

Rédacteur en chef
Sébastien Langevin
Présentation graphique
CGI
Conception graphique
miz'empage
Directeur de la publication
Jean-Pierre Cuq – FIPF



RECHERCHES ET APPLICATIONS

Le français dans le monde
9 bis, rue Abel Hovelacque
75013 Paris
Téléphone : 33 (0) 1 72 36 30 67
Télécopie : 33 (0) 1 45 87 43 18
Mél : fdlm@fdlm.org
<http://www.fdlm.org>

© Clé International 2013
Commission paritaire 0417181661
La reproduction même partielle
des articles parus dans ce numéro
est strictement interdite,
sauf accord préalable.

LE FRANÇAIS DANS LE MONDE
est la revue de la Fédération
internationale des professeurs
de français (FIPF)

Recherches et applications

le français
dans
le monde

N°53
JANVIER 2013
PRIX DU NUMÉRO 18€

Évaluer en didactique des langues/cultures : continuités, tensions, ruptures

Coordonné par Emmanuelle Huver et Aleksandra Ljalikova

Comité de rédaction

Francis Carton (Président du conseil scientifique)
Patrick Chardenet (Président du conseil scientifique)
Jean-Pierre Cuq (Directeur de la publication)
Sébastien Langevin (Rédacteur en chef)

Conseil scientifique

Margaret Bento (Université Paris 5 René Descartes, France) ; Evelyne Bérard (Université de Franche-Comté, France) ; Robert Boucharde (Université Lumière Lyon 2, France) ; Francis Carton (ATILF- Université de Lorraine, CNRS- CRAPEL) ; Patrick Chardenet (Agence Universitaire de la Francophonie, Université de Franche-Comté, France) ; Francine Cicurel (Université Sorbonne Nouvelle – Paris 3, France) ; Fatima Davin-Chnane (Aix-Marseille Université, France) ; Piet Desmet (Université Catholique de Leuven, Belgique) ; Pierre Dumont (Université des Antilles et de la Guyane, France) ; Enrica Galazzi-Matasci (Université Catholique de Milan, Italie) ; Pierre Martinez (Université de Paris 8 – Vincennes – Saint-Denis ; Université nationale de Séoul, Corée) ; Danièle Moore (Simon Fraser University, Canada) ; Samir Marzouki (Université de la Manouba, Tunisie) ; Franz-Joseph Meissner (Justus-Liebig Universität Gießen, Allemagne) ; Jean Noriyuki Nishiyama (Université de Kyoto, Japon) ; Tatiana Zagryazkina (Université d'État de Moscou Lomonossov, Russie) ; Zheng Lihua (Université des Études étrangères du Guangdong, Chine).

Comité varia

Margaret Bento
Francine Cicurel
Franz-Joseph Meissner

Cher lecteur de la revue

Recherches & Applications : Le français dans le monde

Que vous soyez étudiant ou doctorant en didactique du français langue étrangère, enseignant exerçant dans l'enseignement primaire, secondaire ou universitaire, dans un pays francophone ou non, directeur de recherche à l'université, tous soucieux de suivre les évolutions de l'enseignement du français à l'échelle du monde, pour vous la revue *Recherches & Applications : Le français dans le monde* est un repère professionnel incontournable. La revue vous remercie de votre fidélité et de la crédibilité scientifique que vous lui accordez.

Comme elle l'a montré lors de ses derniers congrès, la Fédération Internationale des Professeurs de Français est sensible aux évolutions qui font de l'espace de la connaissance un monde plurilingue, multipolaire, globalisé, mais aussi contextualisé. La revue souhaite y maintenir sa position d'acteur de premier plan, en anticipant, en conduisant ou en accompagnant ces évolutions tout en affirmant la contribution de la langue française à cet espace mondialisé. Pour garantir cette fonction d'excellence, le Comité scientifique de la revue affirme une politique de publication qui reste fidèle à son objectif de toujours : animer le débat en didactique des langues et des cultures, au service d'une diffusion de qualité de la langue française dans le monde, en étant plus que jamais à l'écoute des innovations et des mutations.

Dans cette perspective, la revue s'est progressivement ouverte aux équipes de recherche qui contribuent à cet objectif, en leur confiant la coordination d'un numéro, où qu'elles travaillent dans le monde¹.

Les règles déontologiques du champ scientifique sont rigoureusement respectées : les articles de la revue sont soumis à une double évaluation anonyme prise en charge par un comité de lecture formé de chercheurs reconnus, qui veille à ce que chaque texte s'appuie sur des données de première main, une originalité des analyses, et des références précises des travaux utilisés.

Par ailleurs, la revue a modifié la structure éditoriale jusque-là en usage, pour témoigner de la vigueur des travaux des jeunes chercheurs en y incluant dans sa rubrique *Varia* des articles hors de la thématique générale du numéro, sélectionnés pour leur intérêt et leur qualité².

Enfin, *Recherches & Applications* est engagée dans une collaboration par échanges d'articles avec la *Revue canadienne des langues vivantes / The Canadian Modern Language Review*.

Pour le Comité scientifique, les co-présidents
Francis Carton (ATILF, CNRS-Université de Lorraine, CRAPEL)
Francis.carton@univ-lorraine.fr
Patrick Chardenet (AUF, Université de Franche-Comté – UR ELLIAD)
patrick.chardenet@auf.org

1. Pour soumettre une proposition de numéro, voir <http://fipf.org/publications/recherches-applications> (Instructions aux coordinateurs)

2. Pour soumettre un article, en tant que jeune chercheur (en fin de thèse ou venant de terminer la thèse de doctorat), voir <http://fipf.org/publications/recherches-applications> (Appel permanent à publication d'articles dans le cadre de la diffusion de recherches menées par de jeunes chercheurs)

Évaluer en didactique des langues/cultures : continuités, tensions, ruptures

Présentation

Dimensions épistémologiques, politiques et didactiques

EMMANUELLE HUVER, ALEKSANDRA LJALIKOVA7

Un champ en mouvements et en tensions

L'évaluation en éducation, ses paradigmes et les problématiques identitaires

MICHEL VIAL.....20

La formation à l'évaluation est une entreprise difficile parce qu'elle touche aux convictions, aux « épistémologies profanes » et donc aux dynamiques identitaires. Un travail sur soi d'ordre épistémologique (savoir d'où l'on parle) est nécessaire quand on a accepté que les attitudes de l'évaluateur dépendent d'ingrédients culturels acquis. À ce titre, le concept pragmatique de « paradigme profane » développé dans ce texte occupe une place centrale.

Teaching evaluation is a difficult undertaking as it affects beliefs, "secular epistemology", and therefore the dynamics of identity. To work on his/her own epistemology is necessary (know where it comes from) when it is acknowledged that the evaluator's attitudes depend on culturally acquired ingredients. As such the paradigmatic concept of "secular paradigm" developed in this paper occupies a central position.

Plus de convergences que de divergences : rencontre des conceptions francophone et anglo-saxonne de l'évaluation des apprentissages

DANY LAVEAULT.....31

Les progrès dans le domaine des théories cognitives ainsi que les réformes des programmes d'études dans plusieurs pays occidentaux ont donné lieu à d'importants développements dans la conception et l'utilisation de l'évaluation pour soutenir l'apprentissage, tant dans les pays francophones qu'anglo-saxons. Ces développements ont porté principalement sur la définition et le rôle même de l'évaluation formative ainsi que sur ses modalités. Si des différences principalement reliées aux modalités sont apparues entre les pays, des convergences non négligeables se sont développées autour des besoins de faire porter l'évaluation sur des apprentissages

progressivement plus complexes tout en développant la capacité de l'apprenant à s'autoréguler. Il ressort que l'évaluation est une compétence à développer, tant chez l'enseignant que chez l'apprenant, et que celle-ci peut le mieux s'acquérir dans un environnement où les résultats des étudiants sont utilisés pour réguler les pratiques des enseignants et pour développer le potentiel des étudiants à prendre en charge leurs apprentissages.

Recent advances in cognitive theories and in the reform of curriculum have paved the way for important developments on the conception and use of assessment to support learning in both French and Anglo-Saxon countries. Such developments have born mostly on the definition and function of formative assessments as well as on its modalities. If some discrepancies as to the latest have appeared among different countries, major convergences have developed around the need to assess increasingly complex forms of learning while increasing at the same time learners' abilities to self-regulate. It follows that assessment is a competency to be developed among teachers and student as well and that it can be better acquired when students' results are used to adjust teachers' practice and to develop students' potential to take charge of their own learning.

De quoi l'évaluation peut-elle être le nom ?

PATRICK CHARDENET.....42

En tant que pratique sociale, l'évaluation formelle des activités humaines est fondamentalement une estimation systématisée et non un outil de mesure précise. L'observation des pratiques, quels que soient les niveaux d'intégration et la qualité des artefacts, montre que son fonctionnement repose sur une activité évaluative naturelle à travers des mises en scène de répertoires langagiers. La compréhension de ce paradigme dans un contexte global où pèse le marché des certifications impose des solutions qualitatives à la fois alternatives et cohérentes, dans un système global d'évaluation multiréférentielle.

Como práctica social, la evaluación formal de las actividades humanas es fundamentalmente una estimación sistematizada y no una herramienta de medida precisa. La observación de sus prácticas cualesquiera que sean los niveles de integración y la calidad de los artefactos,

muestra que su funcionamiento se basa sobre una actividad evaluativa natural a través de puestas en escena de repertorios lingüísticos. La comprensión de este paradigma en un contexto global donde pesa el mercado de las certificaciones impone soluciones cualitativas a la vez alternativas y coherentes, en un sistema global de evaluación multirreferencial.

Évolutions des usages d'évaluation : faire avec les tensions

Les descripteurs de compétence en langue : rappels et mise en perspective

PATRICK RIBA56

Nous explorons dans cet article les modalités de conception, les limites théoriques, symboliques et systémiques des échelles de compétence en langue du *Cadre européen*, afin de répondre à une question en apparence simple. Pourquoi les spécifications de niveaux du *Cadre*, pourtant définis sur la base de descripteurs en apparence aussi univoques que «*Peut se débrouiller dans les situations de la vie quotidienne*» (*Cadre*:66), varient-elles autant d'un pays ou d'une méthode FLE à l'autre ?

The description of the Common European Framework for Languages still varies from a country or a textbook to another one, nevertheless it is based on seemingly so unambiguous formulations as "Can manage in the situations of the everyday life" (CEFR:66). This paper analyzes the methods of design and the theoretical, symbolic and systematic limits of scales, and it tries to define why the establishment of a set of common reference points for languages doesn't reach its goal of a consensual interpretation of the Common Reference Levels.

Les tests de langues : une nécessité ?

CLAIRE TARDIEU67

La place croissante de l'évaluation dans le domaine scolaire de l'enseignement-apprentissage des langues pose un certain nombre de questions quant à la légitimité de la demande en tests standardisés et calibrés sur les niveaux du *Cadre Européen Commun de Référence en Langue* (CECRL). Aux tests psychométriques de tradition anglo-saxonne s'ajoutent aujourd'hui des tests de type actionnel qui adoptent une démarche «holistique». Quelle est leur contribution véritable en termes de pertinence, de validité et de fiabilité (Horner 2010) ? Peut-on envisager des modes alternatifs ou complémentaires comme la grille CEF-ESTIM (Tardieu et al. 2011) et le C-Test (Reichert 2011) ?

The increasing importance of evaluation in the field of foreign language teaching and learning at school raises a certain number of questions about the legitimacy of

the demand for tests standardised and calibrated on the Common European Framework of Reference for Languages (CEFR). In addition to psychometric tests belonging to the English tradition, action-oriented tests that take a holistic approach have appeared. What is their true contribution to the world of testing, in terms of relevance, validity and reliability? (Horner 2010) Would it not be possible to profit by alternative or complementary methods such as the CEF-ESTIM grid (Tardieu & al., 2011) and the C-Test (Reichert, 2011) ?

Pour une certification contextualisée du français dans le monde : intégrer les approches actionnelles et les approches portfolio.

MATHILDE ANQUETIL80

L'analyse des dimensions sociales de la validité des tests internationaux standardisés en contexte éducatif hétérogène nous amène à proposer des pistes d'évaluation prenant en compte les contextes, les visées, les voix des sujets impliqués. Mais les approches alternatives doivent se doter d'instruments de légitimation pour s'intégrer dans le champ social de la certification.

L'analisi delle dimensioni sociali della validità dei tests internazionali standardizzati in contesto educativo alloglotto induce a proporre alcune piste per una valutazione certificativa più rispettosa dei contesti, delle finalità e delle voci dei soggetti implicati. Ma gli approcci alternativi devono dotarsi di strumenti di legittimazione per integrare il campo sociale della certificazione.

Pluralités culturelles et linguistiques : déplacer les problématiques

Approche systémique pour évaluer la compétence de communication interculturelle

DENISE LUSSIER92

L'article aborde les enjeux du développement de la compétence de communication interculturelle en éducation aux langues et de la nécessité de se doter d'une approche systémique qui s'appuie sur quatre composantes : 1) l'adoption d'un cadre de référence validé empiriquement pour une compréhension univoque des divers intervenants, 2) une planification des apprentissages centrée sur la compétence interculturelle et langagière, 3) la mise en place d'interventions adaptées à une pédagogie interculturelle, 4) l'élaboration de situations d'évaluation fiables et respectueuses des apprentissages.

The article questions the process in the development of intercultural communicative competence in language education and the necessity to reach for a systemic approach. Such approach is based on four

components: 1) the adoption of a conceptual framework validated empirically to generate a common understanding among educators; 2) the planning of learning, centered on both language and intercultural competences, 3) the set up of learning activities, adapted to an intercultural pedagogy, 4) the development of situations of evaluation in the respect and reliability of the learning activities.

Le graal de la compétence interculturelle : évaluer l'inévaluable en FLE ?

FRED DERVIN 105

Partant d'une légende finlandaise sur un objet magique, cet article interroge l'évaluation de la compétence interculturelle en didactique des langues. L'auteur passe en revue les critiques du concept, les malentendus afférents, et propose de privilégier une approche formative et critique de l'évaluation de la compétence allant au-delà du culturalisme-différentialisme, de l'individualisme, mais aussi de l'a-politique.

The article starts with a parallel between a Finnish legend about a magical object and intercultural competence in order to problematize its potential assessment. Criticisms and misunderstandings in relation to the concept are first reviewed. Then the author proposes to retain a formative and critical approach to the assessment of intercultural competence beyond culturalism-differentialism, individualism, and the a-political.

Dans les valises du plurilinguisme : quelle conception des normes et des erreurs ?

MARTINE MAROULLÓ LARRUY 113

Après quelques rappels historiques, cette contribution évoque les dérives de modalités évaluatives promues par l'économie néolibérale. En didactique des langues, le plurilinguisme bouleverse les représentations et ouvre des pistes de recherche sur l'appropriation de plus de trois langues. Enfin, du point de vue des normes de référence, la culture éducative francophone induit une tension entre purisme et variation qui doit être prise en compte pour l'élaboration des procédures d'évaluation

Després d'alguns recodatoris històrics, aquesta contribució evoca les desviacions de les modalitats avaluatives promolgudes per a l'economia neoliberal. En didactica de llengües, el plurilinguisme modifica les representacions i obre pistes d'investigació sobre l'aprenentatge de més de tres llengües. Finalment, des del punt de vista de les normes de referència, la cultura educativa francòfona induïx una tensió entre purisme i variació que s'ha de valorar a l'hora dels procediments d'avaluació.

Enseignement bilingue et évaluation : réflexions sur la conception de tests de niveaux

LAURENT GAJO, JEAN-MARC LUSCHER
ET CÉCILIA SERRA 126

Cette réflexion s'inscrit dans une recherche menée à l'École Internationale de Genève sur l'évaluation de productions orales d'élèves de 10 à 12 ans en enseignement bilingue. Nous avons élaboré un test complexe, durant lequel l'apprenant interagit successivement en L1 et en L2, quelle que soit sa langue maternelle. Nous observons les différences dans l'exécution d'activités. Notre but est d'observer ce qui se joue dans l'enseignement bilingue, et de l'influencer par la mise au point d'outils d'évaluation adéquats le rendant comparable et plus efficace.

In this paper, we discuss procedures to be used for the purpose of assessing language proficiency in bilingual education. Our fieldwork is the International School of Geneva, in particular bilingual classes at 5 and 6 grade (age 10 to 12). We bring forward an elaborated oral test where pupils interact by turns in L1 and in L2 whatever their own first language may be. We observe differences in accomplishing activities. Our target concerns the elaboration of appropriate assessing tools in order to make the bilingual teaching more comparable and efficient.

D'une histoire à ses perspectives : évaluer après le CECR ?

Petit retour à Rüschtlikon

DANIEL COSTE 140

Le Symposium réuni en 1991 à Rüschtlikon (Suisse) est à l'origine de l'élaboration du Cadre européen commun de référence pour les langues. Ce « retour » rappelle, s'agissant d'objectifs et d'évaluation, une conception autre que les seules échelles de niveaux. Le dispositif alternatif esquissé, non limité aux capacités communicatives, relève d'une conception modulaire.

The Symposium held in Rüschtlikon (Switzerland) in 1991 is at the origin of the development of the Common European framework of Reference for Languages. The purpose of this "retour" is to recall a proposal then formulated, which differed from the well known competence scales. That proposal, not solely focussed on communicative abilities, is based on a modular scheme.

Varia

L'originalité des CALF comme instance de diffusion du français langue seconde/étrangère au Tchad

AMINATA DIOP 150

P

réésentation

*Évaluer en didactique
des langues/cultures :
continuités, tensions, ruptures*

*Dimensions épistémologiques,
politiques et didactiques*

EMMANUELLE HUVER

UNIVERSITÉ FRANÇOIS RABELAIS DE TOURS, FRANCE

ALEKSANDRA LJALIKOVA

UNIVERSITÉ DE TALLINN, ESTONIE

Enseignement, apprentissage et évaluation entretiennent des relations complexes, qui témoignent de mises en lien diversifiées entre objectifs, contenus, modèles, environnements, méthodologie, pratiques, représentations et matériel pédagogique, mais aussi entre cultures d'enseignement-apprentissage et cultures d'évaluation (Bonniol et Vial 1997, Cardinet et Laveault 2001, Puren 2003). De ce point de vue, l'évaluation présente donc des enjeux à la fois méthodologiques, épistémologiques et politiques, que l'ensemble des contributions réunies ici vise à mettre en lumière.

Dans le domaine de la didactique des langues/cultures (désormais DLC), un certain nombre de travaux francophones ont été produits au tournant des années 1990 (notamment : Bolton 1987, Lussier 1992, Monnerie-Goarin et Lescure coord. 1993). La parution du *Cadre européen commun de référence* (COE, 2001 – désormais CECR) engendre un regain de publications, qui visent essentiellement à proposer des pistes méthodologiques d'adaptation de l'évaluation au CECR (Veltcheff et Hilton 2003, Sampson et Noël-Jothy 2006, Tagliante 2005). Enfin, quelques ouvrages ou numéros de revue parus plus récemment s'inscrivent dans une perspective plus explicite de prise de distance et de mise en lien avec d'autres disciplines (RFLA 2010, Huver et Springer 2011). Il faut noter que sur l'ensemble de cette période, la production scientifique anglophone sur l'évaluation en langue (« language testing ») est restée continue, avec ses réseaux (EALTA, ALTE, ILTA¹, etc.) et ses revues (*Language Testing*, *Language Assessment Quarterly*, *Assessing Writing*, etc.).

1. Respectivement : European association for language testing and assessment, Association of Language Testers in Europe, International Language Testing Association.

Cependant, tant dans la production scientifique francophone qu'anglophone, on observe une prééminence des travaux sur la conception et l'optimisation d'outils, de référentiels et/ou de descripteurs/critères, ce qui témoigne d'une forte centration, voire d'une sur-représentation des dimensions méthodologiques. Les articles de ce numéro illustrent d'ailleurs cet état de fait, et ce sous des formes diversifiées : la plupart des contributeurs s'attachent en effet à présenter des outils ou des dispositifs d'évaluation (C. Tardieu, M. Anquetil, D. Lussier ou encore L. Gajo, M. Luscher et C. Serra²), d'autres centrent leur réflexion sur les modalités d'amélioration des descripteurs (Riba) ou sur l'identification des « méthodologies porteuses » (Laveault).

Si les questionnements méthodologiques apparaissent comme étant au premier plan, l'épistémologie de l'évaluation constitue également un thème récurrent dans différentes publications récentes (par exemple Mottier Lopez et Figari 2012) ainsi que dans plusieurs des articles réunis ici (M. Vial, D. Laveault, P. Chardenet notamment). Au-delà de la diversité de leurs dénominations, qui ne se recourent d'ailleurs pas complètement, deux conceptions voire deux paradigmes de l'évaluation (paradigmes mécaniciste et holistique – cf. *infra*) sont généralement mis en miroir et en tension. Se pose alors la question de leur possible articulation : les enseignants, les acteurs du monde éducatif ainsi que les chercheurs s'attachent en effet essentiellement à articuler, voire à concilier, ces paradigmes, pourtant épistémologiquement opposés, selon des modalités diverses, ce qui renvoie à, dessine, génère des tensions tant au niveau de leurs pratiques que de leurs représentations (cf. *infra*, p. 10).

Enfin, il est intéressant de noter que l'argumentaire initial, qui invitait les contributeurs à interroger et à dépasser une analyse focalisée sur la dimension méthodologique, voire techniciste de l'évaluation, a, excepté l'article de M. Marquilló Larruy, peu donné lieu à une mise en perspective socio-politique des tendances actuellement dominantes. Ceci est curieux au vu du nombre de publications récentes dénonçant le pouvoir excessif pris par l'évaluation ainsi que sa collusion avec une idéologie politico-économique néo-libérale – via, pour le domaine de la DLC, une critique très forte du CECR, de ses finalités et de ses effets. Si la publication du CECR rend plus particulièrement visibles et, conjointement, renouvelle les enjeux et les questionnements liés à l'évaluation, ces dimensions – épistémologiques, méthodologiques, politiques – ainsi que leur articulation génèrent, dans le champ de l'évaluation, des prises de position diversifiées qui nous semblent révélatrices de problématiques transversales à la DLC. La question de l'évaluation offre ainsi un détour à la fois pertinent et riche de sens, en ce qu'il grossit, concentre, voire cristallise, les questions qui traversent la DLC dans son ensemble et les interprétations qui sont faites de l'évolution du champ, entre continuités, tensions et ruptures.

2. Les références bibliographiques non datées renvoient aux contributions du présent volume.

Rupture méthodologique ? Continuité politique ?

Certains chercheurs interprètent l'introduction des niveaux et des descripteurs, la reconfiguration des certifications ou encore la conception et la mise en œuvre de nouveaux outils évaluatifs alignés sur l'approche par les tâches ou la perspective actionnelle comme autant de signes de changement, voire de rupture. Pour certains, il s'agit d'une rupture épistémologique (Puren 2012). Pour d'autres, elle est envisagée sous un angle essentiellement méthodologique et ce, à double titre :

- Du point de vue de l'articulation entre enseignement/apprentissage et évaluation d'une part : le CECR étant interprété comme ayant introduit « une nouvelle approche » de l'enseignement / apprentissage (i.e. « la perspective actionnelle », Bourguignon 2006), il s'agirait en conséquence de construire un nouveau cadre évaluatif (nommé ici par C. Tardieu « approche sociolinguistique intégrative de type actionnel »), dont les scénarios d'évaluation constitueraient l'outil prototypique (cf. Diplôme de compétence en langue, Certificat de Compétences en Langues de l'Enseignement Supérieur) ;

- Du point de vue des modalités de conception des outils d'évaluation d'autre part : le CECR est alors présenté comme étant en rupture avec les pratiques antérieures d'évaluation dans la mesure où il aurait introduit et permis de diffuser une « méthode novatrice » (Riba) fondée sur l'étalement quantitatif et les modèles psychométriques.

Ces analyses ne sont cependant pas partagées. Sont ainsi contestés : l'analyse de l'évolution méthodologique en termes de rupture (Besse 2011), la simulation comme condition de validité des évaluations ou encore le traitement psychométrique des outils comme garant d'une objectivité posée comme incontournable (Chardenet).

D'autres notent au contraire que le CECR s'inscrit dans une continuité méthodologique (la perspective actionnelle constituant alors un prolongement de l'approche communicative) et/ou politique, notamment en termes de politique éducative. Dans ce cadre interprétatif, le CECR prolonge et approfondit les travaux initiés par le Conseil de l'Europe depuis la création de la Division des langues vivantes et s'inscrit dans la continuité de la visée politique de construction d'un espace européen commun développée depuis la fin de la Seconde Guerre mondiale (Coste). L'harmonisation des certifications en langues et la construction de descripteurs de performance standardisés communs sont alors essentiellement envisagées comme s'inscrivant dans cette visée en facilitant la mobilité des citoyens (élèves, étudiants, travailleurs) et la reconnaissance inter-européenne de leurs acquis langagiers (même si cela aurait pu prendre d'autres formes, comme l'analyse D. Coste).

Cependant, de fortes critiques de ces finalités et/ou des modalités de leur réalisation sont récemment apparues, qui visent plus particulièrement l'évaluation, ce qui a engendré et ce qui témoigne de tensions importantes dans ce champ et, plus largement, dans le champ de la DLC.

Critiques et conflits épistémologiques et politiques

Certains chercheurs en DLC ont souligné la domination des certifications et des évaluations standardisées, dont le CECR (ou, tout du moins, une lecture techniciste et centrée sur l'évaluation de celui-ci) constitue à la fois le vecteur et le symptôme. Ils ont également montré que cette domination, renforcée et rendue possible par une diffusion rapide et généralisée des descripteurs et des niveaux communs de référence, avait pour conséquence une forte minoration de l'évaluation formative et de l'auto-évaluation (Huver et Springer 2011).

Ces critiques dénoncent ainsi :

- une survalorisation du paradigme «mécaniciste» (Vial 2001), qui renvoie à des pratiques technicistes, objectivistes et psychométriques, au détriment du paradigme «holistique» (*ibid.*), supposant des pratiques plus qualitatives et (re)légitimant des formes de subjectivité ;
- une réduction de l'évaluation à sa dimension certificative, induisant des effets de «bachotage» contre-productifs dans une perspective formative. La certification est alors à la limite envisagée comme une fin en soi, dans un contexte socio-économique où, paradoxalement, elle est de moins en moins vectrice de mobilité sociale (Dubet et Martuccelli 1996).

On voit ici se dessiner un autre type d'argument, d'ordre politico-économique cette fois : l'évaluation (y compris l'auto-évaluation d'ailleurs) est en effet actuellement massivement érigée en instrument d'un contrôle des individus, des systèmes éducatifs et des politiques publiques inscrit dans le cadre des politiques économiques néo-libérales (Martuccelli 2010, Abelhauser et al. 2011). Dans le champ de la DLC, le CECR (et plus largement les travaux de la Division des politiques linguistiques du Conseil de l'Europe) fait alors figure de fer de lance technocratique de cette politique (Lefranc 2008, Maurer 2011). Ces critiques politico-économiques sont à l'heure actuelle les plus répandues et les plus diffusées, d'autant plus qu'elles s'inscrivent dans un contexte de remise en question profonde du processus de construction européenne, mis à mal par les difficultés économiques que l'on sait, mais aussi par la résurgence des souverainismes, voire des nationalismes.

Enfin, plusieurs travaux, majoritairement anglophones d'ailleurs (Shohamy 2001, Spolsky 1995, Madaus 1990), ont mis au jour l'aporie que constitue un traitement uniquement méthodologique et techniciste de l'évaluation, et ce, quel que soit le cadre de référence adopté. Ces travaux visent essentiellement à mettre au jour les enjeux politiques, voire idéologiques, liés à la mise en place des évaluations et notamment des évaluations certificatives. Ils ne rejoignent cependant pas pour autant (en tout cas pas totalement) les critiques de type anti-libéral : les approches dites « critiques » de l'évaluation considèrent en effet que *toute évaluation* est au centre d'enjeux de pouvoir (Shohamy parle d'ailleurs du « pouvoir des tests »³), qu'il s'agit d'analyser prioritairement aux dimensions méthodologiques, dans la mesure où ceux-ci éclairent et conditionnent celles-là.

L'ensemble de ces arguments ont cela de commun qu'ils dessinent des lignes de rupture qui, cette fois, ne sont pas tant méthodologiques (approche communicative vs perspective actionnelle) qu'épistémologique, didactique⁴ et politique (Huver, à paraître). Ces critiques, ces conflits, voire ces points de rupture, diversement situés, ainsi que les contradictions non explicitées entre les diverses interprétations évoquées ici (méthodologiques, didactiques, politiques) induisent des pratiques et des représentations en tension chez les enseignants, mais aussi chez les chercheurs.

Des pratiques et des représentations en tension

Les enseignants sont pris en tensions entre dénonciation de la domination des évaluations standardisées (dont les certifications) et de leurs effets pervers, reconnaissance de la valeur à la fois symbolique et effective de celles-ci, revendication de liberté pédagogique (que les certifications viendraient, de fait, réduire), respect (plus ou moins contraint) des institutions qui les emploient, respect des traditions éducatives, besoin de cadres communs (les descripteurs et les niveaux communs de référence étant à ce titre perçus comme des outils précieux), etc.

Ces tensions se retrouvent également dans les travaux des chercheurs, qui, soit essaient de « faire sans » ces conflits (en centrant leurs travaux essentiellement sur des dimensions méthodologiques) et/ou de « faire avec » (en proposant des modèles ou des pratiques qui tentent de concilier, d'intégrer ou de dépasser ces conflits et les tensions qu'ils génèrent : Vial, Laveault, Anquetil, Tardieu).

3. Test étant à prendre ici dans le sens plus large que nous conférons, en français, au terme évaluation.

4. Au sens large de « réflexion globale sur les modalités de transmission/appropriation de différentes formes de plurilinguisme » (cf. Castellotti à paraître).

Si ces tensions sont particulièrement visibles dans le champ de l'évaluation en langue(s) (Huver et Springer 2011), elles n'en sont pas moins présentes dans d'autres secteurs s'intéressant aux liens entre évaluation et éducation, comme en témoignent les usages et interprétations contradictoires de certaines notions mobilisées entre autres dans le champ de l'évaluation. Ainsi, la notion d'auto-évaluation est concomitamment et contradictoirement interprétée comme participant d'une visée de démocratisation-socialisation et d'une visée néo-libérale. Dans le premier cas, l'auto-évaluation est définie comme favorisant l'autonomisation des apprenants, la conscientisation des critères attendus, la réflexion sur l'apprentissage (cf. les travaux de G. Nunziati sur l'évaluation formatrice, cités d'ailleurs ici par D. Laveault, les travaux du CRAPEL – et notamment d'H. Holec – sur l'auto-évaluation, etc.). Elle permettrait également de valoriser les capacités individuelles des apprenants, y compris celles qui sont moins liées à des savoirs formels et/ou scolaires, et qui sont, de ce fait, souvent ignorées, voire délégitimées. Dans le deuxième, qui participe d'un positionnement épistémologique, voire politique, radicalement différent, l'auto-évaluation s'inscrirait dans une conception normalisante de l'évaluation et dans le pilotage par les standards et les résultats. Elle se réduirait à un auto-contrôle de conformité et relèverait *in fine* d'une conception de l'individu (et notamment de l'individu apprenant) comme seul responsable de ses actions⁵. Dans le cadre de la DLC, les ambigüités qui en résultent s'incarnent de manière particulièrement parlante dans les conceptions et les usages divers et contrastés du *Portfolio européen des langues* (Huver et Springer 2011).

Ces contradictions, qui ressortissent à la fois d'enjeux politiques et de fondements épistémologiques opposés, traversent également les notions de compétence (Ollagnier 2002 ; Marquilló Larruy ici même), d'autonomie, de réflexivité, ou encore d'acteur social, ce qui donne souvent lieu d'une part à des doubles discours dont les fondements ne sont pas toujours explorés ni interrogés et, d'autre part, à des pratiques en tension qui se déclinent à différents niveaux décisionnels (curricula, instructions officielles, pratiques de classe, etc.).

Tout un chantier d'investigation se dessine ici permettant de développer une approche anthropologique et/ou critique et/ou historique des pratiques évaluatives en DLC et intégrant de manière articulée et historicisée les représentations, voire les imaginaires, des acteurs (évaluateurs, évalués, mais aussi institutions), les évolutions du champ de l'évaluation en langues ainsi que les rapports de force qui s'y dessinent. Dans ce cadre, il s'agirait de considérer, comme l'esquisse Vial ici même, que les sujets constituent « le troisième brin » et qu'à ce titre, ils participent de manière active, singulière, évolutive, située, à « tresser les paradigmes » ou, pour le dire autrement, à faire avec, sans, malgré, contre ce qui est perçu, conjointement parfois, comme continuité, tension et/ou rupture.

5. L'exemple le plus significatif est sans doute celui des entretiens annuels d'évaluation, au cours desquels il est demandé aux salariés d'auto-évaluer leur action (en fait leurs performances)... pour qu'ils arrivent eux-mêmes à la conclusion de la nécessité de leur propre licenciement.

Ruptures méthodologiques, politiques... et épistémologiques ?

De ces premières considérations, on peut tirer les constats suivants :

– les recherches actuelles en évaluation sont essentiellement centrées sur la conception d'outils plus fiables et/ou plus adaptés aux « nouvelles » méthodologies et aux enjeux contemporains de la diffusion des langues ;

– si les pratiques certificatives et standardisantes dominent actuellement le champ de l'évaluation (en langue(s) mais pas seulement), on trouve également de virulentes critiques de ce phénomène, l'évaluation étant alors condamnée non seulement pour ses dérives, mais aussi parfois comme vecteur, en elle-même, de diffusion des théories néolibérales, de domination autoritariste et de contrôle social.

