéties sont-elles réalistes ? La rédaction est-elle assez détaillée ? etc.). Les élèves en tant qu'évaluateurs ont ainsi pu apporter un regard neuf et objectif sur la rédaction qu'il leur était demandé de commenter et ont pu dans certains cas émettre des idées supplémentaires. Les apprenants récupéraient ensuite leur rédaction ainsi que les conseils de leur condisciple et étaient amenés à modifier leur production initiale afin de me rendre une version définitive. (...) Grâce à cette tâche, les élèves ont été impliqués dans l'évaluation de leur camarade car ce sont leurs commentaires qui permettaient*

au condisciple d'améliorer son texte. D'autre part, cet exercice pose les élèves en évaluateur, dès lors ils visualisent mieux la manière dont se fait une correction et peuvent ainsi lors de prochains exercices porter un regard plus critique sur leur production (Étudiant en romanes).

La plupart des étudiants n'avaient toutefois pas mis en place d'évaluation formative durant leur stage ; ce sont donc essentiellement des pistes explicitant ce qu'ils pourraient mettre concrètement en œuvre dans la séquence de cours analysée (« si c'était à refaire... »), qui sont décrites dans les travaux réflexifs. S'il faut reconnaître que beaucoup d'exemples sont peu développés, certaines initiatives intéressantes sont à souligner, comme l'illustre l'exemple repris ci-dessous :

- *Une idée serait de mettre à disposition des élèves des outils tels qu'une grille de critères qu'ils devraient rencontrer lorsqu'ils présentent une réalisation d'exercices. (...) Deux grilles seraient disponibles : une pour les exercices systématiques et une pour l'analyse de problèmes plus complexes. (...) Il serait judicieux d'utiliser ce type de grilles pour chaque séquence de façon à habituer les élèves à travailler avec. Il est évident que, la première fois, le professeur devrait prévoir un temps nécessaire pour expliquer comment fonctionnent les grilles, à quoi elles servent, quelles informations elles vont apporter aux élèves,.... (...) Un élève serait alors envoyé au tableau (...). Le professeur prendrait ensuite la première grille (...) et analyserait la réponse proposée avec les élèves pour que ceux-ci puissent voir réellement comment elle fonctionne. Enfin pour certains exercices ciblés, il proposerait aux élèves de remplir la grille par eux-mêmes après la correction au tableau. Les élèves pourraient ainsi plus facilement voir les points qu'ils leur restent à améliorer (...) (Étudiant en chimie).*

5. Discussion et conclusion

Partant du constat d'un décalage entre le cadre légal qui prescrit la mise en œuvre d'évaluations formatives et la réalité des classes où ce type d'évaluation semble être peu pratiqué (Fagnant et Goffin, à paraître), nous avons fait vivre aux futurs enseignants une expérience d'évaluation mutuelle en tant qu'étudiants universitaires dans un cours pluridisciplinaire. Un des objectifs du dispositif mis en place était de les faire réfléchir à la transposition de ces acquis vers leurs pratiques de stage et leurs pratiques professionnelles futures en tant qu'enseignants d'une discipline donnée.

Sur la base des commentaires rédigés spontanément par les étudiants suite à l'expérience d'évaluation mutuelle vécue, l'objectif de « préparer à l'évaluation certificative » est davantage rencontré que celui de « transposer de nouvelles pratiques d'évaluation ». En effet, 162 segments de commentaires portent explicitement sur les qualités du travail réflexif final, contre 40 qui font référence à un possible transfert. Malgré ce décalage entre les deux objectifs, ce résultat est en soi assez encourageant puisqu'un cinquième des étudiants semble donc envisager la possibilité de transposer l'expérience avec leurs futurs élèves.

Rappelons que les réflexions des futurs enseignants quant à la façon dont ils pourraient intégrer une évaluation formative dans la séquence de cours analysée sont dans la plupart des cas peu développées, donnant l'impression que les étudiants ont éprouvé des difficultés à concrétiser de telles pratiques d'évaluation. Par ailleurs, même si certains travaux font état d'exemples intéressants de possibles mises en œuvre d'évaluations formatives, nous ne

devons pas oublier qu'ils ont été proposés dans le cadre d'une évaluation certificative et qu'ils pourraient simplement témoigner d'une volonté de répondre aux attentes des formateurs. Autrement dit, il est possible que ces exemples traduisent davantage une bonne compréhension des concepts impliqués et des objectifs visés par le dispositif mis en œuvre (ce qui n'est déjà pas négligeable), qu'ils ne constituent de réels indicateurs des pratiques professionnelles futures en matière d'évaluation.

Pour tenter d'expliquer le décalage entre les deux objectifs visés par le dispositif, ainsi que la difficulté liée à la concrétisation des exemples, il convient d'interroger la problématique de l'isomorphisme entre la situation vécue et les pratiques de classe en fonction des disciplines mobilisées. Naïvement, il serait louable de penser que le transfert sera plus facile dans une discipline « littéraire » que dans une discipline « scientifique » dans la mesure où la tâche réalisée s'appuie sur l'évaluation d'une production écrite. Toutefois, les commentaires invitant à une transposition de l'expérience vécue vers les pratiques de classe, ainsi que les exemples développés pour imaginer la mise en œuvre d'évaluations formatives en classe, sont rédigés par des étudiants issus de disciplines diversifiées. Même si ce transfert d'un cours général vers une discipline donnée est un élément positif, ne perdons pas de vue l'importance des spécificités et des enjeux disciplinaires pour la mise en œuvre de pratiques d'évaluation efficaces (Detroz, Crahay et Fagnant, 2017 ; Vantourout et Goasdoué, 2014). Ces aspects disciplinaires n'ont pas du tout été traités dans le contexte 2 (cours général pluridisciplinaire) alors qu'ils sont pourtant au cœur des pratiques dans les contextes 3 et 4 (pratiques d'enseignement en stage et, plus tard, dans la pratique professionnelle), ce qui est une limite importante de l'approche proposée (et une difficulté majeure d'un cours regroupant des étudiants issus de 18 disciplines différentes).

Au sein de la formation initiale toujours, l'écart entre les thématiques travaillées dans les cours généraux (contexte 2) et les préoccupations des étudiants lors de leurs stages d'enseignement (contexte 3) peut constituer un autre frein au développement des transpositions escomptées. Concrètement ici, alors que la réflexion sur la mise en œuvre d'évaluations formatives se focalise sur l'apprentissage et le développement des élèves (niveaux 4 et 5 des préoccupations de l'enseignant selon Durand, 1996), les étudiants se préoccupent essentiellement du maintien de l'ordre et de l'engagement des élèves dans la tâche (niveaux 1 et 2 selon Durand, 1996). Si on ajoute à cela l'idée selon laquelle les enseignants qui reçoivent les étudiants en stage dans leurs classes recourent peu aux évaluations formatives, la transposition de ce type de pratiques risque de se heurter à un nouvel obstacle. Dans le même ordre d'idées, même si les futurs enseignants adhéraient à une vision élargie de l'évaluation formative et même si la formation initiale (contextes 2 et 3) avait pu les convaincre des bienfaits et de la faisabilité de telles pratiques, une autre source de tensions pourrait apparaître lors de leur entrée dans le métier (contexte 4) puisqu'ils auront à collaborer avec des collègues qui ne partagent pas nécessairement la même vision qu'eux. Enfin, ils devront aussi chercher à conjuguer harmonieusement les visées formatives et certificatives des évaluations (Mottier Lopez, 2015), ce qui pourrait s'avérer encore plus complexe dans un système scolaire fortement marqué par l'évaluation-sanction et le redoublement (contexte 1).

Entre les préoccupations de survie (Huberman, 1989) et les possibles tensions ressenties lors de leur entrée dans le métier, certains enseignants débutants auraient tendance à revenir à des pratiques plus traditionnelles (De Stercke *et al.*, 2010) voire à laisser resurgir d'anciennes conceptions qui semblaient pourtant avoir évolué durant la formation (Luisoni et Monnard, 2015). À l'instar de Devos et Paquay (2013), nous souhaitons nuancer ce propos qui donne

l'impression que les enseignants débutants délaisseraient tout ce qu'ils ont appris en formation initiale au profit de ce qu'ils ont eux-mêmes connu en tant qu'élèves ou de pratiques en cours dans leur école. En effet, s'arrêter à un tel constat découragerait toute entreprise innovante en formation initiale. De façon plus optimiste, une fois cette période de survie dépassée, les enseignants devraient pouvoir grimper sur l'échelle des niveaux de préoccupation de Durand (1996) pour mettre en œuvre des pratiques se centrant sur l'apprentissage réel et le développement des élèves. S'ils ont vécu positivement certaines pratiques innovantes, ils devraient alors y repenser et chercher à les adapter au mieux à leur contexte de travail. À ce stade, à la fois motivés et conscients des difficultés inhérentes au développement d'une évaluation réellement au service des apprentissages, sans doute seraient-ils plus enclins à participer à des recherches collaboratives dont les apports en termes de développement professionnel semblent être très intéressants (Morissette et Tessaro, 2016).

Références bibliographiques

- Allal, L. et Mottier Lopez, L. (2005). L'évaluation formative de l'apprentissage : revue de publications en langue française. Dans Centre pour la recherche et l'innovation dans l'enseignement (dir.), *L'évaluation formative pour un meilleur apprentissage dans les classes secondaires* (p. 265-290). Paris : Edition OCDE.
- Blais, M. et Martineau, S. (2006). L'analyse inductive générale : Description d'une démarche visant à donner un sens à des données brutes. *Recherches Qualitatives*, 26(2), 1-18.
- Bourgeois, L. et Laveault, D. (2015). Évaluation par les pairs à l'écrit : qualité des rétroactions pour soutenir la phase de révision. Dans P.-F. Coen et L. M. Bélair (dir.), *Évaluation et autoévaluation. Quels espaces de formation ?* (p. 101-117). Louvain-la-Neuve : De Boeck Supérieur.
- Butori, R. et Parguel, B. (2010). Les biais de réponse. Impact du mode de collecte des données et de l'attractivité de l'enquêteur. AFM, France. Disponible en ligne sur : <https://halshs.archives-ouvertes.fr/halshs-00636228/document>.
- Campanale, F. (2007). Évaluation réflexive en formation professionnelle et évaluation interactive dans les classes. Dans L. Allal et L. Mottier Lopez. (dir.), *Régulation des apprentissages en situation scolaire et en formation* (p. 191-206). Bruxelles : De Boeck.
- Conseil National d'Évaluation du Système Scolaire (Cnesco) (2014). *L'évaluation des élèves par les enseignants dans la classe et les établissements : réglementation et pratiques. Une comparaison internationale dans les pays de l'OCDE*. Paris : Cnesco.
- Crahay, M. (dir.) (2013). *L'école peut-elle être juste et efficace ? De l'égalité des chances à l'égalité des acquis*. Bruxelles : De Boeck.
- Crahay, M., Wanlin, P., Issaieva, E. et Laduron, I. (2010). Fonctions, structuration et évolution des croyances (et connaissances) des enseignants. Une revue critique de la littérature anglo-saxonne. Note de synthèse. *Revue française de pédagogie*, 172, 85-129.
- Décret « Missions » : décret de la Communauté française du 24 juillet 1997 définissant les missions prioritaires de l'enseignement fondamental et de l'enseignement secondaire et organisant les structures propres à les atteindre (1997). *Moniteur belge*, 23 septembre, p. 24653.
- De Stercke, J., Renson, J.-M., Cambier, J.-B., Leemans, M., Maréchal, C., Radermaecker, G., Temperman, G., Beckers, J. et De Lièvre, B. (2010). *Dynamiser l'insertion professionnelle des enseignants débutants du secondaire*. Disponible en ligne sur : http://www.enseignement.be/index.php?page=23827etdo_id=8042etdo_check=.
- Detroz, P., Crahay, M. et Fagnant, A. (dir.) (2017). *L'évaluation à la lumière des contextes et des disciplines*. Bruxelles : De Boeck.

- Devos, C. et Paquay, L. (2013). Le choc de la pratique chez les enseignants débutants : Quelle place pour la réflexivité ? Dans M. Altet, J. Desjardins, R. Etienne, L. Paquay et P. Perrenoud (dir.), *Former des enseignants réflexifs. Obstacles et résistances* (p. 229-247). Bruxelles : De Boeck.
- Durand, M. (1996). *L'enseignement en milieu scolaire*. Paris : PUF.
- Fagnant, A., Etienne, R., Mottier Lopez, L. et Hindryckx, M-N. (soumis). L'évaluation comme objet d'apprentissage et comme outil de développement professionnel dans le cadre de la formation d'enseignants. *Évaluer. Journal International de Recherche en Éducation et Formation*.
- Fagnant, A. et Goffin, C. (à paraître). Les conceptions des futurs enseignants du secondaire en matière d'évaluation : entre un accord de principe et une vision limitée de l'évaluation formative. *Mesure et Évaluation en Éducation*, 40(1).
- Huberman, M. (1989). Les phases de la carrière enseignante. Un essai de description et de prévision. *Revue française de Pédagogie*, 86, 5-16.
- Issaieva, E. et Crahay, M. (2010). Conception de l'évaluation scolaire des élèves et des enseignants : validation d'échelles et études de leurs relations. *Mesure et Évaluation en Éducation*, 33(1), 31-62.
- Issaieva E., Yerly, G., Petkova, I., Marbaise, C. et Crahay, M. (2015). Conceptions et prises de positions des enseignants face à l'évaluation scolaire dans quatre systèmes éducatifs européens : quel est le reflet des cultures et politiques éducatives ? Dans P.-F. Coen et L. M. Bélair (dir.), *Évaluation et autoévaluation. Quels espaces de formation ?* (p. 73-98). Louvain-la-Neuve : De Boeck Supérieur.
- Luisoni, M. et Monnard, I. (2015). L'évolution des conceptions de l'évaluation scolaire chez les futurs enseignants en formation initiale. Dans P.-F. Coen et L. M. Bélair (dir.), *Évaluation et autoévaluation. Quels espaces de formation ?* (p. 175-196). Louvain-la-Neuve : De Boeck Supérieur.
- Morissette, J. (2010). Un panorama de la recherche sur l'évaluation formative des apprentissages. *Mesure et Évaluation en Éducation*, 33(3), 1-27.
- Morissette, J. et Tessaro, W. (2016) Apports des recherches collaboratives sur le développement professionnel des enseignants et des formateurs en matière d'évaluation. *Évaluer. Journal International de Recherche en Éducation et Formation*, 1(2), 9-14.
- Mottier Lopez, L. (2015). *Évaluations formative et certificative des apprentissages*. Louvain-la-Neuve : De Boeck.
- Mottier Lopez, L. et Laveault, D. (2008). L'évaluation des apprentissages en contexte scolaire : développements, enjeux et controverses. *Mesure et Évaluation en Éducation*, 31(3), 5-34.
- Ocdé (2005). *L'évaluation formative. Pour un meilleur apprentissage dans les classes secondaires*. Paris : Édition OCDE.
- Ocdé (2013). *Résultats de TALIS 2013. Une perspective internationale sur l'enseignement et l'apprentissage*. Paris : Édition OCDE.
- Parmentier, P. et Paquay, L. (2002). *En quoi les situations d'enseignement/apprentissage favorisent-elles le développement de compétences ? Vers un outil d'analyse : le CompAS*. Louvain-la-Neuve.
- Paquay, L. (1994). Vers un référentiel des compétences professionnelles de l'enseignant ? *Recherche et Formation*, 16, 7-38.
- Rey, O. et Feyfant, A. (2014). Evaluer pour (mieux) faire apprendre. *Dossier de veille de l'Institut Français de l'Éducation*, 94, 1-44.
- Soussi, A., Ducrey, F., Ferrez, E., Nidegger, C. et Viry, G. (2006). Pratiques d'évaluation : ce qu'en disent les enseignants (à l'école obligatoire et dans l'enseignement postobligatoire général). Rapport du service de la recherche en éducation, Université de Genève.

Disponilbe en ligne sur le serveur suisse de documents pour l'éducation et la formation : <http://edudoc.ch/record/3603/>.

- Vanhulle, S. (2002). Des savoirs en jeu au savoir en « je ». Médiations sociales et processus de subjectivation en formation initiale d'enseignants. *Cahiers du Service de Pédagogie expérimentale*, 11-12, 49-62.
- Vantourout, M. et Goasdoué, R. (2014). Approches et validité psycho-didactique des évaluations. *Éducation et Formation*, e-302, 139-156.
- Vause, A. (2010). Les croyances et connaissances des enseignants de l'école primaire à propos de l'acte d'enseigner. *Éducation et Formation*, 294, 14-19.
- Younès, N. et Gaime, E. (2012). L'évaluation formative en contexte : points de vue d'enseignants, points de vue d'élèves. *Diversité, VEI*, 161-166.

Les raisons du faible usage des résultats d'évaluation externe par les enseignants. Étude croisée dans trois contextes éducatifs

Gonzague YERLY

Université de Fribourg, Suisse

Résumé

L'évaluation des apprentissages des élèves n'est aujourd'hui plus uniquement interne à la classe (au niveau micro), elle est aussi externe et à grande échelle (au niveau macro). Avec les nouveaux modes de gouvernance des systèmes éducatifs, ces deux niveaux se rencontrent. Il est attendu des enseignants qu'ils prennent en compte les résultats d'épreuves externes afin d'améliorer leurs pratiques. Toutefois, la littérature montre un faible usage de ces données externes sur le terrain. Notre analyse du regard de 41 enseignants (entretiens semi-directifs) issus de trois juridictions distinctes quant à leur dispositif et à leur politique d'évaluation externe – au Québec, en Ontario et en Suisse – permet d'identifier les raisons de ce « découplage » spécifique. Notre analyse semi-inductive s'inspire du concept de « validité unifiée d'un test » de Messick (1989) pour comprendre le point de vue des enseignants. Notre analyse montre que le faible usage des résultats d'évaluation externe peut s'expliquer par le fait que les enseignants des trois contextes étudiés accordent peu de validité(s) à ces données externes. Les enseignants remettent en cause non seulement la mesure à large échelle, mais aussi le manque d'implication pratique des résultats d'évaluation externe et leurs conséquences sur les acteurs de l'École. Des différences sont identifiées, liées à certaines caractéristiques des dispositifs d'évaluation externe (finalités, suivi, conséquences) et au profil des enseignants (ancienneté professionnelle, confrontation directe aux tests). Finalement, alors que l'évaluation des acquis des élèves est devenue un enjeu aussi bien politique que pédagogique, notre contribution montre que, aux yeux des enseignants, les différentes approches d'évaluation (interne et externe) sont trop éloignées. D'importantes dissonances les séparent tant sur le plan de leurs méthodes, de leurs usages et de leurs finalités.

Mots-clés

Évaluation externe, résultats d'épreuves externes, enseignants du primaire, validité unifiée, découplage.

Abstract

The assessment of students learning is nowadays not only internal to the classroom (at the micro level), but also external and on a large scale (at the macro level). With the new modes of governance of the educational systems, these two levels meet. Teachers are expected to take into account the results of external tests in order to improve their teaching practices. However, the literature shows little use of these external data in the classrooms. Our analysis of the perceptions of 41 teachers (semi-directive interviews) from three different contexts as to their tools and policy of external evaluation – in Quebec, Ontario and Switzerland – helps to identify the reasons for this specific "decoupling". Our semi-inductive analysis uses the concept of "unified validity of a test" by Messick (1989) to understand the point of view of teachers. Our analysis shows that the low use of external testing results is due to the fact that teachers, in the three contexts studied, give little validity to this external data. It is not only the large-scale measure that is questioned, but also the lack of practical implications of the results of external tests and their consequences on school stakeholders. Differences are identified according to certain characteristics of the external evaluation devices (purposes, follow-up, consequences) and to the profile of teachers (professional experience, direct confrontation with the tests). Finally, while the assessment of students' learning has become both a political and an educational issue, our contribution shows that the different approaches of assessment (internal and external) are too different for teachers. Significant dissonances separate them in terms of their methods, their uses and their aims.

Keywords

External testing, tests' results, primary teachers, unified validity, decoupling.