Cependant, l'ensemble de ces critiques, en privilégiant une analyse des idéologies sous-jacentes à l'évaluation, négligent de poser et de débattre de problématiques plus explicitement épistémologiques. En effet, si la domination de l'évaluation par standards est actuellement indéniable (notamment en DLC), et doit à ce titre être interrogée dans la mesure où elle contribue très largement à lisser les pratiques, une critique ne s'appuyant que sur l'analyse des idéologies en présence ne suffit pas à épuiser la question. Les questionnements d'ordre épistémologiques devraient ici être plus largement débattus, d'une part parce qu'ils éclaireraient le champ de la DLC dans son ensemble, et d'autre part parce que les problématiques liées à l'évaluation permettent de les poser avec plus d'acuité.

Poser la question de l'évaluation en langue/culture suppose de conceptualiser (voire de re-conceptualiser autrement) ce que l'on entend par évaluation (cf. les travaux de M. Vial par exemple) en discutant notamment les enjeux liés à la problématique de la subjectivité (cf. Vial ou encore Chardenet). Mais, en DLC, cela suppose aussi de conceptualiser ce que l'on entend par langue et par culture (Huver, à paraître). Or, en l'état actuel, les évaluations contribuent très largement au maintien de la cohésion de la DLC autour de son objet traditionnel, à savoir la langue (au sens de système) et la communication, ce qui contribue à marginaliser d'autres principes pourtant fondamentaux du CECR tels que la médiation (Zarate et Gohard-Radenkovic 2004), la compétence plurilingue et pluriculturelle (Coste, Moore, Zarate 1997), la compétence interculturelle (Lussier 2007 et ici même), ou encore la contextualisation (Coste 2006, Castellotti et Moore 2008).

L'émergence ou la reconnaissance de contextes d'enseignement et de pratiques d'appropriation plus diversifiés, fréquemment collaboratifs, et de toute façon constitutivement plurilingues et pluriculturels, invite pourtant à questionner l'usage d'évaluations certifiantes décontextualisées, présentées comme plus objectives, mais sans doute

moins pertinentes au regard de la complexité et du caractère processuel des phénomènes à prendre en compte (Gajo, Luscher et Serra, Marquilló Larruy). En outre, les politiques linguistiques et éducatives européennes, en impulsant le développement d'une citoyenneté européenne, de la formation tout au long de la vie, de la mobilité, du plurilinguisme et de l'ouverture interculturelle, engendrent de nouveaux besoins en évaluation. Il s'agit en effet de pouvoir rendre compte du caractère à la fois complexe, processuel et singulier des phénomènes désormais placés au cœur de la formation et de l'évaluation, et, par conséquent, d'interroger les normes au fondement des démarches d'évaluation, qui restent souvent centrées sur des dimensions linguistiques et monolingues (Py 2005, Chardenet, Marquilló Larruy).

Plus fondamentalement, un certain nombre de travaux, qui portent sur la pluralité linguistique et/ou culturelle, remettent largement en question les contours de l'évaluation telle qu'elle est actuellement massivement pratiquée (Huver et Springer 2011). En effet, la pluralité (linguistique ou culturelle), dès lors qu'elle est prise au sérieux, déplace les normes (Marquilló Larruy 2003 et ici même) et interroge donc en profondeur les notions de langue (Blanchet, Calvet et De Robillard, 2007) et de culture (Dervin). Par ailleurs, tenter d'articuler pluralité et évaluation suppose de faire une place à l'imprédictible et à l'inobservable (à des formes d'inaccessibles donc)... qui sont comme autant de graines du désordre dans une mécanique qu'on préférerait bien rodée pour être objective. De ce point de vue, la rupture n'est sans doute pas tant dans la nouveauté méthodologique supposée de la perspective actionnelle que dans le déplacement profond des problématiques épistémologiques liées à l'introduction de la pluralité en DLC.

Cette pluralité est déjà intégrée dans quelques outils : outils visant à évaluer les apprentissages bilingues (Gajo, Luscher et Serra) ou encore *Portfolios européens des langues*, qui, au passage, illustrent et posent d'une part la question de la subjectivité en évaluation et, d'autre part, les limites d'une approche strictement méthodologique et fondée sur un « toujours plus de simulation » des situations de communication (Chardenet). On peut toutefois se demander en quoi il est pertinent de penser les notions de langue et de culture, constitutivement faites de complexité et d'imprédictible, par le biais de niveaux, homogènes, linéaires et se répondant terme à terme, quels que soient le domaine envisagé ou les projets des apprenants (Coste), dans la seule perspective d'une comparabilité formelle (qui, de fait, lisse toute aspérité, toute singularité). Poser la question de la pluralité en évaluation revient donc, de ce point de vue, à remettre en question et à déconstruire certaines notions (ce qui constitue donc une forme de rupture épistémologique), mais aussi à en revisiter d'autres, comme le propose D. Coste ici même avec la notion de modularité, les ruptures épistémologiques n'étant pas incompatibles (bien au contraire) avec des formes de continuité dans les questionnements qui se posent en DLC (Besse 2011).

P *arcours de lecture de ce numéro*

On l'aura compris, le présent numéro se propose de partir des problématiques liées à l'évaluation afin d'éclairer plus largement certaines des problématiques qui traversent actuellement le champ de la DLC : le questionnement sur l'évaluation ne constituera donc pas tant un objectif en soi qu'un moyen de réfléchir la place de l'évaluation au regard de l'enseignement/apprentissage, non seulement du point de vue méthodologique, mais aussi, plus largement – et surtout – du point de vue historique, politique, économique, social, épistémologique. Par conséquent, il ne s'agira pas de présenter un panorama censément exhaustif de certifications, d'outils d'évaluation ou de champs d'application. Les contributions rassemblées dans ce volume contribuent plutôt à l'analyse des phénomènes évoqués plus haut, et en constituent également parfois l'illustration.

Les articles réunis dans la **première partie** visent à rendre compte de manière transdisciplinaire des mouvements et des tensions à l'œuvre, non seulement en DLC, mais aussi dans le champ plus large de l'éducation.

L'article de **Michel Vial** propose une réflexion sur l'épistémologie de l'évaluation en sciences de l'éducation en opposant deux paradigmes («mécaniciste» et «biologiste») qui puisent leurs sources dans l'imaginaire des individus impliqués. **Dany Laveault** met en regard les cultures scientifiques anglophone et francophone, qu'il estime au final largement convergentes, les conflits n'étant selon lui pas liés à des cultures d'expression linguistique différente. Il propose également un usage situé des différentes fonctions de l'évaluation, à envisager en termes de pertinence au regard des besoins des élèves et conclut à une nécessité de relativiser un recours systématique à l'évaluation formative. À partir d'une réflexion sur quelques couples terminologiques clés (objectivité/subjectivité, quantitatif/qualitatif, etc.), **Patrick Char-denet** met en évidence les différentes logiques, contradictions et tensions qui traversent le champ de l'évaluation en DLC.

La **deuxième partie** illustre en quoi ces mouvements et ces tensions contribuent à l'évolution et à la structuration des usages évaluatifs, à la focalisation sur certaines dimensions (notamment méthodologiques), à des jeux de négociation, des tentatives de conciliation et d'intégration de phénomènes profondément contradictoires.

Patrick Riba s'attache ainsi à retracer les origines et les principes d'élaboration des descripteurs de compétence en langue du CECR et met également en évidence quelques limites de ce type de démarche.

Claire Tardieu s'appuie sur le projet EBAFLS pour remettre en question la prétendue transparence des niveaux du CECR et propose la mise en œuvre d'une combinaison d'approches évaluatives associant, selon les besoins et les objectifs, une «approche sociointégrative de type actionnel» (i.e. des scénarios d'évaluation) et des tests plus standardisés

(C-Test). **Mathilde Anquetil** part d'une critique du formalisme des certifications externes centralisées pour poser la question du sens des évaluations certificatives pour les individus évalués : elle propose également une combinaison d'approches, toutefois différente de celle de C. Tardieu, puisqu'il s'agit ici d'associer des épreuves de type DELF à des évaluations plus qualitatives et impliquantes (*Portfolios européens des langues* par exemple).

La **troisième partie** traite plus spécifiquement des enjeux et des perspectives offerts par l'intégration des problématiques liées à la pluralité linguistique et culturelle. **Denise Lussier** compare le CECR aux référentiels canadiens et propose un modèle d'évaluation de la compétence de communication interculturelle. **Fred Dervin** interroge les notions de culture et d'interculturel telles qu'elles sont largement mises en œuvre dans le champ de la DLC pour montrer que la compétence interculturelle ne peut pas être évaluée de manière valide avec les outils existants, en grande partie quantitatifs et sommatifs. **Martine Marquilló Larruy** s'attache à la notion de pluralité linguistique et de variation : elle montre que la prise en compte de ces dimensions induit une remise en question profonde des modalités d'évaluation en langues, dont la DLC n'a pas encore tiré toutes les conséquences. **Laurent Gajo, Jean-Marc Lusher et Cecilia Serra** présentent un projet d'évaluation plurilingue mené dans le cadre d'un enseignement bilingue mis en œuvre dans une école internationale suisse et qui vise à évaluer les ressources mobilisées pour accomplir une tâche donnée, notamment les ressources inter- et translinguistiques.

Enfin, l'article de **Daniel Coste** vient, d'une certaine manière, boucler la boucle dans la mesure où sa visée historicisante permet non seulement de « désévidencier » certaines évidences (autour de l'homogénéité des niveaux notamment), mais aussi de tracer quelques perspectives pour l'avenir.

Bibliographie

- ABELHAUSER A., GORI R. et SAURET M.-J. (dir.) (2011), *La folie évaluation. Les nouvelles fabriques de la servitude*, Paris, 1001 nuits.
- BESSE H. (2011), «Six remarques autour et au-delà de la contextualisation du CECR», dans Castellotti, V. et Nishiyama N. (dir.), «Contextualisations du CECR – Le cas de l'Asie du Nord-Est», *Le français dans le monde – Recherches et applications*, n° 50, p. 150-162.
- BLANCHET P., CALVET L.-J., DE ROBILLARD D. (2007), *Un siècle après le Cours de Saussure : la Linguistique en question*, Paris, L'Harmattan.
- BOLTON S. (1987), *Évaluation de la compétence communicative en langue étrangère*, Paris, Hatier.
- BONNIOL J.-J. et VIAL M. (1997), *Les modèles de l'évaluation. Textes fondateurs avec commentaires*, Bruxelles, De Boeck.
- BOURGUIGNON C. (2006), «De l'approche communicative à l'approche communica-tionnelle : une rupture épistémologique en didactique des langues-cultures», *Synergie Europe* n° 1, p. 58-73. <http://ressources-cla.univ-fcomte.fr/gerflint/Europe1/Claire.pdf>.
- CARDINET J. et LAVEAULT D. (2001), «L'activité évaluative en éducation : évolutions des préoccupations des deux côtés de l'Atlantique», dans Figari G. et Achouche M., Barthélémy V. (collab.), *L'activité évaluative réinterrogée. Regards scolaires et socioprofessionnels*, Bruxelles, De Boeck Université, p. 15-35.
- CASTELLOTTI V. (à paraître), «Quelle(s) didactique(s) pour quel(s) plurilinguisme(s) ?», *Hommages à M. Candelier*.
- CASTELLOTTI V. et MOORE D. (2008), «Contextualisation et universalisme. Quelle didactique des langues pour le XXI^e siècle ?», dans Blanchet P., Moore D., Asselah Rahal S. (Coord.), *Perspectives pour une didactique des langues contextualisée*, Paris, Éditions des archives contemporaines, p. 181-201.
- CHARDENET P. (2006), «Évaluer des compétences plurilingues et interlingues», *Synergies Italie*, n° 2, <http://ressources-cla.univ-fcomte.fr/gerflint/Italie2/chardenet.pdf>.
- COSTE D. (2007), «Contextualiser les utilisations du Cadre européen commun de référence pour les langues», *Le Cadre européen commun de référence pour les langues (CECR) et l'élaboration de politiques linguistiques : défis et responsabilités*, Strasbourg, Conseil de l'Europe, http://www.coe.int/t/dg4/Linguistic/Source/SourceForum07/D-Coste_Contextualise_FR.doc.
- COSTE D., MOORE D., ZARATE G. (1997), *Compétence plurilingue et pluriculturelle*, Strasbourg, Conseil de l'Europe.
- DUBET F. et MARTUCCELLI D. (1996), *À l'école : sociologie de l'expérience scolaire*, Paris, Seuil.
- HUVER E., (à paraître), «CECR et évaluation : interprétations plurielles et logiques contradictoires», *Cahiers du GEPE*.
- HUVER E. et SPRINGER C. (2011), *L'évaluation en langues*, Paris, Didier.
- LEFRANC Y. (2008), «La labellisation des centres de FLE : "politique" ou "police" linguistique ?», dans Truchot C. et Huck D., (coord.), «L'analyse des pratiques d'évaluation des politiques linguistiques : une entrée pour l'étude des politiques linguistiques ?», *Cahiers du GEPE*, n° 1, <http://www.cahiersdugepe.fr/index920.php>.
- LUSSIER D. (2007), «Les compétences interculturelles : un référentiel en enseignement et en évaluation», Conférence plénière, Colloque ALTE/CIEP, Paris, 2 avril 2007.
- LUSSIER D. (1992), *Évaluer les apprentissages dans une approche communicative*, Paris, Hachette.
- MADAUS G. (1990), *Testing as a social technology*. Boston, Boston College.

- MARQUILLÓ LARRUY M. (2003), *L'interprétation de l'erreur*. Paris, Clé international.
- MARTUCCELLI D. (2010), « Critique de la philosophie de l'évaluation », *Cahiers internationaux de sociologie*, décembre 2010, p. 27-52.
- MAURER B., (2011), *Enseignement des langues et construction européenne. Le plurilinguisme, nouvelle idéologie dominante*. Paris, Éditions des archives contemporaines.
- MONNERIE-GOARIN A. et LESCURE R. (coord.) (1993), « Évaluation et certifications en langue étrangère », *Le Français dans le monde – Recherches et applications*, août-septembre.
- MOTTIER LOPEZ L. et FIGARI G. (coord.) (2012), *Modélisations de l'évaluation en éducation*, Bruxelles, De Boeck.
- OLLAGNIER E. (2002), « Les pièges de la compétence en formation d'adultes », dans Dolz J. et Ollagnier E. (Éd.), *L'énigme de la compétence en éducation*, Bruxelles, De Boeck Université, p. 183-201.
- PUREN C. (2012), « Des certifications en FLE/LE, pour quoi et pour qui ? ». *Les cahiers de l'ASDIFLE*, n° 23. *Évaluations et certifications*, p. 111-126.
- PUREN C. (2003), « Vers une didactique comparée des langues-cultures », *Études de Linguistique Appliquée* n° 129.
- PY B. (2005), « Quelle place attribuer à la linguistique dans la formation des enseignants de langues ? L'exemple de l'enseignement plurilingue », dans Mochet M.-A. et al., *Plurilinguisme et apprentissages, Mélanges Daniel Coste*, Lyon, École normale supérieure. p. 113-120.
- RFLA (2010), « Évaluer les compétences langagières », *Revue française de linguistique appliquée*, Vol. XV.
- SAMPSONIS B. et NOËL-JOTHY F. (2006), *Certifications et outils d'évaluation en FLE*, Paris, Hachette.
- SHOHAMY E. (2001), *The power of Tests : A critical perspective on the uses of language tests*, Essex, Pearson Education.
- SPOLSKY B. (1995), *Measured words : the development of objective language testing*, Oxford, Oxford University Press.
- TAGLIANTE C. (2005), *L'évaluation et le Cadre européen commun de référence*, Paris, Clé international.
- VELTCHEFF C. et HILTON S. (2003), *L'évaluation en FLE*, Paris, Hachette Français Langue Étrangère.
- VIAL M. (2001), *Se former pour évaluer, Se donner une problématique et élaborer des concepts*, Bruxelles, De Boeck.
- ZARATE G. et GOHARD-RADENKOVIC A. (2004), *La reconnaissance des compétences interculturelles : de la grille à la carte*, Paris, Didier.

U *n champ
en mouvements
et en tensions*

MICHEL VIAL

DANY LAVEAULT

PATRICK CHARDENET

L'évaluation en éducation, ses paradigmes et les problématiques identitaires

MICHEL VIAL

UNIVERSITÉ AIX-MARSEILLE 1, FRANCE

Ce texte est un bilan des recherches effectuées sur le concept de paradigme et ses effets dans le comportement de l'évaluateur en éducation. Le propos sera donc d'abord situé dans le champ des ressources humaines : non pas seulement du scolaire mais dans la relation éducative. Il sera explicité ensuite ce que paradigme peut vouloir dire en évaluation. Enfin il sera pointé les problématiques que ce concept déclenche chez les évaluateurs. La formation à l'évaluation s'avère être une entreprise difficile, en ce sens qu'elle touche aux dynamiques identitaires.

L'évaluation en éducation

Si on fait un bilan des recherches sur l'évaluation en éducation, on peut dire aujourd'hui que l'évaluation n'est plus réductible au scolaire ; elle n'est pas la seule mesure des acquis ; elle n'est pas réductible à la vérification des savoirs, ni à la pratique de la notation. Elle ne se restreint pas à des moments identifiables, spécifiques, appelés « bilans, tests, épreuves ». Nous parlons bien de l'évaluation en éducation, qui institue une relation éducative entre les « évaluant » (tous ceux qui prennent part à un dispositif d'évaluation), que ce soit dans le monde scolaire, en entreprise, dans les soins infirmiers ou dans les institutions publiques. Dans ce contexte, on peut attendre quatre bénéfices avouables d'une évaluation en éducation, quatre apprentissages :

- un apprentissage *cognitif* : évaluer, c'est produire des connaissances sur l'objet d'évaluation,

- un apprentissage *instrumental* : évaluer, c'est construire des stratégies d'amélioration ou de changement,
- un apprentissage de *l'orientation* : évaluer, c'est proposer une visée, une politique pour l'organisation (l'institution, le service, la classe),
- un apprentissage de la *professionnalisation* des acteurs : évaluer, c'est se repérer dans sa pratique et être réflexif, autonome et responsable.

Éducation n'est pas ici synonyme de scolaire. L'école n'est pas le seul lieu où l'éducation se fait : la famille d'abord et d'autres secteurs professionnels s'y consacrent (Vial 2010)... L'évaluation en éducation est une modalité de la relation éducative, dans un champ appelé les ressources humaines.

Parler de ressources humaines ici n'est pas désigner un stock de ressources naturelles dans lequel on puise à volonté et selon ses propres besoins, dont on use et que l'on « gère » pour en profiter. L'évaluation ne peut y être réduite à « estimer la capacité des salariés à atteindre les objectifs qui leur sont assignés » (Site <http://fr.wikipedia.org>). C'est confondre le poste de direction des ressources humaines et le champ social du même nom. « Ressources humaines » désigne l'ensemble des métiers de la relation humaine. C'est dire que ceux-là avec qui on travaille ont des ressources *parce que* ce sont des humains, des ressources incomparables aux ressources de la nature – et que, sans eux, rien ne se ferait. Ils sont l'entreprise qu'ils servent.

Et servir ici ne veut pas dire « être au service de », donc être le domestique, l'inférieur, le serviteur du XIX^e siècle, assujetti par nature et par condition aux races supérieures des chefs ou des clients qu'il doit satisfaire. Servir veut dire ici participer à un projet, un idéal, une entreprise humaine, une aventure. Autrement dit, on ne demande pas aux dits « managers » d'y croire, ni de s'y croire, mais de gouverner... des humains et de ne pas se contenter de les gérer comme des marchandises.

L'évaluation est une pratique greffée sur une autre pratique (par exemple l'enseignement, l'intervention sociale, le soin, la formation, etc.) et elle s'en distingue par un *travail des critères*, une convocation ou une organisation *d'un débat de valeurs* qui traduit le désir de rendre compte ou de rendre intelligible ce que l'on fait. Évaluer implique d'établir un dispositif qui n'est pas mécanique, technique, algorithmique, malgré tous les efforts méthodologiques. C'est une *rencontre* (plus ou moins contrôlée selon le projet d'évaluation que l'on porte, selon que l'on est en posture d'expert conseil ou d'accompagnement). Une rencontre où les émotions et les projections sont inévitables. C'est dire que l'évaluation est comme la relation humaine dont elle est une modalité particulière : « porteuse de doutes, d'incompréhensions, de rapports de pouvoir, de violences, de séductions ; elle provoque, interpelle, fait éclater toute neutralité, entraîne vers des tensions psychiques et des angoisses » (Cifali M. 1994, p. 249). Même l'expert qui s'abrite derrière son objectivité et sa neutralité rationnelle

n'y échappe pas, sauf à s'aveugler sur son désir de maîtrise et de toute puissance. C'est dire que l'évaluation n'est pas réductible au contrôle. L'évaluation en éducation travaille le rapport au savoir (Charlot 1997 - Beillerot, Blanchard-Laville, Mosconi 1996). On ne réduit plus le savoir aujourd'hui au savoir objectivé qui est dans les livres et relève des théories de l'apprentissage : le savoir patrimonial, dont on hérite, qui était là avant et que l'on acquiert dans le versant instruction de la relation éducative. Dans le versant maturation de la relation éducative sont en priorité travaillés les savoirs que l'on a sans le savoir (le savoir insu, informel) et le savoir qui s'éprouve, le savoir expérientiel. En rester à l'instruction, c'est être dans le pilotage de l'autre. S'intéresser à l'appropriation des savoirs, c'est entrer dans l'accompagnement pour la maturation. Ce qui correspond exactement à la distinction (et non pas l'opposition) entre le contrôle comme vérification de l'acquis et le reste de l'évaluation comme interprétation de la valeur de ce qui est fait (Ardoino, Berger 1986). On retrouve ici la dialectique entre « imposer un objectif assorti de sa trajectoire » et « faire en sorte que l'accompagné fasse son chemin ». Avec la difficulté qu'on connaît : « parce que l'éducation se déploie dans la tension d'une contradiction, il est constamment et toujours tentant de réduire les contraires, en les niant ou en les déniait, et en rendant absolue une des deux finalités au détriment de l'autre. » (Beillerot 1998, p. 86).

Donc penser l'évaluation en éducation implique de mettre hors champ :

- L'évaluation des objets fabriqués dans le monde industriel où l'on donne primat aux procédures quand la sécurité est confondue avec qualité et qu'on multiplie les contrôles techniques.
- L'évaluation dans la finance internationale où, sous l'effet de la mondialisation, on impose un modèle de la mesure des coûts, une surveillance des mises en actes des politiques publiques par des standards indiscutables appelés « normes ».
- L'évaluation médicale qui est un décalque de la méthode expérimentale, pour traiter les symptômes, où il s'agit de faire la preuve de l'efficacité du traitement.

Ces trois types d'évaluation obéissent à d'autres impératifs que l'évaluation en éducation. Ce qui n'exclut pas qu'elles puissent avoir des liens, mais « il n'existe de rapport qu'entre des termes radicalement distincts » (Sauret 1995, p. 30).

Zoom sur le concept de paradigme en évaluation

Dans ce contexte, justement, d'une relation éducative dans les métiers des ressources humaines, à quels impératifs obéit l'évaluateur ? À un

ensemble d'ingrédients culturels acquis au hasard de la vie et de la formation, qui définissent son positionnement épistémologique et agissent comme des surnormes déterminant ses choix dans les modèles et les dispositifs d'évaluation existant. C'est le résultat des recherches que j'ai effectuées pendant plus de dix ans (Vial 1997, 2000, 2001, 2003, 2007, 2010). Pour les identifier, il m'a fallu passer par une exégèse de textes sur l'évaluation et de textes comptes rendus d'évaluations ; faire des entretiens avec des concepteurs de dispositifs, des entretiens avec des responsables d'évaluations et des entretiens avec des évalués. Et puisqu'il a été nécessaire de les nommer, ces ingrédients de la posture d'évaluateur sont devenus les concepts de :

- «registre de pensée»,
- «mode de pensée» (terme emprunté à Dufour 1996),
- «logique d'action» (terme utilisé en sociologie : Bernoux, Herreros 2000),
- *paradigme profane* que nous allons développer ici (inspiré par «les savoirs profanes» (Feldman, Le Grand 1996) et
- *vision du monde*, qui existait.

Ces «concepts de l'évaluant» forment système, sont interdépendants. Le tout constituant l'essentiel du processus de référenciation : cette convocation de références pour prendre un certain regard, un angle de prise de vue, en privilégiant certains savoirs sur un objet pour pouvoir en parler.

Je vais m'attacher ici à «paradigme profane». C'est Jacques Ardoino qui le premier en éducation à propos de l'évaluation attira l'attention sur ce concept, faisant le lien avec la pensée d'Héraclite d'Éphèse pour l'évaluation et de Parménide pour le contrôle : «*Tout comme chez les grammairiens, le paradigme est la "loi", la "raison" d'une déclinaison ou d'une conjugaison, le paradigme est, scientifiquement parlant, le noyau de principes et d'hypothèses fondamentaux déclinant et conjuguant les phénomènes d'un "épistème" donné. Les méthodologies en seront la traduction opératoire*» (Ardoino et Berger 1989, p. 219).

Le paradigme est : «Un ensemble de croyances, de conceptions, de valeurs, c'est le modèle conceptuel dont dérivent tous les discours de celui qui en fait son système explicatif» (Berbaum 1982, p. 43) ou : «Je dirai que je conçois le paradigme comme le type de relations institué entre quelques catégories ou notions maîtresses, et qui gouverne les discours, pensées et théories lui obéissant» (Morin 1986, p. 82).

Dans la tradition occidentale sont inculqués deux paradigmes que l'on peut désigner par des étiquettes différentes :

- *l'instrumental et le symbolique* (Riffault 2002, p. 34) : «distinguer deux grandes modalités d'activité de l'esprit tournées vers la connaissance : une modalité qu'on peut dire "instrumentale", celle qui vise les résultats produits dans l'agir ; dans ce cas, il s'agit de connaître en vue d'une plus grande efficacité de l'action ; une modalité qu'on peut dire "symbolique" consistant principalement à ouvrir les fenêtres sur le

monde, à mettre des mots sur les choses, à leur attribuer des significations, à les échanger et ainsi les intégrer dans l'univers humain en tant qu'univers du sens.», ou

– *l'expertise et l'incertitude* (Saint-Arnaud 2001) : l'expertise, c'est quand l'acteur s'appuie sur le savoir homologué (théories, concepts, modèles, grilles d'analyse, etc.) pour expliquer ce qui se passe dans une interaction et planifier son action. Il cherche à modifier son action en fonction des prescriptions de ce savoir. En expliquant son action, il est persuadé de mieux transmettre son expérience. Il ne remet jamais en question la possibilité de contrôler la situation où il se trouve à partir du savoir homologué. Il réduit chaque situation à ce qu'il connaît (à ce qui est contrôlable). Alors que dans le paradigme de l'incertitude, l'acteur s'appuie sur ce qu'il y a d'unique et d'incertain dans chaque interaction pour s'autoréguler dans l'action. Cela suppose d'accepter de vivre un processus par essai et erreurs. L'erreur est anticipée comme inévitable en raison des limites inhérentes au savoir homologué. Il faut anticiper et utiliser l'imprévisible. La réaction subjective, l'affect est plus rapide que toute analyse rationnelle. L'intelligence des émotions, l'intuition y sont valorisées.

– *l'individualisme et le holisme* (Caillé 2000, p. 124-129) : «L'individualisme (méthodologique) prétend faire dériver toutes les actions, règles ou institutions des calculs, plus ou moins conscients et rationnels ; effectués par les individus, posés comme seuls réels. Le holisme [...] pose au contraire que l'action des individus (ou groupes, classes, ordres...) ne fait qu'exprimer ou actualiser une totalité a priori, qui lui pré-existe, et qui apparaît à son tour comme seule réelle.»

L'usage petit à petit a stabilisé deux mots étiquettes pour les désigner : «paradigme mécaniste» et «paradigme biologiste». D'un côté, l'homme est maître de l'outil, le monde est un moteur, le réel existe et se saisit par des procédures adaptées dites scientifiques, l'univocité est possible. De l'autre côté, règnent la poésie, l'évanescence, l'insaisissable, l'émergent, l'imprévisible, voire l'indicible parce que la polysémie est inhérente à l'homme, au monde. Mais peu importe le mot-étiquette dont on se sert pour désigner une lignée paradigmatique : le propre d'un paradigme, c'est que *n'importe quel mot de l'échelle peut à son tour devenir le mot pour désigner l'ensemble*. C'est l'attitude propre à chacune des lignées (puisque tout paradigme «fait ligne» entre ses notions) qu'il importe de reconnaître (et d'abord en soi). Les éléments tiennent les uns aux autres, de telle manière qu'entrer par l'un, c'est avoir à faire avec tous les autres de la même lignée. C'est ce type particulier de lien, qui n'est pas hiérarchique mais celui d'une filiation (d'où l'idée de «lignée»), que la notion de paradigme permet de signifier : comme en linguistique, le paradigme est une déclinaison.

La tradition anthropologique occidentale permet de n'identifier que deux grands types d'attitudes devant le «réel», n'en déplaise aux

partisans de la pensée ternaire ou du triangle (il y a bien, en revanche, «tresse des paradigmes» : le sujet est le troisième brin). D'un côté, le «paradigme analytique et déterministe institutionnel (qui) ne pouvait tolérer la complexité que comme un substitut distingué de la compli-cation» (Le Moigne 1984, p. 50). De l'autre, le paradigme du Vivant, qui survalorise l'impression globale et rejette en arrière-fond l'analyse des parties. On a donc «deux ensembles épistémologiques respective-ment cohérents mais pourtant irréductibles l'un à l'autre» (Ardoino, de Peretti 1998, p. 25), qui sont bien évidemment en tension, ou que l'on met en tension si on les dialectise.

La symétrie entre ces deux attitudes donne lieu à des «constellations sémantiques» (Ardoino, de Peretti 1998, p. 25), des lignées, qui sont alimentées par des idées générales que les gens ont sans trop savoir pourquoi et sur lesquelles ils fonctionnent à l'évidence, entre croyances et opinions. Du côté du mécanique, on croit volontiers à la nature humaine qui postule que l'homme est définissable par des caractères permanents, éternellement vrais, hors du temps. Tandis que de l'autre côté, on adhère volontiers au personnalisme, et on fait l'apologie du vécu et du ressenti dans un cosmos où règnent les énergies et où on postule que n'existent que des passages éphémères, que nous sommes en changement, en éternel devenir. «Il y a d'un côté ce qui est immuable, de l'autre ce qui est passager ; d'un côté ce qui est un, de l'autre ce qui est multiple ; d'un côté ce qui est intelligible, de l'autre ce qui est sensible. Bref, il y a le monde humain du devenir où rien n'est sûr, où tout passe, où tout est éphémère, où il n'y a pas de calcul exact, comme disait Kyoré, le monde de l'à peu près ; et puis il y a le monde des mathématiciens, des philosophes qui essaient de définir des essences immuables : l'un, l'être, l'absolu s'opposant au relatif» (Ver-nant 1999, p. 70).

Alors, la pensée rationalisante, logique, la pensée dite formelle, analy-tique, avec ses principes d'identité, de tiers exclus et de non-contradiction et le positivisme qui en découle, caractérise un certain type d'évaluateur «rigoureux», qui survalorise l'analyse et la distanciation experte, et va s'opposer à celui qui fonctionne dans la pensée du mythe, et qui revendique le relationnel, le qualitatif : les paradigmes sont en lien direct avec les deux logiques d'action en évaluation : le contrôle (le guidage, le conseil) et l'accompagnement.

On a donc en Occident des idées générales, des règles de pensée, des symboles qui alimentent ces deux types d'attitudes. L'évaluateur penche vers le cristal ou vers la fumée (selon le titre d'Atlan 1979), la fermeture ou l'ouverture du sens (selon les termes d'Edgar Morin 1981). Le sujet considère comme évident, soit le presque rien de la taxonomie systématique, soit le n'importe quoi du sens proliférant. Ce sont des investissements symboliques du sujet évaluant, des schèmes sociaux flottants (Berthelot 1990) : il s'agit «d'un dualisme de référents», et non pas de substances. C'est «une dualité sémantique exprimant

une dualité de perspectives», «deux perspectives hétérogènes : non réductibles l'une à l'autre, non dérivables l'une de l'autre» (Ricoeur 1998). Une longue tradition appuie cette dualité dynamique : on peut se rapporter aux courants principaux de la pensée esthétique que Brandi a distingués (1989), aux deux traditions scientifiques exposées par Pourtois et Desmet (1988) ou par Capra (1983)... Au paradigme mécaniciste correspond le «savoir-épistémè» de Lerbet (1984) et au paradigme biologiste son «savoir-gnose». Ainsi la «discipline épistémologique» de Schwartz (2000), qu'il oppose et articule à «la discipline ergologique»...