1. Introduction

Depuis plusieurs décennies, l'évaluation des apprentissages est l'objet de nombreuses réformes dans la plupart des systèmes scolaires. On attend des enseignants qu'ils mobilisent un ensemble grandissant de méthodes afin de faire le bilan des acquis des élèves, mais aussi de les soutenir dans leurs apprentissages. En outre, l'évaluation des apprentissages des élèves n'est aujourd'hui plus seulement interne à la classe, elle est également externe et à grande échelle (Allal, 2009 ; Mottier Lopez, 2009). Il est attendu des enseignants et/ou des établissements scolaires – par des mécanismes divers – qu'ils prennent en compte les résultats de leurs élèves à des épreuves externes (désormais EE) afin d'ajuster leurs pratiques (Laveault, 2009). Ainsi, ces nouveaux modes de régulation institutionnelle (Maroy, 2014) font se rencontrer les différents niveaux contextuels (De Ketele, 2006) autour de l'évaluation des acquis des élèves : le niveau micro (celui de l'enseignement et de la classe), le niveau macro (celui du système éducatif et des politiques), mais aussi le niveau méso (celui de l'établissement).

Les travaux de synthèse traitant de l'impact des EE sur les pratiques des enseignants (Au, 2007 ; Mons, 2009 ; Rozenwajn et Dumay, 2015 ; Yerly, 2014) partagent les mêmes conclusions : indépendamment des contextes, les enseignants confrontés à ces dispositifs ne procèdent généralement qu'à des ajustements mineurs ou symboliques de leurs pratiques et/ou mettent en œuvre des pratiques de préparation intensive aux examens en cas de conséquences importantes liées aux résultats. Surtout, si le format et le contenu des tests leur permettent un certain alignement de leurs pratiques – d'enseignement et surtout d'évaluation – sur certains contenus et/ou sur certaines méthodes attendues, les résultats ne seraient que très peu mobilisés par les enseignants (Yerly, 2017a, 2017b). La présente recherche¹ propose d'identifier, dans le discours d'enseignants, les raisons de ce « découplage » (Coburn, 2004) spécifique observé entre résultats d'EE et pratiques enseignantes. Notre étude qualitative a été menée dans trois contextes scolaires primaires francophones sélectionnés pour leur variété quant à leur politique d'EE : une commission scolaire au Québec, un conseil scolaire en Ontario et un canton en Suisse.

2. Cadrage théorique et empirique

Des sondages menés dans de nombreux contextes européens ou nord-américains ont permis d'interroger le regard des enseignants quant aux EE. La synthèse de ceux-ci montre que l'attitude des enseignants face à ces dispositifs peut être qualifiée de manière générale comme « ambivalente » (Mons, 2009). Une étude, que nous avons réalisée préalablement dans quatre contextes bien distincts quant aux dispositifs et aux politiques de « gouvernance par les résultats » (ou politiques d'*accountability*) en Suisse, en Belgique, en Ontario et au Québec, a permis d'identifier et de comprendre davantage les « ambivalences » des enseignants face à ces politiques et à leurs outils (Yerly et Maroy, soumis). Elle montre que les enseignants accordent une certaine « légitimité » (morale, cognitive et pragmatique) au principe d'une évaluation externe (par exemple, pour renforcer une certaine uniformité au sein des systèmes). Néanmoins, les praticiens sont beaucoup plus critiques – également sur ces trois niveaux de légitimité – sur la mise en place (par exemple, renforce des inégalités, temps perdu) et les effets de ces politiques sur les acteurs du système (par exemple, pression, pratiques de

¹ Cette recherche a été menée dans le cadre d'une bourse postdoctorale financée par le Fonds National suisse de la recherche scientifique (P2FRP1_161705), avec le soutien des professeurs : Dany Laveault (Université d'Ottawa) et Christian Maroy (Université de Montréal). L'investigation empirique a également bénéficié du soutien du Conseil de la Recherche en Sciences Humaines (CRSH) du Canada (projet CRSH Savoir n° 435-012-0701, sous la direction de C. Maroy).

teaching to the test). Ces constats rejoignent également ceux établis par Demailly (2003) quant à l'évaluation externe des établissements scolaire en France. Elle propose « trois cadres explicatifs » qui permettent d'expliquer les résistances des enseignants face à cette forme d'évaluation institutionnelle (Demailly, 2003) : les intérêts de l'acteur (résistance stratégique, lutte de pouvoir), les valeurs de l'acteur (résistance politique, lutte entre différentes visions de l'École, la rationalité de l'acteur (résistance cognitive, distance cognitive avec la mesure statistique).

Les différentes études présentées permettent une compréhension sociologique approfondie des réactions des enseignants face aux politiques d'EE et à leurs outils. Pour cette contribution, nos objectifs sont spécifiques au faible usage des résultats des élèves aux EE par les enseignants. En effet, nos études préalables (Yerly, 2017b ; Yerly et Maroy, soumis) montrent que la plupart des enseignants interrogés relèvent la qualité de ces examens (leur forme et leur contenu). Ils se servent des (anciennes) épreuves notamment pour identifier les contenus à enseigner et pour modifier leurs propres outils d'évaluation sommative. Par contre, les enseignants interrogés en Suisse, en Belgique, au Québec et en Ontario ne mobilisent que très peu les résultats des élèves aux EE pour ajuster leurs pratiques de leur propre initiative (en Ontario, les résultats sont par contre fortement mobilisés par les conseillers pédagogiques lors de formations ; en Belgique certains résultats sont analysés en équipe-école). L'analyse des 53 entrevues montre que la comparaison entre résultats internes (ceux de la classe) et externes (ceux des EE), est réalisée de manière très superficielle par les enseignants et ne les entraînent que peu à s'autoévaluer et à procéder à des modifications de leurs pratiques d'enseignement et/ou d'évaluation. Ils ne consacrent par exemple pas ou très peu de temps à analyser les résultats de leurs élèves et en discutent de manière très informelle et rapide avec leurs collègues (échange d'avis ou de sentiment global sur la réussite de sa classe). À la rigueur, les résultats des EE sont pour eux une sorte de « baromètre » très général pour valider leur travail, s'assurer être globalement dans la « bonne ligne » en comparaison à une norme régionale.

L'objectif de cette contribution est de comprendre davantage quelles sont les raisons du faible usage des résultats d'EE par les enseignants. C'est la question de la validité d'un test – concept que nous empruntons pour un temps à la psychométrie – qui a servi à analyser le point de vue des enseignants. Plus précisément, nous nous inspirons du concept de « validité unifiée » de Messick (1989). Selon ce concept, la validité d'un test n'est pas seulement à considérer quant à la qualité de la mesure, mais aussi quant à ses implications pratiques et à ses conséquences sociales. Ainsi, trois questions ont guidé notre analyse du regard des enseignants : les résultats d'EE sont-ils une mesure pertinente des acquis des élèves (*meaningfulness*) ? Les résultats d'EE permettent-ils des implications pratiques (*usefulness*) ? Les résultats d'EE ont-ils des conséquences sociales (*appropriateness*) ?

3. Contextes de l'étude

Notre analyse porte sur trois contextes de scolarité primaire : une commission scolaire au Québec, un conseil scolaire en Ontario et un canton en Suisse². Depuis les années 2000, l'évaluation externe des acquis des élèves tient une place centrale dans ces trois juridictions, suite à l'instauration de politiques de « gouvernance par les résultats » qui visent à améliorer les résultats du système scolaire. Ces réformes ont été amorcées et sont inspirées en grande partie par les enquêtes internationales sur les acquis des élèves (par exemple, PISA, TIMMS).

² Le contexte belge qui a été étudié dans les travaux précédents n'a pas été mobilisé dans la présente analyse.

Pour atteindre leurs buts, les administrations ont misé principalement sur l'EE. Elle doit offrir aux acteurs des différents niveaux (micro, méso, macro) davantage d'information sur la performance des élèves et ainsi orienter leurs pratiques. C'est ainsi que se sont multipliées les EE évaluant les acquisitions des élèves à divers moments de la scolarité³. Si les finalités générales et les instruments principaux (curriculum uniforme, EE) sont plutôt semblables, les politiques et les dispositifs d'EE des trois juridictions étudiées se différencient sur certains points que nous identifions dans le tableau 1.

Juridiction	Dans un canton en Suisse	Dans un conseil scolaire en Ontario	Dans une commission scolaire au Québec
Modèle (Maroy et Voisin, 2014)	Responsabilisation douce	Reddition de comptes et responsabilisation réflexives	Reddition de comptes néo-bureaucratique
Mécanisme	Responsabilisation interne de l'enseignant	Reddition de comptes interne (équipe-enseignants → direction) et externe (école → conseil scolaire → ministère → public)	Reddition de comptes interne (équipe-enseignants → direction) et externe (école → commission scolaire ministère → public)
Instruments	Curriculum uniforme EE cantonales (en 2 ^{ème} , 4 ^{ème} et en 6 ^{ème} année) ⁴	Curriculum uniforme EE provinciales (en 3 ^{ème} et en 6 ^{ème} année) Plan d'action des écoles et du conseil scolaire	Curriculum uniforme EE provincial (en 4 ^{ème} et 6 ^{ème} année) Contractualisation avec cibles de réussite pour les écoles et la CS
Contenus évalués	Maths, L1, L2, sciences naturelles et humaines (en alternance pour les EE de 2 ^{ème} et 4 ^{ème})	Maths, L1 (écriture et lecture)	L1 (en 4 ^{ème}); L1, L2 et maths (en 6 ^{ème})
Rôle des enseignants	Passation, correction (en partie pour l'EE de 6 ^{ème}) et communication des résultats aux élèves EE de 2 ^{ème} et 4 ^{ème} = faibles (prise en compte libre par l'enseignant – passation parfois facultative) EE de 6 ^{ème} = élevées (critère pour l'orientation des élèves dans les filières du secondaire)	Passation (mais ne peuvent pas voir les copies des élèves)	Passation, correction et communication des résultats aux élèves
Conséquences des résultats pour les élèves	Aucune conséquence formelle. Toutefois, la réputation de l'enseignant est engagée fortement pour l'EE de 6 ^{ème} qui confirme l'évaluation interne des élèves	Les résultats des EE ne comptent pas au bulletin, mais ils sont publiés et servent aux parents dans le choix de l'école.	Modérées (20% de la note du bulletin) Les résultats des tests sont publiés et servent aux parents dans le choix de l'école.
Conséquences pour l'établissement et les enseignants	Aucun – enseignant individuellement (N.B. il n'y pas de direction dans les établissements)	Suivi et support pour les écoles qui performant le moins. Classement des écoles dans les médias selon les résultats d'EE	Aucune conséquence formelle. Toutefois, les écoles sont redevables au niveau supérieur. Classement des écoles dans les médias selon les résultats d'EE Présentation des résultats à l'équipe-école (par la direction) Rédaction d'un plan d'action école par les enseignants (supervisé par la direction et le CS)
Feedback et soutien suite aux résultats des EE	Aucun – enseignant individuellement (N.B. il n'y pas de direction dans les établissements)	Présentation des résultats à l'équipe-école (par la direction) Suivi des écoles par le CS (formation spécifique) sur la base des résultats	Présentation des résultats à l'équipe-école (par la direction) Rédaction d'un plan d'action école par les enseignants (supervisé par la direction et le CS)

Tableau 1 : Description des dispositifs d'EE des trois contextes étudiés

³ Certains dispositifs d'EE au Québec et en Suisse sont antérieurs à ces réformes. Ils servaient alors uniquement à la certification des élèves à la fin de certaines étapes du cursus.

⁴ Correspond aux 4^{ème}, 6^{ème} et 8^{ème} années selon la terminologie HarmoS.

Ces contextes représentent d'ailleurs trois des quatre catégories de politiques *d'accountability* définies par Maroy et Voisin (2014)⁵. Quelques nuances existent sur les contenus et la passation des tests externes. Les principales différences entre les contextes se situent surtout dans le mécanisme des politiques d'EE, leurs conséquences et le feedback et le soutien apportés suite aux résultats d'EE.

4. Méthodologie

Notre étude suit une approche qualitative et compréhensive (Albarello, 2003). Les données sont issues d'un projet de recherche plus large¹ portant sur l'impact des politiques de gouvernance par les résultats sur les pratiques des enseignants. Quarante et un enseignants du primaire ont répondu à un entretien semi-directif individuel de 60 à 90 minutes. Pour chacun des trois contextes, l'échantillon (cf. Tableau 2) a été formé afin de couvrir les différents degrés de scolarité et de varier le profil des enseignants surtout quant à leur ancienneté professionnelle et donc au nombre d'EE passées avec leurs classes.

Juridiction	Dans un canton en Suisse	Dans un conseil scolaire en Ontario	Dans une commission scolaire au Québec
Nombre d'enseignants	16 enseignants	14 enseignants	11 enseignants
Degrés de scolarité	1 ^{ère} à 6 ^{ème} année	1 ^{ère} à 6 ^{ème} année	3 ^{ème} à 6 ^{ème} année
Ancienneté	2 à 35 ans	10 à 27 ans	1 à 25 ans
Nombre d'établissements	12 établissements	2 établissements	2 établissements

Tableau 2 : Description de l'échantillon

Pour tous les enseignants que nous avons interrogés, les thématiques interrogées étaient les mêmes : le profil professionnel de l'enseignant, le profil de son établissement, le récit de la mise en œuvre de la politique d'EE dans son établissement, les perceptions de l'enseignant quant à la politique d'EE, les effets perçus sur ses pratiques. Les entretiens ont été intégralement retranscrits, puis traités au moyen du logiciel Nvivo10 pour mener une analyse de contenu en trois étapes : analyse de cas (chaque enseignant), analyse de chaque contexte, analyse croisée. Notre analyse est « semi inductive » (Maroy, 1995). Elle s'appuie sur l'analyse approfondie d'un certain nombre de cas pour arriver à la formulation d'explications quant à la question étudiée. Nos constats émergent des données recueillies sans hypothèses *a priori* mais avec l'aide d'un modèle d'analyse, le concept de « validité unifiée d'un test » de Messick (1989). Pour ce faire, une grille catégorielle a servi au traitement des données. Elle comprend les trois dimensions de la validité selon ce modèle : pertinence de la mesure ; implications pratiques des résultats d'EE ; conséquences sociales des résultats d'EE. Cet outil a permis d'identifier, dans les entretiens, les raisons évoquées par les enseignants quant à l'usage et leur non usage des résultats d'EE. Des synthèses d'étapes (par enseignant, par contexte, croisées) ont ensuite permis de répondre à notre questionnement spécifique.

⁵ Le quatrième type est celui de « reddition de compte dure » (par exemple, Angleterre et certains états américains) dans lequel des conséquences importantes pour les acteurs sont prévues suite aux résultats d'EE (par exemple, primes, licenciements).

Afin de compléter notre connaissance des contextes et suite à l'analyse de la documentation officielle, des entretiens ont été menés avec des responsables locaux et/ou intermédiaires (directions d'établissement, inspecteurs, conseillers pédagogiques).

5. Résultats

En s'inspirant de la notion de « validité unifiée d'un test » de Messick (1989), nous pouvons dégager trois catégories de raisons qui permettent d'expliquer – souvent simultanément – le faible usage des résultats d'EE par les enseignants : (a) le manque de pertinence de la mesure, (b) les implications pratiques limitées des résultats d'EE et (c) les conséquences indésirables des résultats d'EE. Nous concentrons la présentation des résultats sur les différences entre contextes. Certaines différences dues au profil des enseignants (ancienneté professionnelle, année scolaire) sont simplement soulignées, faute d'espace.

5.1. Le manque de pertinence de la mesure

Les enseignants québécois, ontariens et suisses que nous avons interrogés donnent généralement peu de « valeur » aux données produites par les EE. S'ils considèrent que les résultats sont un indicateur complémentaire « intéressant » (par exemple, un regard de l'extérieur, une autre façon d'évaluer), ils font généralement davantage confiance à leurs propres résultats d'évaluation et à leur connaissance des élèves. Même en cas d'écarts importants entre évaluation interne et externe, la plupart des enseignants (et sans différences observée selon leur profil) disent ne pas remettre en question l'évaluation des apprentissages de leurs élèves ou ajuster leur degré d'exigence par rapport au « standard » régional. Les enseignants justifient les éventuels écarts de performances des élèves – à la hausse comme à la baisse – par divers facteurs :

- *État particulier des élèves lors de la passation.* Les enseignants interrogés justifient généralement des différences de résultats entre évaluation interne et externe par un état particulier des élèves lors de la passation d'une EE (par exemple, la fatigue, le stress). Cette situation d'examen est peu habituelle, trop longue et/ou revêt trop d'importance. Au contraire, en Ontario, les résultats n'ayant aucun impact sur le cursus des élèves, les enseignants justifient des résultats plus bas par le manque de motivation et d'implication des élèves lors des examens uniformes.
- *Degré de difficulté de l'examen.* Le degré de difficulté des examens peut être jugé inapproprié : trop élevé ou trop bas et/ou trop fluctuant d'une année à l'autre. Au Québec, les examens du ministère sont souvent considérés par les enseignants comme non pertinents car leur niveau est trop fluctuant d'une année à l'autre.
- *Trop grande standardisation des EE.* La pertinence de la mesure est souvent critiquée par les enseignants car elle est à leurs yeux « injuste ». Celle-ci ne prend pas en compte les différences entre les élèves et les milieux (par exemple, élèves allophones, écoles de milieux défavorisés) et crée ainsi des inégalités entre les élèves et les établissements.
- *Inégalité des conditions de passation et de correction.* En Ontario et au Québec, les enseignants remettent également en question la pertinence des résultats d'EE quant aux conditions de passation et de correction qui ne seraient pas identiques pour tous (par exemple, aide des enseignants, correction plus clémente). L'importance de ces EE

pour les écoles et les enseignants (par exemple, reddition de comptes aux autorités, palmarès dans les médias) entraînerait des pratiques « frauduleuses » permettant de s'extraire de ces pressions.

5.2. Implications pratiques limitées

Si les enseignants n'utilisent que peu les résultats des élèves aux EE, ce n'est pas uniquement parce qu'ils ne croient pas en la pertinence intrinsèque des données produites. C'est également qu'ils n'y voient pas ou peu d'opportunités pour développer leurs pratiques et/ou que le dispositif ne le permet pas :

- *Information trop générale.* Pour les enseignants interrogés, l'EE est une mesure trop générale et ponctuelle des acquis des élèves pour pouvoir agir. Elle n'est pas aussi approfondie que leur évaluation en continu des apprentissages de l'élève. Les résultats des classes ou des écoles sous forme de moyennes de pourcentage (Québec et Ontario) ou de note (en Suisse) sont des informations trop peu approfondies pour être prises en compte. En Ontario, par exemple, au moment de la réception des résultats par la direction de l'école certains enseignants demandent à voir les résultats individuels de leurs (anciens) élèves. Pourtant, même les résultats individuels – lorsqu'ils sont disponibles – ne représentent pour les enseignants des divers contextes qu'une performance d'un élève à un moment donné sur un ensemble très (trop) vaste de contenus. Le plus souvent, les enseignants souhaiteraient avoir accès à des résultats davantage détaillés (par exemple, par compétence ou par item).
- *Feedback insatisfaisant.* Les enseignants interrogés regrettent pour la plupart le manque de feedback qui accompagne les résultats d'EE, voire remettent en question la pertinence de celui-ci lorsqu'il existe. En Suisse et au Québec, aucun feedback – si ce n'est les résultats chiffrés – n'est proposé aux enseignants sur la base des résultats, ni individuellement ni collectivement. En Suisse, personne ne présente les résultats aux enseignants⁶. En Ontario, le suivi est plus important. Les résultats d'EE sont exploités dans des rencontres de formation (3 à 4 par année) axées sur les domaines dans lesquels les élèves de l'école (ou du conseil scolaire) ont moins bien performé. Mais même là, les pistes didactiques proposées sont souvent critiquées par les praticiens : pas assez personnalisées, peu opérationnelles, peu pertinentes. Certains enseignants ontariens considèrent même ce soutien davantage comme une sanction pour les écoles et les enseignants. Ils ont le sentiment de perdre une partie de leur autonomie.
- *Trop de différences entre les méthodes d'évaluation interne et externe.* Dans les trois contextes, les enseignants reconnaissent que les EE « couvrent » bien le programme. En outre, en Suisse et au Québec, les enseignants reconnaissent de manière générale que les examens externes sont proches des pratiques évaluatives du terrain (par exemple, questionnement, mise en forme). En revanche, en Ontario, les enseignants ne reconnaissent que peu leurs pratiques évaluatives. Les EE y sont composées en partie de questions à choix multiples, type de question que les enseignants jugent peu pertinentes pour évaluer les acquis des élèves (par exemple, aucune analyse possible, recours à la chance). Il n'est donc pas possible pour eux de comparer les résultats issus d'instruments trop différents.

⁶ Pour l'EE de 6^{ème} année, les enseignants reçoivent tout de même un document qui fait la comparaison par discipline entre la moyenne des résultats de leur classe et la moyenne des résultats de tous les élèves du canton.

- *Responsabilité des enseignants.* Les enseignants interrogés remettent en cause leur (trop grande) responsabilité dans les résultats d'EE. Pour eux, d'autres facteurs expliquent autant – si ce n'est davantage – la réussite ou l'échec des élèves : caractéristiques du milieu, niveau de la cohorte d'élèves, compétences intrinsèques des élèves, suivi des parents, apprentissages réalisés – ou non – lors des années scolaires précédentes. Ainsi, les résultats montrent avant tout la performance de l'élève et non celle de l'enseignant. L'implication pratique des résultats d'EE est ainsi très faible surtout que les examens ont lieu en fin d'année (et en fin de cycle), juste avant que les élèves ne changent de classe. Les enseignants ne pourront alors plus avoir d'influence sur les apprentissages des élèves concernés.
- *Destinataires des résultats d'EE.* Souvent, les praticiens interrogés ne se voient pas comme les destinataires des résultats d'EE. Ceux-ci ne servent donc pas à faire leur autoévaluation. Pour eux, les résultats d'EE sont destinés avant tout aux élèves et aux parents (pour se situer, pour leur orientation) ou à l'administration. En Suisse, les enseignants ignorent souvent à qui sont dédiés les résultats⁷ (à un mystérieux « ils » dans leur discours). Ils associent par exemple de manière erronée ces données comme constitutives des résultats PISA. Au Québec et en Ontario, certains enseignants estiment que les EE sont plutôt l'affaire des « gestionnaires » (directions d'écoles, conseil et commission scolaire).

Selon nos analyses, ces constats peuvent être nuancés selon l'ancienneté des enseignants. Les praticiens novices ou peu expérimentés que nous avons interrogés souhaiteraient pouvoir se servir davantage des résultats d'EE. Les plus anciens, eux, sont plus critiques et se contentent bien souvent de leurs propres pratiques qui ont « fait leurs preuves ». Aussi, les plus expérimentés jugent souvent les méthodes évaluatives des EE trop différentes de leurs propres pratiques (par exemple, évaluation des compétences vs des connaissances).

5.3. Conséquences indésirables des résultats d'EE

Le faible usage des résultats d'EE par les enseignants peut également être expliqué par les conséquences indésirables qu'entraînent les résultats d'EE – mais aussi plus largement les politiques d'EE – sur les différents acteurs de la vie éducative et sur toute l'institution École :

- *Pression sur les élèves et les enseignants.* Les élèves et les enseignants ressentent parfois de fortes pressions face aux EE et à leurs résultats. Ce stress se traduit notamment par des pratiques didactiques et pédagogiques contraires à l'idéal des enseignants, peu efficaces à long terme et peu motivantes pour les élèves (*teaching to the test*). Les enjeux liés aux résultats d'EE sont en effet souvent perçus comme importants pour les élèves et les enseignants. Dans le canton suisse, les résultats des EE de 6^{ème} année sont un des indicateurs pour l'orientation des élèves dans les filières du secondaire. Ils confirment (ou non) l'évaluation interne, celle de l'enseignant. Au Québec, les résultats comptent pour 20% dans le bulletin de l'élève. Aussi, au Québec et en Ontario, les résultats des écoles sont publiés sur les sites Internet des écoles, mais aussi dans la presse sous forme de palmarès. Dans ces deux provinces canadiennes, les écoles doivent également « rendre des comptes » au niveau supérieur (conseil/commission scolaire). Surtout au Québec, les écoles ont des objectifs chiffrés de réussite à atteindre. En Ontario, les écoles « moins performantes » font l'objet d'un

⁷ Pour les EE de 2^{ème} et 4^{ème} année.

suivi plus important, parfois considéré comme une « mise sous tutelle » par les enseignants. En Suisse, la pression ressentie par les enseignants est également engendrée par un manque d'informations transmises aux enseignants. Ils craignent ainsi d'éventuelles conséquences.

- *Image faussée de l'École.* Pour les enseignants, l'EE et ses résultats donnent une image trop partielle de l'École. Ce seul résultat quantitatif ne montre pas l'ensemble des missions des enseignants qui ne s'arrête pas à une performance chiffrée à un moment donné. Au Québec et en Ontario surtout, les résultats publiés et utilisés par les parents pour faire le choix d'une école pour leurs enfants donnent selon eux un « mauvais signal » à la société.
- *Impact sur le fonctionnement de l'École.* Pour certains enseignants, l'EE a un impact important sur le fonctionnement de l'École. Elle prend beaucoup de place dans les relations entre les différents acteurs (enseignants, administrateurs, parents). En Ontario et au Québec surtout, les EE et plus largement les politiques de reddition de comptes dans lesquelles elles s'inscrivent sont souvent perçues comme une « intrusion » de l'administration dans la sphère pédagogique autrefois réservée aux seuls enseignants. Certains enseignants interrogés se sentent ainsi atteints dans leur autonomie, étant poussés à axer leurs pratiques sur la réussite aux examens et se sentant jugés voire contrôlés par rapport aux résultats de leurs élèves.

Nos analyses, montrent que ces pressions peuvent être plus importantes pour les enseignants peu expérimentés. Ils doivent encore prouver leurs compétences et connaissent moins les plans d'études, les méthodes attendues, etc. Aussi, ces pressions sont bien plus importantes pour les enseignants qui sont confrontés directement aux EE, c'est-à-dire ceux qui enseignent dans les années concernées par la passation des tests. Ces enseignants se sentent particulièrement mis « sous le spot », contrôlés et jugés par les parents, leurs supérieurs hiérarchiques, voire même leurs collègues.

6. Discussion

Dans les trois contextes étudiés, malgré des différences importantes entre leurs dispositifs et leurs politiques d'EE (cf. Tableau 1), le faible usage des résultats d'EE par les enseignants peut s'expliquer par des raisons plus ou moins similaires. Notre analyse montre que les praticiens interrogés ne leur accordent en effet que peu de validité(s) (Messick, 1989). Premièrement, cette mesure externe n'est pour eux pas une « preuve » pertinente du degré d'acquisition des élèves. Elle est marquée de nombreux biais : état particulier de l'élève lors de la passation des EE, degré de difficulté des EE peu pertinent ou trop fluctuant, iniquités lors de la passation et de la correction des EE. Deuxièmement, c'est l'implication pratique, l'usage (potentiel) de ces données externes, qui est remis en question par les enseignants. Ceux-ci mettent donc plus largement en doute la pertinence du dispositif : information trop générale, feedback insuffisant et/ou peu opérationnel, comparaisons non pertinentes, manque de communication. Troisièmement, les enseignants rejettent les conséquences négatives des résultats d'EE sur les enseignants, les élèves mais aussi sur les établissements scolaires et le système : stress, compétition, transformation de l'École.

Ainsi, notre analyse de l'expérience des enseignants montre que la « validité » accordée – ou non – par les enseignants aux résultats d'EE va au-delà de la simple « valeur » accordée à la mesure, au chiffre. En s'inspirant du concept de « validité unifiée » de Messick (1989), notre

analyse permet de démontrer que la validité des résultats d'EE est également liée à ses implications pratiques et à ses conséquences sociales. Certes, il serait pertinent de quantifier et d'analyser les différences et les corrélations de ces trois « facettes » de la validité, mais il nous paraît impossible de les isoler. Comme l'écrit Messick (1989), elles sont « inséparables ». Il est donc nécessaire de les analyser toutes les trois pour comprendre comment les enseignants « construisent » la validité des résultats d'EE.

Si de nombreuses similitudes sont observées quant aux raisons de ce « découplage » (Coburn, 2004), certaines particularités semblent également se dessiner entre les contextes. Dans les deux juridictions canadiennes (Québec et Ontario), l'usage institutionnel des résultats d'EE est bien plus important qu'en Suisse : ils participent grandement aux mécanismes de « reddition de comptes » et à la compétition entre les écoles ou entre les juridictions à l'intérieur d'une même province (conseils/commissions scolaires). Selon nos analyses, cet important usage des résultats d'EE à l'extérieur de la salle de classe diminue encore leur validité aux yeux des enseignants aussi bien sur le plan de la mesure (la pression/compétition entraîne davantage de biais), de son implication pratique (les enseignants se considèrent encore moins comme les destinataires des résultats) que de ses conséquences (la pression/compétition entraîne davantage d'effets néfastes). Certaines raisons de ce découplage sont également liées au profil des enseignants. En effet, certaines tendances, d'ailleurs communes aux trois contextes, ont été soulignées. Elles sont liées à l'ancienneté professionnelle (quant aux implications pratiques des résultats d'EE et aux conséquences des résultats d'EE) et à la confrontation directe ou indirecte avec les EE, selon le niveau scolaire enseigné (conséquences des résultats d'EE). Ces différences seraient également à investiguer plus finement et avec un échantillon d'enseignants plus large et plus diversifié.

7. Conclusion

L'évaluation des apprentissages des élèves est devenue un outil aussi bien politique que pédagogique. Il semble cependant que les niveaux micro (la classe) et macro (le système) (De Ketele, 2006) soient très hermétiques en ce qui concerne l'évaluation des acquis des élèves. Chaque niveau protège ses pratiques, ses intérêts et ses valeurs, sans forcément reconnaître ceux de l'autre. Alors que les systèmes éducatifs ne font pas (plus) tout à fait confiance à l'évaluation réalisée par les enseignants, il semble que les enseignants de leur côté ne soient pas enclins à partager une compétence qui leur était alors réservée. Notre contribution montre que, aux yeux des enseignants, chaque approche d'évaluation (externe et interne) – propre à chaque niveau du système – est trop différente pour permettre un « dialogue ». Il y a pour eux une dissonance trop importante aussi bien sur les méthodes (évaluation standardisée à grande échelle vs évaluation en continue approfondie pour chaque élève), sur les usages (traitement statistique et comparaisons vs analyse pour chaque élève) et sur les finalités (améliorer les résultats du système vs améliorer l'apprentissage de chaque élève).

En conclusion, à l'instar de ce que propose Allal (2009), il semble que c'est au niveau méso – celui de l'établissement – que se situe le plus grand défi si l'on entend faire « dialoguer » les niveaux micro et macro au sujet de l'évaluation des acquis des élèves et dépasser les dissonances que soulèvent les enseignants interrogés dans cette contribution. C'est à ce niveau intermédiaire que les différences entre ces deux approches évaluatives (interne et externe) peuvent être exploitées. Du côté des enseignants, il paraît important que les rôles, les spécificités et l'apport de chaque approche évaluative leur soient régulièrement exposés. Ensuite, des opportunités – individuelles et collectives – devraient leur être fournies quant à l'usage des résultats d'EE comme un indicateur utile (parmi d'autres) pour développer leurs

pratiques. Encore faut-il que les enseignants soient accompagnés et aient accès – dans un idéal – à des résultats robustes, détaillés, opérationnels et aux conséquences néfastes limitées. Enfin, dans l’optique d’un « dialogue » entre les niveaux micro et macro, une question nous paraît encore importante : comment l’évaluation des acquis des élèves réalisée par les enseignants peut-elle, dans le sens inverse, contribuer au pilotage et à l’amélioration du système ?

Références bibliographiques

- Albarello, L. (2003). *Apprendre à chercher : l'acteur social et la recherche scientifique*. Bruxelles : De Boeck.
- Allal, L. (2009). Pratiques évaluatives des enseignants face aux méthodologies des recherches évaluatives portant sur le système scolaire. Dans L. Mottier Lopez et M. Crahay (dir.), *Évaluations en tension. Entre la régulation des apprentissages et le pilotage des systèmes*. (p. 29-45). Bruxelles : De Boeck.
- Au, W. (2007). High-Stakes Testing and Curricular Control: A Qualitative Metasynthesis. *Educational Researcher*, 36(5), 258-267.
- Coburn, C. E. (2004). Beyond Decoupling: Rethinking the Relationship between the Institutional Environment and the Classroom. *Sociology of Education*, 77(3), 211-244.
- De Ketele, J.-M. (2006). La recherche en évaluation : propos synthétiques et prospectifs. *Mesure et évaluation en éducation*, 29(1), 99-118.
- Demailly, L. (2003). L'évaluation de l'action éducative comme apprentissage et négociation. *Revue française de pédagogie*, 142(1), 115-129.
- Laveault, D. (2009). L'évaluation en classe: des politiques aux pratiques. *Mesure et évaluation en éducation*, 32(3), 3-22.
- Maroy, C. (1995). L'analyse qualitative d'entretiens. Dans L. Albarello, F. Digneffe, J. Hiernaux, C. Maroy, D. Ruquoy et P. Saint-Georges (dir.), *Pratiques et méthodes de recherche en sciences sociales*. Paris : Armand Colin.
- Maroy, C. (2014). L'évaluation externe et les indicateurs comme instruments d'action publique : fonctions instrumentales et symboliques. Dans J. Morrissette et M.-F. Legendre (dir.), *Enseigner et évaluer. Regards sur les enjeux éthiques et sociopolitiques*. (p. 165-184). Québec : Presses universitaires de Laval.
- Maroy, C. et Voisin, A. (2014). Une typologie des politiques d'accountability en éducation : l'incidence de l'instrumentation et des théories de la régulation. *Éducation comparée - Nouvelle série*, 11, 31-57.
- Messick, S. (1989). Meaning and Values in Test Validation: The Science and Ethics of Assessment. *Educational Researcher*, 18(2), 5-11.
- Mons, N. (2009). Effets théoriques et réels des politiques d'évaluation standardisée. *Revue française de pédagogie. Recherches en éducation*, 169, 99-140.
- Mottier Lopez, L. (2009). L'évaluation en éducation : des tensions aux controverses. Dans L. Mottier Lopez et M. Crahay (dir.), *Évaluations en tension. Entre la régulation des apprentissages et le pilotage des systèmes*. (p. 7-25). Bruxelles : De Boeck.
- Rozenwajn, E., et Dumay, X. (2015). Les effets de l'évaluation externe sur les pratiques enseignantes : une revue de la littérature. *Revue française de pédagogie*, 189(4), 105-138.
- Yerly, G. (2014). *Les effets de l'évaluation externe des acquis des élèves sur les pratiques des enseignants. Analyse du regard des enseignants du primaire*. Thèse de doctorat en Sciences de l'éducation. Université de Fribourg, Fribourg.

- Yerly, G. (2017a). Quel impact des politiques de responsabilisation douce sur les pratiques enseignantes ? Dans Y. Dutercq et C. Maroy (dir.), *Professionalisme enseignant et politiques de responsabilisation* (p. 121-141). Bruxelles : De Boeck Supérieur.
- Yerly, G. (2017b). Évaluation en classe et évaluation à grande échelle. Quel est l'impact des épreuves externes sur les pratiques évaluatives des enseignants ? *Mesure et évaluation en éducation*, 40(1), 33-60.
- Yerly, G. et Maroy, C. (soumis). La gouvernance par les résultats est-elle un mode de régulation de l'école légitime aux yeux des enseignants ? Une enquête qualitative dans quatre systèmes scolaires. Article soumis à la *Revue française de pédagogie*.

Analyse sociodidactique des évaluations internationales inspirées du Cadre européen commun de référence pour les langues

M'hand AMMOUDEN

Laboratoire LAILEMM¹ – Université de Bejaia, Algérie

Résumé

Notre étude s'inscrit dans le cadre d'une sociodidactique qui prend en considération la diversité des contextes et des cultures éducatives – et par extension des cultures d'évaluation – qui les caractérisent. Elle porte sur la prise en compte des spécificités des contextes étrangers dans les évaluations internationales qui s'inspirent du Cadre Européen Commun de Référence pour les Langues (C.E.C.R.), et sur les retombées des cultures d'évaluation concernant des candidats vivant dans des contextes géographiques, culturels et éducatifs différents de ceux des concepteurs de ces évaluations sur les résultats qu'ils obtiennent. La première partie montre que plusieurs des descripteurs de l'échelle des niveaux de compétences proposée dans le C.E.C.R. suggèrent la nécessité de la prise en compte de l'environnement immédiat de l'apprenant/usager de la langue étrangère. La deuxième partie montre que ce principe ne semble pas être suffisamment respecté par les concepteurs des épreuves des DELF et DALF, et elle met en évidence quelques-unes des difficultés auxquelles pourraient faire face ceux qui les subissent, quand ils ne sont pas européens, voire français. Nos expérimentations nous conduisent à conclure, d'une part, que ces difficultés découlent essentiellement de la nature des consignes et des thématiques, et, d'autre part, que les contraintes qui en résultent désavantagent les publics qui ne connaissent pas la France et sa culture et remettent en question partiellement la fiabilité et l'équité de ces évaluations.

Mots-clés

Évaluation, culture éducative, C.E.C.R., consigne, DELF, DALF.

Abstract

Our study is part of a socio-didactic research that considers the diversity of contexts and of the educational cultures – and by extension of the evaluation cultures – that characterize them. It focuses on taking into account the specificities of foreign contexts in the international evaluations based on the Common European Framework of Reference for Languages (C.E.F.R.), and on the impacts of these evaluation cultures on the results obtained by candidates who live in geographical, cultural and educational contexts different from those of the designers of these evaluations. The first part shows that several of the descriptors of the scale of competency levels proposed in the C.E.F.R. suggest the need to take into account the immediate environment of the foreign language learner/user. The second part reveals that this principle does not seem to be sufficiently considered by the designers of the DELF and DALF tests, and highlights some of the difficulties faced by those who take those tests when they are not European, if not French. Our experiments lead us to conclude, on the one hand, that these difficulties stem essentially from the nature of the tests' instructions and themes and, on the other hand, that the resulting constraints penalize in particular those who do not know France and its culture and, in part, call into question the reliability and fairness of these assessments.

Keywords

Evaluation, educational culture, C.E.F.R., instructions, DELF, DALF.

¹ Laboratoire de recherches et de formations en Langues Appliquées et Ingénierie des Langues En milieu Multilingue.

1. Introduction

L'évaluation est souvent sujette à l'hétérogénéité, même entre deux enseignants assurant la même matière (ou discipline) dans la même école. Elle suscite souvent plus de débats animés et de positions divergentes que la plupart des autres sujets concernant les situations éducatives telles que les démarches d'enseignement, les contenus, les supports, etc.² Les concepteurs du Cadre Européen Commun de Référence pour les Langues (désormais C.E.C.R.), élaboré à la suite des travaux du Conseil de l'Europe (2001), à qui ce phénomène ne semble pas avoir échappé, ont présenté un précieux outil visant à harmoniser les conceptions et les pratiques d'évaluation en langues : il s'agit de l'échelle des niveaux communs de compétences. Celle-ci peut certainement contribuer considérablement à remédier à la quasi-totalité des problèmes que pose l'évaluation et à réduire les écarts résultant de la coexistence de diverses conceptions d'évaluation entre les enseignants d'une même école, d'un même pays, voire d'un continent, l'Europe³. En effet, tous les États membres du Conseil de l'Europe ont été représentés dans les équipes qui l'ont conçu et on explique dès ses premières pages que le Cadre propose « *une base commune* » pour l'élaboration, entre autres, des examens en Europe (Conseil de l'Europe, 2001 : 3-9). Le fait que le succès du C.E.C.R. ait finalement dépassé les frontières de l'Europe en a fait un outil qui peut contribuer à pallier les problèmes que nous venons d'évoquer même entre des évaluateurs appartenant à deux ou plusieurs continents. Ce sont quelques-unes des raisons qui nous ont conduit, à l'instar de plusieurs autres didacticiens, à voir dans la conception qui sous-tend l'évaluation telle que décrite dans le C.E.C.R., un outil à recommander même au-delà de l'Europe ; et à prôner son utilisation en Algérie (Ammouden et Ammouden, 2010).