Dufour, quant à lui, signale que «lorsqu'on considère le savoir non plus sous le rapport de son usage externe mais sous celui de sa cohérence épistémologique, il apparaît tout de suite qu'à l'intérieur de ce qu'on appelait tout uniment "le savoir", il existe une irréductible fracture. [...] À la place de cette unité, on trouve une division dont je m'empresse de dire qu'elle n'est pas neuve, loin de là, puisqu'elle traverse l'histoire de la pensée depuis l'établissement des sociétés démocratiques. Les premières sociétés démocratiques de la Grèce du v^e siècle de notre ère se sont en effet construites à partir de la différenciation de deux modes de penser, de deux formes de savoirs, impliquant des modalités différentes d'être-soi et d'être-ensemble, qu'on identifie très classiquement sous le nom de mythos d'un côté et de logos de l'autre. [...] Il existe en somme deux grandes modalités de discours : l'une renvoie au monde du mythe et de la croyance et l'autre au monde de la vérification, de l'expérimentation et de la démonstration» (Dufour 1996).

Ces deux paradigmes prennent une tournure particulière en éducation, que Beillerot partage ainsi : «En Occident, le mythe de l'éducation s'est enrichi de l'évolution des connaissances et des savoirs ; non tant la connaissance de soi-même comme dans d'autres cultures, mais la connaissance des autres et de la nature. Il s'agit moins de la formation morale et intellectuelle de soi, dans une conquête cosmique de la sagesse, que de l'éducation des autres. Pour survivre et moins souffrir, l'Occident choisit la maîtrise, y compris bien sûr celle des autres hommes, plutôt que le stoïcisme. Certes, il y a toujours eu une autre tradition, minoritaire, celle d'assumer l'état des choses et en particulier la souffrance [...] Les spiritualités diverses renvoient le sujet à lui-même et la rationalité moderne à la transformation du monde.» (Beillerot 1998, p. 30-31.)

En évaluation ces deux paradigmes dessinent deux mondes : celui des évaluateurs inscrits dans le paradigme mécaniciste : «La conception mécaniciste correspond à une explication du complexe par le simple, à une réduction des phénomènes à un ensemble d'événements élémentaires. La composition des événements élémentaires serait essentiellement additive ; il y aurait juxtaposition des éléments simples. Le déterminisme régissant les faits que l'on observe serait essentiellement de type logique. À une telle conception correspond, du point de vue de

la recherche, un recueil des données selon une démarche analytique en fonction d'un découpage logique de la réalité» (Berbaum 1982, p. 43). Et celui des évaluateurs inscrits dans le paradigme du biologisme : le monde est incertain, plein de bruits et de fureurs, d'énergies ; l'homme est soumis aux forces de la vie, aux flux, aux dynamiques. Il doit ruser pour survivre, avoir des stratégies (et non pas des calculs). Convivialité, écoute de l'autre, partage : l'homme est un parmi d'autres. Il privilégie les traces et leurs interprétations, les synthèses, l'appréhension globale. La compréhension par la clinique est revendiquée.

Les paradigmes ne sont pas des savoirs transmis, appris (comme de considérer la terre ronde et non plus plate) mais des choses sues, ressenties. Ce ne sont pas des savoirs objectivables mais des matrices de mobiles, des certitudes. Les repères invoqués ne sont pas issus d'une solide culture dans les champs de savoirs experts mais des reliquats plus ou moins idéologiques qui se rattachent à des bribes de savoirs approximatifs amalgamés sous la forme d'évidences. Ce ne sont pas des savoirs savants, ce sont des paradigmes profanes.

Il s'agit bien d'une « conceptualisation tacite, en acte dans le langage ordinaire » (Clot 1998, p. 148) : ces allant de soi, ces « théories, souvent implicites, inconscientes, ou même contradictoires » (Sainsaulieu 1987, p. 18) avec lesquelles nous fonctionnons. Les paradigmes sont l'utilisation de théories profanes, depuis des « épistémologies ordinaires », des modélisations implicites, insues mais incarnées, saturées d'investissements symboliques forts. Et non pas des paradigmes « scientifiques » qui sont, eux, des programmes de recherche (Kuhn 1972). Vouloir saisir la réalité par des procédures rationnelles de contrôle ou l'approcher avec un appareil clinique pour comprendre sont deux attitudes de l'évaluateur révélatrices d'un système de références, qui dépendent de choix profonds, trop souvent insus, et que la formation à l'évaluation permet d'élucider.

Problématiques

Le paradigme n'est pas, sauf rares exceptions de travail épistémologique, directement nommé, il constitue les fondations du sujet évaluant, qu'il faut mettre au jour à partir d'indices, de symptômes, de signes : apparaissent au détour d'une phrase l'un ou l'autre des repères de la lignée paradigmatique élue, valorisée par l'auteur ou le déni de l'autre brin, refoulé, déprécié. Car les paradigmes se repoussent la plupart du temps (ils sont antagonistes), le sujet est pris dans un *conflit paradigmatique*. Ce qui peut donner des attitudes intransigeantes, entre partisans du contrôle, de la démarche instrumentée et ceux qui

revendiquent l'évaluation du sens de ce qui se fait. Le positionnement paradigmatique pacifié ne peut être que le résultat espéré d'une formation à l'évaluation.

Les évaluateurs qui arrivent parfaitement à refouler une des lignées sont rares. Tel féru de statistique refusera de passer sous une échelle, on le sait. Et tel inconditionnel de l'empathie ne concevra (par exemple pour refuser de faire des évaluations) l'autoévaluation que par le moyen des «grilles d'évaluation pour mesurer le changement», sans s'apercevoir de l'incohérence. Nous sommes plutôt dans l'un ou plutôt dans l'autre de ces deux paradigmes. Et c'est souvent sur le mode de la mauvaise foi ou du déni. C'est dire combien ces deux paradigmes nagent dans *l'imaginaire*.

Certains évaluateurs ont cru qu'il fallait choisir et survalorisent l'un des fils au détriment de l'autre. D'autres évaluateurs oscillent entre les deux dans des dilemmes difficiles, n'étant jamais à l'aise nulle part, et d'autres espèrent mettre en harmonie les termes du conflit paradigmatique, au risque de tuer la dynamique. Rares sont ceux qui utilisent, selon le projet d'évaluation qu'ils portent, les deux conceptions.

Alors, évaluer en éducation demande une formation qui ne peut pas être qu'instrumentale, le temps d'un travail sur soi et de soi, un travail épistémologique (Vial 2005).

Conclusion

La conception de l'évaluation dépend de savoirs profanes, parmi lesquels la notion de paradigme est particulièrement tenace : selon le paradigme privilégié (mécaniciste ou biologiste), l'évaluateur est conduit à mésestimer, soit la promotion des possibles, soit le contrôle. Sans un travail sur soi que la formation initie, « nous connaissons ces duels entre deux éléments. Il s'agit de l'instruction et de l'éducation ; du savoir et de relation ; de la raison et de l'intuition ; de la technique et de l'amour ; du social et du sujet individuel ; de l'intellect et du cœur... Ils émanent de deux antagonismes plus fondamentaux : cœur et raison ; sujet et social. Tantôt l'un, tantôt l'autre rêve de s'imposer. Nous aboutissons sûrement à une impasse : puisqu'il y a simplification de la réalité, il ne peut résulter de gagnant définitif. Un jour ou l'autre, l'adversaire refait surface et l'emporte » (Cifali 1994, p. 256). Faire un travail sur soi, c'est trouver un moyen de sortir de l'opposition manichéenne entre les paradigmes profanes. Le travail épistémologique sur soi est une expérience de la réflexivité, une condition pour l'évaluation, une condition pour la professionnalisation. C'est d'abord et essentiellement mettre au travail son système de références.

Bibliographie

- ARDOINO J. et BERGER G. (1986), «L'évaluation comme interprétation», *Pour* n° 107, p. 120-127.
- ARDOINO J. et BERGER G. (1989), *D'une évaluation en miettes à une évaluation en actes, de cas des universités*, Paris, RIRELF.
- ARDOINO J. et DE PERETTI A. (1998), *Penser l'hétérogène*. Paris, Desclée de Brouwer.
- ATLAN H. (1979), *Entre le cristal et la fumée*, Paris, Seuil.
- BEILLEROT J. (1998), *L'éducation en débats : la fin des certitudes*, Paris, L'Harmattan.
- BEILLEROT J., BLANCHARD-LAVILLE C. et MOSCONI N. (1996), *Pour une clinique du rapport au savoir*, Paris, L'Harmattan.
- BERBAUM J. (1982), *Étude systémique des actions de formation*, Paris, PUF.
- BERNOUX P. et HERREROS G. (2000), «Le changement dans les organisations : entre compromis et pouvoir», *Gestion* n° 6, p. 42-55.
- BERTHELOT J.M. (1990), *L'intelligence du social*, Paris, PUF.
- BRANDI C. (1989), *Les deux voies de la critique*, Paris, Hazan.
- CAILLÉ A. (2007), *Anthropologie du don*, Paris, La Découverte.
- CAPRA F. (1983), *Le temps du changement*, Paris, Rocher.
- CHARLOT B. (1997), *Du rapport au savoir : éléments pour une théorie*, Paris, Anthropos.
- CIFALI M. (1994), *Le lien éducatif : contre-jour psychanalytique*, Paris, PUF.
- CLOT Y. (1998), *Le travail sans l'homme ?* Paris, La Découverte.
- DUFOUR D.R. (1996), «Les deux sujets de l'éducation», dans Bouchard P., *La question du sujet en éducation*, Paris, L'Harmattan, p. 29-44.
- FELDMAN J. et LE GRAND J.-L. (1996), «Savoirs savants, savoirs profanes», dans Feldman J., Filloux J.-C., Lécuyer B.-P., Slez M., Vicente M., *Éthique, épistémologie et sciences de l'homme*, Paris, L'Harmattan, p. 89-105.
- KUHN T. (1972), *La structure des révolutions scientifiques*, Paris, Flammarion.
- LE MOIGNE J.L. (1984), *La théorie du système général, théorie de la modélisation*, Paris, PUF.
- LERBET G. (1984), *Approche systémique et production de savoir*, Maurecourt, UMFREO.
- MORIN E. (1977), *La méthode*, tome 1, *La nature de la nature*, Paris, Fayard, (1981).
- MORIN E. (1986), *La méthode*, tome 3, *La connaissance de la connaissance*, Paris, Seuil.
- POURTOIS J.-P. et DESMET H. (1988), *Épistémologie et instrumentation en sciences humaines*, Bruxelles, Mardaga.
- RICCEUR P.M. (1998), *Ce qui nous fait penser, la nature et la règle*, dialogue avec Changeux J.-P., Paris, Odile Jacob.
- RIFFAULT J. (2002), «Quelques préalables à l'étude des processus de référence théorique en travail social», *Vie sociale* n° 2, p. 27-36.
- SAINSAULIEU R. (1987), *Sociologie de l'organisation de l'entreprise*, Paris, Presses de la fondation nationale des sciences politiques et Dalloz.
- SAINT-ARNAUD Y. (2001), «La réflexion dans l'action, un changement de paradigme», *Recherche et formation* n° 36, p. 17-27.
- SAURET J.-M. (1995), *La psychologie clinique, histoire et discours, de l'intérêt de la psychanalyse*, Toulouse : Presses Universitaires du Mirail.
- SCHWARTZ Y. (2000), «Discipline épistémique, discipline ergologique, Paideia et politeia», dans Maggi, B., *Manières de penser, manières d'agir en éducation et en formation*, Paris, PUF, p. 33-68.
Site <http://fr.wikipedia.org>
- VERNANT J.-P. (1999), *La volonté de comprendre*, Paris, L'aube.

- VIAL M. (1997), *Les modèles de l'évaluation, textes fondateurs et commentaires*, postface de J.J. Bonniol, Bruxelles, De Boeck Université.
- VIAL M. (2000), *Organiser la formation : le pari sur l'auto-évaluation*, Paris, L'Harmattan.
- VIAL M. (2001), *Se former pour évaluer, se donner une problématique et élaborer des concepts*. Bruxelles, De Boeck Université (2009).
- VIAL M. (2005), *La formation par la recherche contribue-t-elle à une approche réflexive de l'évaluation ?* Conférence à la Journée professionnelle IFCS Montperrin, Aix-en-Provence (voir <http://www.michelvial.com>).
- VIAL M. (sous la dir. de) (2010), *Le travail des limites dans la relation éducative : aide ? guidage ? accompagnement ? Analyses de pratiques*, Paris, L'Harmattan.
- VIAL M. et MENCACCI N. (2007), *L'accompagnement professionnel ? Méthode à l'usage des praticiens exerçant une fonction éducative*, Bruxelles, De Boeck.

Plus de convergences que de divergences : rencontre des conceptions francophone et anglo-saxonne de l'évaluation des apprentissages

DANY LAVEAULT

UNIVERSITÉ D'OTTAWA, CANADA

Depuis une trentaine d'années, la recherche en évaluation scolaire connaît une activité sans précédent, en partie en continuité avec les problématiques qui l'habitent depuis longtemps déjà et aussi en réponse aux réformes récentes des programmes d'études (Mottier-Lopez et Laveault 2008). Devant cette profusion de recherches, il apparaît crucial de faire ressortir les points de convergence, mais aussi un certain nombre de problématiques soulevées par la mise en œuvre de nouvelles conceptions en matière d'évaluation. Cet article se propose d'effectuer de tels rapprochements entre les tendances qui traversent les courants de recherche francophone (France, Belgique, France, Suisse et Québec) et anglo-saxon (États-Unis, Canada, Australie, Nouvelle-Zélande, Royaume-Uni et certains pays d'Europe) en matière d'évaluation afin d'identifier, au-delà des divergences, les idées et les méthodologies qui se sont avérées porteuses.

Les recherches, tant du côté anglo-saxon que francophone, démontrent tout l'impact que peut jouer l'évaluation pour soutenir l'apprentissage des élèves et les faire progresser vers des cibles d'apprentissage ambitieuses. Tout en admettant qu'il s'agit d'une simplification, on peut affirmer que les efforts se sont concentrés, du côté anglo-saxon, sur le rôle et l'efficacité de différents modèles d'évaluation formative pour réguler tant l'enseignement que l'apprentissage. Du côté francophone, les efforts se sont cristallisés autour de l'évaluation d'objectifs d'apprentissage formulés en termes

de compétences. Ceux-ci font appel à la mobilisation des connaissances et des ressources internes et externes de l'élève dans l'exécution de tâches complexes s'inscrivant dans des familles de situations authentiques et hautement contextualisées.

L'apprentissage : au cœur de toute évaluation scolaire

Il n'y a rien d'essentiellement nouveau à affirmer que l'évaluation scolaire porte sur les apprentissages. De tout temps, ce fut le cas. Ce qui a changé, cependant, c'est la conception que nous nous faisons de l'apprentissage et avec elle, notre conception même de l'enseignement et de l'évaluation.

Dans une certaine mesure, l'évaluation scolaire a suivi cette tendance alimentée par les recherches dans le domaine cognitif qui démontrent le rôle actif joué par l'apprenant dans la construction du savoir et du sens. Cette perspective est au cœur des modèles constructiviste et socioconstructiviste de l'apprentissage du côté francophone, mais n'en est nullement l'apanage exclusif. Les modèles de la psychologie cognitive et sociocognitive ont également exercé une influence indéniable, en particulier du côté anglo-saxon. Ils ont provoqué un déplacement du centre d'intérêt de l'évaluation du produit vers les processus cognitifs et métacognitifs en jeu et vers des pratiques qui impliquent davantage l'élève et ses pairs dans le processus de prise d'information et d'évaluation (Cardinet et Laveault 1996).

Les recherches, tant du côté anglo-saxon que du côté francophone, ont révélé le décalage sérieux qui pouvait exister entre les apprentissages scolaires et leur généralisation que ce soit à des situations autres que celles ayant servi à apprendre ou encore à des situations proches du quotidien de l'élève. Les nouveaux programmes d'étude mettent d'ailleurs l'accent sur l'intégration des matières en vue de favoriser l'intégration des apprentissages, leur transfert et leur généralisation. Du côté anglo-saxon, le mouvement associé à l'évaluation dite « authentique » a accompagné ces nouveaux développements en favorisant le développement de tâches riches et complexes, fortement contextualisées et ancrées dans le vécu de l'apprenant. Du côté francophone, c'est autour du besoin d'inscrire les connaissances dans des compétences plus vastes qui leur donnent sens que les efforts se sont déployés (IGEN 2007).

Tous ces développements ont eu pour effet d'élargir les domaines de l'apprentissage et de l'évaluation (Mottier-Lopez et Laveault 2008). La réalisation de tâches complexes, individuellement ou en équipes, est

sans doute plus conforme à la vie quotidienne et au marché du travail où les habiletés d'interaction sociale, la capacité de travailler en collaboration et l'appartenance à des réseaux prennent de plus en plus d'importance. Enfin, la prise en compte des multiples facettes de l'acquisition des compétences a entraîné un intérêt grandissant pour des formes d'évaluation qui tiennent compte de la continuité de l'apprentissage sur la durée et le long terme ainsi que de ses aspects qualitatifs et quantitatifs. Si les connaissances déclaratives se dénombrement facilement, il n'en va pas de même de l'apprentissage de capacités nouvelles et d'attitudes telles que l'ouverture d'esprit, la manifestation d'empathie ou le « leadership » qui consiste en la capacité individuelle ou collective d'exercer une influence et de prendre l'initiative d'une action.

Plusieurs des changements qui surviennent dans le domaine de l'évaluation résultent d'un intérêt accru pour des objets d'apprentissage plus complexes et de notre approfondissement des modèles cognitifs. Mieux l'apprentissage est compris, mieux il est évalué. La réciproque de cette affirmation est également vraie : c'est en développant des méthodes novatrices d'évaluation que s'accroît la validité des inférences que nous formulons à partir des résultats d'apprentissage quant aux processus cognitifs qui leur donnent lieu. Dans la perspective où l'évaluation scolaire peut servir à soutenir l'apprentissage et à piloter les progrès de l'apprenant, plusieurs auteurs (en anglais, voir notamment Black et William 2009 ; en français, voir Allal et Mottier-Lopez 2005) soutiennent que mieux nous évaluons l'apprentissage, plus grandes sont les probabilités que les ajustements ou régulations en support à l'élève seront formateurs.

Évaluation formative, formatrice ou soutien d'apprentissage?

Suite aux développements théoriques sur l'apprentissage, la validité des résultats des méthodes traditionnelles d'évaluation a été progressivement remise en question. En particulier, la réponse apportée par l'évaluation formative à l'adaptation des stratégies pédagogiques est rapidement apparue comme moins satisfaisante.

Du côté francophone, le modèle de régulation s'est avéré particulièrement fécond parmi les chercheurs fortement influencés par les théories socioconstructivistes (Allal 2010 ; Allal et Mottier-Lopez 2005 ; De Ketele 2006 ; Perrenoud 1998). Ces chercheurs francophones ont été davantage préoccupés d'aligner l'évaluation sur l'acquisition de compétences, d'où une emphase marquée sur les processus cognitifs plutôt

que sur les activités, sur les théories d'apprentissage plutôt que sur les contenus d'apprentissage prévus par les programmes d'études par objectifs. Du côté anglo-saxon, la problématique de la régulation s'est cristallisée autour d'un effort pour reconceptualiser l'évaluation formative (*formative assessment*) autour d'une nouvelle appellation : « évaluation soutien d'apprentissage » (*assessment for learning* ; abrégé en français : ÉsA, Allal et Laveault 2009). Pour plusieurs auteurs anglo-saxons, les deux expressions sont interchangeables. Pour d'autres, elles présentent des distinctions importantes.

Pour Stiggins (2005), l'ÉsA consiste à informer prioritairement l'élève sur lui-même et sur sa progression au moment où celle-ci se produit. Cette conception met l'emphase sur le contrôle de l'élève sur ses apprentissages par rapport à une évaluation qui développerait uniquement le contrôle de l'enseignant. Selon William (2011), il importe de juger l'arbre à ses fruits. Il ne suffit pas que l'évaluation ait le projet de faire progresser l'élève vers les objectifs d'apprentissage, encore faut-il qu'elle se traduise par des effets significatifs qui ne se seraient pas produits s'il n'y avait pas eu d'évaluation. Black et William (2009 : 9) préfèrent donc redéfinir l'évaluation formative en des termes plus fonctionnels, celle-ci étant utilisée pour décrire toute évaluation dont la conception et la mise en œuvre ont pour effet d'accroître les chances qu'une décision concernant l'apprentissage soit mieux fondée qu'en l'absence d'une telle évaluation.

Pour Bennett (2011), ces débats de terminologie obscurcissent encore davantage les conclusions que l'on peut tirer de la recherche sur l'évaluation. Les méta-analyses associant l'évaluation formative à des effets expérimentaux forts sont contestables à plusieurs titres et font ressortir la diversité des formes de mises en œuvre de l'évaluation formative à tel point qu'il est devenu difficile d'en attribuer l'impact à l'une ou l'autre de ses composantes. Allal et Mottier-Lopez (2005) font état des avancées des pays francophones sur le plan théorique, mais déplorent tout autant le peu de recherches empiriques sur l'évaluation formative.

Bennett (2011) fait ressortir toute l'importance d'une théorie de l'action pour l'évaluation de l'impact de ce qu'il conviendrait mieux d'appeler « les hypothèses formatives ». Aborder ainsi l'évaluation formative permet de concevoir une multiplicité de méthodes par lesquelles l'évaluation peut soutenir l'apprentissage : que celles-ci soient plus ou moins formelles, plus ou moins instrumentées, plus ou moins continues et finalement plus ou moins régulées par l'élève. C'est cette flexibilité et diversité de formes que peut prendre l'évaluation formative que plusieurs experts internationaux de pays anglo-saxons (États-Unis, Canada, Australie, Nouvelle-Zélande, Royaume-Uni et certains pays d'Europe) ont cherché à traduire en définissant l'ÉsA de la façon suivante : « l'évaluation-soutien d'apprentissage fait partie des pratiques quotidiennes des élèves et des enseignants qui, individuellement et en interaction, réfléchissent sur et réagissent à l'information provenant

d'échanges, démonstrations et observations afin de favoriser les apprentissages en cours» (Allal et Laveault 2009, p. 102). Cette définition globale cherche à établir un équilibre entre le rôle de l'enseignant et le rôle de l'élève, et propose d'étendre l'évaluation jusque dans les questionnements, dialogues et échanges continus entre l'enseignant et les élèves de même qu'entre les élèves.

Le débat terminologique anglo-saxon n'est pas sans rappeler une problématique francophone semblable soulevée par la distinction apportée par Scallon (2000) entre évaluation *formative* et évaluation *formatrice*, distinction apparue nécessaire pour rendre compte des travaux de Nunziati (1990). L'évaluation formatrice pose comme point de départ que les instruments privilégiés de la construction des apprentissages sont : l'appropriation par les élèves des critères d'évaluation des enseignants; la pratique de l'autocontrôle : l'autogestion des erreurs; la maîtrise des processus d'anticipation et de planification des actions.

Vial (1997 in Scallon 2000, p. 266) va jusqu'à remettre en question l'influence que la régulation exerce sur l'autoévaluation : « celle-ci doit s'inscrire dans une démarche de pleine autonomie, loin de toute conformité à des normes préétablies ». Hadji (1997) propose qu'un principe majeur d'une éthique de l'agir évaluatif soit le suivant : « N'exercer son pouvoir d'évaluateur que s'il contribue à la prise de pouvoir sur soi de l'évalué ». Ces positions des auteurs francophones tels que Nunziati, Vial et Hadji sont très proches de la conception d'évaluation en tant qu'apprentissage (*assessment as learning* : Earl 2003) qui élève l'autoévaluation au rang de compétence à développer. Pour Earl, il ne suffit pas de faire de l'étudiant un collaborateur du processus d'évaluation, mais un acteur critique qui fait le lien entre évaluation et apprentissage.

En résumé, les divergences apparaissent davantage dans la modulation des méthodes pour soutenir l'apprentissage, plus particulièrement dans le rôle joué par l'élève lui-même dans la régulation de ses apprentissages et le rôle joué par l'évaluation dans le développement du contrôle de l'élève sur ses apprentissages. La question se pose alors de déterminer quel est le niveau souhaitable d'autorégulation et jusqu'où il faut aller dans la dévolution du contrôle à l'élève.

Le modèle de régulation apporte un élément de réponse à cette question en permettant de situer l'évaluation sur un continuum variant d'une régulation assumée par l'enseignant à une régulation assumée par l'élève, dans un but qui sert principalement l'enseignant ou principalement l'élève. L'enseignant navigue entre perfectionner ses outils d'évaluation à titre d'instruments de régulation externe (« faire de meilleures évaluations ») et faire en sorte que l'évaluation devienne un agent du développement des capacités d'autorégulation des élèves (« faire de meilleurs évaluateurs »). Dans un cas comme dans l'autre, il est difficile de concevoir comment la régulation externe pourrait soutenir l'apprentissage sans que les stratégies d'évaluation soient

davantage diversifiées et adaptées aux besoins de l'élève et sans que soit prévue éventuellement une forme de désétayage favorisant une autorégulation progressive de l'apprentissage visé. L'évaluation joue pleinement son rôle de soutien non seulement lorsqu'elle renseigne l'enseignant et l'élève sur les résultats d'apprentissage, mais aussi lorsqu'elle contribue à développer la capacité de l'élève à exercer un contrôle sur ceux-ci à travers la prise en charge de ses processus cognitifs, métacognitifs, motivationnels ou émotionnels.

Les compétences : des objectifs qui mobilisent

Tant du côté francophone qu'anglo-saxon, il est possible d'affirmer qu'il existe une conviction largement répandue que l'évaluation formative peut jouer un rôle constructif en soutien à l'apprentissage lorsque se trouvent réunies les conditions favorables. Un autre point de convergence est sans doute l'intérêt pour des objets d'apprentissage qui favorisent l'intégration et la mobilisation des connaissances, des capacités et des attitudes des élèves dans des familles de situations complexes et authentiques. Il s'agit ici des compétences en tant qu'objectifs d'apprentissage.

L'évaluation formative et l'approche par compétences relèvent toutes deux de finalités communes et difficilement conciliables en apparence : élever les exigences (« standards ») et accroître la réussite du plus grand nombre d'élèves. Si, dans leur formulation la plus globale, ces finalités remportent l'adhésion, les moyens d'y parvenir font l'objet de nombreuses discussions au niveau scientifique et de controverses tant au niveau des politiques éducationnelles que de la pratique. Le dicton anglais « The devil is in the details » s'applique bien ici.

Le choix des compétences comme objectifs d'apprentissage a entraîné d'importantes réformes de l'enseignement et de l'évaluation dans plusieurs pays francophones telles que l'utilisation de grilles d'évaluation descriptive de performances complexes, le recours à des tâches représentatives de familles de situation ou encore à des modèles d'évaluation en plusieurs phases destinés à évaluer des niveaux variables de mobilisation des compétences (Laveault 2010). L'évaluation d'objectifs en termes de fonctions, tout comme les compétences, suppose que soit établie « une progression de situations » (Cardinet 1984, p. 33) ou « parcours de développement » (Tardif 2006). Tout en admettant qu'il soit possible d'établir une telle progression, Cardinet reconnaît qu'il s'agit de la principale difficulté d'une approche « fonctionnelle ». Le choix d'une approche par compétences dans la conception de programmes d'études comporte son lot d'avantages et d'inconvénients.

Selon Cardinet (1984 : 17), les objectifs sont d'abord l'expression d'un jugement de valeur : « un type de résultats n'est pas plus "vrai" qu'un autre, mais simplement plus "utile" », dans un contexte donné, avec des finalités particulières, etc.

Parmi les impacts recherchés de l'approche par compétences, figure le besoin de concevoir des programmes d'études qui préparent l'étudiant à exercer des fonctions qu'il est difficile de prévoir. Comme le dit Marton (in Bowden 2000) : « In order to become capable of dealing with a varying future we must have met a varying past ». La notion même de compétence est donc intimement associée à la notion de « famille de situations » qui devrait préparer l'élève à faire face à des situations nouvelles (Beckers 2011).

Selon Carette et Dupriez (2009) les notions de « famille de situations » et de « famille de tâches » constituent des ensembles où « une tâche complexe présenterait des invariants qui seraient communs avec d'autres tâches » (page 34). Une telle conception est fortement répandue tant du côté francophone qu'anglo-saxon. Or, selon Bowden (2000) une telle conception limitée de familles de situations conduit à l'apprentissage de « blocs de savoirs spécialisés » isolés les uns des autres et à la formation d'un type d'expert dont les compétences se limitent à la reconnaissance de ces blocs et à la mobilisation des ressources qui y sont rattachées. Le principal défi selon lui consiste à faire en sorte que le changement de contexte ne soit pas qu'un habillage de la tâche mais fasse partie de la solution du problème. Bowden (2000) propose une conception du transfert où l'étudiant s'approprie un concept dans une variété de contextes parce qu'il « a développé la capacité de juger quels aspects du concept sont pertinents à un contexte particulier » (notre traduction). L'évaluation de compétences ne saurait donc se restreindre à la résolution de problèmes dans une variété de contextes mais inclure aussi la capacité à découvrir comment le changement de contexte fait partie du problème.

L'emphase sur des objectifs formulés par compétence a également connu des dérives non souhaitées, l'une des plus fréquemment mentionnée consistant à opposer objectifs de « compétence » et de « connaissance ». Cette opposition s'est avérée stérile et a probablement eu des effets négatifs chez certaines catégories d'élèves pour lesquels les défis se situent au niveau de l'acquisition de connaissances. Difficile de mobiliser ces élèves et leurs connaissances vers l'acquisition de compétences lorsque celles-ci ne sont pas au rendez-vous pour débiter. L'occultation des connaissances par les compétences a eu des effets imprévus, dont celui de contribuer à remettre à l'avant-plan l'importance des premières. Mais s'il faut se soucier des « connaissances inertes » (IGEN 2007) et de la capacité des élèves à acquérir des connaissances conditionnelles qui mobiliseraient connaissances et habiletés dans les contextes appropriés, qu'en est-il des élèves qui n'ont ni acquis ni maîtrisé suffisamment un certain nombre

de connaissances et habiletés préalables ? Un enseignement structuré et explicite avec de tels élèves apparaît mieux adapté dans ces cas qu'une évaluation formative qui submergerait l'élève de feedbacks qu'il peinerait à interpréter et à prioriser (Hattie et Timperley 2007). Si l'enseignement des compétences établit des cibles ambitieuses pour l'élève, il ne doit pas nous détourner des besoins de ceux d'entre eux qui peinent à acquérir certaines connaissances et capacités de base et qui nécessitent une attention particulière. Il est plus important d'adapter l'évaluation des compétences à l'élève que d'adapter l'élève à l'évaluation des compétences.

Des fonctions et des rôles qui se complètent

Le milieu scolaire reconnaît aujourd'hui plusieurs rôles et fonctions à l'évaluation. Dans certaines situations complexes, ces fonctions peuvent parfois être difficiles à différencier et à articuler entre elles. Tant du côté francophone qu'anglo-saxon, des auteurs font remarquer que l'opposition entre évaluation formative et sommative est malheureuse et confond les finalités de l'évaluation avec la démarche employée (De Ketele 2006 ; Figari 2006 ; Harlen 2005). L'une des conséquences négatives de l'opposition entre les deux veut que l'évaluation sommative ne puisse servir à soutenir les apprentissages des élèves, cette fonction étant accaparée par l'évaluation formative. Tourneur (1985) soulignait déjà le tort injustifié que la réaction négative à l'encontre des examens traditionnels faisait subir à l'évaluation sommative. Il déplore la confusion conceptuelle qui en résulte : « L'évaluation sommative se trouve réduite au jugement de valeur sur la personne de l'élève ou à la notation des étudiants en termes de pourcentage ou de classement, quand elle n'est pas jugée responsable des abus liés à l'exercice du "pouvoir magistral" ». (p. 7)

Pour De Ketele (2006), il est préférable de différencier les formes d'évaluation selon leurs fonctions principales : « préparer une nouvelle action ; améliorer une action en cours ; certifier au terme d'une action » (p. 102). Les deux dernières fonctions correspondent à la distinction que Stiggins, Arter, Chappuis et Chappuis (2004) font entre « *assessment for learning* » et « *assessment of learning* », quoique dans leur esprit ces deux formes ne s'opposent pas mais se complètent, pour ultimement soutenir l'apprentissage. Selon De Ketele (2006), l'opposition entre sommatif et formatif a donné lieu à un « réductionnisme historique [qui] a conduit à restreindre la démarche sommative à la fonction certificative, la démarche descriptive à la fonction régulatrice ou formative, la démarche herméneutique ou intuitive à la fonction d'orientation » (p. 103).

Évaluer : une compétence à développer

Des observations récentes sur les débats qui animent les courants francophone et anglo-saxon en matière d'évaluation, il ressort plus de convergences que de divergences. Quant aux divergences, elles ne sont pas plus nombreuses entre les deux courants qu'à l'intérieur de ces courants eux-mêmes. Elles font partie intégrante du débat scientifique contemporain. De plus, les grands principes issus de ces convergences reflètent des consensus généraux à travers les pays membres de l'OCDE. En effet, l'OCDE (2011) identifie clairement l'intégration de l'évaluation formative ainsi que la validité des stratégies d'évaluation selon une approche par compétences comme faisant partie des principaux défis des systèmes éducatifs actuels.

Les attentes envers la profession enseignante en matière d'évaluation sont énormes et la formation initiale de même que le perfectionnement professionnel des enseignants ne sont pas toujours au rendez-vous. Il faudra bien pourtant que l'on s'en préoccupe. De l'avis de l'OCDE, il est capital de compter sur le rôle de l'enseignant pour évaluer la progression d'apprentissages complexes de façon continue. Toutefois, ce rôle va bien au-delà de la construction d'outils d'évaluation et du respect des règles de la docimologie. Par un effet de rétroaction inverse, le feedback que l'analyse des résultats des élèves procure à l'enseignant contribue non seulement au perfectionnement continu de sa compétence à évaluer avec justesse, mais aussi de sa compétence à réguler ses stratégies pédagogiques à travers la prise de conscience des effets de ses choix sur la réussite des élèves. Dans ce sens, l'évaluation des élèves est une source d'information à prendre en compte pour faire progresser non seulement les élèves au moyen des rétroactions des enseignants, mais aussi les enseignants, individuellement et collectivement, dans leur rôle de soutien de l'apprentissage, et ce grâce au feedback que leur procure l'analyse des résultats des élèves. C'est en impliquant l'enseignant et l'élève dans des environnements éducatifs riches en évaluation que les compétences en ce domaine, tant celles de l'enseignant que des élèves, pourront le mieux se développer.

Bibliographie

- ALLAL L. (2010), "Assessment and the regulation of learning", dans Peterson P., Baker E., McGaw B. (éds.), *International Encyclopedia of Education*, 3, Oxford, Elsevier, p. 348-352.
- ALLAL L. et LAVEAULT D. (2009), «Évaluation-soutien d'apprentissage». Prise de position formulée par la Troisième Conférence internationale sur l'Évaluation-soutien d'Apprentissage, *Mesure et évaluation en éducation*, 32(2), p. 99-107.
- 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*, Paris, OECD-CERI Publication (What works in innovation in education), p. 241-264.
- BECKERS J. (2011), *Développer et évaluer des compétences à l'école : vers plus d'efficacité et d'équité*, 2^e éd., Bruxelles, Labor Éducation.
- BENNETT R. E. (2011), "Formative assessment : a critical review", *Assessment in Education : Principles, Policy and Practice*, 18 (1), p. 5-25.
- BLACK P. et WILLIAM D. (2009), "Developing the theory of formative assessment", *Educational Assessment, Evaluation and Accountability*, 21 (1), p. 5-31.
- BLACK P. et WILLIAM D. (1998), "Assessment and classroom learning", *Assessment in Education*, 5 (1), p. 7-74.
- BOWDEN J.-A. (2000), *Competency-based education. Neither a panacea nor a pariah*. Disponible en ligne : <http://crm.hct.ac.ae/events/archive/tend/018bowden.html>
- CARDINET J. et LAVEAULT D. (1996), «Dix années de travaux européens et nord-américains sur l'évaluation : quelles lignes de force ?», *Mesure et évaluation en éducation*, 18 (3), p. 1-25.
- CARDINET J. (1984), «Le découpage des objectifs pédagogiques», *Mesure et évaluation en éducation*, 7 (2), p. 5-48.
- CARETTE V. et DUPRIEZ V. (2009), «La lente émergence d'une politique scolaire en matière d'évaluation des élèves. Quinze ans de transformations en Belgique francophone», *Mesure et évaluation en éducation*, 32 (3), p. 23-46.
- DE KETELE J.-M. (2006), «La recherche en évaluation : propos synthétiques et prospectifs», *Mesure et évaluation en éducation*, 29 (1), p. 99-118.
- EARL L. (2003), *Assessment as learning : Using classroom assessment to maximize student learning*, Thousand Oaks, Corwin Press, Inc.
- FIGARI G. (2006), «L'activité évaluative entre cognition et réponse sociale : nouveaux défis pour les évaluateurs», *Mesure et évaluation en éducation*, 29 (1), p. 5-18.
- HADJI C. (1997), «Pour une éthique de l'agir évaluatif», *Mesure et évaluation en éducation*, 20 (2), p. 7-26.
- HARLEN W. (2005), "Teachers' summative practices and assessment for learning – tensions and synergies", *The Curriculum Journal*, 16 (2), p. 207-223.
- HATTIE J. et TIMPERLEY H. (2007), "The power of feedback", *Review of Educational Research*, 77 (1), p. 81-112.
- INSPECTION GÉNÉRALE DE L'ÉDUCATION NATIONALE (IGEN), (2007), *Les livrets de compétences : nouveaux outils pour l'évaluation des acquis*, Paris : IGEN, Rapport n° 2007-048.
- JAMES M. (2006), «Assessment, teaching and theories of learning», dans Gardner J., (éd.), *Assessment and Learning*, London, Sage, p. 47-60.
- LAVEAULT D. (2010), "Classroom assessment in a policy context (French sources)", dans Peterson P., Baker E., McGaw B., (éds.), *International Encyclopedia of Education*, London, Elsevier, volume 3, p. 432-437.

- MOTTIER-LOPEZ L. et LAVEAULT D. (2008), «L'évaluation des apprentissages en contexte scolaire : 30 ans de travaux. Développements, enjeux et controverses», *Mesure et évaluation en éducation*, 31 (3), p. 5-34.
- NUNZIATI G. (1990), Pour construire un dispositif d'évaluation formative, *Cahiers pédagogiques*, 280, p. 48-64.
- OECD (2011), *Evaluation and assessment frameworks for improving school outcomes. Common policy challenges*, Paris, OECD, 10 pages. [Disponible en ligne à : http://www.oecd.org/document/56/0,3746,en_2649_39263231_44567992_1_1_1_1,00.html (Accédé pour la dernière fois en janvier 2012)].
- PERRENOUD P. (1998), "From formative evaluation to a controlled regulation of learning processes : Towards a wider conceptual field", *Assessment in Education*, 5 (1), p. 85-102.
- SCALLON G. (2000), *L'évaluation formative*, St-Laurent, Québec, Éditions du renouveau pédagogique.
- STIGGINS R.J. (2005), *Assessment For Learning Defined*, An international forum on promoting sound assessment practice in every classroom. Portland, OR, Assessment Training Institute. [Consulté le 18 janvier 2012 dans : <http://www.docstoc.com/docs/14728977ASSESSMENT-FOR-LEARNING-DEFINED>]
- STIGGINS R. J., ARTER J. A., CHAPPUIS J. et CHAPPUIS S. (2004), *Classroom assessment for student learning. Doing it right – Using it well*, Portland, OR, Assessment Training Institute.
- TARDIF J. (2006), *L'évaluation des compétences : documenter le parcours de développement*, Montréal, Chenelière-Éducation.
- TOURNEUR Y. (1985), « La certification des compétences », *Mesure et évaluation en éducation*, 7 (4), p. 5-20.
- WILLIAM D. (2011), "What is assessment for learning?", *Studies in Educational Evaluation*, 37, p. 3-14.

D*e quoi l'évaluation peut-elle être le nom ?*

PATRICK CHARDENET

UNIVERSITÉ DE FRANCHE-COMTÉ, FRANCE

L'évaluation, en tant que domaine de connaissances sur des pratiques sociales et en tant que pratique sociale largement distribuée dans les sociétés où la compétition et la compétitivité dominent, peut désigner des segments de réalité très différents. En éducation et dans l'enseignement des langues en particulier, on a pu voir se développer, au niveau européen, des instruments puissants articulés à des politiques linguistiques éducatives, des outils construits sur les différents terrains, mais paradoxalement, une relative absence de travaux de recherche empirique. Avec Jean Cardinet (1983), on adoptera comme principe que les recherches dans le domaine de l'évaluation doivent étudier comment les acteurs procèdent et chercher ensuite à perfectionner les procédures et les procédés dans leur agencement systémique que l'on peut définir comme système global d'évaluation (Chardenet 2010 a). Ce qui conduit ici à envisager la production des données de la recherche par la description des situations d'évaluation, qui se matérialisent par des comportements discursifs. Cette démarche repose sur l'observation et l'analyse des manifestations langagières dans l'opération d'activité évaluative et dans la mise en acte d'évaluation formelle (Chardenet 1999 b), en deçà de toute considération docimologique. Le postulat établi à la suite d'enquêtes sur les pratiques (procédures et procédés¹) suggère que l'on ne peut répondre à la question « qu'est-ce que l'évaluation ? » en s'arrêtant à une simple liste phénoménologique d'actes situés car l'activité évaluative en jeu dans la parole commande d'une certaine façon le discours qui instaure l'évaluation.

Interrogé par une forte demande sociale pour la production d'outils, le chercheur peut répondre de deux façons : celle qui considère l'acte d'évaluation comme un ensemble de techniques dont il s'agit d'améliorer l'efficacité et d'adapter chacune à des contextes variables du marché des langues (supports, publics, objectifs d'apprentissage, objectifs d'enseignement, buts) ; celle qui interroge le fonctionnement cognitif et social de l'activité évaluative et de l'acte d'évaluation afin de mieux en connaître les fondements, les mises en œuvre, les possibilités et les limites pour instruire une production critique et constructive de procédures d'évaluation des apprentissages en langue, socialement acceptables.

1. On trouvera un ensemble de définitions des notions utilisées dans Blanchet P., Chardenet P. 2011 (<http://wiki.auf.org/glossairedlc/Index>, 07/01/2012).

Ayant déjà été amené à m'exprimer sur la construction notionnelle (Chardenet 1999 a, 2011) et sur les effets de la circulation de la notion d'évaluation (Chardenet 1993, 1999 b), j'aborderai ici d'autres mouvements et tensions. En particulier, comment, en prenant pour objet l'évaluation dans le discours, qui organise le passage de l'activité évaluative (partagée par tout être humain naturellement doté du langage et des compétences de caractérisation des objets du monde) à l'acte d'évaluation, on peut construire un système global d'évaluation multi-référentielle comme espace de dialogue pédagogique.

Les tensions entre la mesure et le grain : sens de l'évaluation et évaluation du sens

Dans l'avant-propos de son ouvrage sur la sémantique de corpus, François Rastier (2011, p. 9) explique qu'il interprète (peut-être abusivement) un vers de poème persan mis en exergue et qui distingue le grain du sens et la mesure. La sémantique s'opposant ainsi à une « linguistique de fauteuil » (idem, p. 13) autant formelle qu'hypothétique, elle révèle une dualité entre des méthodes indicielles simplement quantitatives et des méthodes qualitatives complexes de caractérisation : « L'homme intelligent prend le grain du sens, il ne s'arrête pas à la mesure » (Rûmi, Mathnawi, II, 3622). La question qui se pose ici est de savoir si l'on peut mesurer les performances du vivant et si oui, sur quels référentiels stables ? établir des critères, des barèmes et relever des indicateurs d'objectifs atteints ou non, en fonction d'intentions identifiées.

Il ne viendrait à l'idée de personne de mesurer des masses, des distances, du temps avec des instruments d'approximation. Les instruments utilisés dans ces mesures sont en fait constitutifs des référentiels inventés pour les découper et les borner, ce qui leur garantirait toute objectivité et précision. Le temps est représenté par le temps solaire moyen basé sur un soleil moyen fictif qui se déplacerait à une vitesse constante tout au long de l'année. La durée d'un jour solaire moyen est de 24 heures quel que soit le jour de l'année. Ce référentiel posé arbitrairement est mesuré par une horloge atomique qui utilise la pérennité et l'immuabilité de la fréquence du rayonnement électromagnétique émis par un électron. L'impression que le temps tourne rond et que l'horloge bien réglée le mesure parfaitement contribue à une représentation de la juste mesure devenue un rêve docimologique.

L'évaluation fonctionne de la même façon dans sa conception de contrôle : un objectif attendu, des critères qui rendent compte de l'accomplissement, des indicateurs (d'erreur le plus souvent), un

barème arithmétique dont on ne sait pas vraiment expliquer pourquoi il est établi sur une base, 5, 10, 20, 100... Autant dire que, dans ce cas, si cela ne tourne pas vraiment rond, les choses sont très carrées et le sens relégué en second plan. Or, si l'évaluation formelle en langue se construit sur des cadres référentiels standardisés d'attendus pré-construits fermés, on peut s'interroger sur le rendement effectif de l'efficacité de la mesure des compétences des sujets évaluataires (l'évaluataire n'est pas seulement le destinataire de l'acte d'évaluation - notion passive, c'est également le sujet producteur d'inférence - notion active, de ce que l'évaluateur attend de lui à travers une procédure d'évaluation), eu égard à une norme dont la sociolinguistique montre qu'elle est constamment remise en cause (Marcellesi, Gardin 1987 : 146-147). Dans un monde où les flux de communication réels et virtuels se tissent entre le local et le global, les langues internationalisées comme le français se prêtent à des variations complexes qui touchent jusqu'aux modèles langagiers introduits par l'enseignant (Chardenet 2010 b).

De l'activité évaluative à l'acte d'évaluation : une sémiotique différentielle pour l'objectivation

L'évaluation en tant que pratique sociale est fondamentalement une estimation systématisée et non un outil de mesure précise. Ce qui veut dire que l'univers sémiolinguistique en constitue les conditions de fonctionnement par transcodage entre : (1) une activité évaluative naturelle fondée sur le jugement et la catégorisation, et (2) l'acte d'évaluation qui cherche la mesure.

De ce point de vue, on peut autant expliciter la différence d'appréciation d'une même production, par les effets de variation établis par la docimologie critique², que par les effets d'une recherche de différence (Peytard 1999) ou de distinction (Chardenet 1999 c) entre les discours de jugement. Ce qui questionne de fait l'objectivité comme objectif et comme moyen de l'évaluation.

Schématisé comme un ensemble de discours, l'acte d'évaluation qui met en scène une institution (actant qui légitime socialement l'acte par des discours sur l'évaluation : buts, règles et fonctionnement des procédures), un ou des évaluateurs (acteurs soumis à une obligation d'évaluation incluse dans le service d'enseignement ou de formation qui instrumente la procédure : procédés formatifs et/ou sommatifs), un évaluataire (acteur qui soumet un discours réponse aux procédés mis en place par l'évaluateur en s'interrogeant sur les attentes de celui-ci),

2. Étude systématique des examens (modes de notation, variabilité interindividuelle et intra-individuelle des examinateurs, facteurs subjectifs...) fondée par H. Piéron (1929).

peut être décrit (Chardenet P. 2011), à partir de relevés au sein d'une unité d'enseignement ou de formation (classe, établissement, réseau d'établissement, système éducatif).

Par leur spécificité due aux enjeux psychosociologiques qu'ils engagent, ces discours constituent un sous-ensemble particulier de l'ensemble des discours d'éducation ou de formation, qui ne peuvent se réduire aux interactions de classe, aux discours des manuels, aux discours des administrations, des établissements, à ceux des systèmes éducatifs, tous pris séparément. Interreliés, ces discours forment un tout dans lequel ceux de l'évaluation tiennent une place particulière eu égard aux enjeux de sélection articulés à la compétition sociale. Mais cette interrelation constitutive, qui semblerait pouvoir dire qu'elle assure une cohésion entre chacun et une cohérence de l'ensemble, est plutôt vécue dans la pratique de l'évaluation comme une « déhiscence », pour reprendre le terme médical utilisé par Jean Peytard (1992, p. 148), pour désigner une ouverture ou une rupture anormale. Plusieurs études et comptes rendus d'expériences de formation (Barlow 1992) signalent comment les représentations mettent en évidence les micro-violences symboliques de ce vécu (Cocoradă, Luca, Cliniciu et Pavalache-Ilie 2008) et le questionnement de l'estime de soi (Hame-line 1996).

Dans le cadre d'un enseignement « Évaluation des apprentissages en didactique des langues »³, j'ai pu constituer, depuis 2006, un corpus de remue-méninges distants asynchrones autour de la notion d'évaluation. Répétée chaque année au début de la formation, cette activité clinique, qui conduit ensuite à des classements et à des analyses de réseaux de termes et d'expressions, montre l'existence d'un socle symbolique fondamental dans les constituants de la représentation, résumé par une étudiante :

« Enfin, nous voyons déjà bien ici l'influence qu'a eue l'expérience des participants (qu'on peut imaginer de longue date) en tant qu'élèves. En effet de nombreux termes (Examen, contrôle, test, diplôme, note(s), niveau(x), critères, stress, erreurs, jugements, correction, appréciations, progrès-progresser-progression) peuvent sembler particulièrement révélateurs du fonctionnement de l'évaluation dans le milieu scolaire et universitaire, à savoir un processus par étapes, échelonné par des examens ayant pour but de valider la progression des élèves et étudiants, et vecteur d'un certain sentiment de stress lié à la notion de jugement. » (E.C., 2010, TD1 : Analyse des données à partir du remue-méninges sur la notion d'évaluation)

L'expérience scolaire imprime sa marque autant sur le modèle qui semble immuable (*examen, contrôle, test, diplôme*) que sur les sentiments négatifs à l'égard de l'évaluation (*stress, erreurs, jugements*). Évaluer, c'est, au moins au niveau de l'activité évaluative, simplement estimer comme le définissent les dictionnaires non spécialisés, c'est-à-dire juger par approximation. Ce qui tend à renvoyer l'innovation dans l'acte formel d'évaluation, soit du côté d'une transformation radicale

3. Université Stendhal Grenoble 3, Master Sciences du langage, spécialité FLE, 2006-2011.

de cette activité par des artefacts, soit à en assumer ce qui en fait le fondement et peut-être l'inéluçabilité constitutive, la subjectivité.

Depuis que les activités humaines sont mesurées, la subjectivité a toujours été considérée comme un phénomène à corriger, l'évaluation devant contribuer à produire des informations ordonnées en termes quantitatifs, relativement indexables, ce qui a conduit de nombreux travaux à revendiquer son exclusion par des artefacts technologiques cherchant à comprimer l'appréciation, le jugement. Mais paradoxalement, les technologies développées ne modifient pas les représentations inquiètes. La lutte est vaine et mène à une impasse par rejet de ce qui fait que l'activité évaluative est porteuse de connaissance qualitative si l'on distingue objectivité et objectivation comme construction empirique multiréférentielle et plastique⁴.

L'objectivité réduit le potentiel des observables (avec des performances classées A1 on peut réaliser telles opérations langagières et uniquement celles-là, avec C2 on peut réaliser n'importe quelle opération langagière) ; l'objectivation l'étend (avec un ensemble de compétences et la mobilisation d'un répertoire langagier, souvent multilingue, on peut développer un nombre non connu de stratégies permettant de faire plus ou moins aboutir un projet de communication).

« Lorsqu'un chimiste veut expliquer la composition de l'eau distillée, il soumet cette eau à l'expérience du voltamètre ; si les conditions requises sont toutes remplies, le courant électrique se chargera lui-même de l'analyse et le chimiste obtiendra infailliblement de l'hydrogène et de l'oxygène ; l'eau distillée étant la même dans tous les laboratoires, les chimistes de tous les laboratoires obtiendront tous le même résultat précis. Travaillant sur une matière qui est partout et toujours la même, le chimiste laisse la nature produire ses réactions et parvient heureusement à ses fins. Voilà, dira-t-on, un homme objectif ! » (Trudel 1951, p. 1).

La mobilisation du répertoire langagier, avec des outils variables dans des contextes et des situations variables, est totalement à l'opposé de la manipulation chimique dont les paramètres sont toujours identifiables comme référentiels de la mesure. La formalisation de lois générales en évaluation est rendue difficile par la circonstance que les observables sont souvent influencés par beaucoup de facteurs qu'il est très difficile d'évaluer séparément. Or, par une sorte de manipulation du rapport objectivité vs subjectivité, la docimologie et la psychométrie tendent à donner l'illusion de la juste mesure.

Il est paradoxal que le constructivisme en éducation ait porté tant de fruits dans l'approche de l'apprentissage et que sa conséquence logique n'ait pas conduit à une plus grande radicalité en évaluation : l'image de la réalité et les notions qui structurent cette image sont le produit de l'esprit humain en interaction permanente avec cette réalité. Autrement dit, l'évaluation devrait y être abordée comme un processus d'auto-socio-construction.

En langue, la performance est toujours une prise de parole et d'inférence, dans des codes et des supports variables selon les situations de

4. Il ne s'agit pas ici de l'objectivation comme processus de réification (constituant un objet fini).

communication. La contrainte d'évaluation dans l'apprentissage n'échappe pas au fait qu'il s'agit bien d'un contexte actionnel de communication dans la communication (l'évaluataire répond autant à la problématique posée par l'épreuve, qu'à une question jamais posée mais toujours présente : qu'est-ce que l'évaluateur attend de moi ?), de paroles sur paroles et d'altération qui caractérise l'appropriation d'un discours. Sauf à en revenir à des approches de simples opérations déclencheur/stimulus type + structure a = réponse + structure b, où les attendus se présenteraient en listes fermées, on ne peut que très rarement prédéterminer la forme et le degré de complétude communicationnelle du tour de parole de l'évaluataire. Le sens, dans les discours d'évaluation, relève d'une cohérence complexe entre le sens de l'acte d'évaluation (ses objectifs) et le sens du discours simulé, entre :

- des énoncés directeurs (procédés déclencheurs) destinés à provoquer des énoncés réponse ;
- des énoncés réponse dont le défi est de comprendre l'attendu évaluatif, autant que de participer au scénario.

D'une certaine façon, on peut dire que les situations de communication en évaluation sont spécialisées et se révèlent bien plus complexes que nombre de situations de communication dans le réel ordinaire où le sujet aurait un nombre de recours supérieur à ceux autorisés en évaluation (éléments de communication non verbale ; répertoire multilingue ; ajustements multiples avec interrogations et reformulations).

L a mise en scène du répertoire langagier : la tâche et l'agir

Sollicités par l'effet d'une dynamique de la compétition sociale, ces artefacts sont-ils vraiment garants de ce qu'ils avancent ? Les discours de promotion des tests vantent leur qualité en termes de *fiabilité* :

« Le TCF est un instrument de mesure fiable. Le candidat est assuré d'être positionné au niveau qui correspond réellement à ses compétences, quelle que soit la session à laquelle il se présente et donc la version du test qui lui est proposée. Cette garantie est obtenue grâce aux choix techniques et aux étapes de l'élaboration du TCF, test standardisé et calibré », CIEP (<http://www.ciep.fr/tcf/fiable.php>, 10/12/2011). Le Test d'Évaluation de Français, test délivré par la CCIP [...], atteste avec fiabilité du niveau en langue française d'un candidat », CCIP (<http://www.cours-de-francais.fr/france/cours/TEF.php>, 10/12/2011).

C'est une manière de remotiver en valeur de marché le sens de cette notion prélevée dans les travaux de psychométrie et d'éduométrie sur lesquels ils se fondent. Les tests calibrés fonctionnent dans le strict cadre des conditions imposées pour leur passation, en visant des

connaissances et non des compétences jugées impossibles à mesurer. On s'en tient donc à un référentiel fermé d'attendus autour du code et d'actes de langage répertoriés comme acceptables. Mais peut-on mesurer le langage autrement que par rapport à des échelles de normes qui renvoient à des contextes de connaissance particuliers (diagnostic thérapeutique, sociographie de populations) et rarement universalisables⁵ ? Or, ce que recherche l'évaluateur des apprentissages en langue, c'est bien à pouvoir déterminer le potentiel d'un sujet à manier son répertoire langagier dans un ensemble de situations où plus l'évaluataire sait dire, moins il s'avère facile de prédire l'ensemble possible des façons de dire (les performances classées A1 correspondant à un ensemble restreint, celles classées C2 à un ensemble non fini).

Comme le rappellent Françoise Demaizière et Jean-Paul Narcy-Combes (2005, p. 46), les points de vue théoriques qui coexistent sur le fonctionnement de la production langagière s'imbriquent certainement de manière complexe. Pour certains, elle suit des règles, pour d'autres, la parole est produite en mémorisant des modèles en relation avec leurs contextes d'emploi, l'origine et la finalité de cette fonction pouvant, *in fine*, revenir au politique : être mobile et organiser la vie dans des groupes, des entités de plus en plus élargies serait l'un des fondements de l'apparition du langage et du développement des langues. Homo mobilis et homo politicus, avant d'être homo loquens, le petit de l'homme apprend le mouvement et l'action avec le langage. Raison pour laquelle la notion de tâche semble si évidente aux didacticiens, même si son origine et sa conceptualité restent floues (Chiss 2005 : 39).

La meilleure mise en scène de l'évaluation de compétences en langue étant celle de l'épreuve de la vie réelle et non de l'épreuve formelle de l'acte d'évaluation, nombre de propositions de procédés se présentent, depuis les années 1990, sous la forme de scénarios (Lussier 1992 ; Huver 2007) et de tâches depuis les années 2000 avec l'irruption de la perspective actionnelle⁶. Évaluation de tâches (cognitives, langagières), évaluation de comportements d'apprentissage face à la tâche (éducation) et tâches d'évaluation (psychologie clinique) se mêlent.

« Il y a "tâche" dans la mesure où l'action est le fait d'un (ou de plusieurs) sujet(s) qui y mobilise(nt) stratégiquement les compétences dont il(s) dispose(nt) en vue de parvenir à un résultat déterminé. » (CECR, 2001 : 15).

Mais ces tâches exécutables à partir d'une consigne, dont le sens ne consiste qu'à jouer un rôle en acceptant le jeu, peuvent-elles permettre de sortir de la mesure de surface et d'entrer dans le grain ? La tâche n'est que réaliste, ancrée dans le monde réel de l'évaluation qui n'est qu'une toute petite partie du monde dans lequel l'évaluataire, pourvu d'un certificat qui le classe, aura à mobiliser ses connaissances et activer ses compétences. Son algorithme standardisé est finalisé dans le temps et l'espace, avec un début, une structure, un contenu et un résultat qui ne donne à voir que ce qu'il peut donner. Même fiable sur le plan

5. Voir, à cet égard, les détournements du QI.

6. Les scénarios de tâches pour l'évaluation langues existent en France depuis 1995 (DCL- Diplôme de compétence en langue et 2000 (CLES- Certificat de Compétences en Langues de l'Enseignement Supérieur).

psychométrique, aucune certification en langue aujourd'hui ne peut garantir le potentiel qualitatif de l'évaluataire au-delà des composantes de la tâche. Or, davantage encore que l'évaluation des connaissances/compétences dans les autres domaines, l'évaluation en langue repose sur un pari de dépassement. Si on attend de l'élève évalué compétent dans l'appropriation des quatre opérations arithmétiques qu'il sache les utiliser à la fois pour d'autres apprentissages opératoires plus complexes et pour des applications, le potentiel d'exploitation de ces opérations, si large soit-il, reste inférieur à celui du langage. Le décalage est encore plus évident avec des savoirs et des savoir-faire plus restreints, qui restent relativement mesurables. L'agir avec le langage, pour être évalué, suppose de sortir de la croyance en la mise en scène comme source intègre qui autorise à sélectionner en fonction du jeu.

Cela peut paraître bien complexe au plan de l'ingénierie de formation, où l'on trouvera peut-être plus simple, moins dangereux, moins coûteux de calibrer des épreuves identiques pour un marché élargi, en gommant les contextualisations. C'est ainsi que le DELF, diplôme d'étude en langue française, et le DALF, diplôme approfondi de langue française, créés en 1985, ont connu, à partir des années 1990, un succès important (360 000 unités vendues chaque année, 900 centres d'examens dans 150 pays), ce qui a amené ses responsables à prendre une décision gestionnaire en supprimant, dans les années 2000, les commissions nationales de concepteurs, au profit d'une conception centralisée au Centre international d'études pédagogiques de Sèvres. Cela peut sembler également complexe d'un point de vue politique. L'Europe, avec des effets bien au-delà de ses frontières, s'est engagée dans un processus de compatibilisation de ses certifications qui, avec *Un Niveau Seuil* (1975) puis le *Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer* (2001), a permis une production de connaissances partagées jamais vue, dans un but de démocratisation de l'apprentissage des langues afin d'encourager la mobilité des personnes et des idées et de répondre aux défis de la cohésion sociale et de l'intégration. Mais dans la mesure où elle penche aujourd'hui vers une dynamique de réduction des coûts comme principe, la complexité est plutôt niée. Autant dire qu'il est difficile d'imaginer promouvoir autre chose qu'une évaluation selon des standards et des calibres pré-établis.

Un système global d'évaluation multiréférentielle

Dans la troisième partie de leur ouvrage, Emmanuelle Huver et Claude Springer (2011, p. 206-319) proposent d'articuler un ensemble de

paradigmes composant les procédures et les procédés d'évaluation, en définissant des postures qualitatives, alternatives à celles qui s'imposent aujourd'hui comme dominantes avec le marché des certifications. Ce qui est en jeu est de réussir à articuler la variété des moyens langagiers que l'on classe sous la notion de *compétence* (qui, fondés sur le triple savoir-faire, -être, -connaître, s'organisent entre ressources générales et ressources fines), avec la variété des procédures acceptables par les systèmes, les établissements, les acteurs évaluateurs et évaluataires (qui se distribuent entre procédures certificatives et procédures formatives). On ne reviendra pas sur un marché des certifications établi qui, reposant sur de nombreux et puissants intérêts matériels et symboliques, joue un rôle dans la reconnaissance et la valorisation de l'apprentissage des langues. Il s'agit de construire son équivalent matériel et symbolique au plan qualitatif, rendre sinon solidaires, au moins cohérents, la mesure et le grain associés dans un système global.

Pour faire sortir les démarches qualitatives d'une position alternative vers une position associée, nous avons besoin de développer la production de savoir dans ce domaine, de concentrer des moyens de recherche aujourd'hui faibles, dispersés et trop souvent bornés par un cadrage *a priori* sur des standards (combien d'équipes, de laboratoires travaillent cette problématique ? combien d'étudiants-chercheurs y consacrent leur thèse ?). Dans la préface de l'ouvrage de Murielle Molinié (2011), Blanchet (p. 13) précise « qu'un portfolio n'est pas un portefeuille, que les ressources linguistiques et culturelles, plurilingues et interculturelles, ne sont ni uniquement ni forcément ni prioritairement des valeurs économiques (ce qui n'empêche pas qu'elles le soient aussi parfois) ». Employant ici *ressources pour compétences*, il affirme les pertinences éthique et politique du portfolio, auxquelles j'ajouterais la pertinence scientifique par le fait qu'il peut formaliser l'espace dialogal entre l'évaluateur et l'évaluataire sur la base d'une reconnaissance de l'activité évaluative comme compétence. Évidemment, il ne s'agit pas d'un instrument transparent (l'évaluataire y indique ce qu'il a envie de dire), mais d'un support formel d'expression de la subjectivité qui peut fournir une base de dialogue évaluatif. On ne croira pas davantage à son objectivité, qu'à celle d'une épreuve d'examen qui n'est qu'une variante de conception behavioriste, stimulus (consigne)/réponse devoir. La question ne se situe plus à ce niveau mais à celui de la construction d'une conscience évaluative partagée telle qu'elle se développe entre deux interlocuteurs dans une communication exolingue où chacun évalue sa performance à travers l'autre par une série d'ajustements progressifs du sens. Une approche constructiviste de l'évaluation qui met l'accent sur l'activité évaluative du sujet.

La démarche scientifique en évaluation des apprentissages en langue ne se trouve pas seulement du côté de la rigueur annoncée des approches psychométriques et des procédures de tests calibrés qui

contournent la compétence, mais également et surtout, du côté de celles qui en affrontent l'objectivation, c'est-à-dire la production d'informations contextualisées et variables dans le processus d'évaluation. Que les tâches soient construites (dans le cas d'épreuves préparées), ou déconstruites (dans le cas de l'analyse de moments de biographie interlinguistique et interculturelle), la compréhension de la performance réalisée passe par un agir communicationnel. L'activité évaluative qui se manifeste aussi en langue étrangère comme compétence en mobilisant les ressources des opérations de caractérisation, de catégorisation et de jugement devrait pouvoir se constituer en objectif d'apprentissage et prendre pour objet la biographie interlinguistique et interculturelle. En traduisant le vécu global d'appropriation (acquisition informelle et apprentissage formel), cette démarche confère un statut à l'expérience interlinguistique et interculturelle en tant que théorie : le triple savoir des acteurs évaluateur/évaluataire, comme théorie de l'acte d'évaluation. Dans le sillage de l'ethnométhodologie, on utilise en ethnologie, des grilles ethno-critériées d'évaluation participative (Absil 2007), qui assument l'intersubjectivité (que je range ici sous la notion d'objectivation), dans le cadre des critères de qualité de l'agir communicationnel définis par Habermas (1992), qui assurent la co-construction de l'évaluation (enjeux, objectifs, critères, indicateurs). Adaptée à la didactique des langues et des cultures, cette approche lie les évaluateurs et l'évaluataire dans un rapport de formalisation de répertoires de connaissances/compétences en relation à des référentiels évolutifs (Vial 2001) en langues et entre les langues : conscience interlinguistique et interculturelle (sources de l'expérience) ; acceptabilités (conformité aux normes variables) ; intégrité (intentionnalité/actionnalité) ; légitimité (compromis entre points de vue variés). On ne peut bien entendu minorer le poids symbolique des représentations de l'évaluation, qui par les cultures d'enseignement et d'apprentissage fossilisées, peuvent freiner l'extension du domaine dynamique du portfolio. Mais dans les mouvements de ce champ, cet enjeu relève aussi d'un élargissement probable du nombre de candidats à l'apprentissage des langues.