Notons que, de toute manière, les publics algériens ont déjà affaire à des évaluations fondées sur la logique des six niveaux au moins à travers les évaluations qui concernent les examens des DELF⁴, DALF⁵ et TCF⁶, proposées par les Centres culturels français en Algérie, et par les tests de niveaux qui se font dans certains Centres d'enseignement intensif des langues⁷ et dans quelques écoles privées algériennes. Néanmoins, cela ne va pas sans poser de nombreux problèmes. L'examen des contraintes auxquelles les confrontent ces évaluations est le principal objectif de la présente étude.

Cet objectif situe cette étude dans le champ de la sociodidactique qui accorde une importance primordiale à l'étude des spécificités des contextes (Abbes-Kara, Kebbas et Cortier, 2013 ; Cortier, 2009 ; Rispaïl et Blanchet, 2015). Il la place particulièrement dans le champ de celles qui insistent sur l'inévitable prise en considération des particularités des cultures éducatives et d'évaluation qui caractérisent ces contextes (Ammouden et Cortier, 2009 ; Beacco, 2015 ; Cortier, 2009 ; Mottier Lopez et Laveault, 2008). En traitant de cette problématique, Beacco

² Les raisons qui sont à l'origine de ces phénomènes, que les limites de cet article ne permettent pas d'exposer, sont diverses : diversité des objectifs assignés à l'évaluation par les enseignants et par les évaluateurs, incidences de leurs résultats sur les taux de réussite, réactions des évalués et de leurs parents, etc.

³ Quelques-uns des avantages qui peuvent découler du recours à cette grille sont expliqués dans l'ouvrage de F. Noël-Jothy et B. Sampsonis (2006).

⁴ Diplôme d'Études en Langue Française.

⁵ Diplôme Approfondi de Langue Française.

⁶ Test de Connaissance du Français.

⁷ Les C.E.I.L sont des établissements universitaires qui proposent des formations en langues générales (usuelles), en langues de spécialités ou en langues sur objectifs spécifiques. Celles-ci sont facultatives et souvent payantes, ce qui n'est pas le cas des autres formations universitaires. Les apprenants qui s'inscrivent pour poursuivre ces formations peuvent être des extra-universitaires. Les tests de niveaux permettent de classer les apprenants dans des groupes pédagogiques selon leur niveau de compétences (A1 à C2).

(2015 : 55) explique que c'est le fait de considérer les situations dans lesquelles se déroule un enseignement « *non comme des réceptacles neutres où sont activés des processus universaux (psycho et sociolinguistique par exemple), mais comme des éléments déterminants de ceux-ci* » qui devrait conduire à « *instaurer en objets de recherche les éléments constitutifs des cultures éducatives* ».

Dans quelle mesure les évaluations qui s'inspirent de la logique du C.E.C.R. prennent-elles en considération les cultures d'évaluation auxquelles étaient habitués auparavant les candidats étrangers ? Nous supposons que, dans la mesure où le C.E.C.R. concerne plusieurs langues et plusieurs contextes, il offrirait probablement cette possibilité. Les résultats d'une étude menée en 2010, nous permettent par contre de supposer que les évaluations proposées par le CIEP⁸ pour les candidats concernés par les épreuves des DELF et DALF ne tiendraient pas compte suffisamment de ces possibilités qu'offre le C.E.C.R. et donc de l'hétérogénéité des contextes, qui implique celles des cultures éducatives qui les caractérisent, et de là des cultures d'évaluation. Or, cela peut avoir de lourdes conséquences sur la manière dont des apprenants étrangers reçoivent les contenus de ces évaluations internationales en lien avec le contexte des concepteurs. Pour vérifier ces hypothèses, nous nous appuyons sur l'examen des corpus suivants :

- l'échelle des six niveaux de compétences proposée dans le C.E.C.R. ;
- les programmes de langue française des 10 niveaux scolaires concernés par cette langue (de la 3^{ème} année du primaire à la 3^{ème} année du secondaire) ;
- des sujets d'évaluation pratiqués par les enseignants de français algériens ;
- des sujets de DELF, pris sur le site du CIEP⁹ et/ou dans les ouvrages de Chevallier-Wixler, Dayez, Lepage et Riba (2006) ou de Dupoux, Havard, Martial et Weeger (2006) ;
- les résultats de plusieurs évaluations à travers lesquelles nous avons utilisé des sujets du DELF et du DALF.

2. La possibilité de contextualisation dans le C.E.C.R.

Les concepteurs du C.E.C.R. encouragent-ils à prendre en compte les spécificités des contextes ?

2.1. Évaluation dans le C.E.C.R. : entre homogénéisation et unification

L'objectif du C.E.C.R. est de proposer un cadre commun pour l'enseignement/apprentissage et l'évaluation des langues. Cela est clairement expliqué dès les premières lignes du document, comme en témoignent les exemples d'énoncés suivants :

- « Le Cadre européen commun de référence offre une base commune pour l'élaboration de programmes de langues vivantes, de référentiels, d'examens, de manuels, etc. » ;
- « (...) Donner des critères objectifs pour décrire la compétence langagière facilitera la reconnaissance mutuelle des qualifications obtenues dans des contextes d'apprentissage divers et, en conséquence, ira dans le sens de la mobilité en Europe » ;
- « Le Cadre européen commun de référence est conçu pour que soient surmontées les difficultés de communication rencontrées par les professionnels des langues vivantes et qui proviennent de la différence entre les systèmes éducatifs » (Conseil de l'Europe, 2001 : 9).

⁸ Centre International d'Études Pédagogiques.

⁹ Voir la page « Langue française, évaluation et certifications » : <http://www.ciep.fr/langue-francaise-evaluation-certifications/expertise/evaluation-certifications>.

Qu'en est-il des spécificités des contextes et des « cultures » évaluatives ? On précise dans le C.E.C.R. que les échelles diffèrent d'un pays à un autre et qu'on note « *quelquefois de 1 à 6 avec 4 comme note d'admission, quelquefois de 1 à 20, avec 10 (la « moyenne ») comme note d'admission* ». On souligne aussi qu' « *il y a une myriade de normes puisque chaque classe dans chaque type d'école dans chaque académie se constitue une norme différente* » ; soucis auxquels s'ajoute « *la question de la forme de l'évaluation mise en œuvre et le degré d'accord sur l'interprétation des notes auquel parviennent les enseignants dans un contexte donné* » (Conseil de l'Europe, 2001 : 38).

Si ces éléments sont cités notamment pour justifier la nécessité d'une homogénéisation, ou mieux d'une harmonisation, des échelles de niveaux, ils peuvent également permettre de supposer que cette prise de conscience imposerait une prise en compte des différences pour éviter des résistances de la part des utilisateurs.

En effet, dès le début du chapitre 9, consacré à l'évaluation, on précise que « *l'appréciation et l'évaluation des connaissances et des savoir-faire devraient pouvoir tenir compte de l'ensemble des circonstances et expériences où ces connaissances et savoir-faire se mettent en place* ». On suggère aussi en traitant du Portfolio l'importance d'y mentionner non seulement les résultats obtenus lors des évaluations officielles, mais aussi les « *expériences plus informelles de contact avec des langues et cultures autres* » (Conseil de l'Europe, 2001 : 132).

Même si le souci de tracer les contours d'un certain cadre homogénéisant des pratiques d'enseignement/apprentissage (et surtout de l'évaluation) est à l'origine du C.E.C.R. (Conseil de l'Europe, 2001 : 9-10), la lecture de nombreux passages de ce document permet de constater que ce souci n'implique pas une uniformisation qui ne tiendrait pas compte des spécificités des contextes, et des usagers et apprenants qui y vivent.

2.2. Le place du contexte dans les descripteurs du C.E.C.R.

La lecture de l'échelle globale des six niveaux communs de compétences (Conseil de l'Europe, 2001 : 25) suffit amplement pour confirmer que le C.E.C.R. accorde une grande importance au contexte dans lequel se trouve l'utilisateur ou l'apprenant d'une langue. En décrivant les compétences qui caractérisent le niveau A1 (Introductif ou découverte), on explique que l'utilisateur ou apprenant qui possède les compétences de ce niveau :

Peut comprendre et utiliser **des expressions familières et quotidiennes** ainsi que des énoncés très simples qui visent à satisfaire des besoins concrets. Peut se présenter ou présenter quelqu'un et poser à une personne des questions la concernant – par exemple, **sur son lieu d'habitation, ses relations, ce qui lui appartient**, etc. – et peut répondre au même type de questions (...).

Retenons que les expressions que nous mettons en évidence¹⁰ dans ce descriptif indiquent clairement que ce niveau est intimement lié à l'utilisation de la langue dans le contexte qui concerne l'apprenant. Nous retrouvons des indications similaires à celles que nous venons de mentionner dans la définition du niveau A2 (Intermédiaire ou de survie) :

Peut comprendre des phrases isolées et des **expressions fréquemment utilisées en relation avec des domaines immédiats de priorité** (par exemple, informations personnelles et familiales simples, achats, **environnement proche, travail**). Peut communiquer lors de tâches simples **et habituelles** ne demandant qu'un échange

¹⁰ C'est nous qui mettons en évidence (par le gras) dans les descriptifs des six niveaux de compétences, et cela tout au long de l'article.

d'informations simple et direct sur des sujets **familiers et habituels**. Peut décrire avec des moyens simples sa formation, **son environnement immédiat** et évoquer des sujets qui correspondent à **des besoins immédiats**.

En ce qui concerne le niveau B1 (Niveau seuil), on trouve certes un descriptif qui peut (mais pas forcément) concerner un contexte autre que celui de l'apprenant/usager, « *Peut se débrouiller dans la plupart des situations rencontrées en voyage dans une région où la langue cible est parlée* », mais ce n'est pas le cas des autres :

Peut comprendre les points essentiels quand un langage clair et standard est utilisé et s'il s'agit de **choses familières dans le travail, à l'école**, dans les loisirs, etc. (...). Peut produire un discours simple et cohérent **sur des sujets familiers et dans ses domaines d'intérêt**. (...).

Dans le descriptif du niveau B2 (Avancé ou indépendant), la première partie ne comporte pas d'indices clairs sur le lien avec le contexte, mais ne contient également pas ceux qui suggèreraient le contraire. La seconde partie, permet davantage de lier le niveau au contexte de l'apprenant : « *Peut s'exprimer de façon claire et détaillée sur une grande gamme de sujets, émettre un avis sur un sujet d'actualité et exposer les avantages et les inconvénients de différentes possibilités* ». Le descriptif du niveau C1 (Autonome) comprend aussi au moins un descriptif qui renvoie à la même logique : « *Peut utiliser la langue de façon efficace et souple dans sa vie sociale, professionnelle ou académique* ». Pour ce qui est enfin du niveau C2 (Maîtrise), on n'y décèle ni indice explicite qui indique la prise en compte de l'environnement, ni celui qui suggérerait le contraire.

Nous pouvons donc conclure que la lecture des définitions des six niveaux de compétences, dont notamment celles des niveaux de l'utilisateur élémentaire (A1 et A2), de l'utilisateur indépendant (B1 et B2) et de l'utilisateur autonome (C1), montre clairement que ces niveaux sont toujours définis par des situations de communication qui concernent généralement le contexte et l'environnement de l'apprenant. Qu'en est-il des évaluations internationales qui s'appuient sur le C.E.C.R. ?

3. Les publics algériens face aux épreuves des DELF

Pour apporter des éléments de réponse à cette question, nous nous interrogerons dans ce qui suit sur le degré de prise en compte des spécificités de publics étrangers par les concepteurs des épreuves des DELF et DALF, dirigés par le CIEP de Paris. Ces diplômes ont été, selon Noël-Jothy et Sampsonis (2006 : 65), « *revus de façon à correspondre aux niveaux du CECR du Conseil de l'Europe : depuis le 1^{er} septembre 2005* ». Des données issues d'enquêtes nous permettent pourtant de mettre en évidence quelques-unes des contraintes auxquelles font face les candidats algériens, dont un nombre très important passe chaque année ces tests.

3.1. L'oral et son évaluation

L'enseignement des langues dans le contexte algérien s'est fait durant longtemps, aussi bien sur le plan des discours théoriques que sur celui des pratiques de classes, de la manière « *scripto-centrée* ». Suite à la réforme amorcée progressivement, il y a une quinzaine d'années, les documents théoriques émanant du Ministère de l'Éducation nationale algérien « *décroissent* » de plus en plus la nécessité d'accorder une meilleure place à l'enseignement de l'oral. Néanmoins, les pratiques peinent à adopter ces indications : « *l'oral est souvent pratiqué par "défaut" – avant l'université – grâce à la technique de la question/réponse* » (Ammouden, 2010, 2013).

Aujourd'hui, les efforts du ministère commencent à porter leurs fruits. On trouve, dans tous les inventaires des enseignables proposés dans les programmes de français, des contenus et indications qui orientent clairement vers le développement des compétences de réception et de production orales. On consacre, notamment dans les manuels du primaire et du collège, une rubrique consacrée au développement de ces compétences au début de chaque séquence didactique. On met fréquemment à la disposition des enseignants concernés des supports audio à utiliser.

Force est de constater toutefois que ces efforts ne suffisent pas à conférer une bonne place à l'oral. Pourquoi ? D'abord parce que, si les nouveaux programmes du lycée réservent une place à l'oral, les manuels eux n'ont pas été actualisés. Or, il y a des enseignants qui accordent plus d'importance aux manuels qu'aux programmes. Ensuite, vu que les salles de classe ne sont généralement pas dotées de moyens audiovisuels, seuls les enseignants qui possèdent des outils adéquats font écouter des supports audio. Enfin et surtout parce qu'on n'enseigne que très rarement des genres de discours oraux. On se contente de travailler uniquement la compréhension et la production. On place ces contenus aux débuts des séquences didactiques, mais on ne programme pas encore des ateliers qui prennent vraiment en charge les spécificités de l'oral. Ainsi, les contenus réservés à l'oral jouent-ils davantage le rôle d'une « mise en situation » ou d'une sorte d'évaluation diagnostique dans la séquence qui se focalisera sur l'écrit. Ajoutons à cela le fait qu'avant l'université (en licence de français), les évaluations (notées) ne concernent jamais l'oral.

Cette situation fait que le fait même d'être confronté à des épreuves orales est suffisant pour déstabiliser les candidats algériens, mais aussi ceux issus des nombreux autres contextes dans lesquels l'oral, son enseignement/apprentissage et son évaluation ne bénéficient pas encore de toute l'importance qu'ils méritent.

3.2. La nature des activités et des consignes

L'autre contrainte importante qui déstabilise les apprenants découle de la nature des activités et des consignes qui s'avèrent être parfois très différentes de celles auxquelles ils sont habitués. Dans l'enseignement et l'évaluation du français dans le contexte de notre étude, les activités et exercices qui dominent sont ceux consistant à :

- identifier le type auquel appartient un texte (narratif, argumentatif, explicatif, etc.) ;
- donner ou choisir un titre à un texte (écrit surtout) ;
- répondre par vrai ou faux ;
- repérer un synonyme, un antonyme ou une expression ayant le même sens, etc. ;
- utiliser le paratexte pour identifier la situation (schéma) de la communication ;
- relever des informations (arguments, personnages, etc.) ;
- classer dans un tableau (causes/conséquences, avantages/inconvénients, etc.) ;
- relier par des flèches des éléments présentés dans deux colonnes d'un tableau ;
- compléter des textes à trous (souvent le résumé du texte support) ;
- identifier le rapport logique (cause, conséquence, but, etc.) exprimé dans des phrases ;
- compléter en choisissant l'articulateur, la conjonction, le relatif, ... qui convient ;
- réécrire des phrases en effectuant les transformations qui s'imposent (passage du singulier au pluriel ou inversement, passivation, transposition au style indirect, etc.) ;
- conjuguer des verbes donnés entre parenthèses à l'infinitif dans des phrases ;
- reconstituer des textes dont les parties sont données dans le désordre ;
- établir le plan, le schéma narratif, le schéma actanciel, ou le résumé d'un texte ;
- rédiger un court texte obéissant souvent au plan suivant : introduction, développement, conclusion.

Quand les apprenants sont confrontés à d'autres types d'activités ou de consignes, ils sont déstabilisés. Une expérience à travers laquelle nous avons fait subir aux mêmes étudiants deux tests concernant le niveau B1, l'un conçu par nous et l'autre tiré de l'ouvrage de Chevallier-Wixler, Dayez, Lepage et Riba (2006 : 106-115), a révélé – à la suite des entretiens avec ces apprenants – que quelques-unes des consignes de ce deuxième test leur ont posé problème et les ont empêchés de trouver la bonne réponse (Ammouden et Ammouden, 2010 : 44). Il s'agissait notamment des questions n°3 et n°4 (p. 111) dans les items desquels l'alternative « Vrai/Faux » que connaissaient les apprenants se trouvait accompagnée de l'option « On ne sait pas » que les enquêtés ont découvert pour la première fois. Certains d'entre eux nous ont alors confié qu'ils avaient interprété cet item comme « Je ne sais pas » et donc comme une non réponse et qu'ils avaient alors opté pour l'une des deux autres options (Vrai/Faux) pour avoir une chance sur deux de trouver la bonne réponse.

D'autres expériences menées plus tard nous ont permis de confirmer que les questions qui diffèrent de celles que nous venons d'énumérer font partie de celles qui posent le plus de problèmes aux apprenants.

Les questions n°4 et n°5 et qui comportent l'item « On ne sait pas » dans l'épreuve de compréhension des écrits (DALF C1)¹¹ sont celles qui ont posé le plus de problèmes à trente-quatre étudiants de 3^{ème} année de licence de français. Le dépouillement des résultats obtenus par nos enquêtés à la suite de la question n°5, notée sur 3 points, donne les statistiques suivantes :

Notes obtenues	Nombre d'étudiants	Pourcentages
0/3	11	32,35 %
1/3	10	29,41 %
2/3	09	26,47 %
3/3	04	11,77 %

Nous constatons que 60,76 % ont obtenu moins de la moitié de la note complète (0 ou 1/3). Par ailleurs, pour répondre correctement à cette question, qui consistait à choisir à trois reprises entre « Vrai », « Faux » et « On ne sait pas », il fallait répondre une fois sur trois par « On ne sait pas ». Mais comme cet item a été mal interprété, il n'a été utilisé que par 5 étudiants sur les 34 (dont les 4 qui avaient obtenu 3/3). Cela montre que l'échec des étudiants est essentiellement dû à leur refus d'utiliser « On ne sait pas », qui, tel qu'ils l'ont interprété, les priverait assurément d'une note, nous avaient confié plusieurs d'entre eux, puisqu'il s'agirait d'avouer qu'ils ne pouvaient pas répondre à la question.

Les deux exercices que comporte l'épreuve de compréhension du test du DELF B2 ont révélé que les questions qui portent sur le ton ou sur la position de l'auteur – auxquelles nos enquêtés, qui sont pourtant des étudiants en 3^{ème} et dernière année de licence de langue et littérature françaises, sont confrontés pour la première fois – posent problème :

- la réponse à la question : « *Quel est le ton des journalistes quand elles concluent : Surtout éviter d'y penser ? (Polémique/ironique/dubitatif)* » n'a été trouvée que par 9 étudiants sur 18 ;
- la réponse à la question « *D'après vous, l'auteur de cet article : prend position/reste neutre/se veut allusif/cherche la confrontation* » n'a été donnée que par 6 étudiants ;

¹¹Sujet disponible sur: https://www.delfdalf.ch/fileadmin/user_upload/Unterlagen/Kapitel_1.6/C1/c1_s_1-exemple1_correcteur.pdf

- dans les deux cas, quelques étudiants n'ont même pas répondu à la question (ils sont 2 dans le cas de la première et 4 dans celui de la seconde) ;
- il est possible qu'au moins certains, parmi ceux qui ont trouvé les deux réponses, l'aient fait simplement par hasard, vu qu'il s'agit de questions à choix multiples.