Bibliographie

- ABSIL G. (2007), *Le savoir des acteurs comme théorie de l'action*, <http://www.apes.be/documentstelechargeables/powerpoint/07040%20Ethno-évaluation%20Absil.pdf>
- BARLOW M. (1992), *L'évaluation scolaire. Décoder son langage*, Chronique Sociale, Lyon.
- BLANCHET P., « Un portfolio n'est pas un portefeuille... Réflexions sur l'humanisation des "compétences" plurilingues et interculturelles », dans Molinié M., (dir.), *Démarches portfolio en didactique des langues et des cultures. Enjeux de formation par la recherche action*, Encrepage/ Université de Cergy-Pontoise.
- BLANCHET P. et CHARDENET P. (2011), « Guide pour la recherche en didactique des langues et des cultures. Approches contextualisées, Éditions des archives contemporaines » ; *Ressources pour la recherche en didactique des langues : glossaire des notions et concepts*, Agence universitaire de la Francophonie [En ligne]. <http://wiki.auf.org/glossairedlc/Index/>.
- CARDINET J. (1983), *Évaluer les conditions d'apprentissage des élèves plutôt que leurs résultats*, Institut romand de recherches et de documentation pédagogiques, Neuchâtel.
- CHARDENET P. (1993), « De l'acte d'enseigner à l'acte d'évaluer, analyser des discours », *Le français dans le monde/Recherches et applications, Évaluation et certifications en langue étrangère*, Monnerie-Goarain, A. et Lescure, R. (coord.), 08-09/1993, Hachette, p. 51-56.
- CHARDENET P. (1999a), « Les notions de [langue apprise], [langue enseignée] et leur détermination », *ENSAIO, revista do mestrado em Letras, Universidade Federal de Santa Maria, Brésil*, p. 51-60.
- CHARDENET P. (1999b), *De l'activité évaluative à l'acte d'évaluation*, L'Harmattan.
- CHARDENET P. (1999c), « L'enjeu de la distinction », dans *Le français dans le monde/Recherches et applications, Apprendre les langues étrangères autrement*, Caré, J.-M. (coord.), Hachette, p. 27-32.
- CHARDENET P. (2009), « Pour un interlinguisme méthodologique », conférence au CUSO, École doctorale en sciences du langage – Bloc de formation 3, Université de Genève, 26-27 mars, (<http://patrickchardenet.unblog.fr/files/2011/08/pouruninterlinguismemethodologique.doc>, 15/12/2011).
- CHARDENET P. (2010a), « Peut-on encore questionner l'évaluation », conférence inaugurale, Troisième journée d'études du comité scientifique du TEF, CCIP, Université Paris IV (http://www.fda.cciip.fr/documents/pdf/tef/je_2010/chardenet.pdf, 15/12/2011).
- CHARDENET P. (2010b), « Observer les espaces d'interlocution plurilingues et les pratiques langagières dans des langues associées », dans Blanchet P., Martinez P., *Pratiques innovantes du plurilinguisme* (dir.), Éditions des Archives Contemporaines Éditions Scientifiques GB, Paris, p. 121-130.
- CHARDENET P. (2011), « De l'analyse du discours à l'analyse des discours en situation comme outil de recherche et d'intervention : l'exemple de recherches sur l'évaluation », dans *Guide pour la recherche en didactique des langues et des cultures. Approches contextualisées*, sous la direction de Blanchet P. et Chardenet P., Éditions des Archives Contemporaines Éditions Scientifiques GB, Paris.
- CHISS J.-L. (2005), « Réflexions à partir de la notion de tâche : langage, action et didactique des langues », dans Marie-Anne Mochet, Marie-José Barbot, Véronique Castellotti, Jean-Louis Chiss, Christine Develotte, Danièle Moore, *Plurilinguisme et apprentissages : Mélanges Daniel Coste*, ENS, p. 39-48.

- COCORADĂ E., LUCA M., CLINCIU A. et PAVALACHE-ILIE M. (2008), « Le clivage des représentations sur l'évaluation et ses conséquences », *Actes du 20^e colloque de l'ADMEE-Europe*, Université de Genève.
<https://plone.unige.ch/sites/admee08/communications-individuelles/v-a3/v-a3-1>, (15/12/2011).
- DEMAIZIÈRE F. et NARCY-COMBES J.-P. (2005), « Méthodologie de la recherche didactique : nativisation, tâches et TIC », dans ALSIC, Volume 8, p. 45-64, <http://alsic.revues.org/index326.html>, (15/12/2011).
- HABERMAS J. (1992), *De l'éthique de la discussion*, Flammarion, Paris.
- HAMELINE D. (1996), « De L'estime », dans Jean-Michel Zakhartchouk (coord.), *La Motivation*, Cahiers pédagogiques Hors série, Paris, CRAP.
- HÜVER E. (2007), *Évaluer autrement : le scénario d'évaluation*, (http://cravie.ac-strasbourg.fr/Scenario_evaluation/scenario_evaluation.pdf, Centre de Ressources Alsace Ville (05/12/2011).
- HÜVER E. et SPRINGER C. (2011), *L'évaluation en langues*, coll. « Langues et didactique », Didier, Paris.
- LUSSIER D. (1992), *Évaluer les apprentissages dans une approche communicative*. Collection autoformation. Hachette, Paris, 126 p.
- MARCELLESI J.-B. et GARDIN B. (1987), *Introduction à la sociolinguistique*, Cahiers de Linguistique sociale, I.R.E.D., Mont-Saint-Aignan
- MOLINIÉ M., (2011) dir., *Démarches portfolio en didactique des langues et des cultures. Enjeux de formation par la recherche action*, Encrage/Université de Cergy-Pontoise.
- PEYTARD J. et MOIRAND S. (1992), *Discours et enseignement du français. Les lieux d'une rencontre*, Hachette.
- PEYTARD J. (1999), « Écriture et pointillés de sens : lecture-analyse de deux pages de Proust (La Fin de la jalousie) », dans Semen [En ligne], 11 I, mis en ligne le 31 mai 2007, consulté le 10 novembre 2011.
<http://semen.revues.org/2911>.
- PIÉRON H., PIÉRON H. (Mme) et LAUGIER H. (1929), « Étude critique de la valeur sélective du certificat d'études et comparaison de cet examen avec une épreuve par tests. Contribution à une docimastique rationnelle », compte rendu de la IV^e Conférence internationale de psychotechnique, Paris, Institut international de coopération intellectuelle, 10 au 10 octobre 1927, Paris, Félix Alcan, 1929, p. 499.
- RASTIER F. (2011), *La mesure et le grain*, Honoré Champion.
- TRUDEL M. (1951), « L'objectivité en histoire », *Revue d'histoire de l'Amérique française*, vol. 5, n° 3, 1951, p. 315-31
<http://id.erudit.org/iderudit/801713ar> (15/12/2011)
- VIAL M. (2001), *Se former pour évaluer. Se donner une problématique et élaborer des concepts*, De Boeck.

Évolutions
*des usages
d'évaluation :
faire avec
les tensions*

PATRICK RIBA

CLAIRE TARDIEU

MATHILDE ANQUETIL

Les descripteurs de compétence en langue : rappels et mise en perspective

PATRICK RIBA

UNIVERSITÉ DES ANTILLES ET DE LA GUYANE, FRANCE

North B. et Schneider G. (North, Schneider 1998 : 217-262) situent l'engouement pour les échelles de descripteurs de compétences dans les décennies 1980 et 1990. Alors que pendant de nombreuses années c'est l'échelle dite de Wilds (1975), utilisée par le *Foreign Service Institute, FSI*, des États-Unis d'Amérique, qui semble la plus modélisante pour les pays qui ont recours à ce type d'outil, apparaissent durant les deux décennies suivantes une échelle du *British National Language Standards* en 1992, la grille Eurocentres, *Scale of Language Proficiency*, en 1983 revue en 1992, celle du *Finnish Scale of Language Proficiency* en 1993, ou encore le *ALTE Framework* de l'Association des certificateurs européens en langue (ALTE ; cf. sitographie), en 1994, pour n'en citer que quelques-unes.

L'échelle du FSI se centrait sur l'examinateur et fonctionnait sur la base de listes de contrôle de la compétence attendue. Elle assurait une évaluation relativement fiable, et permettait de compléter des tests souvent basés sur une appréciation globale ou sur la maîtrise des formes linguistiques et des compréhensions orale et écrite, en se consacrant à l'observation du fait communicatif. En ce sens, elle répondait déjà à une exigence sociétale en assurant une validité de construct, entendue comme « une validité interne qui indique le degré d'adéquation entre la structure d'un instrument de mesure et les modèles théoriques, l'habileté ou les composantes sous-jacentes qu'il prétend mesurer » (Lussier, Turner 1995 : 180), plus en phase avec l'évolution d'une société américaine soumise à davantage de règles de transparence, soucieuse de répondre à des besoins d'échanges croissants avec d'autres pays (Bachman, Palmer 1981). C'est ainsi par exemple que les concepteurs du *Test of English as a Foreign Language, TOEFL*, s'intéressant à l'évaluation de la production orale dès

les années 1970, mirent en place une épreuve d'entretien en face à face sur la base du *FSI Criterion Interview* (Clark, Swinton 1979).

La centration de ces échelles évolua peu à peu, délaissant la perspective de l'évaluateur pour se focaliser dans les années 1990 sur les capacités de faire des apprenants (Alderson *et al*, 1991), tout d'abord sur un plan strictement fonctionnel avec reprise des actes de langage (ex. : *Can ask and answer questions on topics very familiar to him* - Clark, Swinton 1979 : 5), puis dans des dimensions plus précises (ex. : *est capable de poser des questions factuelles et de comprendre les réponses exprimées dans une langue simple* - spécifications ALTE, in Conseil de l'Europe, COE, 2001 : 179).

À l'initiative du Conseil de l'Europe, un projet fut lancé en Suisse en 1991 pour concevoir une nouvelle échelle de descripteurs basée sur un nouveau modèle didactique de l'objet langue qui donnera naissance en 2001 au *Cadre européen commun de référence pour les langues*, désormais le Cadre. Ce projet prévoyait la rédaction de descripteurs de compétence traduisibles et intelligibles dans toutes les langues, sur 6 niveaux, et dans un ensemble d'activités langagières, compréhension, production, interaction et médiation, orales et écrites, ainsi que dans leurs dimensions sous-jacentes, linguistiques (étendue lexicale, orthographe, etc.), stratégiques (planification) ou pragmatiques (parler avec un locuteur natif, ou obtenir des biens et des services, par exemple).

La traduction du Cadre dans de nombreuses langues, puis le projet DIALANG (cf. sitographie), et enfin l'élaboration et l'utilisation en classe de portfolios européens de compétences permirent la mise en pratique de ces travaux que ne tardèrent pas à reprendre les certificateurs, les méthodologues et auteurs de manuels scolaires, ainsi qu'un nombre toujours croissant d'enseignants. Les échelles de descripteurs sont aujourd'hui assez communément utilisées à des fins multiples, d'auto-évaluation en ligne comme avec DIALANG (qui reste la première application immédiate du Cadre dans le domaine), d'auto-évaluation ou de co-évaluation grâce aux portfolios de compétences (Schärer 2007), et dans tous les cas dans une perspective d'évaluation guidée et critériée, sommative bien sûr, mais aussi formative ou prédictive. Elles ont aussi dépassé le champ initial de l'évaluation, car elles sont utilisées pour l'élaboration de séquences d'enseignement apprentissage.

Dix ans après la sortie du Cadre, il peut paraître légitime de s'interroger sur ces descripteurs et leur répartition dans des échelles, sur leurs apports, leurs atouts mais aussi leurs limites. Ils participaient d'un projet plus vaste, décrit en de nombreux chapitres d'un ouvrage qui aspirait à doter l'ensemble des pays, les institutions et les individus qui souhaitaient y adhérer, d'un outil commun de référence. Le Cadre se proposait de donner un nouveau sens social à l'enseignement apprentissage des langues, la perspective actionnelle (Cadre 2001 : 15), sur la base de modèles didactiques plus ou moins convergents, initialement inspirés des travaux de Bachman (1990). Pourtant, certains (dont l'auteur de ces

lignes ou encore Beacco J.-C., FLE.fr ; cf. sitographie) considèrent que seules ces échelles semblent avoir retenu l'attention, et affirment que la perspective des évaluateurs et surtout celle des certificateurs en ont imposé l'usage, comme si l'enseignement et l'apprentissage passaient au deuxième rang par rapport à l'évaluation.

Force est de constater qu'elles ont séduit, rassuré et structuré des dispositifs éducatifs. Le baccalauréat français, dans sa toute-puissance régalienne, s'en est officiellement inspiré, et même au-delà de l'Europe, des pays en ont fait une référence explicite, comme par exemple le Mexique dans le cadre de sa «Certificación nacional de nivel de idioma» (CENNI ; cf. sitographie).

Ce n'est pas le propos de cet article de se pencher sur leur impact dans l'enseignement et l'apprentissage ; mais, dans une réflexion sur l'évaluation en langue au début du XXI^e siècle, nous pouvons nous demander si elles ont vraiment permis de mettre en place ce système de transparence que tout le monde semblait appeler de ses vœux.

L *'élaboration des descripteurs de compétence : rappels et précisions*

Le Cadre rappelle les principes d'élaboration de descripteurs de compétence dans son Annexe A (Cadre 2001 : 147-153). Ceux-ci reposent sur une méthodologie qui combine une phase dite intuitive durant laquelle des experts (enseignants, méthodologues, chercheurs) élaborent une première série de descripteurs en relation avec le contexte éducatif dans lequel ils seront utilisés ; puis, un panel plus vaste de praticiens les valide dans une approche qualitative de plus en plus rigoureuse, les premières étapes permettant d'écarter des descripteurs par trop éloignés des finalités attendues, les dernières se penchant par exemple sur le respect de consignes de rédaction strictes. Pour être retenus et pour que le système soit cohérent dans sa globalité comme dans sa spécificité, ces descripteurs doivent être indépendants les uns des autres, précis, clairs et brefs. Ils doivent également être utilisables dans toute langue possible, et donc éviter toute référence à l'une d'entre elles en particulier. Les auteurs du Cadre ajoutent enfin un dernier critère qui est, lui, un parti-pris pédagogique, ils doivent être formulés selon un principe d'affirmation positive, ce qui leur permet par exemple d'écarter le descripteur «possède un répertoire de discours étroit qui l'oblige à reformuler et à chercher ses mots constamment» au profit du descripteur «possède un vocabulaire suffisant pour satisfaire les besoins communicatifs élémentaires».

Enfin, et en ce sens, le Cadre s'est doté d'une méthode novatrice ou tout au moins peu connue du grand public, ils sont soumis à un étalonnage quantitatif basé sur des modèles psychométriques. Les concepteurs du Cadre se sont aperçus que la plupart de leurs descripteurs avaient recours à des mots clés, des verbes (se faire comprendre, se débrouiller, participer à), des adjectifs (restreint, étendu, complexe) ou des adverbes (simplement, suffisamment, clairement), etc. La méthode quantitative leur a permis d'une part d'identifier avec précision ces termes (analyse discriminante), puis de tester leur valeur isotopique dans une série de tests, c'est-à-dire la récurrence selon laquelle les utilisateurs les classaient à un niveau donné (étalonnage multidimensionnel). Ils ont ainsi pu établir une valeur prédictive pour chaque mot-clé, calculée selon le principe de l'analyse théorique de la réponse aux items. C'est cette valeur d'ancrage des termes clés qui a permis d'établir la cohérence interne des échelles les unes par rapport aux autres. En complément de l'annexe citée *supra*, le Conseil de l'Europe a publié deux ouvrages sur ce thème, *Relier les examens de langues au Cadre européen commun de référence pour les langues : Apprendre, enseigner, évaluer (CECRL), un manuel* (traduction française Breton, Georges, Tagliante 2009), qui offre dans une lecture aisée une méthodologie qualitative de conception, puis un supplément beaucoup plus technique et principalement consacré à l'approche psychométrique du sujet (Takala et al., 2009). Cet étalonnage permet enfin la mise en banque des descripteurs et leur gestion informatisée (Lenz, Schneider 2004). Ce haut niveau de technicité a conféré au dispositif final une cohérence et une intelligence pressenties et appréciées par tous, il a contribué au succès de ces échelles, mais il en a également fixé les limites.

Les trois valeurs de l'étalonnage

Tout autant que l'émergence elle-même des descripteurs, la méthode retenue pour leur validation est représentative à plus d'un titre de certains courants épistémologiques de notre époque, qui explique sans doute en partie d'ailleurs leur succès et celui du *Cadre*.

Nous en repérons au moins trois. Le premier est d'ordre scientifique. En convoquant les mathématiques dans l'univers de la didactique des langues et des cultures, celle-ci semble s'être dotée, à l'instar de sa grande sœur la linguistique, et pour reprendre les mots de Serres M., de « cet outil qui dans sa simplicité abstraite [...] prend comme en écharpe l'espace complet du savoir » (Serres 1982 : 277). Mais les mathématiques invitées ici ne sont pas ces sciences qualitatives des formes et du rapport possible des choses qui ont tant aidé Chomsky N.,

ce sont celles de la mesure et de la quantité qui séduisent l'école et le docimologue. En sériant les descripteurs selon un indice de difficulté déterminé par le modèle de Rasch (Le Cadre, Annexe A, 2001 : 152) les concepteurs de descripteurs se sont dotés d'un argumentaire qui faisait défaut aux méthodes traditionnelles pour identifier la variation significative de l'interprétation des descripteurs par rapport à différentes variables (régions linguistiques, secteurs éducatifs, etc.). Il en résulte un indice de stabilité élevé qui explique en partie le taux d'adhésion du grand public. Cette méthode possède également une force de prédictibilité importante. C'est elle qui assure l'unité conceptuelle d'un tout composé d'éléments indépendants les uns des autres : chaque descripteur peut être lu, entendu et utilisé de manière indépendante, il n'en reste pas moins intrinsèquement dépendant de celui qui le précède et de celui qui le suit.

Le deuxième facteur est d'ordre sémantique, voire symbolique. En offrant aux utilisateurs une approche par compétences formulée en termes observables, les concepteurs de ces échelles ont sans doute scellé l'importance d'une pédagogie centrée sur l'apprenant, souvent annoncée et souvent malmenée, dont les portfolios sont sans doute la meilleure expression. Le Cadre et ses travaux adjacents ont opérationnalisé de manière relativement simple la graduation de parcours d'apprentissages basés sur les acquis des apprenants, dont la congruence semblait assouvir une volonté universelle de cohérence méthodologique aussi bien au plan international, donc dans plusieurs langues, qu'au simple niveau d'une équipe d'enseignants, donc dans une langue donnée.

C'est sans doute pour cela que les chapitres consacrés à ces échelles ont rencontré un tel succès aux dépens d'autres tout aussi fondamentaux en termes d'innovation et de réflexion pédagogique, comme le sont par exemple celui sur les stratégies d'apprentissage ou celui consacré à la compétence interculturelle ; et l'on a tort d'en rendre responsables les seuls certificateurs et les évaluateurs en oubliant que le succès remporté par le Cadre n'est pas strictement lié à celui rencontré par les diplômés en tous genres que l'on rencontre au sein d'ALTE et ailleurs.

La dernière valeur de ces étalonnages est d'ordre systémique. Cette approche semblait en effet fournir un regard objectivant sur la notion d'aboutissement des apprentissages, dépassant le cadre biaisé de l'évaluation normative interne (qui compare les étudiants les uns par rapport aux autres) grâce au recours à des illustrations et des ancrages qui semblent en garantir l'impartialité. En créant des kits de formation destinés à illustrer les attendus didactiques en production écrite (CIEP 2005), en mettant à la disposition des usagers un cd-rom illustrant les 6 niveaux du Cadre en production orale dans 5 langues différentes (CIEP, cf. sitographie), ou en créant des référentiels qui spécifient les contenus de chaque niveau dans chaque langue (pour leurs versions

françaises, Beacco et al., 2004, 2007, 2008, 2010), les équipes qui ont travaillé avec la division des Politiques linguistiques du Conseil de l'Europe ont mis à disposition des enseignants des outils méta-didactiques dont ils ne disposaient pas jusqu'ici, dans une approche qui dépassait le simple fait évaluatif.

Les limites du projet

Ces trois facteurs expliquent sans doute le succès des descripteurs de compétences, mais ils n'en assurent pas pour autant la validité, et plusieurs biais en limitent la portée. Le premier est à mettre au corps défendant de ses auteurs ; le succès du Cadre est tel que ces échelles aux vertus modélisantes revêtent désormais un caractère sacralisant. Plus de méthode, plus de certification, (presque) plus d'article qui n'y fasse directement allusion ou référence, au moins en Europe ; si l'on y prend garde, cet ouvrage destiné à la réflexion pourrait bien devenir le grand livre rouge de la didactique des langues, ce qui conduirait inéluctablement à sa proche condamnation. Les auteurs du Cadre et leurs promoteurs au sein du Conseil de l'Europe ont peut-être sous-estimé la force normalisante d'une institution multinationale représentative. Un projet politique fort articule cet ouvrage, délibérément présenté en introduction de l'ouvrage (défense du plurilinguisme, ouverture à la mobilité, etc.). Dans un contexte où existe une forte attente de régulation, et au moment où les systèmes éducatifs sont soumis à des comparaisons de performance qui intéressent beaucoup les élèves, les parents, les contribuables et les citoyens, un ouvrage qui propose des niveaux décrits par plusieurs dizaines de descripteurs ne pouvait échapper à cette tendance.

Mais d'autres limites doivent nous alerter au-delà de toute rançon du succès.

Cette recherche du consensus est réductiviste pour au moins deux raisons. L'étalonnage quantitatif conduit au rejet de toutes les formulations qui ne satisfaisaient pas aux normes de la plus grande univocité. En recherchant les plus petits dénominateurs communs, les concepteurs de descripteurs ont donc retenu les formulations les plus simples, laissant de côté des champs entiers de la didactique que l'on ne peut pourtant pas oublier, comme par exemple la dimension interculturelle qui se plie mal à la formulation de descripteurs observables de compétence. Par ailleurs, ils ont produit des outils généralistes qui font oublier un peu vite que la contextualisation est l'une des bases de l'enseignement.

Enfin, en imposant une méthode aussi rigide, ils ont ôté au Cadre toute capacité d'adaptation. Les descripteurs ne prenant leur sens que dans leur

degré d'acceptabilité, personne ne semble en mesure d'en proposer de nouveaux dix ans après leur publication, sauf à pouvoir mettre en œuvre un nouveau processus de conception dont on conviendra qu'il est lourd et coûteux. Pourtant, des besoins se font jour pour corriger par exemple ce que l'on peut appeler aujourd'hui des erreurs de formulation, comme par exemple celle retenue pour la maîtrise de l'orthographe au niveau C2, «les écrits sont sans faute d'orthographe» (Cadre : 93), qui rend invalide le descripteur, sauf à le limiter à de trop rares locuteurs pour qu'il conserve une pertinence sociale discriminante.

Au-delà de ces limites fonctionnelles, et après près de dix années d'utilisation, les descripteurs ont-ils réellement apporté à leurs utilisateurs la cohérence didactique attendue ?

Jamais peut-être un argumentaire de validité de *construct* théorique n'avait été élaboré dans une dynamique aussi consensuelle, et l'on perçoit cependant toujours presque autant de différences dans les approches qu'avant la sortie du Cadre. Sur le plan international, le projet de description des niveaux de référence du Conseil de l'Europe et les référentiels qui en sont issus ont mis en évidence de substantiels écarts de conception du niveau B2 entre équipes britannique, *English Profile* (cf. sitographie), allemande, *Profile Deutsch* (Glaboniat et al., 2005) et française, *Un référentiel pour le français, niveau B2* (Beacco et al., 2004-2010) cette dernière proposant des attendus de maîtrise des formes largement supérieurs aux attendus décrits par les deux autres. L'étude que nous avons pu mener sur l'évaluation et la certification du niveau C2 en français (Riba 2010) confirme que les certificateurs sont dans la même tendance, le DALF C2 français, ou son équivalent dans le Test de connaissance du français, TCF, décrivant des attendus sans doute plus complexes que ceux du *Proficiency* anglais.

Même si peu de travaux permettent encore de tirer une définition à caractère scientifique sur ce sujet, il semble que cette remarque reste valable si l'on observe le fonctionnement d'équipes d'enseignants au sein d'une même institution ou d'une même équipe. En termes d'évaluation, les travaux menés par le Centre international d'études pédagogiques, CIEP pour former les équipes de correcteurs examinateurs du Diplôme d'études en langue française, DELF, et du Diplôme approfondi de langue française, DALF (CIEP 2005), puis les études de corrélation inter jurys menées par exemple aux niveaux A1 et B2 (CIEP 2007) ont mis en évidence des disparités encore assez importantes, même si une forte tendance à la réduction des écarts se notait après les stages de formation. En termes didactiques, les choix programmatiques faits par les méthodologues ne font toujours pas l'unanimité au sein de la profession, et l'on sait que les enseignants continuent d'accorder une importance propre à chaque dimension du langage, les uns privilégiant la dimension pragmatique, les autres la maîtrise des formes linguistiques par exemple.

Alors en quoi des descripteurs énoncés en termes aussi observables, aussi stables et aussi univoques que ceux qui suivent, peuvent-ils donner lieu à autant d'écarts d'interprétation ?

*A2 INTERACTION ORALE, OBTENIR DES BIENS ET DES SERVICES (:66)
Peut se débrouiller dans les situations de la vie quotidienne telles que les déplacements, logement, repas et achats.*

*INTERACTUAR PARA OBTENER BIENES Y SERVICIOS (:84)
Se desenvuelve en aspectos comunes de la vida cotidiana como son los viajes, el alojamiento, las comidas y las compras.*

*TRANSACTIONS TO OBTAIN GOODS AND SERVICES (:80)
Can deal with common aspects of everyday living such as travel, lodgings, eating and shopping.*

Une réponse élémentaire s'impose. En grande partie écrit en anglais et traduit dans les autres langues, le Cadre présente des difficultés de traduction. Tout locuteur averti peut mesurer les différences qui existent entre «Peut se débrouiller», «Se desenvuelve» ou «Can deal with», mais Silva (2012) a également mis en évidence qu'au-delà de ces apparents problèmes de terminologie se dissimulaient en fait des curricula cachés d'habitus et de cultures éducatives, ce qui est parfaitement latent dans un concept aussi fondateur que «l'approche actionnelle», «el enfoque orientado a la acción» ou «an action-oriented approach».

Mais cette réponse en cache une autre, qui se pose dans chaque langue, indépendamment des autres. La relation entre ce que nous appellerons les dimensions didactiques du langage, les compétences générales et langagières et leurs relations avec les activités et les stratégies langagières, n'est peut-être pas suffisamment claire pour devenir univoque. Cette description entre des niveaux, linguistiques, pragmatiques et sociolinguistiques selon le schéma de Bachman L. (Bachman 1990 : 41) ne tranche pas en effet sur les rapports de priorisation didactique entre ces dimensions ; or, l'enseignant qui élabore son unité didactique tout autant que celui qui évalue ne peuvent prétendre les enseigner et les évaluer sur un même plan pour tous les niveaux. Il n'existe *in extenso* aucune raison d'affirmer que la réalisation de l'acte «can deal with», «peut se débrouiller» ou «se desenvuelve» n'implique qu'une maîtrise approximative de formes impératives ou locutionnelles, mais il y a pour l'enseignant comme pour l'évaluateur une impérieuse nécessité à les définir.

La rédaction d'un descripteur tel que «peut communiquer dans le cadre d'une tâche simple et routinière ne demandant qu'un échange d'information simple et direct» (Cadre 2001 : 67) semble porter en elle une valeur d'unicité holistique observable et évaluable. C'est oublier d'une part qu'en matière d'évaluation la notion de compétence, implicitement véhiculée par des verbes comme «peut», «est capable de», «can» ou «puede», que l'on retrouve en amorce de tous les descripteurs, n'est qu'inférable sur la base d'un acte performatif nécessairement limité au

contexte et aux contingences de cette évaluation, mais aussi et peut-être surtout que cette évaluation repose sur l'observation d'un ensemble de traits multidimensionnels dès lors que l'on ne satisfait plus d'un simple avis binaire «peut»/«ne peut pas».

Afin de procéder à une lecture multidimensionnelle qui permettrait d'inférer la compétence citée *supra* sur la base d'une performance (par exemple un jeu de rôle en face à face avec un enseignant ou un pair), on a désormais souvent recours à des grilles d'interprétation qui listent les dimensions qui seront observées dans la performance et attribuent une série de valeurs correspondantes. Celles que le CIEP a conçues pour son DELF sont parmi les plus connues (CIEP ; cf. sitographie). Elles reprennent les descripteurs de compétence du Cadre et permettent pour une même performance une attribution synoptique de notes selon différentes dimensions, fonctionnelles, sociolinguistiques et linguistiques. Elles confèrent à l'évaluation une meilleure fiabilité, mais rien ne permet par contre de justifier théoriquement la relation qui existe entre la somme des valeurs attribuées sur chaque dimension et l'inférence finale d'un acte de langage performatif.

onclusions

Ainsi, si les descripteurs pris individuellement peuvent donner l'illusion d'une vision globale, leur multiplicité génère et conditionne des liens encore mal connus et mal explicités entre eux.

Le travail fourni par North, et ses collègues il y a dix ans, puis celui des nombreuses équipes qui se sont penchées sur le Cadre, reste d'une puissance étonnamment novatrice, et l'on pourra nous reprocher avec juste raison d'en revisiter un peu vite les limites et les malentendus une décennie après. Puisse le lecteur ne pas se méprendre sur nos intentions.

Cet article relativement eurocentré ne doit pas nous faire oublier les nombreux travaux précurseurs ou en cours aux États-Unis, au Canada ou ailleurs, et si la didactique des langues et des cultures tente d'adopter un langage commun, cela ne signifie nullement qu'elle doive s'orienter vers une approche uniforme. Les limites évoquées nous rappellent qu'en matière éducative la valeur normative est souvent inversement proportionnelle à ses ambitions, et l'on doit s'en réjouir car s'il convient de n'évaluer que ce qui a été enseigné, cela ne signifie nullement que nous ne devons enseigner que ce qui sera évalué.

L'évaluation n'est un processus réductif que si l'enseignement décide de l'être.

Bibliographie

- ALDERSON J.-C., BANDS, SCORES, dans Alderson J.-C. et North B. (sous la direction de), (1991), *Language testing in the 1990s*, Londres, British Council/Macmillan, Developments in ELT, p. 71-86.
- BACHMAN L.F. et PALMER A.S. (1981), *The Construct Validation of the FSI Oral Interview*, dans *Language Learning*, vol 31, p. 67-86.
- BACHMAN L.F. (1990), *Fundamental Considerations in Language Testing*, Oxford, Oxford University Press.
- BEACCO J.-C., BOUQUET S. et PORQUIER R. (2004), *Niveau B2 pour le français, un référentiel*, Paris, Didier.
- BEACCO J.-C. et PORQUIER R., (2007), *Niveau A1 pour le français, un référentiel*, Paris, Didier.
- BEACCO J.-C., LEPAGE S., PORQUIER R. et RIBA P. (2008), *Niveau A2 pour le français, un référentiel*, Paris, Didier.
- BEACCO J.-C., BLIN B., HOULES E., LEPAGE S. et RIBA P. (2010), *Niveau B1 pour le français, un référentiel*, Paris, Didier.
- BRETON G., GEORGES S. et TAGLIANTE C. (2009), traduction française, *Manuel pour relier les examens au Cadre européen commun de référence pour les langues*, 2009, Strasbourg, Conseil de l'Europe éd.
- CONSEIL DE L'EUROPE, COE, (2001), *Cadre commun de référence pour les langues, apprendre, enseigner, évaluer*, Paris Strasbourg, Didier.
- CENTRE INTERNATIONAL D'ÉTUDES PÉDAGOGIQUES, CIEP, (2005), *Kit de formation des correcteurs examinateurs du DELF et du DALF*, Sèvres, CIEP éd.
- CENTRE INTERNATIONAL D'ÉTUDES PÉDAGOGIQUES, CIEP, (2007), *Étude de corrélation inter jurys de la production écrite du DELF A1 et du DELF B2*, Sèvres, CIEP éd.
- CLARK J.D. et SWINTON S.S. (1979), *An exploration of Speaking Proficiency Measures in the TOEFL Context*, dans *TOEFL Resarch Report n° 4*, Princeton, New Jersey, Educational Testing Service.
- GLABONIAT M., MÜLLER M., RUSCH P., SCHMITZ H. et WERTENSCHLAG L. (2005), *Profile Deutsch*, Munchen, Langenscheidt Verlag.
- LENZ P. ET SCHNEIDER G. (2004), *Introduction à la banque de descripteurs pour l'auto-évaluation créée pour le Portfolio européen des langues*, Strasbourg, Conseil de l'Europe éd.
- LUSSIER D. et TURNER C.E. (1995), *Le point sur l'évaluation en didactique des langues*, Anjou, Centre éducatif et culturel éd.
- NORTH B. et SCHNEIDER G. (1998), *Scaling descriptors for language proficiency scales*, *Language Testing*, Cambridge, Cambridge University Press, vol. 15 n° 2, p. 217-262.
- RIBA P. (2010), *La spécification pour le français des niveaux C1 et C2 du Cadre européen commun de référence pour les langues*, thèse de doctorat, Paris, Sorbonne Nouvelle Paris III.
- SCHÄRER R. (2007), *Portfolio européen des langues : Rapport intérimaire 2007*, Strasbourg, Conseil de l'Europe éd.
- SERRES M. (1968, rééd 1982), *Le système de Leibniz et ses modèles mathématiques*, Paris, Presses universitaires de France.
- SILVA H. (2011), *De l'illusion d'univocité à la revendication de la polyphonie : pour des lectures plurielles du Cadre européen commun de référence pour les langues*, dans *Synergies Europe*, n° 6.
- TAKALA S., KAFTANDJIEVA F., VERHELST N., BANERJEE J. et ECKES T. (2009), *Reference Supplement to the Manuel for Relating Language examinations to the Common European Framework of Reference for Languages: Learning, teaching, assessment*, Strasbourg, Conseil de l'Europe éd.
- WILDS C. P. (1975), *The oral interview test*, in Jones R., Spolsky B., (Eds.), *Testing language proficiency*, Arlington, Center for Applied Linguistics, p. 29-44.

Sitographie

(dernière consultation, 25/9/2012)

NORME MEXICAINE CENNI,

<http://www.cenni.sep.gob.mx/>

DIALANG (via le site de l'université de Lancaster),

<http://www.lancs.ac.uk/researchenterprise/dialang/about.htm>

PORTFOLIOS DU CONSEIL DE L'EUROPE,

http://www.coe.int/t/DG4/Portfolio/?L=FetM=/main_pages/portfoliosf.html

ALTE, Association of Language Testers in Europe,

http://www.alte.org/attachments/files/framework_french.pdf

ENGLISH PROFILE,

<http://www.englishprofile.org/>

CIEP, centre international d'études pédagogiques :

Diplôme d'études en langue français, diplôme approfondi
de langue française, DELF DALF,

<http://www.ciep.fr/delfdalf>

ILLUSTRATION DES NIVEAUX DE COMPÉTENCE EN PRODUCTION ORALE DANS CINQ LANGUES,

[http://www.ciep.fr/publi_evalcert/dvd-productions-orales-ccercl/
index.php](http://www.ciep.fr/publi_evalcert/dvd-productions-orales-ccercl/index.php)

FLE.fr,

COMPTE RENDU DU FORUM DES CENTRES DE FLE, LA DYNAMIQUE DU CADRE EUROPÉEN
DE RÉFÉRENCE POUR LES LANGUES, CAVILAM Vichy, 2006

<http://www.fle.fr/Vichy/compterenduvichy2006.htm>

L es tests de langues : une nécessité ?

CLAIRE TARDIEU

SORBONNE NOUVELLE PARIS 3, FRANCE

I ntroduction

Pour commencer, interrogeons-nous sur cette réflexion de Gardner : «Quand l'évaluation se fond progressivement dans le paysage, elle n'a plus besoin d'être distinguée des autres activités de la classe. À l'instar de ce qui se passe dans un bon apprentissage, les enseignants et les élèves sont toujours en train d'évaluer. Il n'y a plus alors nécessité d'enseigner pour l'évaluation parce que l'évaluation est omniprésente ; le besoin de tests formalisés pourrait bien disparaître» (Gardner 1996 : 115). L'image suscitée ici est celle d'une évaluation de type fondu-enchaîné avec l'apprentissage conçu comme une action mesurable sur le terrain ; autrement dit une évaluation formative, accompagnant la progression comme son ombre, et rendant de fait caduc tout autre type de mesure distanciée. Pure utopie ou simple réalisme ? La question se pose aujourd'hui d'autant plus que des outils d'évaluation en langues distanciés se développent en réponse à une demande sociale et politique effective et pressante. La suggestion de Gardner, qui provient d'une remise en cause des tests psychométriques à l'anglo-saxonne, se justifie-t-elle encore avec l'apparition d'un nouveau type de test, de quatrième génération que l'on pourrait appeler «l'approche sociolinguistique intégrée de type actionnel» ?

En premier lieu, nous présenterons les trois premières approches de l'évaluation formalisée en montrant leurs points forts et leurs points faibles. Puis nous nous demanderons si la quatrième approche, contemporaine, pallie effectivement les défauts des précédentes. Enfin nous proposerons une définition souple de l'évaluation et des pistes possibles pour un usage modéré et utile.

L es tests et les 3 premières approches

Spolsky (1981) divise l'histoire du *testing* en langues en trois grandes approches : pré-scientifique, structuraliste psychométrique et sociolinguistique-intégrative (cf. Reichert : 38-39).

L'APPROCHE PRÉ-SCIENTIFIQUE

L'approche pré-scientifique s'inscrit dans une logique d'examens et de juges, mais manque d'objectivité et de fiabilité, la fiabilité étant, selon Horner, « la mesure selon laquelle on retrouvera le même classement des candidats dans deux passations (réelles ou simulées) des mêmes épreuves » (Horner 2010 : 17). Elle manque aussi de validité – à savoir, « la garantie que ce qui est effectivement testé est ce qui, dans le contexte donné, doit être évalué et que l'information recueillie donne une image exacte de la compétence des candidats en question » (Horner, 2010 : 17). L'approche pré-scientifique est celle pratiquée au Baccalauréat, par exemple, qui ne s'appuie ni sur des théories du langage ou de l'apprentissage, ni sur une méthodologie scientifique. Dès la première moitié du xx^e siècle, la docimologie (en particulier la commission Carnegie de 1932) a montré qu'il faudrait 127 correcteurs pour obtenir la valeur stable d'une dissertation de philosophie, 78 pour une composition française, 28 pour un devoir d'anglais, 19 pour une version latine et 13 pour un devoir de mathématiques... (Laugier et Piéron 1938 ; Piéron 1963 : 23)

L'APPROCHE STRUCTURALISTE-PSYCHOMÉTRIQUE

L'approche structuraliste-psychométrique intervient pour répondre à la demande d'objectivité et de fiabilité, dans un souci d'évaluer la maîtrise du système linguistique (Mc Namara 2000), avec en contrepartie un effet d'atomisation et de décontextualisation. On trouve trace de cette approche dans une partie de l'épreuve du baccalauréat appelée « épreuve de compétence linguistique » instaurée en 1984 (B.O.E.N. n° 27 du 7 juin 1983) et abolie en 2002 (B.O.E.N. n° 23. Note de service 2001-091 du 7 juin 2001).

L'APPROCHE SOCIOLINGUISTIQUE INTÉGRATIVE

À partir des années 1960, la demande est en faveur d'une approche intégrée, testant à la fois la compétence linguistique fine et les savoir-faire (CO, CE, PO, PE) malgré les difficultés en termes de coût et d'attribution de scores en productions orales ou écrites.

Le TOEFL (Test Of English as a Foreign Language) est un bon exemple de test qui associe l'approche structuro-psychométrique avec des items discrets et l'approche intégrée introduisant les *skills*, soit une évaluation des compétences de réception et de production en contexte social plausible. Ainsi, aux exercices de compétence linguistique pure s'ajoutent des items sous forme de QCM, par exemple, qui testent la capacité à comprendre un texte oral ou écrit « authentique » ou à produire un discours adapté à un type d'interlocuteur ou à une forme particulière (lettre, article, etc.) Ce type d'évaluation semble à la fois valide, fiable, et faisable au sens où « les examinateurs ne disposent que d'un temps

limité, ne voient qu'un échantillon limité de performances et qu'il y a des limites au nombre et à la nature des catégories qu'ils peuvent manipuler comme critères (Horner 2010 : 17).»

Alors pourquoi une quatrième approche ? Quels besoins nouveaux se sont-ils fait jour ?

LES LIMITES DE L'APPROCHE SOCIOLINGUISTIQUE INTÉGRATIVE

Limites quant à la compétence évaluée

Une critique, celle de Bachman (1990 : 82), porte sur le fait que, dans ce type de tests alliant des savoir-faire et des éléments discrets, il est impossible de déterminer la nature du lien entre les deux. Ensuite, ce type d'évaluation ne mesure que certains aspects de la compétence à communiquer langagièrement. On peut citer à cet égard la synthèse réalisée par Springer (2000) dans sa «marguerite de la communication» (reprenant l'évolution de Chomsky à Canale et Swain en passant par Hymes et Charaudeau) et bien sûr les travaux de Oller (1976) et Bachman (1990). Oller affirme que la compétence langagière n'est pas seulement constituée d'éléments séparés. D'où l'idée d'une compétence intégrative : «The Unitary Competence Hypothesis». Selon Oller, la production écrite ou orale n'est pas séparée de la réception, car le locuteur doit en permanence contrôler sa production pour s'assurer d'être compris. L'inverse est par ailleurs tout aussi vrai : en phase de réception d'un message, l'auditeur ou le lecteur doivent formuler des hypothèses sur la suite du message. D'autre part, les tests formalisés ne mesurent pas, en particulier, les dimensions véritablement sociologiques et pragmatiques de la communication, ni les dimensions cognitive et intellectuelle, ni encore moins la capacité de communication réelle chère à Bachman. Bachman fait en effet la distinction entre la compétence à communiquer langagièrement (*communicative language ability*) et la communication véritable (*actual communication*). La première inclut une dimension liée aux savoirs («language knowledge dimension») comprenant elle-même des compétences langagières organisationnelles et pragmatiques, des compétences stratégiques et des mécanismes psycho-physiologiques ; la seconde se caractérise par sa dimension «d'utilisation réelle» (Bachman 2007 : 54). Puren souligne ce problème en écrivant :

«Alors que la recherche d'objectivité voire de "scientificité" dans l'évaluation implique une approche très analytique – avec au minimum une dissociation des différentes compétences de compréhension/expression orale/écrite –, l'évolution récente de la didactique des langues a réactivé l'idée ancienne d'une indissociabilité entre les trois objectifs fondamentaux. D'une part l'objectif actuel de compétence de communication inclut une "composante socioculturelle" (la maîtrise de la langue suppose la maîtrise des codes qui régissent son emploi en situation). L'approche cognitive, d'autre part, met l'accent sur l'importance des activités réflexives (et donc de la formation intellectuelle) dans le processus d'apprentissage linguistique. Enfin l'approche par les tâches, vers laquelle semble s'orienter actuellement l'enseignement scolaire des langues en France, ne va pas

simplifier le problème puisque c'est principalement le produit final d'une activité personnelle ou en petits groupes qui devra désormais être évalué, alors que jusqu'à présent l'enseignant pouvait dans une certaine mesure observer et évaluer directement et en temps réel l'activité même d'apprentissage parce qu'elle était étroitement liée à son enseignement : comment faire la part, dans l'exposé collectif d'un groupe d'élèves, entre les compétences langagières et les compétences culturelles de chacun, la maîtrise des ressources documentaires, ou encore l'efficacité du travail de groupe ? » (Puren 2001 : 14)

Il semble que Puren poursuive ici dans la droite ligne des travaux de Bolton (1991) qui, en retraçant l'évolution de l'objet évalué, constatait déjà voici plus de vingt ans, avec l'avènement de l'approche communicative dans l'enseignement-apprentissage des langues, qu'il est beaucoup plus facile d'évaluer des connaissances que des compétences, a *fortiori* lorsque ces dernières s'exercent dans un travail collaboratif et, conviendrait-il d'ajouter, « proche de la vie réelle ».

Gardner a raison sur ce point, et Bachman tout autant : les tests formalisés ne certifient pas la compétence réelle à communiquer langagièrement.

Limites quant à la fiabilité des items

L'expérimentation menée au sein du projet européen EBAFLS (Building a European Bank of Anchor Items For Language Skills) a montré que des items corrélés aux niveaux du CECRL par des experts de chaque pays et ayant fait l'objet d'une validation psychométrique ne fonctionnaient pas de la même manière d'un pays à l'autre, avec des élèves de huit pays européens censés être au même niveau du CECRL et qu'il était par conséquent impossible, pour l'instant tout au moins, de créer une banque d'items d'ancrage (Tardieu 2010). Comment expliquer les causes de l'échec de ce projet mené de 2005 à 2007 ?

La taxonomie de Van de Vijver et Hambleton (1996) établit trois biais possibles, à savoir :

- un premier biais conceptuel, soit la non-équivalence des concepts d'un groupe culturel à l'autre ;
- un second, méthodologique, résultant de la manière d'administrer les instruments ;
- un troisième, lié aux items eux-mêmes (des traductions inadéquates par exemple).

Au cours du projet EBAFLS, ces différents biais ont pu être observés (voir Tardieu 2010).

Concernant le biais lié aux items, une différence essentielle est apparue à l'issue du projet EBAFLS. Les analyses des résultats ont révélé un important fonctionnement différentiel à l'item (DIF) rendant impossible l'établissement d'échelles concomitantes de pays à pays. En d'autres termes, un même item ne se comporte pas de la même manière selon le pays. Ce qui revient à dire que les compétences d'un élève au niveau B1 dans un pays ne sont pas équivalentes à celles d'un élève au niveau B1 dans un autre pays. Cette différence de fonctionnement de l'item s'exprime à la fois en termes de difficulté et en termes de discrimination.

En matière de biais lié à l'administration du test, on s'est aperçu que le type de consignes n'était pas le même selon les pays. Certains adressant des consignes concernant l'ensemble du test, d'autres des consignes par tâches, ou pour chaque item.

Le choix de la langue de passation relève aussi d'un biais lié à l'administration du test, comme l'a montré le pilote effectué en utilisant les deux circonstances. Un protocole fut établi en compréhension de l'écrit pour le français, et administré dans deux pays. Ce pilote a fait apparaître que les scores étaient plus élevés lorsque le test était administré en L1. Néanmoins, il fut décidé d'administrer le test en L2, en raison de l'objectif même du projet : la création d'une banque d'items communs.

On peut se demander si le choix de la langue d'administration du test ne relève pas également d'un biais conceptuel, certains pays ne se posant pas la question, d'autres y accordant une grande importance. Enfin, des différences de jugement ont été observées au moment de la spécification et de la standardisation des items, en amont de la passation, et également en aval par le retour des professeurs ayant fait passer les items dans les différents pays.

Ainsi rien ne permet d'affirmer l'existence d'un lien ferme entre EBA-FLS et le CECRL. On peut légitimement se demander comment les tests contemporains de type actionnel parviennent à lier leurs composantes aux paramètres du CECRL.

L'*approche sociolinguistique intégrative de type actionnel*

LES FONDEMENTS

Pour compléter la classification de Spolsky, une quatrième approche, non psychométrique, devrait être ajoutée : « l'approche sociolinguistique intégrative de type actionnel ». Cette approche s'appuie sur une conception socio-pragmatique de la communication appliquée à l'enseignement dont on trouve les fondements chez Nunan (1989), Widdowson (2002[1990]), Ellis (2003), entre autres. Ces chercheurs ont posé les bases de l'approche actionnelle, dans la poursuite de l'approche communicative : on ne communique pas seulement pour agir sur autrui mais aussi et surtout pour agir avec autrui. L'accent est mis sur l'utilisation de la L2 aux fins d'accomplir des tâches sociales qui vont de la simple déclinaison d'identité à la discussion de haut vol dans des domaines de spécialité. La nouveauté de cette approche est de mettre l'accent sur l'ancrage dans la réalité de la communication, et de concevoir les activités de réception, de production, d'interaction et de médiation de manière

holistique, concourant toutes à la réalisation d'un projet. Ce type d'enseignement appelle un nouveau type d'évaluation, plus en phase avec des scénarii réels, et davantage apte à mesurer des compétences indépendamment de tout contexte d'apprentissage.

Pour reprendre la taxonomie d'Alderson (2002), il s'agit ici de faire la distinction entre le test de type « achievement » relié à un apprentissage en amont et le test de type « proficiency », indépendant d'une formation et tourné vers l'avenir. Ainsi fera-t-on mieux la distinction entre l'évaluation du savoir (ou du niveau) qui est « l'évaluation de l'atteinte d'objectifs spécifiques », « centrée sur le cours », correspondant à « une vue de l'intérieur » et l'évaluation de la capacité (mise en œuvre de la compétence ou performance), c'est-à-dire « de ce que l'on peut faire ou de ce que l'on sait en rapport avec son application au monde réel » (CECR, 2001 :139), l'évaluation de la capacité correspondant « à une vue de l'extérieur ». Certains certificats, comme le *Deutsches Sprachdiplom* (certificat d'allemand) ou le Certificat en langue de l'enseignement supérieur (désormais CLES), se réclament de cette approche.

LES PROPOSITIONS DE L'INSTITUTION

Développer les certifications

L'article 3 du décret n° 2005-1011 du 22 août 2005 relatif à l'organisation de l'enseignement des langues vivantes étrangères dans l'enseignement scolaire français prévoit la mise en place de certifications des connaissances et compétences acquises en langues vivantes étrangères¹. Ce qui signifie que les élèves devraient pouvoir se voir offrir à différents moments du cursus définis par l'institution des évaluations en langues indépendantes de leurs cours de langues, validant leurs performances selon les échelles du CECRL.

Depuis le décret de 2005, au sein même de l'institution, des efforts sont effectivement menés pour « crédibiliser » les diplômes. L'examen de langue pour l'obtention du Diplôme National du Brevet (DNB – B.O. n° 22 du 7 juin 2007), qui s'est adapté à l'évaluation de 5 compétences (compréhension de l'oral, expression orale en continu, interaction orale, compréhension écrite et expression écrite), donne lieu à la délivrance non d'un certificat à proprement parler, mais d'un diplôme, où la langue étrangère est affectée d'un coefficient 1 et validée au niveau A2 du CECRL. Dans ce cas cependant, l'évaluation n'est pas conçue par un Institut extérieur du pays de la langue étudiée, mais par les professeurs eux-mêmes ou par une équipe de professeurs.

Il en va de même pour les nouvelles épreuves du Baccalauréat entrées en vigueur à la session 2013, avec une validation à B2 (LV1) ou B1 (LV2). On devine les limites de ces diplômes délivrés de l'intérieur, tout au moins en termes de validité et de fiabilité.

Un type de certification indépendant aurait plusieurs avantages : non seulement il posséderait une valeur non limitée au territoire français et

1. JO du 25 août 2005, Titre 1, articles 3 à 7.

à ses traditions éducatives, mais il permettrait aussi à terme de « libérer » les professeurs d'une partie de leur tâche (qui les oblige à être à la fois formateur et évaluateur-certificateur) et de redonner ses lettres de noblesse à leur mission d'évaluation formative propre à accompagner l'apprentissage, selon la vision de Gardner. On pourrait ainsi d'un côté évaluer pour accompagner l'apprentissage en fonction des objectifs requis à chaque niveau, et de l'autre proposer des tests en langue déconnectés de l'apprentissage. Une telle proposition qui se dessine dans l'environnement institutionnel paraît bien adaptée au contexte contemporain.

À titre d'exemple, une évaluation en allemand du Goethe Institut indexée sur les niveaux du CECRL est proposée aux élèves volontaires en fin de Troisième ou de Seconde depuis 2006 sur la base du volontariat. Les épreuves portent sur la compréhension écrite, la compréhension orale (écoute d'un document sonore et réponses par écrit à un questionnaire) et l'expression écrite. Les épreuves orales prennent la forme d'entretiens visant à évaluer l'expression orale (la capacité à converser sur des sujets familiers pour le niveau A2 par exemple). La certification est acquise à un niveau précis du CECRL quand les candidats ont obtenu ce niveau dans chaque activité langagière. Ils reçoivent alors le diplôme DSD (Deutsches Sprachdiplom) avec le relevé de leurs notes par activité langagière. Tous les élèves qui ont réussi au moins une activité se voient remettre une attestation signée par l'autorité académique.

Les limites de ces certificats

Le projet EBAFLS ayant montré la difficulté de standardisation des items, on peut en premier lieu douter de la valeur de ces certificats ou diplômes en termes d'attestation de niveaux.

Ensuite, il apparaît que les composantes de la compétence à communiquer langagièrement sont évaluées de manière séparée et non véritablement holistique.

Enfin, ces évaluations sont sans doute moins valides et moins fiables que des évaluations de type approche sociolinguistique intégrée (avec volet psychométrique).

D'autres certificats, comme le DCL ou le CLES, sont également proposés, voire requis à des niveaux supérieurs du cursus. Offrent-ils une pertinence plus grande ? On prendra ici l'exemple du CLES2 dont l'obtention est en principe requise à la sortie du Master (en France en tout cas).

LE CLES

Le certificat en langue de l'enseignement supérieur, niveau 2 n'est pas exempt d'intérêt : il entend proposer une évaluation qualitative de la compétence de communication langagière conçue dans sa globalité. Il se déroule pendant 3 heures, suivant une phase de compréhension

de l'oral prenant appui sur des documents audio ou vidéo authentiques d'environ 800 mots au total, puis d'une phase de compréhension de l'écrit (environ 2000 mots à traiter en compréhension globale et sélective), puis une production écrite contextualisée de 300 mots, et enfin une phase d'interaction orale de 10 min en présence d'un examinateur, le tout articulé autour d'une thématique commune et les différentes phases fonctionnant en synergie (B.O. n° 20 du 17 mai 2007).

Ce certificat a le mérite de solliciter les multiples facettes de la compétence à communiquer langagièrement (CO, CE, PE, IO) en une tentative d'évaluation holistique autour d'une thématique suivie devant être traitée de façon multimodale à travers les différents supports. L'innovation majeure consiste à introduire l'interaction orale entre deux candidats, au lieu de celle entre l'examinateur et le candidat, interaction qui repose sur une conception de la tâche pouvant se rapprocher des types définis par Prabhu ou Pica, Kanagy et Falodun. Prabhu (1987) distingue trois grands types de tâches fondés sur le genre d'activité cognitive impliquée (déficit d'information, déficit de raisonnement, déficit d'opinion) (Ellis, *op. cit.*, 214). Pica, Kanagy et Falodun (1993) opèrent une classification psycholinguistique fondée sur des catégories interactionnelles qui envisagent la relation entre les interactants – par exemple, qui détient l'information nécessaire à l'accomplissement de la tâche ? Ou encore, la tâche requiert-elle que les participants parviennent à un accord final (cf. Ellis, *op. cit.*, 215) ? Mais dans quelle mesure les sujets du CLES proposés au niveau 2 empruntent-ils véritablement à de tels fondements théoriques ?

Tout d'abord on constate que même si la thématique est la même tout au long de la passation du test, et même si l'interaction orale requiert que l'on recherche un consensus par rapport à un problème évoqué tout au long de l'épreuve, les différentes composantes de la compétence à communiquer langagièrement n'en sont pas moins évaluées séparément, comme dans l'approche antérieure non actionnelle. Et cela est d'autant plus vrai que les performances ne sont pas compensables (ce qui cause d'ailleurs un taux d'échec assez élevé au CLES2 – plus de 50 %). Et peut-il en être autrement ? L'intérêt réel du CLES nous semble résider plutôt dans le fait qu'il peut modéliser l'apprentissage en amont d'une manière plus satisfaisante d'un point de vue didactique que les tests habituels, ce qui est un atout majeur, par rapport à d'autres certificats.

Mais si ces tests distancés, non psychométriques, ne testent pas la compétence langagière d'une manière plus globale que les tests de l'approche précédente, sont-ils vraiment utiles en tant que tels ? Ne peut-on se contenter, comme le préconise Gardner, d'une évaluation au plus près des progrès de l'élève ? Tout dépend en fait de l'information que l'on recherche.

Quelles perspectives ?

ÉVALUER LA PERFORMANCE DU SYSTÈME

En termes institutionnels, le test formalisé semble avoir sa place pour évaluer la performance du système. Ainsi, les évaluations PISA ou celles menées par la Division de la l'Évaluation de la Prospective et de la Performance (DEPP) créée en 1987² renseignent utilement les professionnels de l'enseignement et les politiques sur la pertinence du système. En termes d'évaluation ou de certification individuelle, d'autres modalités nous semblent pouvoir être envisagées.

ÉVALUER LES PERFORMANCES INDIVIDUELLES

Estimer le niveau d'un cours

La première modalité consiste à valider le niveau individuel par l'estimation du niveau d'un cours. C'est ce que font déjà, d'une certaine manière, de nombreux instituts d'évaluation (Cambridge Institute, Goethe Institut, CIEP, etc.), en proposant des thématiques différentes selon le niveau de la formation et du test qui la clôt. Par exemple, à A2, le candidat est interrogé sur sa famille, ses hobbies, etc., en accord avec les descripteurs de niveaux du CECRL. Au niveau B1 on lui demande de raconter un film ou un livre. La grille CEF-ESTIM (Tardieu et al. 2011) est un outil permettant aux professeurs de toute langue et enseignant à tous les niveaux du cursus de situer le niveau de leur cours en relation avec les descripteurs du CECRL. Ce faisant, il devient possible de valider le niveau des élèves en fonction du niveau estimé du cours ou de la formation en question.

Le C-Test

La seconde modalité consiste à utiliser un test simple, fiable, valide, faisable et très économique, le C-Test. En fait, le C-Test est une forme améliorée du test de closure. Il s'agit de fournir un texte écrit dans lequel on efface la deuxième moitié du mot tous les 2 mots, tout en laissant intacte la première et la dernière phrase du texte. Chaque mot tronqué vaut pour un item d'évaluation. Un C-Test est généralement constitué de 4 à 6 textes et chaque texte contient 20 à 25 items. Ce qui fait en tout environ 80 à 100 items par C-Test (Grotjahn 2006, Reichert 2011 : 40). Le candidat doit s'efforcer de reconstituer le texte au mieux. (À titre indicatif un locuteur natif réalise une performance de 90 % en général)

Dans sa thèse, Reichert explique que l'on s'est aperçu que les items avaient une forte corrélation entre eux, et donc que la compétence linguistique pouvait être unitaire conformément à la théorie développée par Oller (Reichert : 41). Mais surtout, Reichert a montré le haut degré de corrélation entre les résultats au C-Test et ceux obtenus par les mêmes étudiants au Test de Connaissance du Français (TCF), de type TOEFL amélioré comprenant l'évaluation en compréhension écrite et orale, en

2. La DEP fut créée en 1987 sous ce nom mais existait précédemment sous diverses appellations en tant que service statistique du ministère. C'est en 1975/1976 seulement qu'est conçue et réalisée la première évaluation d'élèves en France. Par la suite, la DEP est devenue DPD (Direction de la programmation et du développement) en 1997, a repris le nom de DEP en 2003. C'est en 2006 que le deuxième P fut ajouté avec la restructuration du ministère en avril, passant sous la coupe du nouveau Secrétariat général.

expression écrite et orale, et en compétence linguistique. Ce qui revient à dire qu'un test très facile à créer et très peu coûteux à faire passer fournit une information aussi fiable que d'autres tests beaucoup plus lourds sur la compétence langagière. Bien évidemment le C-Test ne fournit pas de renseignements précis sur chacune des compétences, ni sur les connaissances, et ne peut donc remplacer d'autres tests dont tel est l'objet, mais il peut tout à fait valider un niveau général.

ÉVALUATION ET CERTIFICATION

Fonction sociale de l'évaluation

De Ketele ajoute une dimension importante, au cœur de notre problématique, à savoir, sa fonction sociale. Selon lui, évaluer c'est confronter un ensemble d'informations à un ensemble de critères (référentiel) en vue de prendre une décision (De Ketele 1980). De quelle décision parle-t-on ici ? S'agit-il d'une décision personnelle de l'apprenant, de celle du professeur ou du formateur évaluateur, d'un jury d'examen, d'un employeur ? Par définition, une évaluation distanciée de type certification se situe hors de la sphère familière de la classe, et implique une prise de décision qui échappe au sujet évalué. Plus le besoin d'un regard distancié est affirmé, plus le sujet semble devoir être mis à distance de la prise de décision, avec des effets potentiellement nuisibles pour lui.

Si l'on veut vraiment privilégier l'aspect individuel et social de l'évaluation, il conviendrait d'afficher clairement une éthique. À cet égard, cette seconde réflexion de Gardner paraît appropriée : « Évaluer, c'est recueillir des données sur les compétences et les potentiels des individus dans le double objectif de leur en faire retour utilement et de procurer des informations indispensables à la communauté environnante (Gardner 1996 : 114). Ajoutons cependant que si l'évaluation doit s'avérer utile et profitable à l'individu et à la communauté, elle devrait avant tout permettre la mesure positive du potentiel de chacun pour le bénéfice de tous.

En définitive, évaluer, c'est (s') informer à des degrés divers, dans des domaines divers, avec des intentions diverses émanant de personnes ou d'instances également diverses. On peut résumer cet éventail de combinaisons possibles par le schéma ci-contre.

Il paraît ainsi important de situer toute forme d'évaluation ou de certification et de définir en amont les paramètres qui lui donneront à la fois forme et sens.

La proposition de Gardner citée en entrée en matière constitue une option pragmatique qui pourrait suffire si l'on utilisait en parallèle des référentiels d'apprentissage. Le temps que l'on consacre à évaluer serait utilisé à concevoir des activités d'apprentissage reliées à ces référentiels. D'une certaine manière, c'est ce que le CECRL a tenté de faire en proposant des descripteurs de niveaux de compétences. La difficulté à relier directement les compétences individuelles à ces

descripteurs nous engage à imaginer une troisième voie qui est la suivante : prévoir des sessions d'apprentissage (en classe ou en situation réelle d'apprentissage) constitués de projets et d'activités reliés à ces descripteurs, en une sorte de balisage initial et offrir en parallèle des occasions de tester son niveau, à certains moments du cursus, indépendamment de l'apprentissage.

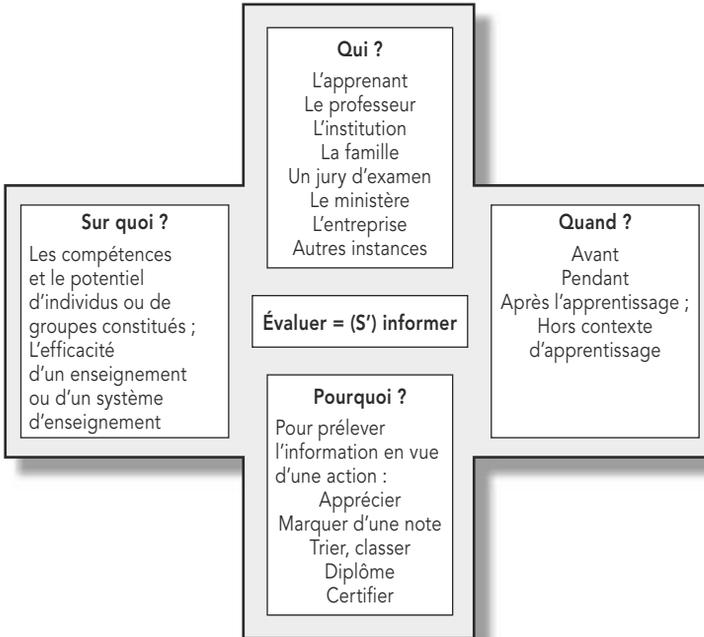


Schéma de l'évaluation située

C onclusion

Dans ces conditions, la proposition de Gardner n'est plus utopique mais s'inscrit au contraire dans un réalisme pragmatique. Les informations recueillies sur chaque apprenant lui sont retournées sur le mode de son choix (appréciation, note, couleur, diplôme, etc.) : elles lui appartiennent de droit en tant que sujet ; et la société quant à elle peut se satisfaire de l'information pourvue par le référentiel correspondant à la session d'apprentissage.

Si l'on veut absolument répondre à la demande du monde de l'entreprise, on pourrait effectivement offrir la possibilité à certains moments du cursus d'obtenir une certification individuelle par le biais d'un C-Test (peu coûteux) ou d'une évaluation-certification de type CLES ou Cambridge Exams, plus coûteuse et pas forcément plus fiable – sans oublier d'indiquer une date de péremption...

Bibliographie

- ALDERSON C., FIGUERAS N., KUIJPER H., NOLTE G., TAKALA S. et TARDIEU C. (2006), "Analysing Tests of Reading and Listening in Relation to the *Common European Framework of Reference*, The experience of the Dutch CEFR Construct Project", in *Language assessment Quarterly*, Volume 3, n° 1. p. 3-30.
- BACHMAN L.F. (1990), *Fundamental Considerations in Language Testing*, Oxford, UK, Oxford University Press.
- BACHMAN L.F. (2002), "Some reflections on task-based language performance assessment", *Language Testing*, Vol. 19, n° 4, p. 453-476.
- BACHMAN L.F. (2007), "What is the construct? The dialectic of abilities and contexts in defining constructs in language assessment", in Fox J. et al. (Eds.), *Language testing reconsidered*, Ottawa, CA, University of Ottawa Press, p. 41-72.
- CANALE M. et SWAIN M. (1980), "Theoretical bases of communicative approaches to second language teaching and testing", *Applied Linguistics*, Volume 1, p. 1-47.
- Conseil de l'Europe (2001), *Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer*, 1^{re} éd. 1996, 2^e éd. corr. 1998, Paris, Didier.
- ELLIS R. (2003), *Task-based Language Learning and Teaching*, Oxford : Oxford University Press.
- ESOL Examinations, University of Cambridge (2004), *Samples of oral production, illustrating for English, the levels of the Common European Frame of Reference*, Strasbourg, Council of Europe.
- GARDNER H. (1996), *Les intelligences multiples*, Paris, Retz.
- GILLE E. et NOJONS J. (2010), "Issues in constructing an international item bank : Differences in item behaviour across countries", *Revue française de linguistique appliquée*, vol. 15, n° 1.
- GROTJAHN R., KLEIN-BRALEY C. et RAATZ U. (2002), "C-Tests: an overview", in J.A. Coleman, R. Grotjahn et Raatz U. (Eds.), *University Language Testing and the C-Test*, Bochum, Germany, ASK-Verlag, p. 93-114.
- HYMES D. (1972), "On Communicative Competence", in J.B. Bride and J. Holmes (Eds.), *Sociolinguistics*, Reading, Penguin Modern Linguistics.
- LAUGIER H., PIÉRON H., Mme H. PIÉRON, TOULOUSE E. et WEINBERG D. (1934), *Études docimologiques sur la perfection des examens et concours*, Paris, Conservatoire des arts et métiers.
- LAUGIER H. et WEINBERG D. (1938), *Recherches sur la solidarité et l'indépendance des aptitudes intellectuelles d'après les notes des examens écrits au baccalauréat*, Paris, Chantenay.
- MC NAMARA T. (2000), *Language Testing*, Oxford, UK, Oxford University Press.
- NUNAN D. (1989), *Designing Tasks for Communicative Classroom*, Cambridge, Cambridge University Press.
- OLLER J.W. (1976), "Evidence for a general language proficiency factor: an expectancy grammar", *Die Neueren Sprachen*, n° 75, p. 165-174.
- PIÉRON H. (1963), *Examens et docimologie*, Paris, PUF.
- QUIVY M. et TARDIEU C. (2002), *Glossaire de didactique de l'anglais*, Paris, Ellipses.
- REICHERT M. (2011), *The validity of the C-Test Revisited: Findings from a Multilingual Environment*, Thèse de Doctorat en psychologie de l'université du Luxembourg, sous la direction de R. Martin.
- SCRIVEN M. (1967), "The methodology of evaluation", in *Perspectives of curriculum evaluation*. Chicago, Rand Mc Nally, p. 39-83.