3.3. Les supports et les thématiques

La plus importante des contraintes auxquelles font face les candidats algériens – mais sans aucun doute aussi ceux de plusieurs autres contextes – qui ont affaire aux épreuves du CIEP, résulte de la nature des supports et des thématiques dont il s'agit. Ceux-ci concernent davantage l'environnement français et sa culture, ce qui ne va pas sans poser de nombreux problèmes aux étrangers qui se trouvent de ce fait désavantagés.

Dans l'enseignement du français au moyen (collège) et au secondaire (lycée) les supports qui dominent dans les leçons qui travaillent la compréhension sont des textes ou extraits de textes (littéraires, journalistiques, encyclopédiques, procéduraux, etc.), approchés conformément à la logique des types de textes. Or, nous constatons qu'il arrive souvent que les épreuves proposées par le CIEP portent sur des genres de discours qui n'y sont pas enseignés (e-mails, petites annonces, enseignes, publicités, etc.). Sur ce plan, nous pensons que les genres sur lesquels portent ces évaluations correspondent souvent très bien à ceux que suggèrent le plus les descripteurs des échelles de niveaux du C.E.C.R. L'erreur incomberait davantage aux concepteurs des programmes qui adoptent une logique aujourd'hui dépassée ou du moins qui n'est pas conforme aux pistes issues du C.E.C.R.

Par ailleurs, dans la conclusion de notre première étude (Ammouden et Ammouden, 2010), il a été mentionné que la majorité des tests du CIEP propose des supports qui ne s'inscrivent pas dans le contexte social des publics auxquels ils sont destinés : bien que ces tests soient essentiellement destinés à des publics étrangers au contexte français, les supports portent essentiellement sur ce contexte.

Pour confirmer cette contrainte, nous avons mené cette fois une autre expérience, en nous appuyant sur un test de DELF A1, pris sur le site du CIEP. L'exercice n°3 de la compréhension orale (p. 3 du test) consiste à entendre six courts dialogues, puis à associer chacun d'eux à une image. Les dialogues concernent respectivement une visite en autobus, un requin dans un aquarium, un parc, les déguisements de deux clowns, l'achat de fruits et le métro. Les situations qui ont posé problème sont la première et la sixième :

- l'image à associer au premier dialogue comporte essentiellement trois indices : l'autobus, des personnes se trouvant sur sa « terrasse » (mais qu'on ne distingue pas bien à cause des constructions qui apparaissent dans le fond de la photo) et l'écriteau « L'Open Tour ». Ces indices pouvaient être mis en relation avec la réplique « mais quand il ne pleut pas » par laquelle se termine le dialogue. Le fait que nos enquêtés ne savent pas que ce genre d'autobus existe et qu'ils n'ont pas pu interpréter l'écriteau les a conduits à associer le dialogue à l'image qu'ils devaient associer au dialogue concernant le métro ;
- dans le cas du sixième dialogue, il s'agissait notamment de rentabiliser l'écriteau (ou mieux la silhouette) de l'enseigne contenant le mot Métropolitain (mais qu'on ne peut deviner que si l'on a déjà vu cette enseigne). Le fait que nos enquêtés n'aient jamais visité la France a fait que la réponse n'a souvent pu être trouvée que par élimination.

Il arrive certes que les concepteurs des sujets des évaluations qui nous concernent pensent à des consignes qui peuvent convenir à des publics étrangers. C'est par exemple le cas des consignes de productions écrites de DELF B1 qui consistent à décrire ou à raconter sa relation avec l'un des grands-parents ou d'une autre personne amie de la famille (B1 junior) ou à répondre à la question « *À votre avis, quels ont été le ou les changements les plus importants des vingt dernières années dans votre pays ? (...)* » (Chevallier-Wixler, Dayez, Lepage et Riba, 2006 : 112). C'est également le cas de la production écrite du DELF A2 junior (exercice 2), à travers laquelle on demande de rédiger une lettre pour décliner une invitation à visiter Paris. La prise en compte du candidat étranger est explicite. On lit dans la lettre « reçue », la phrase suivante : « *Je sais que tu ne connais pas Paris, alors je te propose de venir passer quelques jours chez moi* ».

Cela dit, on a fréquemment affaire à des consignes qui s'avèrent très contraignantes et désavantageuses pour les candidats qui ne connaissent pas le contexte français et sa culture. C'est le cas de la consigne de production écrite ci-après qu'on trouve dans l'épreuve du DELF A2 junior (exercice 1) : « *Vous avez fait un voyage d'une semaine à Paris, vous racontez dans votre journal personnel ce que vous avez fait* ». Même si la consigne est accompagnée de trois images au bas desquelles sont indiqués les écriteaux « La Seine », « Cathédrale de Paris » et « Château de Versailles », l'activité s'avère très contraignante. Les difficultés découlent du fait qu'on demande à ces candidats :

- de produire un texte qui appartient à un genre de discours qu'ils ne connaissent pas et n'étudient pas, à savoir le journal personnel ;
- d'« imaginer », un voyage qu'ils n'ont pas fait dans une ville (Paris) qu'ils ne connaissent pas ;
- de parler d'endroits (et d'une culture) au sujet desquels ils ne savent pas forcément beaucoup, ce qui est notamment le cas de la « Cathédrale » de Paris ;
- de mettre en relation au moins les trois endroits mentionnés, sans savoir par exemple s'ils sont géographiquement proches ou non.

4. Conclusion

Nous nous sommes intéressé à l'une des problématiques qui est au cœur de la sociodidactique, à savoir la prise en compte des spécificités des contextes dans les outils didactiques qui se veulent universels : les certifications internationales. Ainsi, avons-nous choisi de nous interroger d'une part sur les possibilités qu'offre l'échelle des niveaux de compétences proposée dans le C.E.C.R., et adoptée dans les évaluations de plusieurs pays ; et d'autre part, sur la place accordée à ces possibilités dans quelques-unes des évaluations émanant du CIEP de Paris.

La première partie de l'étude a révélé que ces possibilités ne sont pas seulement offertes par l'échelle des 6 niveaux de compétences du C.E.C.R., mais sont même au cœur de la logique qui la sous-tend, dans la mesure où les descripteurs qui définissent ces niveaux sont souvent intimement liés à l'environnement (immédiat) de l'apprenant ou de l'utilisateur de la langue.

La seconde partie nous a permis de mettre en évidence, grâce à des expérimentations réalisées avec des étudiants à qui nous avons proposé divers sujets de DELF et de DALF, que ces évaluations ne prennent en considération que très rarement les spécificités de ces publics et de leurs contextes. Ils sont, de ce fait, confrontés à de nombreuses contraintes découlant particulièrement de la nature des consignes et des thématiques auxquelles ils ont affaire.

Ces résultats devraient permettre de considérer – ou du moins de supposer – que le niveau des Algériens en langue française peut s'avérer être plus élevé que celui dans lequel ils sont classés à la suite de ces évaluations. En définitive, tant que les concepteurs des évaluations internationales continuent à se focaliser sur des exercices qui sont davantage destinés à des publics qui connaissent l'environnement et la culture français, ils ne parviendront pas à adopter vraiment les recommandations du C.E.C.R. Ils ignoreront notamment l'importance qui y est accordée à la prise en compte des spécificités du contexte et de l'environnement immédiat de l'utilisateur ou de l'apprenant (cf. supra sous-titre 2.2.). Ils continueraient à pénaliser les candidats étrangers, et leurs évaluations ne seraient pas très conformes à deux des critères professionnels de l'évaluation, à savoir la fiabilité et l'équité (Noël-Jothy et Sampsonis, 2006 : 42).

Il est temps que les concepteurs des évaluations dont nous traitons pensent à contextualiser leurs épreuves pour tenter de tenir compte davantage de la diversité des contextes, des didactiques et des cultures éducatives des différents candidats qui les subissent. Associer des didacticiens étrangers qui connaissent les spécificités de leurs contextes peut y contribuer largement.

Références bibliographiques

- Abbes-Kara, A.-Y., Kebbas, M. et Cortier, C. (2013). Aborder autrement les pratiques langagières plurilingues en Algérie ? Vers une approche de la complexité. Dans R. Colonna, A. Becetti et P. Blanchet (dir.), *Politiques linguistiques et plurilinguistiques : Du terrain à l'action glottopolitique* (p. 177-195). Paris : L'Harmattan.
- Ammouden, A. et Ammouden, M. (2010). Le Français à l'université et l'échelle des compétences du Cadre Européen Commun de Référence pour les Langues. *Synergies Algérie*, 9, 37-44.
- Ammouden, M. (2010). L'apprentissage de l'oral : pour une approche «globalisante» et intégrée. *Didacstyle*, 3, 55-64.
- Ammouden, M. (2013). Place de l'articulation oral/écrit dans l'enseignement du français dans le secondaire algérien. *Synergies Mexique*, 3, 31-44.
- Ammouden, M. et Cortier, C. (2009). L'enseignement du FLES et de l'arabe dans le contexte algérien : réflexions pour une didactique comparée à l'épreuve de la complexité et de l'interculturalité. Dans F. Leutenegger et al. (dir.), *Actes du 1^{er} Colloque International de l'ARCD « Où va la didactique comparée ? Didactiques disciplinaires et approches comparatistes des pratiques d'enseignement et d'apprentissage »*, [CDROM]/ Edition : Université de Genève FPSE-SSSED et ARCD.
- Beacco, J.-C. (2015). Contextualiser les savoirs en didactique des langues et des cultures. Dans P. Blanchet et P. Chardenet (dir.), *Guide pour la recherche en didactique des langues et des cultures* (p. 53-62). Paris : Archives contemporaines.
- Chevalier-Wixler, D., Dayez, Y., Lepage, S. et Riba, P. (2006). *Réussir le DELF : niveau B1 du Cadre européen commun de référence*. Paris : Didier/CIEP.
- Conseil de L'Europe (2001). *Un cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer*. Paris : Didier.
- Cortier, C. (2009). Propositions sociodidactiques face à la diversité des contextes. Vers une didactique convergente des langues collatérales et de proximité : éducation bi/plurilingue et projets interlinguistiques. *Synergies Italie*, 5, 109-118.
- Dupoux, B., Havard, A.-M., Martial, M. et Weeger, M. (2006). *Réussir le DELF : niveau B2 du Cadre européen commun de référence*. Paris : Didier/CIEP.

- Halté, J.-F. (2005). Intégrer l'oral : pour une didactique de l'activité langagière. Dans J.-F. Halté et M. Rispaïl (dir), *L'oral dans la classe : Compétences, enseignement, activités*, (p. 11-31). Paris : L'Harmattan.
- Mottier Lopez, L. et Laveault, D. (2008). L'évaluation des apprentissages en contexte scolaire : développements, enjeux et controverses. *Mesure et évaluation en éducation*, 31(3), 5-34.
- Noël-Jothy, F. et Sampsonis, B. (2006). *Certifications et outils d'évaluation en FLE*. Paris : Hachette.
- Rispaïl, M. et Blanchet, P. (2015). Principes transversaux pour une sociodidactique dite « de terrain ». Dans P. Blanchet et P. Chardenet (dir), *Guide pour la recherche en didactique des langues et des cultures* (p. 101-106). Paris : Archives contemporaines.

Pierre angulaire de l'enseignement des langues étrangères dans un environnement d'apprentissage virtuel : l'évaluation

Ana Maria de JESUS FERREIRA NOBRE

LE@D¹ – Universidade Aberta, Portugal

Résumé

Le concept d'évaluation en langues étrangères nécessite une redéfinition liée non seulement au modèle traditionnel (évaluation des connaissances) mais aussi à l'intégration des technologies tant dans l'enseignement/apprentissage que dans l'évaluation (évaluation des compétences). Évaluer les langues étrangères dans ses modalités orales dans l'enseignement supérieur à distance au Portugal revêt un caractère particulièrement sensible. Le passage de l'évaluation **des** apprentissages à l'évaluation **pour** les apprentissages est une expérience qui a été vécue à l'Universidade Aberta. Des *podcasts* aux *showcasts*, le chemin est tracé. Nous passons de l'approche d'évaluation des connaissances à une approche qui vise le développement des compétences, si nécessaires aux étudiants de ce siècle. Le modèle que nous présentons a fait ses preuves.

Mots-clés

Évaluation, langues étrangères, enseignement / apprentissage à distance.

Abstract

The concept of evaluation in foreign languages learning needs to be redefined in relation not only to the imperative revision of the traditional model (knowledge assessment) but also to the integration of technologies in both teaching and assessment practices (skills assessment). Assessing oral skills in a foreign language, in the context of distance higher education in Portugal, is a sensitive matter. The transition from the evaluation **of** learning to the evaluation **for** learning has been experienced at the Universidade Aberta (Portugal). From podcasts to showcasts, the way forward has been marked out. We are moving from a knowledge assessment approach to an approach that seeks the development of the skills that are crucial to students in the 21st century. The model presented in this article has already proven its worth.

Keywords

Assessment, foreign languages, distance learning / teaching.

¹ Laboratório de Educação a Distancia e eLearning.

1. Introduction

« Les technologies sont comme des instruments de musique joués par toute une culture à un moment précis » (Kerckhove, 1997 : 213). La première étape dans la production d'un son harmonieux est de connaître les instruments, les utiliser et les mettre au service de la communauté scientifique, à tous les niveaux du monde éducatif. Selon Figueiredo (2000), le nouveau défi est de construire des situations en contexte où les apprenants construisent leurs propres connaissances, leurs propres savoirs. Il est donc nécessaire de changer l'évaluation où le « critère d'excellence, hiérarchique, compétitif, basé sur l'agressivité par rapport à l'autre » (Kerckhove, 1997 : 100) se traduira par une appréciation de l'interactivité et de la collaboration. Compte tenu de la nécessité d'un changement de vision à propos de l'évaluation et d'une adaptation progressive aux nouvelles technologies, nous avons défini l'évaluation des langues étrangères selon le Modèle Pédagogique Virtuel (MPV) de l'Universidade Aberta (UAb) du Portugal et selon les compétences du Cadre Européen Commun de Référence pour les Langues (CECRL). Ainsi, cet article est constitué de quatre points. Dans un premier point, nous souhaitons ouvrir le débat sur l'évaluation des compétences en montrant les changements qui concernent les méthodes d'évaluation utilisées dans le processus d'apprentissage. Au point suivant, nous présenterons les répercussions du changement d'objet d'évaluation sur l'apprentissage : le passage de l'évaluation des connaissances vers l'évaluation des compétences ; le passage des évaluations des apprentissages vers l'évaluation pour les apprentissages. Dans le troisième point, nous décrirons l'impact des évaluations sur le processus d'apprentissage des étudiants en langue étrangère et les modes d'évaluation utilisés dans une classe virtuelle. Finalement, dans le quatrième et dernier point, nous montrerons que les évaluations ainsi dégagées conduisent à un jugement sur les apprentissages, qui permettent de recueillir, d'analyser et d'interpréter des données en vue de répondre à des décisions pédagogiques.

2. L'évaluation des compétences

Le nouveau système éducatif résultant de la construction de l'Espace Européen de l'Enseignement Supérieur (EEES) valorise la mobilité des étudiants et la compatibilité des cursus universitaires. Entre autres, ce scénario européen vise à former des professionnels et des citoyens, au-delà des connaissances techniques et scientifiques. Ces citoyens du siècle naissant doivent acquérir un ensemble de compétences à mettre en œuvre face à des situations professionnelles complexes, certaines étant encore à découvrir. Dans ce contexte, l'accent est mis sur des compétences générales. Habituellement, on préconise le développement de compétences liées au travail d'équipe : organiser, communiquer, rechercher et traiter l'information, l'analyse et la synthèse et l'utilisation d'outils technologiques. Dans une perspective holistique et globale, le projet DeSeCo (*Directorate for Education, Employment, Labour and Social Affairs*, 2002) met en évidence trois catégories de compétences clés :

- agir de manière autonome ;
- lier l'autonomie aux traits identitaires de chaque étudiant ;
- utiliser des outils physiques et socio-culturels de manière interactive et « travailler » dans des groupes socialement hétérogènes.

Cette catégorisation de compétences oblige à des changements profonds dans l'organisation des cursus proposés par les universités. L'accent est mis sur les étudiants et leur apprentissage, ainsi que sur les compétences qu'ils doivent développer. Tout ceci remet en question le modèle centré sur l'enseignant. Cette option d'orientation opère non seulement une modification sur le choix des pédagogies utilisées, mais aussi les stratégies d'évaluation à

mettre en place. Il est reconnu qu'une évaluation unique ne suffit pas à évaluer toutes les compétences d'un individu. Comme « une compétence est un savoir agir complexe prenant appui sur la mobilisation et la combinaison efficaces d'une variété de ressources internes et externes à l'intérieur d'une famille de situation » (Tardif, 2003), il est nécessaire de remplacer, la « culture de test » par la « culture de l'évaluation » (Birenbaun, 1996 ; Dierick et Dochy, 2001). La « culture de l'évaluation » a un regard critique sur l'utilisation des tests² comme la seule forme valable d'évaluer l'apprentissage et sur l'importance attribuée à ces derniers dans la planification de l'enseignement et de l'apprentissage. Elle met aussi l'accent sur les acteurs et les processus d'évaluation et sur l'utilisation de multiples techniques, de multiples moments. L'évaluation de compétences requiert une approche où connaissances, capacités et attitudes doivent être mobilisées pour des savoir-agir complexes. Ainsi, le recours à une variété de stratégies d'évaluation se revêt d'évidence. Les environnements d'apprentissage dans l'enseignement supérieur à distance exigent un déplacement dans l'axe paradigmatique évaluatif : concevoir des modèles numériques évaluatifs différents de ceux de l'enseignement traditionnel. C'est dans ce contexte qu'apparaît la Culture de l'Évaluation instaurée à l'Universidade Aberta.

L'Universidade Aberta (UAb) est une université publique de l'enseignement supérieur, dont le siège principal se situe à Lisbonne. Elle est l'institution pionnière de l'enseignement à distance (enseignement online et e-learning) au Portugal. En 2007, l'UAb adopte un modèle pédagogique (Pereira *et al.*, 2007) et tous les cours formels ou non-formels sont enseignés virtuellement sur la plateforme Moodle : classe et salle virtuelle. En termes généraux, le modèle de l'UAb préconise, entre autres, un apprentissage centré sur l'apprenant, la flexibilité de l'apprentissage, l'inclusion des technologies de l'information et recommande fortement aux étudiants de privilégier un travail régulier. L'étudiant a « [...] un rôle actif dans la gestion temporelle de ses activités, dans le suivi des apprentissages réalisés, dans les objectifs de travail et dans la communauté d'apprentissage modérée par les étudiants eux-mêmes » (Pereira *et al.*, 2007). Ainsi, trois principes façonnent la pédagogie online « qui prend en charge les activités éducatives du modèle de l'UAb » (Pereira *et al.*, 2007) à savoir : la primauté de la flexibilité, la prééminence de l'interaction et le principe de l'inclusion numérique. La flexibilité est entendue par les auteurs de ce modèle comme « [...] l'étudiant peut apprendre, quand il le souhaite, indépendamment de la distance, n'importe où, sans l'embarras d'un horaire ; le profil de l'étudiant de l'UAb : adultes, avec des responsabilités professionnelles, familiales, civiques, des citoyens actifs et acteurs de la société ». Enseigner et apprendre à l'UAb, c'est accepter des défis. Dans le contexte des langues étrangères, le champ de notre action pédagogique et de réflexion, les provocations sont peut-être plus grandes, en particulier dans les modalités orales. Alors, comment évaluer les compétences de communication orale ?

3. Culture de l'évaluation : système conservateur, structuré, normalisé ou système centré sur les étudiants ?

La Culture du Test qui remonte aux pratiques mises en place à la fin du XVIII^{ème} siècle dans les institutions scolaires publiques, exprime aujourd'hui son insuffisance. Pour démontrer l'apprentissage, la réalisation d'une épreuve écrite (examen, test, etc.) préfigure un des moyens utilisés pour vérifier les acquisitions. D'une façon générale, ces épreuves écrites servent à légitimer la propre institution dans le sens où il existe la matérialisation de

² La culture du test renvoie au choix des interrogations écrites classiques et autres modalités de vérification écrites d'acquisition de connaissances.

l'évaluation et ainsi une certaine reconnaissance de la société. Force est de reconnaître que la Culture du Test évalue simplement des connaissances et non des compétences. La Culture de l'Évaluation, quant à elle, n'évalue pas seulement des connaissances, elle va plus loin, par sa perception plus large et intégrante, elle bouleverse les méthodologies et les instruments. Au-delà de leur participation aux activités, il est demandé à chaque étudiant d'intégrer une réflexion sur les démarches mises en place. Dierick et Dochy (2001) soulignent justement ce changement de paradigme, quand :

Several developments in society have slowly but surely changed the way we value the importance of having a firm knowledge store. Surely, highly knowledgeable individuals are a starting point of education but cannot be its goal ; other aspects, such as problem-solving skills, professional skills and being able to function in authentic, i.e., real-life contexts are important (Dierick et Dochy, 2001 : 307).

Selon ces auteurs, la Culture du Test implique que les étudiants assimilent des connaissances et les mémorisent dans le but de les montrer au moment de la réalisation de tests spécifiquement conçus pour vérifier les résultats de leurs efforts. La Culture du Test fournit des résultats ponctuels, mais elle ne nous donne qu'une idée incomplète de toute la dynamique mise en place pour que l'apprentissage s'effectue réellement.

De leur côté, et dans le cadre de la Culture de l'Évaluation, les professeurs cherchent à analyser non seulement les produits, mais aussi les processus. Pour être compétent, un étudiant doit démontrer qu'il peut mobiliser ses ressources internes (connaissances, habiletés, attitudes) et externes (ouvrages de références, collègues, etc.) pour agir dans différents contextes. Ainsi, l'emphase est placée sur l'observation des compétences à atteindre grâce à la façon dont l'étudiant applique les connaissances assimilées pour résoudre des problèmes dans des contextes différents de ceux qu'il connaît déjà et comment il fait face aux situations qu'il rencontre au quotidien. En fonction de ces deux représentations de l'évaluation, le rôle du professeur et de l'étudiant change. L'apprentissage cesse ainsi d'être centré sur le professeur et évolue vers une centration sur l'étudiant qui devient responsable de la gestion de son processus d'apprentissage. Le professeur doit guider et orienter l'étudiant dans ce processus en lui fournissant des tâches susceptibles de le défier. Mclellan (2004) affirme que :

The dominance given to students' interpretations of their world is well suited to formative assessment, which is concerned with the facilitation of learning. However, higher education must also be concerned with summative assessment for reasons of accountability and certification (Mclellan, 2004 : 319).

L'évaluation continue tendancielle à privilégier la composante sommative au détriment de la composante formative (voir Tableau 1).

	Évaluation sommative	Évaluation formative
Fonction / Pourquoi ?	Dresse un bilan des connaissances	Intègre toutes les situations d'apprentissage
Moment / Quand ?	Fin de l'apprentissage	Effectuée au cours des apprentissages Participation active des étudiants
Moyen / Comment ?	Attribution d'une note par rapport aux critères	Des occasions de production individuelle ou collective librement choisies
Note ?	Oui	Non. Une note peut être donnée mais à titre indicatif

Tableau 1 : Critères d'évaluation sommative et formative

Procéder à une analyse continuée et réajuster sa démarche pédagogique et l'évaluation en fonction des savoirs déjà assimilés requièrent plus de temps et plus de respect pour le serment de Socrate³ que prêtent les professeurs en Belgique. Il est traditionnellement plus facile de faire correspondre une valeur numérique finale à un ensemble d'acquisitions de connaissances. Dans ce contexte, s'agit-il alors de mettre en place une évaluation **pour les** apprentissages au lieu d'une évaluation **des** apprentissages ? Birenbaum *et al.* (2006) considèrent que :

current assessment practices tend to focus on the assessment, or testing, of learning. These assessments largely fail to address Assessment for learning, i.e. assessment as a means to measure learner progression and to inform the learners about their progression. In other words, a paradigm shift from Assessment of learning towards Assessment for learning is required (Birenbaum *et al.*, 2006 : 63).

La dénomination « Évaluation **des** Apprentissages » indique un processus terminé qui se maintient de façon pérenne dans le temps. Elle implique une évaluation postérieure et une position extérieure au processus. En contrepoint apparaît la perfectibilité inachevée de l'« Évaluation **pour les** Apprentissages » qui semble orienter vers une action en cours et continue dans le temps tout au moins tant que la situation d'apprentissage persiste. Ainsi, de nouveau selon Birenbaum *et al.* (2006) :

accordingly, assessments tend not to be designed as assessment for learning, but as assessment of learning. In other words, assessment systems tend to be summative and not formative. These assessment systems do not allow learners (and their teachers) to develop a clearer understanding of how they can improve in their learning. Instead, they tend to be considered as an endpoint instead of a beginning or a step forward (Birenbaum *et al.*, 2006 : 63).

L'Évaluation **des** apprentissages fait ainsi référence à un processus posthume et nous semble bien incomplète. En réalité, ce qui est évalué n'est qu'une proposition d'enchaînement de connaissances sans offrir des solutions pour combler leur non-développement dans le cas où les résultats obtenus ne correspondraient pas au niveau espéré.

L'Évaluation **pour les** apprentissages – évaluer **pour** apprendre (évaluer pour évoluer) – nous semble donc être un processus bien plus vaste et complexe. Elle présuppose un procédé continu ou une attention continuée de façon à permettre un accompagnement tout au long de la situation. Des rétroactions régulières sur le niveau de développement des compétences sont déduites afin que d'éventuelles corrections puissent être incorporées. Birenbaum *et al.* (2006 : 63) réfèrent que « *Assessment for learning are assessment systems that allow both learners and teachers to gain information about learning progression* ».

Au niveau des apprentissages, il doit exister un ensemble de principes (une implication accrue de l'étudiant dans les processus d'apprentissage et d'évaluation ; des situations d'apprentissage « authentiques » significatives pour les étudiants) qui valorisent les acquisitions et les productions des étudiants et ceci dans un environnement propice aux processus de développement personnel et social afin de favoriser la réussite des étudiants. À cet ensemble de principes s'en ajoutent d'autres plus sous-jacents : le respect des différences individuelles ; la valorisation des expériences antérieures ; la considération des intérêts des étudiants ; les motivations et les besoins individuels ; la promotion des interactions ; le partage des savoirs ; l'accession à l'autonomie et à la décision. De tout cela découlent des stratégies d'évaluation

³ « Je m'engage à mettre toutes mes forces et toute ma compétence au service de l'éducation de chacun des élèves qui me sera confié ». Voir : <http://www.agers.cfwb.be>.

qui régulent les processus d'apprentissage tout en définissant une fonction régulatrice de la relation pédagogique.

Les évaluations doivent ainsi être ajustées aux contenus et/ou aux dimensions à évaluer, diversifiées dans leur forme et leur nature et porter sur les processus, les compétences et les acquisitions. Une évaluation excessivement centrée sur les produits de l'apprentissage, majoritairement sommative ne sert qu'à certifier l'acquisition des connaissances de l'étudiant mais n'informe pas, ne guide pas et n'ajoute aucune valeur à l'apprentissage proprement dit. Il est pourtant souhaitable qu'il existe des stratégies d'évaluation qui orientent, tant l'étudiant que le professeur, c'est-à-dire, des techniques qui leur indiquent, à tous les deux, les progrès réalisés par l'étudiant et qui servent d'aiguillage au professeur pour orienter les apprentissages.

4. Évaluation dans un environnement numérique

Le scénario éducatif dans un environnement numérique est marqué par la croissante évolution des nouvelles technologies obligeant les personnes, les entreprises et les universités à communiquer différemment entre elles. Il faut donc créer des stratégies de production, des outils technologiques et promouvoir aussi le développement des capacités qui permettent l'adaptation à ces changements. L'utilisation croissante du e-Learning, les perspectives constructivistes accompagnées par des modèles d'apprentissage distribués sur la toile (par exemple les *Massive Open Online Course*) incitent l'étudiant à avoir un rôle plus actif au sein des communautés d'apprentissage et conduisent à l'introduction de techniques d'évaluation numérique, à savoir l'évaluation de tâches accomplies dans et par les groupes, l'évaluation par les pairs et l'auto-évaluation. L'inévitabilité temporelle nous oblige à imaginer et à concevoir de nouvelles évaluations dans les environnements virtuels de l'enseignement supérieur.

L'utilisation de nouveaux outils de communication en ligne qui allient le son et l'image – skype, hangout du google... – et le recours à de nouvelles techniques d'évaluation numérique – les « tests » (compréhension et production écrite), l'enregistrement de podcast (production orale) et le *Colibri* (interaction orale) sur la plateforme Moodle – représentent un défi pour les professeurs de l'Universidade Aberta (UAb) dans le domaine scientifique des langues étrangères. La mise en œuvre de ces outils numériques impose un processus d'actualisation constant. Dans le contexte de l'enseignement des langues étrangères à l'UAb, dont les piliers de base sont le Cadre Européen Commun de Référence pour les Langues (CECRL) et l'actuel Modèle Pédagogique Virtuel (MPV) de l'institution (Pereira *et al.*, 2007), nous présentons quelques-unes de ces stratégies et solutions d'évaluation numérique de la compétence orale, la compréhension, la production et l'interaction.

Le *Colibri* est un environnement collaboratif multimédia pour la communauté académique et scientifique du Portugal qui permet de réaliser des vidéoconférences et de conserver les enregistrements disponibles à tout moment.

Le CECRL – Apprendre, Enseigner, Évaluer (2001) est un document qui émane du Conseil de l'Europe dans le cadre du Projet « Politiques Linguistiques pour une Europe Plurilingue et Multiculturelle ». Ce document présente dans un premier temps le contexte politique et éducatif de sa conception, il définit aussi des lignes d'orientation et l'adoption de l'approche méthodologique, mais surtout, il propose la définition de six niveaux communs de référence pour trois grands types d'utilisateurs : l'utilisateur élémentaire, l'utilisateur indépendant et l'utilisateur expérimenté. Il s'agit d'un document de référence dont la prétention est

d'appuyer la création d'environnements qui facilitent un apprentissage motivant et proche des situations réelles de communication. D'autre part, suivant le Modèle Pédagogique Virtuel de l'UAb (Pereira *et al.*, 2007) mis en place à partir de l'année scolaire 2007-2008 et intégré dans le nouveau paradigme dynamique et innovateur de l'institution, toutes les unités de valeur, et en particulier celles de langue étrangère, ont pu inclure et mettre effectivement en pratique l'évaluation numérique. Quand nous nous penchons sur la problématique de l'évaluation, il nous semble pertinent de nous interroger sur les différences et les ressemblances entre l'évaluation dans l'enseignement à distance et l'enseignement présentiel, sur les instruments et les méthodes les plus adéquats, ainsi que sur la meilleure façon de déterminer les indicateurs des compétences développées ou à développer. En fait, nous ne pouvons pas ne pas réfléchir à la nécessité d'allier l'évaluation formative, il est vrai plus motivante pour les étudiants, à une autre facette tout aussi indispensable, la certification des compétences acquises qui, dans un environnement d'apprentissage à distance, relève de plusieurs challenges.

Dans l'enseignement à distance où nous nous insérons, il existe des possibilités qui ne se vérifient pas en présentiel :

- l'enregistrement de toutes les interventions des étudiants autant en compréhension écrite que orale est nécessaire dans l'enseignement à distance et globalement peu utilisé dans l'enseignement en présentiel ;
- l'élaboration d'activités et de tâches diversifiées : exercices « Vrai-Faux, Complétez..., Développer la phrase ou développer l'idée » ;
- l'utilisation d'outils numériques sous forme de logiciels pour diversifier les modes, les processus et les produits relatifs à l'évaluation ;
- recours à des supports numériques en ligne : exemple TV5 monde.

Le principe qui oriente l'évaluation devrait être à la fois un moyen d'encourager l'apprentissage et une façon de l'atteindre. Les étudiants engagés et inspirés dans leurs activités d'évaluation, et qui reçoivent un retour rapide et significatif du professeur, s'engagent plus facilement dans les défis proposés et dépassent plus aisément aussi leurs difficultés.

À l'Universidade Aberta, les cursus formels de Licence en Langues, Littératures et Cultures – Langues Etrangères, Langues Appliquées, Études Européennes, Sciences Humaines dont Histoire – intègrent dans leurs plans d'études les langues étrangères qui suivent : du niveau I au niveau VI en allemand, espagnol, français et anglais. Ces six niveaux suivent le Cadre Européen Commun de référence pour les Langues Etrangères du Conseil de l'Europe (CECRL) déjà mentionné : le I et le II correspondent au niveau B1, le III et le IV correspondent au niveau B2 et le V et le VI correspondent au niveau C1.

5. Évaluation dans les Unités de Valeur de langue étrangère

L'évaluation en langues étrangères, dénommée contrôle continu, s'effectue en deux temps et la classification finale est le résultat de la somme des deux notes obtenues. Le premier correspond à la réalisation, sur la plateforme d'apprentissage, de deux tâches évaluatives (évaluations formatives) en support numérique appelées *e-Fólios*. Pour le deuxième, il s'agit de la réalisation d'une épreuve écrite en présentiel en fin du semestre et appelée *p-Fólio*. À chaque *e-Fólio* est attribué un maximum de 4 points. Comme le contrôle continu est obligatoire, il a été décidé que chaque étudiant devait obtenir une note minimum (supérieure ou égale à 3,5) dans les deux évaluations numériques, c'est-à-dire les deux *e-Fólios* réalisés

sur la plateforme Moodle, pour pouvoir réaliser le *p-Fólio* et ainsi conclure l'Unité de Valeur (discipline) de langue étrangère avec succès. Le *p-Fólio*, c'est-à-dire l'épreuve écrite en présentiel, est réalisé pendant la période officielle d'examens de l'UAb et représente 12 points sur 20. Les évaluations formatives numériques demandées dans les *e-Fólios* ont permis à l'UAb de mettre en place le caractère obligatoire d'une évaluation numérique orale en langues étrangères. Cette expertise orale est réalisée en compréhension et en production dans tous les cursus formels et à tous les niveaux de langue étrangère.

Pour faciliter cette mise en place, des *Kits* d'orientation ont été élaborés pour les professeurs et pour les étudiants. La durée (de 4 à 12 heures) des *e-Fólios* est définie, ainsi que l'agencement des dates de leur réalisation dans les différentes langues et différents niveaux, afin d'éviter les superpositions. La concrétisation d'une évaluation orale numérique a ainsi pu être mise en place « partout dans le monde » puisque la distribution géographique des étudiants de l'UAb est planétaire.

L'énoncé d'un *e-Fólio* est normalement composé de deux parties distinctes. La première partie contient un *podcast* (fichiers audio) ou un *showcast* (fichiers multimédia didactiques) dont l'auteur est l'enseignant ou alors une autre ressource choisie sur la toile en accord avec les thèmes abordés dans la classe virtuelle. L'évaluation de la compréhension orale et de la production écrite est réalisée à partir de ces ressources. Dans la deuxième partie, il est demandé à l'étudiant de réaliser un document (audio et/ou vidéo) sur un thème d'actualité et qui permet l'évaluation de la production orale.

Si pour l'enseignement en général les *podcasts* sont des ressources très utiles, pour l'enseignement des langues étrangères leur intérêt est encore plus grand. La plupart des étudiants a peu de contact avec la langue étudiée. Au Portugal, même si les films, feuilletons et séries télévisées ne sont pas doublés en portugais, la cinématographie en anglais domine. Le public en général et les étudiants en particulier ont ainsi un contact raisonnable avec la langue anglaise. Ce n'est pas le cas pour le français ou les autres langues étrangères enseignées à l'UAb et les *podcasts* peuvent donc jouer un rôle essentiel. Sur la toile numérique, on peut trouver bon nombre de documents audio authentiques en français par exemple. Comme ils sont très différents des documents fabriqués ou didactisés, ils permettent une approche plus juste de la réalité des locuteurs natifs dans différents contextes. Cette authenticité que nous appelons « français véritable » permet à l'étudiant de langue de comprendre que le natif peut, comme lui, se tromper, chercher un mot, hésiter et ceci non seulement le rassure pour la suite de son apprentissage, mais lui permet aussi d'accéder au langage non formel utilisé par le locuteur natif. Pour ce qui est des *showcasts*, en tant que ressources didactiques, il faut reconnaître leur usage récent dans l'enseignement des langues étrangères et en particulier dans l'enseignement à distance. Suivant la définition élaborée par le « Projet d'Enseignement/Apprentissage des Langues Etrangères dans l'enseignement à distance » du centre « Laboratório de Educação a Distância e eLearning » (LE@D), le terme *showcast* désigne spécifiquement un document *PowerPoint* à finalité didactique et qui intègre texte, image et audio/vidéo. Dans le concept de *podcast*, nous pouvons considérer que le *showcast* est un type spécifique de *podcast*, tout comme les *audio podcasts*, les *enhanced podcasts*, les *vodcasts* et les *screencasts*. Comme le *showcast* utilise des caractéristiques propres, dont celle d'être gratuit, nous le situons dans une autre catégorie.

Le *showcast* est une ressource très utile dans l'enseignement à distance des langues étrangères car il assure le développement de différentes compétences linguistiques et technologiques. Son utilisation permet l'accès à une multitude d'exercices et de tâches numériques, de la reproduction écrite d'une information orale, identique à une dictée (comme dans l'activité où les étudiants doivent écrire les questions d'interprétation qu'ils entendent dans le *podcast* proposé à l'argumentation, souvent demandé dans les *e-Fólios*), en passant par la présentation de réponses à des questionnaires ou l'exposition d'opinions et de points de vue. Ses caractéristiques en font donc un instrument qui peut être utilisé tant dans la pratique de la langue que dans les moments plus formels de l'évaluation numérique de l'oralité.

Comme nous l'avons montré, l'enseignement d'une langue étrangère relève d'un processus complexe qui recouvre un ensemble de compétences de compréhension et de production écrites et orales. Définir d'autres stratégies d'évaluation facilite la pratique systématique de l'oralité et une évaluation numérique adéquate. Notre objectif principal à l'UAb, a été d'ouvrir le chemin de l'Évaluation **pour** les Apprentissages. À l'aide de *podcasts* et de *showcasts*, il est possible de faire apprendre les langues étrangères en créant des situations d'apprentissage. En 2015-2016, un questionnaire a été soumis à tous les étudiants de langues étrangères et les résultats obtenus démontrent qu'ils ont confiance en leur savoir et en leurs compétences pour gérer eux-mêmes leurs apprentissages. Ils développent leurs compétences en matière d'auto-évaluation et d'évaluation entre pairs et affinent l'adaptation de leurs stratégies d'apprendre à apprendre.

Nous avons pu vérifier que l'utilisation de ces outils requiert un entraînement constant qui mène à un perfectionnement des apprentissages grâce à la proximité des contextes vécus au quotidien et au développement de compétences significatives pour le futur de ces étudiants, tant dans le domaine des technologies comme celui des langues étrangères, en particulier dans le domaine de l'oralité. La méthodologie adoptée à l'UAb a permis d'évaluer les compétences de l'oralité⁴, de promouvoir des activités de pratique orale de la langue (par l'utilisation de *podcasts* et de *showcasts*), de donner un nouvel essor au développement des compétences de l'oralité, de garantir la centration sur l'étudiant lors du processus enseignement/apprentissage et du processus d'évaluation des langues étrangères tout en assurant qu'il assume l'autorégulation de son apprentissage, en définissant des stratégies pour le dépassement des difficultés et en développant des pratiques effectives d'auto- et hétéro-évaluation de l'oralité. En stimulant ces attitudes, nous désirons que les étudiants prennent connaissance de leurs méthodes d'apprentissage. En démontrant leur intérêt pour l'apprentissage des langues étrangères, les étudiants construisent leur autonomie dans l'apprentissage. Ils vivent alors l'évaluation numérique, non comme un obstacle ou une étape obligatoire, mais comme un chemin parcouru activement où ils laissent leur témoignage et leur marque. Cette qualité est fondamentale pour l'intégration de tous dans la société connectée où nous vivons actuellement. En réalité, les apprenants en langues étrangères à l'UAb sont, en même temps, utilisateurs et producteurs de contenus numériques. L'innovation mise en place pour la conception, l'analyse et le test de nouvelles stratégies – dans l'optique « Évaluer pour Évoluer » – nous mène à croire qu'il s'agit véritablement du chemin à suivre pour l'évaluation numérique de l'oralité des étudiants de *l'Universidade Aberta*.