- SPRINGER C. (2000), « Les multiples facettes du concept de “compétence” : que peut-on en retenir pour l’enseignement et l’apprentissage des langues ? », dans *Cognition, langue et culture, éléments de théorisation didactique*, Actes de la journée d’étude du 17 mars 2000, GEPED, Université Paris VII - Denis Diderot.
- TARDIEU C. (2005), *L’épreuve de didactique aux concours. L’évaluation en langues*, Paris, Ellipses, Collection CAPES/Agrégation langues.
- TARDIEU C. (2009), « Corriger ou évaluer », dans *La correction dans l’enseignement des langues de spécialité, Les Cahiers de l’APLIUT, Pédagogie et recherche*, vol. 28, n° 3, p. 9-25.
- TARDIEU C. (2010), « Votre B1 est-il mon B1 ? L’interculturel dans les tests d’évaluation en Europe », dans *Recherches en didactique des langues - Les langues tout au long de la vie, Cahiers de l’ACEDLE*, vol. 7, n° 2, 2010, p. 225-239. Version complète disponible en ligne : <http://acedle.org/spip.php?article3103>
- TARDIEU C., RAILI H., REICHERT M. et LEHMANN M. (2011), *CEF-ESTIM*. Disponible en ligne : <http://cefestim.ecml.at/>
- TYLER R. (1950), *Basic principles of curriculum and instruction*, Chicago, University Press.
- VAN DE VIJVER F.J.R, et HAMBLETON F.K. (1996), “Translating Tests : Some practical Guidelines”. *European Psychologist*, Vol. 1. n° 2, 89-99.
- WIDDOWSON H. G. (2002), *Aspects of language Teaching*, Oxford University Press (5th impression, 1st published 1990).

β

- B.O.E.N. n° 27 du 7 juin 1983 (Baccalauréat).
- B.O.E.N. n° 23. Note de service 2001-091 du 7 juin 2001 (Baccalauréat)
- B.O.E.N. n° 20 du 17 mai 2007 (CLES).
- B.O. n° 22 du 7 juin 2007 (DNB).
- JO du 25 août 2005, Titre 1, articles 3 à 7 (Certifications en langues étrangères dans le cadre institutionnel)

Sitographie

- Building a European Bank of Anchor Items for Foreign Language Skills (EBAFLS), (2007), http://www.cito.com/research_and_development/participation_international_research/ebafsl.aspx

Pour une certification contextualisée du français dans le monde : intégrer les approches actionnelles et les approches portfolio

MATHILDE ANQUETIL

UNIVERSITÀ DI MACERATA, ITALIE

Comme l'analyse Emmanuelle Huver (Huver, Springer 2011 : 23-32), il existe actuellement un réel clivage entre les pôles mécanicistes et holistiques¹ dans le domaine de l'évaluation en langues. D'un côté, des certifications de plus en plus mécanisées, légitimant leur valeur essentiellement sur la psychométrie, mais qui gagnent de plus en plus de terrain : les tests à choix multiples en ligne en sont l'exemple paradigmatique ; de l'autre, les évaluations dites alternatives : des recherches, des projets pilotes, des expérimentations, qui tentent de faire valoir la multidimensionnalité sociale de l'évaluation formative : les évaluations sur Portfolio sont alors l'exemple paradigmatique de la co-construction de l'évaluation entre l'instance d'enseignement, le candidat et l'évaluateur.

Mais si les critiques contre l'évaluation standardisée se multiplient parmi les chercheurs, pourquoi cet engouement du public pour les tests internationaux organisés avec la participation des acteurs de terrain ? Pourquoi à l'inverse une initiative d'envergure européenne comme le lancement du *Portfolio Européen des Langues* (désormais PEL) est-elle dans une phase d'essoufflement, si ce n'est de désengagement ? En réalité l'offre d'évaluation certificative est plus diversifiée qu'il n'y paraît et mérite sans doute d'être intégrée dans un système institutionnel coopératif ouvert avec des ponts entre les options de sorte qu'au-delà des divergences de perspective des diplômes, les

1. Pour une définition de ces termes, voir l'introduction ainsi que l'article de Vial dans ce numéro.

réels bénéficiaires soient les candidats à la certification, avec leurs visées contextuelles, ainsi que la diffusion du français dans le monde. Nous nous attacherons à promouvoir la voix des apprenants et des enseignants/évaluateurs, comme acteurs sociaux de la communauté francophone, à la recherche d'une politique de certification qui rende honneur à ceux qui renouvellent avec créativité la pratique, et dans les meilleurs des cas, le sentiment d'une langue partagée.

Malaise dans la certification

EXAMINATEUR EXTERNE, CANDIDAT ET SUJET D'EXAMEN CENTRALISÉ : UNE INTERACTION EN PORTE-À-FAUX.

Une anecdote en exergue qui illustrera le locus d'expérience d'où parle l'auteur de cet article.

« Il y a quelques années, j'ai été envoyée en mission dans un lycée d'une bourgade de l'Italie centrale dans les fonctions d'examinateur DELF B2. J'y suis reçue avec tous les honneurs par l'enseignante de français locale et le proviseur. Devant le bureau figurait en bonne place l'attestation tricolore de l'Ambassade de France en Italie reçue par l'établissement comme centre de passation des examens DELF-DALF. Le premier candidat, un jeune homme brillant, se lance avant même le rituel de l'épreuve à proposer un sujet d'examen : le roman social français (Zola) qu'il se propose de mettre en relation avec son programme de philosophie. Je me trouve dans l'obligation institutionnelle de lui rappeler que nous devons d'abord respecter les termes de notre contrat pour la certification, et donc tirer un sujet au sort parmi ceux qui ont été proposés par la commission centrale. Le candidat tire un sujet d'une banalité affligeante sur les méfaits du téléphone portable dans les lieux publics, sujet que nous traitons hâtivement dans des termes minimaux (les critères du barème sont largement respectés) pour en venir à ce qu'il voulait communiquer à cette sorte d'ambassadeur français auquel il pensait avoir affaire (ou retenait utile de donner à penser). »

C'est suite à l'accumulation de ce genre d'épisodes vécus ou relatés, que nous avons été amenée à repenser l'acte évaluatif dans sa contextualisation globale, pour en venir à avancer des arguments qui peuvent sembler désuets comme l'écoute de la parole, le respect de la voix (Kern 2008 : 31) que le candidat se crée dans une langue étrangère choisie, comme la reconnaissance des enseignants dans leur travail en amont de la certification et comme l'honneur rendu à ceux qui, de par le monde, s'engagent dans l'enseignement/apprentissage d'une langue qui ne leur est pas toujours immédiatement utilitaire.

Les sujets d'examen ne sont pas soumis au débat démocratique de l'opinion publique, comme les sujets du bac par exemple, car ils ne touchent pas les citoyens français mais les étrangers qui ont confiance en leurs institutions ; ils sont d'ailleurs inaccessibles aux enseignants de FLE eux-mêmes (et aux chercheurs) hors les quelques exemples

publiés sur le site, alors que les sujets des premiers DELF avant 2005 (réforme du DELF pour l'adapter au CECR) étaient du ressort de commissions bi-nationales qui pouvaient opérer une certaine médiation et adaptation des thèmes au public du pays, devenant des experts en certification pour leur pays (Dayez 1999). Le malaise déjà signalé par Patrick Chardenet (1999) s'accroît :

Les enseignants manifestent [...] une attitude inquiète face à des procédures d'évaluation qui semblent ne pas satisfaire à la connaissance du processus d'acquisition, mais qui s'imposent ou sont imposées comme incontournables. Face à l'ensemble de ces interactions langage/sujet/social, on aboutit à la description de processus d'évaluation qui fondent un ordre social et dont chacun dépend des instances qui gèrent cet ordre. (Chardenet 1999 : 27)

LE LIEN ENTRE VALIDITÉ ET DIMENSION SOCIALE

C'est la validité même de la certification qui est mise en défaut par le manque d'attention à la dimension sociale de la performance, car les certifications tendent à reprendre à leur compte de façon réductive les dichotomies parole/langue, performance/compétence : la consigne d'examen ne servirait qu'à déclencher une activité langagière pour faire produire de la matière verbale que l'examineur doit recueillir sans interférence puis traiter pour la dépouiller de son aspect contextuel, en déduire la compétence sous-jacente et la classer sur l'échelle qualitative par palier.

L'indication du CECR dans le chapitre sur l'évaluation – « pour rendre compte de la compétence, l'évaluation ne doit pas se focaliser sur une performance particulière, mais tendre plutôt à juger les compétences généralisables mises en évidence par cette performance » (CECR 2001 : 137) – est si souvent réaffirmée que les conditions de production de la performance sont désormais très sous-estimées dans le processus. On le voit en particulier lors de l'opération de calibrage inter-langues des productions orales organisée en 2008 au CIEP² qui servent désormais de référence dans les documents européens de formation des certificateurs (Noijons 2011).

Or la linguistique socio-interactionnelle nous fournit deux indications à prendre en compte (Mondada, Pekarek 2000 ; Pekarek-Doehler 2007) :

- les compétences langagières ne sont pas indépendantes des conditions socio-interactionnelles où elles sont mobilisées pour la production d'une performance,

- tout discours se définit comme co-construction où s'organise le lien social, d'où le caractère à la fois local et socio-historique de l'interaction.

En 2006, nous menions une expérience de comparaison entre les performances des étudiants selon les épreuves du DELF B2 et selon une épreuve créée en lien avec leur formation dans notre centre linguistique universitaire : à partir d'une version minimale du PEL, les étudiants étaient invités à présenter une biographie linguistique et des

2. http://www.ciep.fr/publi_evalcert/dvd-productions-orales-cecrl/

travaux rassemblés dans leur dossier, l'interaction portant sur une petite recherche qu'ils avaient effectuée selon leurs intérêts. L'étude de l'interaction verbale montre un réel décalage en faveur de ce type d'épreuve co-construite dans un triangle entre l'enseignant, le candidat et l'examineur ; en particulier, la compétence discursive en cohérence et cohésion y apparaissait de meilleure qualité, tandis que la parataxe domine les productions hésitantes de sujets en détresse communicative sur les thèmes non pertinents lors des épreuves officielles. On vérifiait ainsi que :

la compétence linguistique n'est pas indépendante de la connaissance thématique en particulier pour les niveaux avancés (Smith 1989 ; Zuengler et Bent 1991) ; que la complexité de la tâche améliore la qualité linguistique (Robinson 2001), mais surtout que l'orientation personnelle envers le *topic* (Nunan et Keobke 1995), sa représentation comme *topic* pertinent pour la situation communicative, influence notablement la qualité communicative. (Anquetil 2010-2012 : 9)

McNamara (2010) reprend la grille suivante d'analyse des différentes facettes de la validité dans la théorie unifiée de la validité de Messick (1989) que nous traduisons ainsi en 4 critères à prendre en compte.

VALIDITÉ	INTERPRÉTATION DU TEST	USAGE DU TEST
<i>Base des preuves</i>	1. Validité du <i>construct</i>	3. Pertinence contextuelle/Utilité
<i>Base des conséquences</i>	2. Valeurs impliquées	4. Conséquences sociales

La psychométrie se focalise avec ses instruments de mesure uniquement sur le premier critère : la validité du *construct* de l'examen. Ces exigences sont importantes mais à pondérer par les trois autres critères à prendre en compte pour une décision informée en matière d'évaluation certificative.

Pour reprendre notre anecdote initiale : les critères de validité psychométriques étaient garantis par le CIEP, mais les autres critères invalident l'épreuve sur le plan :

- des valeurs impliquées : la soumission au *topic* imposé par l'institution française et donc le non-respect des valeurs démocratiques et interculturelles portées par la Francophonie ; le déni des aspirations culturelles des apprenants de français (la francophonie comme humanisme) ;
- de la pertinence : le thème imposé à l'interaction, fruit d'une banalisation extrême de la notion de culture partagée provoque une baisse de la qualité de la production discursive à laquelle s'ajoute un conflit interculturel quant à la culture discursive, en particulier sur l'argumentation « à la française » (Anquetil 2007) ;
- des conséquences sociales : la dépréciation des institutions chargées de la diffusion de la langue française avec une retombée négative

sur les représentations de la France ; la marginalisation des enseignants de français de terrain évacués de la scène de l'évaluation des compétences qu'ils ont travaillé à construire ; la délégitimation des systèmes scolaires nationaux en matière de pouvoir de certification ; à terme, comme nous l'avons analysé dans l'exemple italien (Anquetil, à paraître), on va vers l'externalisation non seulement de la certification mais aussi de l'enseignement des langues au niveau scolaire, en particulier pour la deuxième langue étrangère qui tend à devenir matière optionnelle, puis périscolaire aux frais des familles.

Cependant les certifications gagnent du terrain, Huver et Springer (2011 : 141) parlent de « l'irrésistible ascension de l'évaluation institutionnelle certificative », témoin qu'elles répondent néanmoins à une demande que le système éducatif ne satisfait pas. Une fois que l'on a pris conscience de la puissance des tests (Shohamy 2001) et de leur lien avec des politiques linguistiques à forts enjeux sociaux (McNamara 2010), quelle voies trouver pour proposer des dispositifs répondant mieux aux exigences socio-culturelles et éthico-éducatives mises en relief ?

SORTIR « L'ÉVALUATION ALTERNATIVE » DE L'IMPASSE

« Qu'est-ce que l'évaluation alternative ? » s'interroge Eija Suomela-Salmi (2010). La définition s'inscrit indéniablement en négatif face aux tests à choix multiples, mais elle ouvre aussi sur un ensemble de pistes :

The term alternative assessment is by far the most common term used today in assessment literature dealing with qualitative assessing at various institutional levels. It is an umbrella term covering a wide range of different forms of assessment practices such as group-assessment, collaborative assessment, performance based assessment, task-based assessment and portfolio-assessment. (Suomela-Salmi, 2010 : 209)

Deux voies s'affirment dans ce secteur de la recherche critique en évaluation, pour ce qui est du segment final des processus formatifs, la validation des acquis : l'évaluation actionnelle et l'évaluation sur Portfolio. Qui peut porter institutionnellement ces approches, comment les insérer dans un secteur en tension ?

DES TESTS³ AUX CERTIFICATIONS PAR TÂCHES

Il revient au système éducatif de doter ses élèves de preuves validées de leurs compétences linguistiques, qu'ils puissent indiquer dans leur CV pour leur employabilité. Mais en quoi la certification ou le diplôme de langue se différencient-ils du test ? Le test devrait s'en tenir à sa valeur d'orientation à titre gratuit, alors que la certification s'intègre dans la durée, elle est l'aboutissement d'une formation longue, caractéristique de l'apprentissage des langues si difficilement compatible avec les temps de la sur-modernité. La certification ne procède pas par sondage d'habiletés ponctuelles mais atteste d'une compétence sur

3. Nous renvoyons pour les références de tous ces tests à : http://ressources.campusfrance.org/catalogues_recherche/diplomes/fr/1le_fr.pdf.

plusieurs activités langagières mises au service de tâches à réaliser en manipulant plusieurs genres textuels, elle est plus enracinée dans des pratiques socio-langagières contextuelles authentiques.

C'est le pari des certifications sur scénario, telles que le DCL⁴, qui propose un scénario professionnel, mais aussi le CLES (Maurer 2012) qui propose un scénario académique : par exemple, pour le niveau B1 un parcours de tâches sur le thème de la mobilité Erasmus. Le récent *Cahier des charges pour l'évaluation en langues à l'université* du programme européen GULT⁵ systématise cette approche en proposant un modèle d'examen pour les langues à l'université, orienté sur des tâches intégrées. Le scénario est à contextualiser selon le pays, la faculté, la langue, le niveau. GULT offre donc une matrice standard d'épreuves, mais dont les contenus sont adaptables pour aller au plus près des besoins des usagers. Ce modèle, proche de la certification linguistique adoptée dans les universités allemandes UNICert (Jaeger 2012), associe donc les enseignants à la définition des modalités de validation des acquis et encadre le système avec une démarche de label.

DES PISTES POUR LES DÉMARCHES ACTIONNELLES EN FRANÇAIS DEPUIS L'ÉTRANGER

La certification sur démarche actionnelle doit pouvoir prouver la capacité pragmatique d'un candidat à accomplir une tâche requérant l'usage du français. On atteste ainsi une compétence linguistique liée à un savoir agir en situation. Mais où se situe cette action, au service de qui ?

Doter les jeunes Français des cursus professionnels et technologiques de ces diplômes actionnels en langues étrangères, comme le propose le DCL pour les lycées technologiques et les IUT en collaboration avec le ministère de l'Éducation nationale et la CCIP, serait véritablement un atout de carrière, mais le DCL en FLE n'est pas disponible à l'étranger. Il est d'ailleurs conçu pour un apprentissage en immersion pour des candidats à l'intégration en France, ce qui ne correspond pas forcément à la situation du jeune apprenant à l'étranger.

Les Diplômes de Français Professionnel de la Chambre de Commerce et d'Industrie de Paris incluent des épreuves de médiation linguistico-culturelles (planifier une visite de sa ville pour des touristes francophones, synthétiser pour un interlocuteur francophone un texte professionnel écrit dans sa langue maternelle..) qui supposent un travail de collaboration entre l'institution centrale, garante de la valeur internationale du titre, et les agences locales, garantes de la pertinence sociale des épreuves qui ne situent plus dans le monolinguisme mais dans un cadre bilingue et biculturel. Ce modèle nous semble fructueux, il jouit d'ailleurs du soutien du Forum Francophone des Affaires, car les certifications génériques basent leur promotion sur leur valeur d'ouverture à la mobilité et à l'intégration en France, mais

4. <http://www.education.gouv.fr/cid55748/le-diplome-de-competence-en-langue-dcl.html>

5. <http://gult.ecml.at/GULT/tabid/2311/language/fr-FR/Default.aspx>

elles valorisent fort peu les situations d'usage de la langue française de l'apprenant en tant que médiateur inséré dans son propre milieu pour exercer des fonctions de correspondant commercial, d'opérateur du tourisme, en somme d'agent social qui ne soit pas *a priori* candidat à l'insertion en France. On sous-évalue d'ailleurs trop souvent l'importance de l'enseignement du français professionnel à l'étranger, au profit d'un enseignement générique dont les apprenants et décideurs hâtifs finissent par se détourner s'il n'aboutit pas à des capacités opératives traduisibles en termes d'emploi au service du pays d'origine.

LES APPROCHES PORTFOLIO ET LEUR VALIDATION

Malgré l'imposant soutien thématique et financier du Conseil de l'Europe autour du PEL, force est de constater que cette démarche n'a pas encore su trouver de véritable entrée sur le secteur de la validation des acquis et qu'elle s'essouffle dès que les fonds viennent à manquer (Little 2011).

La proposition que nous avons faite d'associer le Portfolio à l'épreuve du DELF commence à circuler (Anquetil 2012). Elle représenterait un lieu tangible de collaboration entre le système scolaire local et la certification, ce qui valorise l'ensemble des acteurs en jeu dans un dispositif intégré. Dans le Portfolio, en particulier dans le Dossier, l'apprenant de français a l'occasion d'exprimer sa voix dans sa nouvelle « langue adoptive » ; nous croyons qu'il s'agit là d'une voie privilégiée pour le français comme langue de culture.

Les approches Portfolio basées sur le PEL se sont trop souvent enlisées dans la manipulation fastidieuse de listes interminables de descripteurs abscons, demandant à l'apprenant de réaliser le travail même de l'évaluateur qui est de traduire la performance en termes de compétences graduées, travail que l'apprenant vit comme un déni de sa parole contextuelle et créatrice. Il faudrait donc revenir aux sources mêmes du concept de Portfolio comme mise en valeur de ses meilleurs travaux, de son cheminement vers la maturité (comme dans le domaine des compétences artistiques) et de ses performances. Le Portfolio deviendrait ainsi un recueil personnel d'expériences, de traces de projets de valeur, de parcours dans une langue-culture selon ses propres finalités et capacités, un recueil que le candidat, guidé par son enseignant, propose comme base de dialogue avec l'examineur, dans la co-construction d'une certification de valeur internationale.

Les approches qualitatives doivent cependant travailler à se soumettre à des critères de validation afin d'en tirer la force de conviction suffisante pour s'introduire dans le champ des certifications. Il n'est pas dit que la standardisation soit la seule voie à la validité mais il reste à démontrer le contraire. De nombreux documents du Conseil de l'Europe mettent à disposition des concepteurs de certification des outils

libres d'accès, il s'agit donc de s'en saisir pour construire peu à peu une légitimation sociale (Lenz 2010).

Les outils mathématiques proposés dans le *Manuel pour relier les examens de langue au CECR* (North et al., 2009) sont de nature à décourager les non-spécialistes et sont quelquefois utilisés par les adeptes de la technicité comme argument d'autorité « scientifique » pour délégitimer toute approche holistique, mais ils concernent essentiellement les tests et ne sont aucunement obligatoires pour la validation des certifications. Il n'en reste pas moins que les procédures de spécification, exposées de façon plus accessible dans le manuel de synthèse et de formation (Noijons et al., 2011), constituent une étape indispensable dans un processus de transparence des outils d'évaluation authentique pouvant les mener à une labellisation de qualité. Les fiches d'analyse de spécificité des épreuves en annexe du *Manuel* (op. cit., 138-180) offrent des outils de standardisation, non pas du contenu des épreuves mais de leur analyse et de leur présentation en vue d'un contrôle de qualité externe ou concerté. On consultera aussi avec profit l'étude préliminaire du *Manuel* au sujet des méthodes d'analyse qualitatives (Banerjee 2004), ainsi que les indications pragmatiques dans les travaux CELV pour *Relier l'évaluation au PEL* (Hodel 2010).

C onclusion : pour un système de certification diversifié mais unifié dans l'esprit de la Francophonie

La place du français dans le monde est si fragilisée aujourd'hui (Maurais et al., 2008) qu'il serait néfaste que ses promoteurs dispersent leurs énergies en querelles d'école quant à la certification idéale. Il est bien évident que chacune a ses atouts et que les concurrences sont aussi signe de vitalité du domaine, elles incitent à se remettre en question et à perfectionner les outils au plus proche des besoins. En conclusion nous souhaiterions proposer une piste pour un système intégré du moins pour ce qui est de l'offre de certification du français à l'étranger. D'un point de vue administratif cette offre est déjà regroupée auprès des Instituts français, Alliances françaises, services de coopération culturelle universitaire et scolaire des Ambassades qui servent de relais fiables et disposent d'un réseau hors pair d'une grande compétence par l'accumulation de critères de qualité qui ne se construisent que dans le temps : la connaissance du terrain, la construction d'un réseau d'interlocuteurs locaux, les compétences bi-plurilingues, la compréhension des réalités socio-culturelles locales. Ce sont là les agents de l'indispensable médiation entre les agences de certification centrales,

essentiellement parisiennes, et les instances d'enseignement du français de terrain (Alliances et Instituts, systèmes éducatifs locaux). Il revient à ces dernières de faire valoir les exigences de contextualisation de leurs « porteurs d'intérêt » locaux, en particulier pour les pays où le français est langue seconde.

Conjuguer dans le DELF des épreuves centralisées et des épreuves contextualisées serait l'occasion de diffuser une véritable culture de l'évaluation par co-construction avec les acteurs locaux, une occasion pour le candidat de *faire reconnaître sa voix en français* comme nouveau citoyen de la Francophonie, agent du dialogue des cultures. Ce diplôme renouvelé en concertation avec les agences existantes pourrait faire l'objet d'un projet des institutions francophones. Dominique Wolton ne déclarait-il pas au sein de la Cellule de Réflexion Stratégique de la Francophonie :

« Afin de valoriser le français comme langue mondiale, il faut aussi rationaliser et unifier les différents diplômes utilisés (DALF, DELF, DILF). Avoir ainsi une simplicité et une visibilité [...]. Mais il faut aussi "défranciser" ces diplômes. En faire des diplômes "francophones" susceptibles d'être accordés par des États membres de l'OIF, à partir d'une *charte commune*. Cette prise en charge par tous de l'avenir de la langue française correspond bien aux enjeux de la Francophonie. » (Wolton 2008 : 87-88.)

Bibliographie

Pour une certification contextualisée du français dans le monde : intégrer les approches actionnelles et les approches portfolio

- ANQUETIL M. (à paraître), «Reconnaître et certifier les compétences dans une perspective francophone : comment conjuguer décentralisation, qualification et valorisation ?», dans *Actes du II^e Forum Mondial des centres universitaires de langues HERACLES*, Besançon 25-26 mai 2011.
- ANQUETIL M. (2012), «DELFB2 et Portfolio, une expérience de coopération pour une certification en contexte» dans *Language Learning in Higher Education, Journal of the European Confederation of Language Centres in Higher Education*, n° 1, Berlin, De Gruyter.
- ANQUETIL M. (2010-2012), «L'interaction en situation de certification de FLE, un regard critique» en ligne dans *Colloque international «Spécificités et diversité des interactions didactiques : disciplines, finalités, contextes»*, Université de Lyon - ICAR - CNRS - INRP, Lyon 24-26 juin 2010. http://hal.archives-ouvertes.fr/hal-00534429/édition_papier : dans Rivière V. (dir.) *Spécificités et diversités des interactions didactiques*, Paris, Éditions Riveneuve.
- ANQUETIL M. (2007), «Communication interculturelle ou mimétisme obligé : l'interaction en situation d'examen de langue standardisé DELF» dans *Actes du colloque international «Les enjeux de la communication interculturelle, compétence linguistique, compétence pragmatique, valeurs culturelles»*, IUFM Université Paul Valéry, Montpellier III, 5-7 juillet 2007, Montpellier. Actes en ligne : http://www.msh-m.fr/article.php3?id_article=507
- BANERJEE J. (2004), Reference Supplement to the Preliminary Pilot version of the Manual for *Relating Language examinations to the Common European Framework of Reference for Languages: learning, teaching, assessment Section D: Qualitative Analysis Methods*, Strasbourg, Conseil de l'Europe. http://www.coe.int/t/dg4/linguistic/Manuel1_FR.asp
- CHARDENET P. (1999), *De l'activité évaluative à l'acte d'évaluation*, Paris, L'Harmattan.
- CONSEIL DE L'EUROPE, (2005), *Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer*. Paris, Didier.
- DAYEZ Y. (1999), *Guide du concepteur de sujet DELF-DALF*, Paris, Édition Didier.
- HODEL H.-P. (2010), *Relier l'évaluation au PEL/CECR*. http://archive.ecml.at/mtp2/Elp_tt/Results/PagEF/f10.html
- HUYER E. et SPRINGER C. (2011), *L'évaluation en langues*, Paris, Didier.
- JAEGER C. (2012), «Le système de certification de langues étrangères UNICERT®», dans *Repères-DORIF autour du français : langues, cultures, plurilinguisme*, n° 1, «Le français dans le contexte plurilingue des Centres linguistiques universitaires italiens» (coord. Anquetil M.), www.dorif.it
- KERN R. et LIDDICOAT A. J. (2008), «De l'apprenant au locuteur/acteur», dans Zarate G., Levy D., Kramsch C., *Précis du plurilinguisme et du pluriculturalisme*, Paris, Éditions des archives contemporaines.
- LENZ P. et BERTHELE R., (2010), *Prise en compte des compétences plurilingue et interculturelle dans l'évaluation*, Strasbourg, Conseil de l'Europe, http://www.coe.int/t/dg4/linguistic/Source/Source2010_ForumGeneva/Assessment2010_Lenz_FR.pdf
- LITTLE D., GOULLIER F. et HUGHES G. (2011), *The European Language Portfolio: The Story So Far (1991-2011)*, Strasbourg, Conseil de l'Europe, http://www.coe.int/t/dg4/education/elp/elp-reg/Source/Publications/ELP_StorySoFar_July2011_Final_EN.pdf
- MCMANARA T. et ROEVER C. (2006), *Language Testing : The Social Dimension*. Oxford, Blackwell Publishing.

- MCNAMARA T. (2010), "The use of language Tests in the service of policy: issues of validity" dans *Revue française de Linguistique Appliquée*. Volume XV-1, p. 7-23.
- MAURAS J., DUMONT P., KLINCKENBERG J.-M., MAURER B. et CHARDENET P. (2008) *L'avenir du français*, Paris, Éditions des archives contemporaines.
- MAURER M.-T. (2012), « Le CLES, certification des universités françaises et le programme GULT, Cahier des charges pour l'évaluation des langues du CELV » dans *Repères-DORIF autour du français : langues, cultures, plurilinguisme*, n° 1, « Le français dans le contexte plurilingue des Centres linguistiques universitaires italiens » (coord. Anquetil M.). www.dorif.it
- MESSICK S. (1989) « Validity » dans Linn R. L., *Educational Measurement*, New York : American Council on Education et Macmillan, p. 13-103.
- MONDADA L. et PEKAREK DOEHLER S. (2000), « Interaction sociale et cognition située : quels modèles pour la recherche sur l'acquisition des langues ? » dans *Acquisition et Interaction en Langue Étrangère*, n° 12. <http://aile.revues.org>
- NOIJONS J., BÉREŠOVÁ J., BRETON G. et SZABÓ G. (2011), *Relier les examens de langues au Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer (CECR)*, Les points essentiels du Manuel, CELV, Conseil de l'Europe. <http://www.ecml.at/tabid/277/PublicationID/67/Default.aspx>
- NORTH B. et al. (2009), *Relier les examens de langues au Cadre européen commun de référence pour les langues : Apprendre, enseigner, évaluer (CECRL) Un manuel*, Strasbourg, Conseil de l'Europe. http://www.coe.int/t/dg4/linguistic/Manuel1_FR.asp#P89_7669
- NUNAN D. et KEOBKE K. (1995), "Task difficulty from the Learners Perspective" dans *Hong Kong Papers in Linguistics and Language Teaching*, n° 18, p. 1-12.
- PEKAREK DOEHLER S. (2007), « Démythifier les compétences : vers une pratique écologique d'évaluation » dans *Actes du colloque international CERC, IRFFLE « Construction du sens et acquisition de la signification linguistique dans l'interaction »*, Nantes, 22-24 novembre 2007. <http://gramm-fle.ulb.ac.be/fichiers/colloques/Nantes2007/PEKAREKDOEHLER.pdf>
- ROBINSON P. (2001), "Task Complexity, Task Difficulty and Task Production : Exploring Interactions in a Componential Framework" dans *Applied Linguistics*, 22/1. Oxford University Press, p. 22-57.
- SHOHAMY E. (2001), *The power of Tests, a critical perspective on the uses of language tests*. Essex, Pearson Education.
- SMITH J. (1989), "Topic and variation in ITA oral proficiency : Speak and field-specific tests" dans *English for Specific Purposes*, volume 8, Issue 2, p. 27-57.
- SUOMELA-SALMI E. (2010), "What is alternative assessment ?" dans Dervin F., Suomela-Salmi E., *New Approaches to Assessing Language and (Inter-)Cultural Competences in Higher Education. Nouvelles approches de l'évaluation des compétences langagières et (inter-)culturelles dans l'enseignement supérieur*, Language Testing and Evaluation volume 19, Frankfurt, Peter Lang.
- ZUENGLER J. et BENT B. (1991) "Relative Knowledge of Content Domain : An Influence on Native-Non.native Conversations", dans *Applied Linguistics*, Vol. 12, n° 4, Oxford University Press, p. 397-415.
- WOLTON D. (2008). *Francophonie et migrations internationales*. Cellule de réflexion stratégique de la francophonie. www.francophonie.org

P *luralités
culturelles et
linguistiques :
déplacer les
problématiques*

DENISE LUSSIER

FRED DERVIN

MARTINE MARQUILLÓ LARRUY

LAURENT GAJO, JEAN-MARC LUSCHER

ET CECILIA SERRA

A *pproche systémique pour évaluer la compétence de communication interculturelle*

DENISE LUSSIER

UNIVERSITÉ MCGILL ET UNIVERSITÉ DU QUÉBEC
À MONTRÉAL, CANADA

I *ntroduction*

Le présent article est consacré plus particulièrement à l'évaluation de la « compétence de communication interculturelle » (CCI) puisque, de conception plus récente, elle comprend des zones obscures dans les pratiques d'interventions pédagogiques et, par conséquent, dans l'évaluation de l'ensemble de ses composantes.

Dans un premier temps, il faut se rappeler qu'en éducation des langues comme dans les autres domaines, de nouvelles recherches viennent enrichir les anciennes et de nouvelles théories se développent avec les années : l'objectif de toute science étant en effet de faire avancer l'état des connaissances et des pratiques, de déceler et d'éliminer les paradoxes ou les ambiguïtés résultant d'une représentation inadéquate de la réalité tout en analysant les effets positifs et négatifs de tout changement. C'est ainsi qu'une définition plus fine de certains concepts mène, de façon constante, à de nouvelles théories, à de nouvelles lignes de pensée et à des transferts de connaissances. Aussi devient-il difficile pour les divers intervenants de travailler en vase clos. Une approche systémique et cohérente s'impose entre « théorie » et « pratique ».