⁴ La grille d'évaluation considère les trois grandes catégories de critères suivants: compétences pragmatiques et sociolinguistiques ; compétences et orthographe lexicales ; compétences et orthographe grammaticales.

5. Conclusion

Tous les enseignants de langues étrangères à l'UAb intègrent les technologies numériques dans leurs activités et dans leur évaluation qui vise le développement des compétences. Utiliser des technologies numériques dans l'évaluation est possible et les étudiants s'y prêtent volontiers car ils se repèrent facilement dans ces contextes évaluatifs. Pour maximiser la motivation intrinsèque, l'estime de soi et les compétences reliées au savoir, l'évaluation au sein de l'apprentissage doit, bien sûr, évaluer l'apprentissage de l'étudiant afin de voir ce qui est compris et appris, mais doit particulièrement fournir des rétroactions aux étudiants pour qu'ils puissent prendre connaissance de leurs méthodes d'apprentissage. C'est pourquoi, l'évaluation dans les unités de langues étrangères reflète les compétences nécessaires à la vie réelle et professionnelle. L'apprentissage et l'évaluation dans les unités de valeur de langues étrangères à l'*Universidade Aberta* sont conceptualisés comme un processus par lequel les étudiants construisent activement leurs connaissances et leurs compétences.

Références bibliographiques

- Alves, J. (dir.) (2001). *Quadro Europeu Comum de Referência para as Línguas – aprendizagem, ensino, avaliação*. Conselho da Europa. Porto : Edições Asa.
- Baartman, L. K. J., Bastiaens, T. J., Kirschner, P. A. et Vleuten, C. (2007). Evaluating assessment quality in competence-based education : A qualitative comparison of two frameworks. *Educational Research Review*, 2, 114-129.
- Birenbaum, M. (1996). Assessment 2000: Towards a pluralistic approach to assessment. Dans M. Birenbaum et F. J. R. C. Dochy (dir.), *Alternatives in assessment of achievement, learning processes and prior knowledge* (p. 3-29). Boston : Kluwer Academic Publishers.
- Birenbaum, M. (2003). New insights into learning and teaching and their implications for assessment. Dans M. Segers, F. J. R. C. Dochy et E. Cascallar (dir.), *Optimising new modes of assessment: In search of qualities and standards* (p. 13-36.). Dordrecht, The Netherlands : Kluwer Academic Publishers.
- Birenbaum, M., Breuer, K., Cascallar, E., Dochy, F., Dori, Y., Ridgway, J. et al. (2006). EARLI position paper. A learning integrated assessment system. *Educational Research Review*, 1, 61-67.
- Blanco, A. (2009). *Desarrollo y Evaluación de Competencias en Educación Superior*. Madrid : Narcea, S.A. de Ediciones.
- Dierick, S. et Dochy, F. J. R. C. (2001). New lines in edumetrics : new forms of assessment lead to new assessment criteria. *Studies in Educational Evaluation*, 27, 307-329.
- Directorate for Education, Employment, Labour and Social Affairs (2002). *Definition and Selection of Competences (DESECO) : Theoretical and Conceptual Foundations Strategy Paper*. OCDE.
- Figueiredo, A. (2000). Novos Media e Nova Aprendizagem. Dans A. Carvalho, A. Figueiredo, E. Morin, G. Delacôte, J. Silva, J. Pinheiro, J. Antunes, J. Tribolet, J. Azevedo, L. Resnick, M. Ambrósio, R. Carneiro, R. Filho et S. Papert (dir.), *Novo Conhecimento, Nova Aprendizagem* (p. 71-81). Lisboa : Fundação Calouste Gulbenkian.
- Kerckhove, D. (1997). *A Pele da Cultura*. Lisboa : Relógio d'Água.
- MacLellan, E. (2004a). How convincing is alternative assessment for use in higher education. *Assessment and Evaluation in Higher Education* ?, 29(3), 311-321.
- MacLellan, E. (2004b). Authenticity in assessment tasks : a heuristic exploration of academics' perceptions. *Higher Education Research and Development*, 23(1), 19-33.

- Mc Connell, D. (2006). *E-learning Groups and communities*. Berkshire : Open University Press.
- Pereira, A. *et al.* (2007). *Modelo Pedagógico Virtual da Universidade Aberta para uma universidade do futuro*. Lisboa : Universidade Aberta.
- Tinoca, L., Pereira, A. et Oliveira, I. (2014). A Conceptual Framework for E-Assessment in Higher Education: Authenticity, Consistency, Transparency and Practicability. Dans S. Mukerji et P. Tripathi (dir.), *Handbook of Research on Transnational Higher Education* (p. 652-673). USA : ICI Global.

Revue « *Contextes et Didactiques* », n°9, Juin 2017

REVUE SEMESTRIELLE EN SCIENCES DE L'ÉDUCATION



RUBRIQUE *VARIA*

Contextualisations didactiques et didactologiques. Suite du débat
Réponse à l'article de Philippe Blanchet
dans le n°7 de la revue « Contextes et Didactiques »¹

Véronique CASTELLOTTI, Marc DEBONO et Isabelle PIEROZAK
DYNADIV² (EA 4428) – Université François Rabelais de Tours

Résumé

Cette réponse propose de poursuivre un débat, initié par plusieurs textes des présents auteurs questionnant les fondements d'une démarche de contextualisation en didactologie-didactique des langues (DDdL). Ces textes « débattants » ont fait l'objet d'une réponse de Philippe Blanchet, qui sera la principale base de la continuation de cet intéressant dialogue contradictoire. Pour camper en quelques mots les termes de celui-ci, s'oppose une défense des « notions / concepts de *contexte* et de *contextualisation* » défendue comme « constructive en didactique et en didactologie des langues dans une perspective plurielle » (P. Blanchet), à une réflexion sur les « limites » à reconnaître en la matière, limites qui ne sont pas de superficielles imperfections, mais surtout liées aux choix épistémologiques offerts, en ce qu'ils sont différents, voire opposés. Nous souhaitons donc amener le débat essentiellement sur ce terrain, dont découlent des orientations disciplinaires, éthiques et politiques dont il nous apparaît crucial d'explicitier les différences.

Mots clés

Contexte, contextualisation, didactologie-didactique des langues (DDdL), épistémologies.

Abstract

This article offers to continue an active debate, initiated by several papers by the present authors, questioning the very foundations of a contextualization approach in the field of Language teaching and didactology. These previous “debating” papers were responded to by Phillippe Blanchet, whose reply shall be used as the basis for the continuation of a fruitful debate among conflicting views. If this debate had to be summarized in a few words, the main opposition lies between the defense of the “notions / concepts of context and contextualization” considered as “constructive for Language teaching and didactology in a pluralistic perspective” (P. Blanchet) and the reflection on the necessary acknowledgement of their “limits”. These limits are, rather than superficial imperfections, relative to different – if not opposite – epistemological choices. We therefore intend to bring the debate mainly on this ground, from which arise diverse disciplinary, ethical and political orientations – whose differences deserve a clarification.

Keywords

Context, contextualization, language teaching and didactology, epistemologies.

¹ Ce texte a bénéficié des apports de publications antérieures d'autres collègues, notamment de D. de Robillard (2008, 2014) et de E. Huver (2014) ainsi que des discussions menées de longue date dans l'équipe DYNADIV à propos de la notion de contexte ainsi que d'un séminaire « didactique et diversité » tenu à partir de la lecture des différents textes dont il sera question ici (ceux des auteur-e-s et celui de P. Blanchet notamment).

² DYNAMiques et enjeux de la DIVersité linguistique et culturelle.

1. Textes en jeu dans le débat

Philippe Blanchet s'étend longuement, dans son article (Blanchet, 2016) et à travers les références citées, sur les nombreuses publications qu'il a consacrées, seul et / ou avec d'autres, aux questions de contexte et de contextualisation, aux définitions qu'il en donne et aux problématisations qu'il propose en lien plus ou moins étroit avec ces notions, en sociolinguistique et en didactologie-didactique des langues.

Le débat qu'il lance, et dont nous ne pouvons que nous réjouir pour l'avoir nous-mêmes appelé de nos vœux précédemment, choisit de cibler en priorité deux textes interrogeant les enjeux recouverts par ces notions. Même si ce « ciblage » aurait pu être moins partiel (d'autres textes de l'équipe Dynadiv évoquent des aspects importants de ce débat), nous ne discuterons pas ce choix de lecture, tout en essayant de rétablir un « tableau » plus complet.

Le premier (en ordre chronologique de parution) de ces articles discutés fait partie d'un panel réunissant trois communications dans le cadre d'un même colloque organisé par l'équipe du DILTEC (Didactique des langues, des textes et des cultures, université Paris 3) en 2014 et portant précisément sur ces questions³ (cf. Castellotti, 2014 ; Debono, 2014 ; Huver, 2014). Ces trois textes ont aussi fait l'objet ultérieurement de la rédaction d'un seul chapitre synthétisant leurs principales thèses (Castellotti, Debono et Huver, 2016).

Le deuxième article, paru en 2015, à la suite d'une présentation effectuée lors d'un colloque du DILTEC dès 2012⁴, choisit de se focaliser non pas sur le(s) « contexte(s) » – notion plus communément et aisément critiquable – mais revient sur le processus de « contextualisation », généralement au second plan des critiques qui lui sont pourtant également adressables, et en se plaçant simultanément sur le plan didact(olog)ique.

Ces éléments saillants sont par ailleurs à resituer dans les travaux de l'équipe Dynadiv (Dynamiques et enjeux de la diversité linguistique et culturelle), lorsque celle-ci et l'ancien PREFics (Plurilinguisme, représentations, expressions francophones – information, communication, sociolinguistique) étaient parties prenantes d'une seule et même équipe. Ainsi l'équipe PREFics-Dynadiv avait mené une réflexion sur ces notions, en particulier dans ses séminaires et assemblées générales de l'année 2013-2014, à l'initiative de D. de Robillard.

C'est donc au regard de l'ensemble de ces éléments qu'il convient d'apprécier les réflexions proposées dans les textes mentionnés dans la réponse de P. Blanchet : si toute lecture / réception est nécessairement partielle, la confrontation des points de vue dans un débat scientifique a précisément pour vertu de discuter des complexités qui auraient pu être occultées par les protagonistes. En particulier, pour le texte publié par V. Castellotti, il constitue l'entrée en matière historique d'une réflexion en trois volets, dont nous reprendrons ici un certain nombre de développements pour expliciter les points de vue en jeu.

³ Colloque *Contexte global, contextes locaux. Tensions, convergences et enjeux en didactique des langues*, organisé par l'équipe du DILTEC en janvier 2014.

⁴ Il s'agit du colloque *Langages – Cultures – Sociétés : interrogations didactiques* organisé par l'équipe du DILTEC en juin 2012. Le texte a par la suite été publié, dans un autre contexte éditorial que celui des Actes de ce colloque.

2. Tonalité épistémologique : de quoi le contexte est-il le symptôme ?

Il est un premier point clef sur lequel un choix est à faire. Peut-on poser, comme le fait P. Blanchet, la nécessité de la notion, pour peu qu'on réfléchisse à comment la rendre opératoire conceptuellement, comme il est précisé dans son texte ? Si toute élaboration théorique peut en effet viser une opérabilité *in fine*, celle-ci est-elle pour autant définitoire, et à un horizon assez court, de l'activité de recherche ? Nous proposons de passer de la question du *comment*, qui n'est pas première, à la question, plutôt, du *pourquoi* ce concept est donné comme indispensable⁵. Passer du plan du *comment* à celui du *pourquoi* permet en effet de poser la question des choix épistémologiques, parce que ceux-ci existent, même si des effets de mode en la matière pourraient laisser envisager qu'il n'en est rien.

Le « contexte » et la « contextualisation » ont probablement été « utiles » à un moment donné de l'histoire de la didactologie-didactique des langues (DDdL), principalement dans une logique d'opposition à une perspective universalisante et uniformisante, ce que développe le texte cité de V. Castellotti. Mais cette utilité située dans le temps ne permet pas à elle seule de poser la nécessité actuelle voire intangible de la notion de contexte / de la démarche de contextualisation, sauf à nier l'histoire (ce qui ne serait pas le moindre des paradoxes pour une démarche « contextualiste »).

Dès lors, *pourquoi* ce concept ? Une partie de la réponse à cette question se trouve sans doute dans la préservation qu'il permet : le « contexte » impose, si on y réfléchit, un choix quant au statut de ce qui fait « texte ». Ce « texte » (entendu très largement comme un ensemble de signes, qu'il s'agisse de paroles / discours / pratiques, etc.) est posé comme le plus important, le « contexte » étant « avec » lui, pour faire sens, même s'il lui est reconnu, ce faisant, une forme d'incertitude (le texte seul ne suffit pas au sens)⁶.

Dans le cas didactique qui nous intéresse ici, on pourra objecter que cela semble différent, mais il n'en est rien. Ainsi, à condition de l'adapter au contexte d'intervention, et tel que le définira le didacticien, le « texte » (par exemple des orientations méthodologiques – dont les recommandations du CECRL –, des pratiques pédagogiques, ou un manuel de langue) demeure fondé : en d'autres termes, le texte continue d'être « sauvé » dans ce qui demeure l'essentiel⁷. On aura beau le diversifier contextuellement, c'est au fond le maintenir dans son principe de priorité. L'un est aussi prioritaire que l'autre devient alors nécessaire. On pourrait également fournir d'autres exemples, au plan didactologique.

Or cette priorité dont découle la nécessité du contexte n'a pour nous rien d'évident : plutôt que le contexte, l'*expérience*⁸ antépédicative (Romano, 2010) ou « pré-linguistique » est bien prioritairement ce qui fonde le « texte » (au sens large), et celle-ci n'est pas « apprivoisable »

⁵ Sur le *comment*, au demeurant, P. Blanchet opte pour une conceptualisation constructiviste, dont il précise qu'il est important de l'explicitier et donc de l'assumer en tant que telle, ce qui, dans son épistémologie, est important et parfois reste encore trop souvent implicite.

⁶ Ainsi, P. Schmoll utilise-t-il la métaphore d'un « contexte [...] roue de secours : le contexte se présente comme un lieu où l'analyse vient puiser l'information qui manque pour établir le sens » (1966 : 235-236). Il va également plus loin ensuite dans le raisonnement en soulignant que : « La conception bouche-trou du contexte procède d'une appréhension objective du sens, lequel est contenu dans le texte » (ibid. : 244).

⁷ D'un point de vue épistémologique, ces notions contribuent à « sauver l'importance des signes » (Robillard, 2014)

⁸ Au sens de « présence perceptive du monde » (Merleau-Ponty, 1964 : 48).

par une technologie de la contextualisation (et les conséquences de cette différence radicale ne sont pas minces, notamment du point de vue de la formation des enseignants⁹).

Dans ce qui précède nous avons volontairement passé sous silence ce que les perspectives contextualistes discutent finalement peu, non sans que cela pose de vrais problèmes. Pour P. Blanchet, la constitution du contexte consiste à sélectionner ce que l'on considère comme « efficient » « pour comprendre les éléments et les phénomènes sur lesquels on focalise l'observation » (Blanchet, 2016 : 9), ce que l'on retrouve aussi chez R. Porquier et B. Py :

« La situation serait une notion primitive (« étique »), le contexte, une notion technique (« émique ») pouvant procurer des critères et des outils d'analyse, et retenant ce qui des situations serait jugé pertinent dans une perspective d'observation ou d'analyse. Ainsi, dans la complexité empirique objective d'une instance d'acquisition ou d'apprentissage, seraient retenus comme composantes du contexte, comme traits contextuels, ce qui de la situation peut être considéré comme critères contextuels pertinents, selon le point de vue adopté » (Porquier et Py, 2004 : 50).

La question principale posée par ces citations renvoie à qui est ce « on », et plus précisément quel chercheur / didacticien ? Qui sélectionne et choisit les « éléments et phénomènes », qui « retient » les composantes qu'il juge pertinentes pour constituer le contexte, et, précisément, à partir de quel « point de vue adopté », pourquoi il effectue ce choix, à partir de quelles interprétations, fondées sur quelles anticipations, quelle(s) histoire(s), quel(s) projet(s) il veut déterminer ce qui « fait contexte »¹⁰ ?

Tout se passe donc comme si une méthode (même si P. Blanchet l'entend au sens d'E. Morin) était envisageable pour ce faire. En effet, parler de critères, paramètres, quoique diversement mobilisables, ne sauve pas de la tendance objectivante de fond¹¹, au détriment du relationnel (que P. Blanchet a pourtant ailleurs mobilisé concernant la notion de « terrain »), de l'histoire, et de l'imaginaire en jeu, qui là non plus ne sont pas mis au premier plan. Le risque encouru, ce faisant, est d'envisager que les compréhensions soient plus ou moins « bonnes » / « fiables », en perdant de vue ce qui fait histoire(s) / projet(s) lorsque chacun contextualise.

Il apparaît en définitive que le contexte devient le symptôme d'une difficulté à ne pas vouloir / pouvoir faire autrement, en dépendant d'une épistémologie sauvant coûte que coûte le « texte », ou la dimension sémiotique plus globalement. Et cette difficulté sera d'autant plus entière qu'elle sera assortie de la propension à faire du contexte un nouveau matériau, compensateur, dans lequel résiderait une improbable clef pour « le » sens « adéquat ». Tenter de répondre à la question du pourquoi de cette nécessité concernant ces concepts de

⁹ Il ne s'agit pas ici d'oublier des éléments aujourd'hui qualifiés de « contextuels » (dont, par exemple, les diverses et diversifiées « traditions » d'enseignement / apprentissage) pour un retour à une perspective surplombante et universalisante, mais bien plutôt de ne pas « tromper » sur le pouvoir de la contextualisation : l'enjeu d'une didactique des langues étant la rencontre d'autres, elle doit avertir les futurs enseignants en formation que les autres ne sont pas « décodables », fussent-ils « correctement contextualisés ». Voir aussi ci-dessous.

¹⁰ Nous passons ici sous silence la « primitive » essentialisation qui est faite de la « situation » chez Porquier et Py.

¹¹ Il en va ainsi de même de la position d'un Schegloff par exemple, pour lequel la question de la pertinence des critères trouve une résolution, au-delà du « theoretical imperialism » des chercheurs, par l'analyse de comment les interactants construisent eux-mêmes cette pertinence. En d'autres termes, notre interrogation ne porte pas ici sur les relations de pouvoir en jeu, selon que l'on adopte une posture émique ou étique (à distinguer bien sûr de l'emploi qu'en font Porquier et Py, dans la citation supra), mais elle porte plutôt sur les enjeux de la contextualisation, en ramenant cette dernière à l'histoire / aux projets de son auteur, quel qu'il soit.

« contexte / contextualisation » nous conduit aussi, dans la continuité de ces réflexions, sur le terrain politique et éthique.

3. Tonalité politique et éthique : enjeux d'une démarche de contextualisation

Par-delà la notion même de contexte examinée ci-dessus, c'est le processus de contextualisation qui est dans un deuxième temps à interroger : quelles en sont les justifications et argumentations à l'appui ? Qu'est-ce qui serait contextualisé, par qui et dans quelles perspectives ?

Dans les approches contextualistes, il s'agit d'« adapter »¹² un matériau préalablement existant (que nous avons appelé « texte » ci-dessus), constitué de contenus à (faire) acquérir associés à des choix méthodologiques fixés a priori :

« La contextualisation didactique poursuit et complète en la transformant une dynamique ouverte par la « révolution communicative » des années 1970-80. La question du contexte en didactique des langues se révèle à cette époque en posant comme objectif et comme moyen d'enseignement-apprentissage des usages effectifs dans des situations de communication (contextes « authentiques ») produites ou imitées en situations de classes (contextes pédagogiques) de façon réaliste (contexte social) » (Rispaill et Blanchet, 2011 : 68).

Un positionnement comme celui-là s'appuie sur l'idée d'une forme de supériorité non argumentée d'une approche méthodologique (communicative, fondée sur des présupposés d'ordre pragmatique) qu'il ne resterait plus qu'à adapter, à « implémenter » à l'aide d'une contextualisation bien menée. Or, cette « révolution » qu'il faudrait « poursuivre » et « compléter » en l'étendant sur la surface du globe grâce à la « dynamique » contextualiste possède un point de départ bien situé et non discuté, à la fois dans son origine philosophique / épistémologique et dans sa pertinence pour les « autres » qu'il faudrait convertir / « révolutionner ».

Poursuivons donc la réflexion dans cette direction : les approches didactiques pensées à partir d'évidences pragmatiques (la « révolution communicative » évoquée ci-dessus) subordonnent tout à l'action réaliste (à différents niveaux : des pratiques de classes à la politique diffusionniste de la langue) ; le contexte a alors pour fonction principale d'orienter la meilleure façon de faire passer ces « bonnes orientations », dans une perspective s'affichant comme « humaniste ». Cela a cependant comme corollaire de dissoudre la responsabilité du chercheur¹³ : le contexte lui dicterait les meilleurs moyens d'adapter outils et dispositifs, afin d'assurer un progrès dont on ne peut douter (Castellotti, Debono et Huver, 2017).

C'est au nom de telles « évidences » épistémologiques (qui, pour le coup, sont très « situées ») que se développe toute une politique de diffusion des langues, relayée notamment en direction des pays « du Sud » par des organismes internationaux comme la Banque mondiale ou le Fonds monétaire international (FMI), imposant des méthodologies¹⁴ prétendues innovantes et censément plus efficaces que les précédentes (qui elles-mêmes avaient déjà été importées dans cette même logique). Tout l'enjeu sera alors de masquer ces

¹² Dans certains travaux, « contextualisation » et « adaptation » sont d'ailleurs employés indifféremment, comme des synonymes (voir par exemple : Vilpoux, 2013).

¹³ Nous reviendrons sur cette question de la responsabilité du chercheur en conclusion.

¹⁴ Notamment l'approche par compétences, articulée, pour les langues, à l'approche communicative / actionnelle.

origines épistémologiques de la « révolution communicative », poursuivie et complétée par la « révolution contextualiste », en les naturalisant : elles ne semblent aujourd'hui plus relever d'un choix, mais relever « naturellement » des progrès des sciences modernes. Or, tout cela nous semble au contraire *très* politique, ce qui n'est bien entendu pas un problème en soi : une science conçue comme « non positive » est nécessairement orientée par de tels choix, qu'on peut expliciter et donc discuter... sauf s'ils ne sont pas présentés comme tels !

D'un point de vue politique, les usages qui sont faits de ces notions permettent ainsi de continuer à diffuser le français, non plus de façon explicitement universaliste, comme au temps de la colonisation et des dynamiques post-coloniales qui l'ont suivie, mais avec des oripeaux beaucoup plus « progressistes », qui permettent à ce diffusionnisme d'avancer masqué. Comme nous l'explicitons de manière quelque peu provocatrice ailleurs (Castellotti, Debono et Huver, 2016), nous (les didacticiens / didactologues du FLE) exportons, avec les meilleures intentions du monde, le « français contextualisé » de façon peu différente de ce que fait McDonald avec ses burgers à la sauce locale. La base du produit est la même, elle est conçue et pensée par la « maison mère », mais sa réalisation *feint de prendre en compte* des habitudes, traditions, coutumes (souvent par ailleurs stéréotypées) locales pour tenter d'occulter une politique commerciale trop voyante. Cette *adaptation* superficielle en permet une meilleure *consommation*, car donnant l'air de tenir compte de spécificités locales.

Or, à l'inverse, si on part de l'orientation selon laquelle ce n'est pas (seulement) la pluralité qui est constitutive de l'humain, mais plus fondamentalement la diversité, il devient politiquement responsable de *partir* de cette diversité dans l'élaboration des contenus, mais aussi des approches et méthodes, et non de la prendre en compte a posteriori et de manière presque compensatoire pour adapter / contextualiser des démarches pensées ailleurs et a priori.

Il ne s'agit pas en faisant cela de « renvoyer dos à dos [...] un applicationnisme universaliste et une contextualisation qu'on estime insuffisante » (Blanchet, 2016 : 12), mais bien de considérer que la contextualisation, telle qu'elle est pensée *et* mise en œuvre de façon très dominante en DDdL, y compris au sein d'une perspective sociodidactique, contribue à orienter le domaine de la DDdL de façon prioritairement pragmatique, et au profit de la diffusion des langues, diffusion qui s'effectue au moyen des « bonnes pratiques / méthodologies / ingénieries » communicatives qui, aussi contextualisées soient-elles, s'imposent aux « locaux contextualisés » qu'elles ambitionnent d'aider.

Une question corollaire que nous souhaitons poser est donc : comment choisissons-nous d'orienter ce domaine ?

4. Tonalité (inter)disciplinaire : questions de « pré-carré » ou de dépendance ?

Le fait que la DDdL ait des frontières poreuses, qu'elle soit un domaine carrefour est une grande force si les relations entre les disciplines concernées sont pensées sur le mode de l'articulation¹⁵ (ce que la logique *interdisciplinaire* réclame) et non sur celui de la hiérarchisation : l'histoire de la discipline s'écrit dans cette tension entre des types de relation. Exemples à ce titre, les propos de P. Blanchet (2016 : 11) : « Pour élaborer une

¹⁵ Dans la mesure où ces articulations sont épistémologiquement possibles. En effet, tout n'est pas articulable épistémologiquement, certains points de désaccord n'étant pas réductibles.

didactologie / didactique¹⁶ de quelque chose, il faut avoir une théorie de ce quelque chose (en l'occurrence une théorie des langues [...]). C'est là un point de vue que nous pourrions partager, si le texte de P. Blanchet ne continuait de la manière suivante : « et une théorie sociolinguistique est possible et même appropriée » (Ibid.), ajoutant in fine : « Si l'on considère les situations de didactique des langues comme des situations sociolinguistiques »¹⁷.

Loin de vouloir défendre un quelconque « pré-carré », cette position ne correspond néanmoins pas du tout à notre conception d'une interdisciplinarité « articulée », pour peu qu'il y ait compatibilité épistémologique, mais bien plutôt d'un retour en force d'un serpent de mer : l'applicationnisme de théories qui ne pourraient se construire que dans un « en-dehors » de la DDdL, sans l'idée même d'une possible circularité. Et c'est précisément là que nous divergeons : cela signifierait, comme nous l'avons déjà évoqué (Castellotti, 2016) subordonner la DDdL à des « disciplines de références » (dont la sociolinguistique), et la cantonner à une application stérilisante, quoique s'en défende P. Blanchet.

Le rappel de la porosité positive des frontières de la DDdL est donc l'exact opposé d'une tentative d'exclusion : très loin de nous l'idée d'« exclu[re] ceux et celles avec qui on est en désaccord en leur assignant une identité étrangère, celle de “sociolinguistes” » (et, du reste, notre propre équipe a aussi cet ancrage sociolinguistique), il s'agit uniquement de rappeler qu'une théorie de la langue peut s'élaborer de manière interdisciplinaire, en considérant les situations didactiques comme des situations... didactiques, c'est-à-dire : prenant en compte leurs aspects sociolinguistiques, mais aussi psychologiques, philosophiques, anthropologiques, politiques, éducatifs, etc. (toujours dans la limite des compatibilités épistémologiques, bien sûr, il faut y insister) et en conceptualisant les « langues » en fonction d'interprétations in situ informées par ces éclairages multiples et non à partir du seul regard sociolinguistique. En effet, compte tenu des histoires et des projets de leurs secteurs respectifs, il est probable que la « langue » des didacticiens ne puisse être tout à fait superposable à la « langue » des sociolinguistes (qui, en outre, n'ont pas tous non plus la même théorie de la « langue », cf. infra).

Ainsi, l'idée d'une « sociodidactique » qui semble élargir le champ en altère conjointement la porosité : la logique de « bornage » l'emporte sur la logique interdisciplinaire et la rend potentiellement excluante.

¹⁶ Pour notre part, et à l'instar de R. Galisson, nous accolons didactologie-didactique avec un trait d'union, pour marquer précisément une relation d'articulation intrinsèque entre les deux aspects, et non pas de « distinction » comme le propose P. Blanchet.

¹⁷ Bien que P. Blanchet souligne ici une possibilité et une adéquation en matière de référence à la sociolinguistique – avis que nous respectons et que l'histoire de l'équipe Dynadiv permet d'ailleurs d'illustrer aussi, quoiqu'autrement – son insistance à trois reprises quant à cette référence, qui légitime le préfixe « socio » dans « sociodidactique », questionne. En effet parler de « théorie », « analyse » et enfin de « méthodes de recherche sociolinguistiques », « possibles et appropriées », pose la question du statut de cette sociolinguistique pour la sociodidactique. En d'autres termes cela n'apparaît pas comme une simple entrée « socio », éventuellement à placer à égalité avec une autre, c'est en faire plus fondamentalement une entrée épistémologique. Et ce faisant, il paraît quelque peu contradictoire d'affirmer au final qu'il serait possible d'en dire autant « d'autres disciplines connexes et contributives » : si elles sont telles en effet, pourquoi alors la sociolinguistique devrait-elle prédominer (au point d'identifier une « sociodidactique », qui semble signifier pour ses promoteurs sensiblement la même chose que la préfixation « socio- » pouvait signifier pour les fondateurs de la sociolinguistique : pas de « linguistique » qui ne soit pas « socio »), si ce n'est parce qu'elle n'a pas ce (seul) statut précisément ? On pourrait ensuite s'interroger sur l'identité de ladite théorie et des méthodes sociolinguistiques convoquées (et leur cohérence), car sur le plan épistémologique, les histoires de la sociolinguistique sont pour le moins diverses (Pierozak, Robillard, Razafimandimbimanana et Debono, 2013).

La question n'est donc pas du tout, en effet, de considérer ou non comme « seule valable » « une approche didactologique et didactique en lien étroit avec une sociolinguistique », ni de la « distinguer d'autres approches moins ou non socio- » (Blanchet, 2016 : 11), mais d'imaginer une DDdL qui pense, définit, fait évoluer ses questionnements de recherche et d'intervention non pas comme une application de conceptualisations exogènes mais bien, prioritairement, en fonction des appropriations linguistico-culturelles qui constituent « l'espace social historicisé » (Ricoeur) de sa réflexion et de son action.

Cela ne signifie pas qu'il faille se passer des travaux menés en sociolinguistique qui peuvent avoir des liens avec ces appropriations, au contraire¹⁸, pas plus que ceux émanant des recherches des autres domaines mentionnés ci-dessus (selon les sensibilités disciplinaires qu'il importe d'explicitier en tant que choix). Mais cela n'implique pas, nous le réaffirmons, de faire « découler directement¹⁹ les interventions didactiques de l'analyse des contextes-terrains sociolinguistiques » (Castellotti, 2014 : 18 ; Castellotti, Debono et Huver, 2016 : 57), pas plus qu'il ne s'agirait de mobiliser des « transpositions » (Blanchet, 2015), ce qui reposerait sur l'idée que le « texte » reste fondateur (cf. *supra*).

Enfin il faut rappeler l'existence de diversités épistémologiques importantes (cf. note 17 ci-dessus), qui ne sont pas signalées dans le texte de P. Blanchet : les « langues » seront par exemple conçues comme des *systèmes* de signes, en plus ou moins grande corrélation avec les usages et changements sociaux ; ou comme des *moyens* de communication en interrelation à des contextes construits / définis ; ou comme des *expériences* du monde, ni entièrement prévisibles, ni réduites aux seuls signes qui en constituent des manifestations matérielles, etc. A ce titre, il importe donc de prendre en compte épistémologiquement les imaginaires, diversifiés (Robillard, 2007, 2008 notamment), donnant lieu à des déclinaisons théoriques et méthodologiques, très différentes. Et c'est plutôt sur ce plan – et pas tellement au plan disciplinaire – que devrait se poser la nécessaire problématique des évolutions de la DDdL.

En cohérence avec les orientations épistémologiques rappelées au point 2, des choix de réponses apportées à ces questions dépend largement la conception de l'appropriation *en* langues, qui ne peut être pas « déterminée » par les contextes ou terrains, mais seulement reliée à l'expérience des personnes concernées, donc à leur histoire et à leur projet.

5. Éléments de conclusion

Nous avons essayé, dans ce texte, de replacer le débat sur un plan épistémologique, entre des positions en présence, pour montrer que les oppositions ne sont pas superficielles, qu'elles ne sont pas liées à de bons ou de mauvais usages de la contextualisation, mais à de véritables *choix* éthiques et politiques difficilement conciliables. Or, nous pensons que la tendance à présenter la contextualisation didactique dans une dynamique progressiste et humaniste auto-justificatrice (donc inéluctable) contribue : 1) à occulter certains de ses aspects, plus problématiques (nous avons largement développé cet aspect), et 2) à occulter le fait que la « révolution communicative » dont elle est censée découler n'a rien de « naturel » ni d'inéluctable. Discuter ces deux points – avec les (dés)accords qu'ils induisent – participe

¹⁸ Ceux-ci sont d'ailleurs très présents dans la plupart de nos réflexions et publications.

¹⁹ C'est-à-dire en les adoptant et les *appliquant* telles quelles.

d'une démarche d'explicitation et de prise de position que nous estimons responsable, qui évite de verser dans la simplification²⁰.

Cette responsabilité dans des prises de positions épistémologiques assumées comme porteuses d'enjeux (et non comme la poursuite d'une forme de déterminisme historique, dont l'innovation serait le moteur : l'actionnel suit le communicatif, prolongé / complété par le contextualisme, et ainsi de suite) rend inéluctable la dose de conflictualité intellectuelle présente dans ce débat en cours, positive à nos yeux. Si l'on s'inscrit dans une conception de l'activité scientifique comme non positiviste, traversée certes par l'éthique, le politique, mais aussi l'expérience antéprédicative des chercheurs et des formes de sensibilité, argumenter *contre* n'a rien d'étonnant, ou plutôt ne devrait rien avoir d'étonnant. Pourtant, on constate aujourd'hui que les prises de position un peu marquées quant aux travaux de recherche, qui devraient constituer le ferment habituel de nos disciplines, sont parfois considérées comme d'insupportables remises en question d'une bienséance académique paralysante. Nous faisons ici un pari inverse, sans « politiquement correct » donc, et en espérant uniquement et modestement que cette attitude soit heuristique, qu'elle (nous) fasse (tous) réfléchir observateurs-lecteurs et participants plus directs à ces échanges.

Références bibliographiques

- Blanchet, P. (2016). Contextualisations didactiques et didactologiques. Questions en débat. *Contextes et Didactiques*, 7, 8-14. Disponible en ligne sur : http://www.espe-guadeloupe.fr/wp-content/uploads/2016/06/2_Blanchet-2016.pdf
- Blanchet, P. (2015 [2011]). Les transpositions didactiques. Dans P. Blanchet et P. Chardenet (dir.), *Guide pour la recherche en didactique des langues. Approches contextualisées* (p. 247-252). Paris : Éditions des Archives contemporaines / AUF.
- Castellotti, V. (2016). Idées sociolinguistiques et orientations didactiques. Histoires croisées, projets à repenser. *Glottopol*, 28, 53-69.
- Castellotti, V. (2014). Contexte, contextualisation, cultures éducatives. Quels usages ? Pour quelles orientations de la recherche en DDL ? Dans S. Babault, M. Bento, L. Le Ferrec et V. Spaeth (dir.), *Contexte global, contextes locaux. Tensions, convergences et enjeux en didactique des langues* (p. 111-124). Actes du colloque international, 23-25 janvier 2014, FIPF.
- Castellotti, V., Debono, M. et Huver, E. (2017). Une « tradition de l'innovation » ? Réflexion à partir du corrélat innovation / créativité en didactique des langues. *TRANEL - Travaux neuchâtelois de linguistique*, 65, 113-129.
- Castellotti, V., Debono, M. et Huver, E. (2016). Contexte, contextes, contextualisation en DDL. D'une didactique contextualisée à une didactique diversitaire. Dans S. Babault, M. Bento et V. Spaeth (dir.), *Tensions en didactique des langues : entre enjeu global et enjeux locaux* (p. 49-76). Berne : Peter Lang.
- Debono, M. et Pierozak, I. (2015). Contextualisation et distance en didactique diversitaire des langues-cultures. Exemples de réflexion à partir de Diffodia. *Recherches en didactique des langues et des cultures : les cahiers de l'Acedle*, 12(1), 23-53. Disponible en ligne sur : <http://acedle.org/old/spip.php?article4367>.
- Debono, M. (2014). La *contextualisation*, une dynamique *glocalisante* ? Tribulations de deux notions, de leurs reprises et détournements. Dans S. Babault, M. Bento, L. Le Ferrec et V. Spaeth (dir.), *Contexte global, contextes locaux. Tensions, convergences et enjeux en*

²⁰ A noter que les termes de « progrès » ou d'« humanisme » sont pour le moins très investis dans la littérature scientifique, et de manière parfois très différente.

- didactique des langues* (p. 125-139). Actes du colloque international, 23-25 janvier 2014, FIPF.
- Huver, E. (2014). Quand contexte homogénéise la diversité. Ou : parler de la diversité sans « contexte » ? Dans S. Babault, M. Bento, L. Le Ferrec et V. Spaeth (dir.), *Contexte global, contextes locaux. Tensions, convergences et enjeux en didactique des langues* (p. 140-54). Actes du colloque international, 23-25 janvier 2014, FIPF.
- Merleau-Ponty, M. (1964). *Le visible et l'invisible*. Paris : Gallimard.
- Pierozak, I., Robillard D. de, Razafimandimbimanana, E. et Debono M., (2013). Vers une sociolinguistique française qualitative ? Perspectives historiques critiques sur des processus de reconnaissance. *Recherches qualitatives*, 32(1), 107-131.
- Porquier, R. et Py, B. (2004). *Apprentissage d'une langue étrangère, Contexte et discours*. Paris : Didier/CREDEF, Collection Essais.
- Rispail, M. et Blanchet, P. (2015 [2011]). Principes transversaux pour une sociodidactique dite « de terrain ». Dans P. Blanchet et P. Chardenet (dir.), *Guide pour la recherche en didactique des langues. Approches contextualisées* (p. 101-105.) Paris : Éditions des Archives contemporaines / AUF.
- Robillard, D. de (2014). La magie des signes : *Elémentaire mon cher Watson !* Réflexivité, « pratiques réelles », « corpus », « interactions » et autres « données ». Dans G. Forlot et F. Martin (dir.), *Regards sociolinguistiques contemporains. Terrains, espaces et complexité de la recherche* (p. 39-56). Paris : L'Harmattan, collection Carnets d'atelier de sociolinguistique.
- Robillard, D. de (2008). *Perspectives alterlinguistiques*. Paris : L'Harmattan.
- Robillard, D. de (2007). La linguistique autrement : altérité, expérimentation, réflexivité, constructivisme, multiversalité : en attendant que le Titanic ne coule pas. Dans P. Blanchet, L.-J. Calvet, D. de Robillard (dir.), *Un siècle après le Cours de Saussure : la linguistique en question. Carnets d'Ateliers de Sociolinguistique, 1*, 81-228.
- Romano, C. (2010). *Au cœur de la raison, la phénoménologie*. Paris : Folio Essais.
- Schmoll, P. (1996). Production et interprétation du sens : la notion de contexte est-elle opératoire ? *Scolia*, 6, 235-255.
- Vilpoux, C. (2013). *La rénovation de l'enseignement du français dans les universités en Ukraine : une analyse didactique contextualisée*. Thèse de doctorat sous la direction de P. Blanchet, Université Rennes 2.