En ce qui concerne le développement de la compétence langagière, nous pouvons référer au premier modèle canadien de Canale et Swain (1980) et au modèle américain de Bachman (1999) considérés comme les deux pierres d'assise du Cadre européen commun de référence pour les langues (2001). Quant au développement de la compétence de communication interculturelle, trois modèles retiennent l'attention

(Byram 1997 ; Lussier 1997 ; Landis, Bennett et Bennett 2004). Par contre, un seul cadre commun de référence a été validé empiriquement auprès de plus de 1 500 jeunes adultes (Lussier 2009a ; 2009b ; 2009c). Trois niveaux de réflexion en résultent : l'obligation de bien cerner les enjeux de la CCI (Partie 1) pour savoir quoi enseigner (Partie 2) et quoi évaluer (Partie 3).

Partie 1 : Enjeux du développement de la « compétence de communication interculturelle »

Lorsque nous abordons le développement de la compétence de communication interculturelle, plusieurs interrogations d'ordre conceptuel viennent à l'esprit : Comment définir la « culture » en éducation des langues ? Qu'est-ce qui différencie « culture » et « interculture » ? Quelles valeurs sont rattachées à ces deux derniers concepts ?

CULTURE ET INTERCULTURE : UNE QUESTION DE VALEURS

L'UNESCO (2001 : 49) définit la culture « comme l'ensemble des traits distinctifs spirituels et matériels, intellectuels et affectifs qui caractérisent une société ou un groupe social ». Elle se manifeste par les interactions qui existent entre les individus. Il semble donc que, pour mieux comprendre une autre culture en tant qu'autre culture, il importe de prendre en considération la découverte de la spécificité culturelle de l'Autre. De ce fait, il devient de plus en plus rare d'aborder la dimension culturelle sans aborder l'interculturalisme puisque l'interculturalisme englobe tous les aspects de la culture et que l'inverse n'est pas nécessairement valable (Lussier et Lebrun 2009).

De plus, avec la mobilité des populations et la mondialisation des communications, il se crée de nouveaux défis pour les sociétés d'accueil. Vouloir ou devoir intégrer de nouveaux arrivants de façon aussi harmonieuse que possible représente un défi d'intercompréhension des groupes en contact. Dans les deux cas, il se crée des frictions, des incompréhensions, des mésententes. Une question surgit : comment peut-on apprendre de part et d'autre à s'ouvrir aux autres cultures et à la diversité pour susciter le respect de l'autre et un mieux vivre ensemble ? Nous touchons au développement d'une *compétence de communication interculturelle* : le but ultime de son enseignement étant de former de nouvelles générations de jeunes de tout groupe ethnique qui développent une meilleure conscientisation de leur propre culture et de celle des autres cultures basée sur le respect

mutuel (Lussier 2009b). Il s'agit donc d'en arriver à comprendre d'autres modes de pensée, de croyances et de comportements. Les enjeux sont importants, d'où la nécessité de se donner un référentiel commun et une compréhension univoque de ses composantes. Notre vision de l'éducation aux langues tente de s'adapter.

UN CADRE CONCEPTUEL POUR UNE COMPRÉHENSION UNIVOQUE

Un cadre conceptuel a pour but de mettre en lien les dimensions qui constituent, délimitent et précisent les concepts liés au domaine étudié. Sa mise en place s'avère essentielle pour assurer une compréhension univoque des enjeux poursuivis, des interventions pédagogiques à privilégier et des modalités efficaces pour les évaluer. Tout cadre conceptuel sert donc de référentiel aux divers intervenants et vise à faciliter une compréhension commune même s'ils œuvrent dans des milieux différents. Il sert également de cadre théorique aux chercheurs et permet d'assurer le transfert des données dans des perspectives interreliées. Il limite la subjectivité de l'interprétation des résultats des apprentissages. De ce fait, il répond aux valeurs de rigueur, transparence et équité : trois concepts essentiels de l'évaluation des apprentissages. Dans une telle perspective, l'enseignement, l'apprentissage et l'évaluation demeurent complémentaires les uns aux autres.

En ce qui concerne la CCI, il semble que, très souvent, l'éducation à la diversité culturelle se limite à la compréhension des éléments superficiels des représentations culturelles telles que les stéréotypes, les artefacts et les aspects folkloriques (Tomalin et Stempleski 1993 ; Lussier 1997). Pour en arriver à dépasser ce stade, il devient essentiel de se doter d'une approche systémique, ce qui permet d'effectuer l'étude globale d'une réalité conceptuelle en mettant l'accent sur les interrelations et interdépendances des éléments entre eux (Legendre 2005). Dans le cas de la CCI, l'approche comprend la mise en place de trois volets : le volet « langue », rattaché au développement des composantes linguistique, sociolinguistique et pragmatique de la compétence langagière ; le volet « interculture », relié aux interactions sociales entre les personnes d'affinités différentes ; le volet « pensée » en lien avec l'apport des identités culturelles et des croyances autres que celles de sa propre culture (Rubinfeld, Clément et Lussier 2006).

Tout comme il existe un *Cadre commun européen de référence pour les langues* (Conseil de l'Europe, 2001), il existe un *Cadre commun de référence pour l'enseignement de la CCI* (Lussier 1997, 2009, 2011). Il a servi de référentiel pour deux études réalisées par le Centre européen des langues vivantes, portant sur *Les représentations des Autres et des autres culturels dans le contexte de la formation initiale et continue des enseignants* en contexte européen (Lussier, Auger, Urbanicova et al., 2003) et *les Lignes directrices pour évaluer la compétence en communication interculturelle* (Lussier, Golubina, Ivanus et al., 2007). S'ouvrir

aux autres cultures et devenir un intermédiaire culturel deviennent des incontournables dans des sociétés multiculturelles (Rubinfeld, Clément, Vinogard et Lussier 2007). L'éducation aux langues devient un moment privilégié par son contact avec une autre culture. De plus en plus, il semble évident que l'enseignement, l'apprentissage et l'évaluation ne peuvent porter seulement sur l'acquisition des «savoirs» et des «savoir-faire», mais également sur le développement des «savoir-être».

Le domaine des connaissances ou des «savoirs» relève du domaine cognitif et se regroupe selon trois approches (Lussier 2009b) : l'approche humaniste, c'est-à-dire celle de la culture avec un grand «C», la culture cultivée, cet héritage, dit de civilisation, qui constitue la mémoire collective d'un peuple ; l'approche anthropologique, liée à la formation des relations avec les membres d'autres sociétés et de cultures différentes ; l'approche sociologique, liée au contexte socio-culturel et à la présentation de faits, de statistiques et de données sociales, économiques et politiques.

Le domaine des savoir-faire relève de l'application des connaissances et de l'utilisation de la langue en situation de communication. Les apprenants développent des habiletés pour «fonctionner», linguistiquement parlant, dans la langue cible. Ils développent avec l'expérience de pratiques pluriculturelles et plurilingues des habiletés pour «savoir-ajuster» adéquatement leurs comportements à un environnement social et culturel afin de «savoir-interagir» efficacement. Ils en arrivent à décoder des messages porteurs de diverses interprétations, ce qui leur permet d'«intégrer» les concepts de langue et interculture pour, finalement, devenir aptes à «négocier» adéquatement des situations d'incompréhension, de malentendus ou même conflictuelles.

Le domaine des savoir-être porte sur le développement des attitudes et des représentations culturelles, lesquelles façonnent notre vision du monde. Trois types de compétence sont à considérer : la «conscientisation culturelle», liée à la sensibilisation de l'existence des autres cultures et souvent limitée à un savoir-comprendre ; «l'appropriation critique» liée au «savoir-accepter» et «savoir-interpréter» les modes de pensée, les croyances et les valeurs associées à sa propre culture et aux autres cultures ; la «compétence transculturelle» qui se définit par l'intégration et le respect des valeurs des autres découlant de la coexistence de divers groupes ethniques et de cultures en présence dans une même société ou dans des sociétés distinctes tout en prônant l'enrichissement identitaire de chacune des cultures en contact (Lussier 2011). Plus spécifiquement, il s'agit de centrer nos actions sur une conception de la CCI attentive aux situations d'incompréhension ou conflictuelles, centrée sur des interventions pédagogiques réalistes visant à permettre aux éducateurs en langues d'entrevoir leur rôle en tant que facilitateurs et médiateurs dans le développement de la CCI.

Le cadre conceptuel commun de référence en CCI, tel que décrit précédemment, a fait l'objet d'une étude empirique et a été validé de 2004 à 2010 auprès de plus de 1500 jeunes adultes (Lussier, Auger, Clément et Lebrun-Brossard 2004-2010). De nouvelles analyses, réalisées en 2011, ont permis la modélisation du cadre conceptuel, soit de concevoir le processus dans le développement de la CCI. Les analyses montrent qu'un individu, dans ses relations avec les autres, peut se situer à quatre niveaux distincts : l'antipathie, l'indifférence, la sympathie et l'empathie. Le premier niveau se veut un stade négatif puisque l'*antipathie* repose sur l'intolérance envers les autres cultures et le rejet de l'Autre. L'individu est actif. Il manifeste, mais en rétroaction. Il est sur la défensive et soumis à une perturbation visant à provoquer une réaction de sens contraire à l'empathie. Il rejette les autres cultures et se situe dans le déni social (Landis, Bennett et Bennett (2004), c'est-à-dire le refus de reconnaître une réalité dont la perception est possiblement traumatisante. Le deuxième niveau porte sur l'*indifférence*. Celle-ci repose sur la tolérance et non plus sur l'intolérance des autres cultures. L'individu accepte les autres cultures et s'abstient de les renier. Il demeure passif et évite les conflits culturels. De fait, il minimise l'importance et la présence des autres cultures (Landis, Bennett et Bennett 2004). Le troisième niveau, basé sur la *sympathie*, est l'expression chaleureuse et spontanée de sentiments et de compassion envers les autres cultures qui résulte de la neutralisation des émotions négatives et d'une disposition positive dans les relations interpersonnelles (Baumgratz-Gangl 1993). L'individu est actif. Il accepte les autres cultures et s'adapte à des croyances et des comportements autres que les siens. Le dernier niveau, l'*empathie*, relève de la faculté de s'identifier à quelqu'un, de ressentir ce qu'il ressent. L'individu est proactif. Il se met à la place de l'autre, essaie de le comprendre dans ses valeurs, ses croyances et ses comportements. Il se donne le rôle de médiateur culturel dans des situations d'incompréhension ou conflictuelles.

Partie 2 : Enjeux en enseignement/ apprentissage

Comme on a pu le constater, un cadre de référence est généralement neutre, sans attache contextuelle, sans affinités politiques et sans être sujet à plusieurs interprétations. Par contre, il existe différentes écoles de pensée et plusieurs approches dans la façon d'aborder l'enseignement et l'apprentissage de la CCI. Il revient donc aux intervenants de l'éducation en langues de faire des choix. Il est possible de greffer la « compétence interculturelle » à la « compétence langagière ». L'inverse

est aussi possible. Toutefois, quatre types de questionnement s'imposent : Quel cadre conceptuel adopté devrait répondre aux attentes de la société ? Que doit comprendre un programme d'études intégrant « langue » et « interculture » ? Comment doit-on définir les objectifs d'apprentissage, le « quoi enseigner » ? Quelle approche et quels contenus de l'enseignement doit-on privilégier ? Une telle réflexion permet d'illustrer les dimensions de l'enseignement de la CCI. La planification de l'enseignement et de l'apprentissage doit reposer sur une approche systémique, telle qu'illustrée dans le tableau 1.

Tableau 1 : Exemple de planification de l'enseignement / apprentissage de la CCI

VOLETS	PROGRAMMES D'ÉTUDES : – ORIENTATIONS – OBJECTIFS D'APPRENTISSAGE	ENSEIGNEMENT/ CONTENUS	TÂCHES LANGAGIÈRES
Compétence langagière*			
a) linguistique	Utilise adéquatement les structures syntaxiques simples.	Présent des verbes	Compréhension d'une narration
b) sociolinguistique	Utilise les formes de politesse élémentaires.	Emploi des mots clés (merci...)	Jeu de rôle en situations de communication
c) pragmatique	Utilise le discours argumentatif.	Justifie son opinion.	Rédaction d'une lettre...
Compétence interculturelle**			
a) Savoirs	Identifie les symboles, rituels...	Sens d'un hymne	
b) Savoir-faire	Utilise les conventions sociales...	Règles d'hospitalité	
c) Savoir-être	Reconnaît les différences entre deux cultures à partir d'un texte...	Distingue le « nous » du « eux »	

* La compétence langagière : *Cadre européen commun de référence pour les langues* (Conseil de l'Europe, Didier, 2001)

** La compétence interculturelle : *Enseigner et évaluer la « compétence de communication interculturelle », enjeux et complémentarité* (Lussier 2009b)

Dans une approche axée sur le développement de la CCI, rien ne peut être laissé au hasard ou se faire de façon aléatoire. La planification de l'enseignement et de l'apprentissage exige l'intégration des volets de la compétence langagière et de la compétence de communication interculturelle, le choix des orientations qui sont privilégiées, la définition d'objectifs d'apprentissage en termes de compétences et de sous-compétences, la spécification des contenus de même que les tâches que devront réaliser les apprenants.

P artie 3 : Enjeux de l'évaluation de la CCI

Nous devons aborder le domaine de l'évaluation de la CCI avec circonspection puisqu'il existe présentement une large part de subjectivité dans son enseignement et plus encore dans son évaluation. La culture et l'inter-culture sont des concepts et des facteurs de médiation qui ne sont pas compris par tous de la même façon. Il faut aussi comprendre que la planification de l'évaluation de la CCI demande deux modalités d'intervention 1) l'évaluation des apprentissages qui relève d'un programme de formation, 2) l'évaluation institutionnelle, plus globale, représentative de l'ensemble d'une population spécifique et liée à des profils de compétence. Ces deux modalités d'intervention se complètent. Par contre, leurs finalités diffèrent.

L'ÉVALUATION DES APPRENTISSAGES D'UN PROGRAMME DE FORMATION

L'évaluation des apprentissages liée à des programmes de formation est généralement sous la responsabilité de l'enseignant. Il lui revient de suivre la progression des apprentissages, de porter un jugement sur le degré de maîtrise des objectifs d'apprentissage, et éventuellement de rendre une décision d'échec ou de réussite au terme des apprentissages. Trois questions s'imposent : Que signifie évaluer les « savoirs » ? Que signifie évaluer les « savoir-faire » ? Que signifie évaluer les « savoir-être » ? Trois domaines d'évaluation inter-reliés mais de nature différente puisqu'il est impossible d'évaluer des savoirs, des habiletés et des comportements selon les mêmes caractéristiques. Il faut recourir à des stratégies d'évaluation alternatives selon le champ d'évaluation.

En ce qui concerne *l'évaluation des savoirs*, les méthodologies d'évaluation indirectes, effectuées par le biais de contrôles dits « objectifs » et relevant des tests traditionnels papier-crayon, sont toujours valides. Les apprenants répondent à des questions à choix multiples, fournissent des réponses courtes à des questions de compréhension, associent des informations ou les regroupent par paires ; l'objectif étant de mesurer l'acquisition de faits culturels. Ils identifient des éléments culturels semblables ou différents, réorganisent, regroupent et comparent différents types de caractéristiques ou d'expressions (physiques, mentales, morales, affectives) à partir de textes écrits ou oraux. De plus, la recherche a pu démontrer que trois domaines de connaissances doivent néanmoins être pris en compte : les connaissances liées à la mémoire collective en termes de culture et civilisation (approche humaniste) ; les éléments de connaissance liés à la diversité des modes de vie des différentes cultures (approche anthropologique) ; enfin, les

éléments en lien avec les contextes socioculturels des sociétés ciblées (approche sociologique).

L'évaluation des savoir-faire porte sur la réalisation de tâches en situations réelles de communication présentant des contextes interculturels variés. Elle touche les aspects linguistiques de la CCI visant à montrer dans quelle mesure les apprenants sont capables de fonctionner et d'interagir dans la langue cible, de s'ajuster à divers environnements sociaux et interculturels, et d'intégrer les expériences acquises pour utiliser efficacement leur compétence langagière en tant que locuteurs et médiateurs interculturels. Le tout n'est pas seulement de savoir fonctionner dans une langue, mais de savoir adapter son discours, interagir, intégrer, interpréter et négocier les interactions sociales dans divers contextes interculturels. Cette modalité de l'évaluation est fortement utilisée par les enseignants puisque l'ensemble des programmes d'études développés avec l'approche communicative porte désormais sur le développement de compétences, ce qui se traduit par la maîtrise de fonctions langagières et de types de discours en situations de communication.

L'évaluation des savoir-être, tributaire des comportements affectifs et psychologiques, est très souvent centrée sur le développement d'une « conscientisation culturelle », axée sur l'acquisition de connaissances et l'expression des différences entre les cultures. De nos jours, la recherche reconnaît l'importance de ce premier niveau de sensibilisation aux savoir-être, mais considère primordial que les apprenants puissent dépasser ce stade pour développer une compétence « critique » et tendre vers une compétence « transculturelle » qui leur permet de reconsidérer leurs propres valeurs et de devenir des intermédiaires interculturels dans des situations de tension, d'incompréhension ou conflictuelles. L'enseignant partage l'évaluation avec l'apprenant. Il devient un observateur du processus de développement de la CCI et non pas d'un produit. L'évaluation se veut formative et privilégie l'auto-évaluation, l'inventaire des attitudes, les fiches des comportements, les études de cas, les enquêtes, les grilles d'observation, le journal de bord, l'inventaire des centres d'intérêt, l'expression écrite des perceptions de l'autre, le carnet de route et le dossier d'apprentissage (portfolio). L'administration de tests normalisés, l'analyse des résultats en termes de moyennes et la comparaison des groupes par rapport à une norme perdent leur signification. Aussi, une évaluation de type sommatif apparaît moins pertinente, puisque l'objectif n'est pas de certifier par une note le niveau de savoir-être atteint par l'apprenant. L'appréciation finale doit reposer sur l'appréciation du processus et fournir aux apprenants des informations précises sur les progrès accomplis en CCI et une conscientisation des faiblesses perçues.

L'ÉVALUATION INSTITUTIONNELLE ET LES PROFILS DE COMPÉTENCE

L'évaluation institutionnelle découle de l'évaluation dite « systémique » telle que définie antérieurement. Elle relève des institutions mandatées à cet effet. Elle couvre également l'évaluation des savoirs, des savoir-faire et possiblement des savoir-être, mais dans le respect des apprentissages. Elle a pour but de rendre compte de la qualité des apprentissages de l'ensemble d'une population spécifique. Elle doit se faire à partir de données fiables et significatives tirées d'une échelle de mesure, en plus de fournir aux apprenants des normes qui leur permettent de se comparer à d'autres populations et d'évaluer leur compétence, langagière et interculturelle, en fonction de profils de compétences validés auprès de populations similaires.

Les échelles de mesure précisent dans un langage accessible ce que tout apprenant devrait savoir ou réaliser en termes de tâches à des moments précis de l'apprentissage ou au terme de celui-ci. Ces échelles comprennent des descriptions de niveaux qui définissent de façon globale les compétences attendues au-delà des programmes de formation. De ce fait, elles procurent un référentiel commun entre divers intervenants et divers contextes, ce qui facilite l'établissement de normes. Les profils de compétence regroupent généralement trois éléments ; une définition, une description générale et des indicateurs/descripteurs de performance reflétant les tâches spécifiques aux divers niveaux de complexité de l'échelle de mesure.

Nous pouvons considérer diverses échelles de référence, telles que celle du *Cadre européen commun de référence pour les langues* (Conseil de l'Europe 2001). Elle est élaborée en fonction de trois niveaux (élémentaire, indépendant et expérimenté) pouvant se redistribuer en six échelons. Quant aux échelles développées au Canada (Citoyenneté et Immigration Canada 2006 ; 2012) et au Québec (Ministère de l'Immigration et Communautés Culturelles (2000 ; 2010), elles comprennent douze niveaux de difficulté croissante validés auprès des populations ciblées. On retrouve trois stades (débutant, intermédiaire et avancé) de quatre niveaux chacun, et ce afin de satisfaire le cursus canadien, comprenant de onze à treize années de scolarité.

Nous présentons un exemple de profils de compétences de communication interculturelle pour le domaine des « savoir-être », tiré des *Lignes directrices pour l'évaluation de la compétence de communication interculturelle* (Lussier, Golubina, Ivanus et al., 2007). Le document contient aussi des définitions générales et des indicateurs de performance pour les savoirs, savoir-faire et savoir-être en plus de présenter des exemples de tâches d'évaluation pour chacune des dimensions et des niveaux de complexité.

Tableau 2 : Profils de compétence pour les « savoirs-être »

PROFIL MINIMUM	PROFIL MÉDIUM	PROFIL ÉLEVÉ
<p>Niveau de compréhension</p> <ul style="list-style-type: none"> • Est tolérant • A une conscientisation culturelle • Est sensible aux autres cultures • Identifie de nouvelles croyances, attitudes et valeurs 	<p>Niveau d'acceptation et interprétation</p> <ul style="list-style-type: none"> • Démonstre de la sympathie envers... • Démonstre une appropriation critique • Démonstre une ouverture à d'autres cultures • Démonstre des réactions positives à d'autres croyances, attitudes et valeurs 	<p>Niveau d'intégration et internalisation</p> <ul style="list-style-type: none"> • Démonstre de l'empathie • Fait des transferts entre cultures • Démonstre un sens de l'altérité • Agit avec respect envers les autres cultures (croyances, attitudes et valeurs)

C onclusion

Nous savons que l'enseignement/apprentissage d'une langue incarne la présence de l'autre culture et le contact avec l'altérité dans le développement des représentations culturelles et conséquemment dans le développement de la CCI. Pour en connaître la nature même, nous avons questionné des enseignants de divers pays (Lussier, Auger, Urbanicova *et al.*, 2003) puisqu'ils sont des acteurs sociaux. Nous avons constaté que l'interaction des enseignants auprès des apprenants est fortement liée à leur formation et leurs expériences. Deux considérations méritent notre attention : l'ouverture à l'enseignement de la culture et au développement de la CCI semble acquise, par contre, celle-ci ne peut exprimer, dans une démarche pédagogique et une logique cohérentes, les moyens qu'il faut privilégier pour couvrir toutes les dimensions des concepts en interaction (p. 221). Pour les enseignants, les interrogations ne portent plus sur le « Pourquoi enseigner le développement d'une compétence de communication interculturelle » ou sur le « Quoi enseigner » mais bien sur le « Comment l'enseigner ». Ils perçoivent la « langue » comme intimement liée au développement des représentations culturelles et aux interactions sociales entre les individus. Si nous voulons intégrer le développement d'une compétence de la communication interculturelle au développement de la compétence langagière, nous devons donc repenser la formation initiale et continue des enseignants afin de leur fournir des pistes d'interventions pédagogiques adéquates en salle de classe.

En ce qui concerne l'évaluation de la CCI, nous savons que celle-ci ne sera jamais exempte de subjectivité. Aussi, il importe d'assurer son encadrement à partir d'une compréhension commune de la CCI et de ses composantes en fournissant un cadre commun de référence. Cette première étape de conscientisation est d'autant plus essentielle dans

le contexte de l'éclosion des sociétés plurilingues et de la complexité des interactions sociales entre les individus d'une même société et de cultures différentes. Elle s'inscrit d'abord dans une démarche d'appropriation des concepts de l'évaluation de la CCI ; une démarche devant permettre de recueillir des données fiables. Celles-ci devront être analysées et interprétées de façon circonstancielle puisqu'il s'agit de juger de la situation d'un apprenant en certains domaines de son développement en vue de prendre les meilleures décisions possibles relatives à son cheminement ultérieur. Dans cette perspective, elle est nécessairement partie intégrante du processus d'enseignement et d'apprentissage. Par conséquent, elle s'inscrit dans une démarche d'évaluation formative.

Inscrire l'évaluation dans une perspective pédagogique a pour but d'exploiter le potentiel de l'évaluation pour qu'elle contribue à la mise en place de conditions d'apprentissage optimales susceptibles d'aider l'apprenant dans son cheminement. Dans cette perspective, diverses orientations doivent guider les pratiques évaluatives (Ministère de l'éducation du Québec 1998). 1) L'évaluation doit être intégrée à la dynamique des apprentissages. Elle ne constitue pas une fin en soi. Elle s'ajoute à l'ensemble des moyens utilisés pour soutenir l'apprenant dans ses apprentissages. Elle permet à l'enseignant de recueillir des informations de façon continue en vue d'intervenir efficacement. 2) L'évaluation repose sur le jugement professionnel de l'enseignant. Une telle orientation confirme sa responsabilité à l'intérieur de certaines balises empreintes de rigueur et reposant sur la cueillette d'informations pertinentes, valides et suffisantes. 3) L'évaluation doit s'effectuer dans le respect des différences. Cette orientation implique que tous n'évoluent pas au même rythme ni de la même manière. Ces différences découlent aussi des caractéristiques socioéconomiques et culturelles des milieux d'où ils proviennent. L'enseignant insère dans sa pratique la différenciation pédagogique qui permet à un apprenant de cheminer tout en empruntant des voies diverses. 4) L'évaluation doit favoriser le rôle actif de l'apprenant dans les activités d'évaluation, augmentant ainsi sa responsabilité. Il s'agit de responsabiliser l'apprenant progressivement et de lui faire prendre conscience de l'impact de ses actes et de ses comportements dans ses interactions sociales. En ce sens, la modélisation du parcours d'un apprenant devrait permettre aux enseignants d'ajuster les pratiques d'intervention pédagogiques en tenant compte du positionnement d'un apprenant dans ses relations avec les autres cultures. C'est en agissant auprès des apprenants, sur leurs *savoirs*, leur *savoir-faire* et leur *savoir-être*, que l'école peut agir sur le monde social, le pouvoir constituant du langage étant basé sur des schémas de perception et de pensée (Bourdieu 1982). L'école devient une entrée dans la culture (Bruner 1996). L'intention est évidente : toutes les personnes qui interviennent en éducation aux langues deviennent, par le fait même, des passeurs culturels.

Bibliographie

- BACHMAN L. F. (1990), *Fundamental Considerations in Language Testing*, Oxford, Oxford Press.
- BAUMGRATZ-GANGL G. (1993), *Compétences transculturelles et échanges éducatifs*, Paris, Hachette.
- BOURDIEU P. (1982), *Ce que parler veut dire. L'économie des échanges linguistiques*, Paris, Fayard.
- BRUNER J. (1996), *L'éducation, entrée dans la culture*, Paris, Retz.
- BYRAM M.S. (1997), *Teaching and Assessing Intercultural Communicative Competence*, Clevedon, Multilingual Matters.
- CANALE M. et SWAIN M. (1980), Theoretical bases of communicative approaches to second language teaching and testing, *Applied Linguistics*, 1:1, p. 1-47.
- CITOYENNETÉ ET IMMIGRATION CANADA (2012), *Canadian Language Benchmarks, English as a Second Language for Adults*, Ottawa, Centre for Canadian Language Benchmarks, 3rd edition.
- CITOYENNETÉ ET IMMIGRATION CANADA (2006), *Niveaux de compétence linguistique canadiens 2006, Français langue seconde pour adultes*, Ottawa, Centre des niveaux de compétence linguistique canadiens.
- CONSEIL DE L'EUROPE (2007), *Niveau A1 pour le français, un référentiel*, Paris, Didier.
- CONSEIL DE L'EUROPE (2001), *Cadre européen commun de référence pour les langues*, Paris, Didier.
- LANDIS D., BENNETT J.M. et BENNETT M J., ÉdS (2004), *Handbook of Intercultural Training*, Thousand Oaks : Sage Publications, 3rd edition.
- LEGENDRE R. (2005), *Dictionnaire actuel de l'éducation*, Montréal, Éditions Guérin.
- LUSSIER D. (2011), « Language education as the entry to intercultural communicative competence (ICC) », dans *Canadian Issues on Diversity and Education for Liberation: Realities, Possibilities and Problems*, Special Edition, Spring, p. 60-66.
- LUSSIER D. (2009a), Theoretical Bases of a Conceptual Framework of Reference to Intercultural Communicative Competence, *Journal of Applied Linguistics*, 4-3, 2007, p. 285-308.
- LUSSIER D. (2009b), « Enseigner et évaluer la "compétence de communication interculturelle", enjeux et complémentarité », dans *Le langage et l'homme, Revue de didactique du français*, Bruxelles, E.M.E., p. 145-155.
- LUSSIER D. (2009c), « Enseigner "la compétence de communication interculturelle" : une réalité à explorer », dans *Revue pédagogique*, Novembre 2008, Québec, Ministère de l'Éducation du Loisir et du Sport.
- LUSSIER D. (1997), Domaine de référence pour l'évaluation de la compétence culturelle en langues. *Revue de didactologie des langues-cultures*, ÉLA, 106, Paris, Didier Érudition, p. 232-246.
- LUSSIER D., AUGER R., CLÉMENT R. et LEBRUN-BROSSARD M. (2004-2010), *Représentations culturelles, identité ethnique et compétence de communication interculturelle*, Gouvernement fédéral, Ottawa, Conseil de la recherche en sciences humaines.
- LUSSIER D. et LEBRUN M. (2009), *Le Projet Héritages – Carrefour linguistique, culturel et communautaire. Cadre de référence et recherche action pour la Commission scolaire de Montréal*, Montréal, Gouvernement du Québec, ISBN 978-2-9811727-09.
- LUSSIER D., GOLUBINA K., IVANUS D. et al. (2007), « Lignes directrices pour évaluer la compétence de communication interculturelle », dans *Développer et évaluer la compétence de communication interculturelle. Un Guide à l'usage des enseignants de langues et des formateurs d'enseignants*, Lazar I., Krieglér M., Lussier D., Matei G., Peck C., éditeurs, Strasbourg, Centre européen pour les langues vivantes/Conseil de l'Europe.

- LUSSIER D., AUGER R., URBANICOVA V. et ARMENGOL M. (2003), « Les représentations des autres et des autres cultures dans le contexte de la formation initiale et continue des enseignants », dans *Médiation culturelle et didactique des langues*. Zarate G., Gohard-Radenkovic A., Lussier D., Penz H., éditeurs, chapitre 7, p. 191-220. Centre européen pour les langues vivantes, Strasbourg, Éditions du Conseil de l'Europe.
- MINISTÈRE DE L'IMMIGRATION ET COMMUNAUTÉS CULTURELLES DU QUÉBEC (2010), *Échelle québécoise des niveaux de compétence en français pour les personnes immigrantes adultes*, Québec, Gouvernement du Québec.
- MINISTÈRE DES RELATIONS AVEC LES CITOYENS ET DE L'IMMIGRATION DU QUÉBEC (2000), *Niveaux de compétence en français langue seconde pour les immigrants adultes*, Québec, Gouvernement du Québec.
- MINISTÈRE DE L'ÉDUCATION DU QUÉBEC (2003), *Politique d'évaluation des apprentissages*, Québec, Gouvernement du Québec.
- RUBENFELD S., CLÉMENT R., LUSSIER D. et al. (2006), "Second Language learning and cultural representations: Beyond competence and identity", *Language Learning Journal*, 56, 4, p. 609-632.
- RUBENFELD S., CLÉMENT R., VINOGRAD J., LUSSIER D. et al. (2007), "Becoming a cultural intermediary: A further social corollary of second language learning", *Journal of Language and Social Psychology*, 26, 2, p. 182-203.
- TOMALIN B. et STEMPLSKI S. (1993), *Cultural awareness*, Oxford: Oxford Press.
- UNESCO (2001), *Déclaration universelle sur la diversité culturelle*, Paris : UNESCO.

L e graal de la compétence interculturelle : évaluer l'inévaluable en FLE ?

FRED DERVIN

UNIVERSITÉ D'HELSINKI, FINLANDE

I ntroduction

Je commencerai par une allusion à l'épopée finlandaise *Kalevala*, composée par Elias Lönnrot au XIX^e siècle, qui servit quelques décennies plus tard à poser les bases de l'imaginaire de la nation finlandaise. À travers plus de 200 000 vers, ce «bricolage» de poèmes recueillis par l'auteur dans les campagnes met en scène un objet intrigant : le *Sampo*. Fabriqué par l'un des personnages de l'épopée, Ilmarinen, le *Sampo* est décrit comme étant un objet extraordinaire, qui apporte (ré)confort, bonheur et fortune à celle¹ qui le possède. Cependant, personne ne semble savoir à quoi ce graal ressemble : un astrolabe qui représente le mouvement des astres sur la voûte céleste ? Une relique chrétienne ? Un moulin qui produit du sel, de la farine et même de l'or ? Ou bien un compas ? Dans cet article, je soutiens que le concept de compétence interculturelle constitue lui aussi un *Sampo*. En effet, de nombreux domaines et contextes ambitionnent ce graal : l'éducation, la santé, le monde des affaires, etc. Comme le *Sampo*, il a connu aussi une véritable course aux interprétations. En outre, depuis des décennies, les chercheurs du monde entier tentent d'évaluer ses effets. L'un des personnages centraux de cette épopée interculturelle est sans aucun doute le chercheur britannique Michael Byram (1997), qui a amplement contribué à problématiser le concept dans la didactique des langues et a conduit les chercheurs et praticiens à s'interroger sur son éventuelle évaluation. D'ailleurs, au moment où je commence à rédiger cette contribution, se tient à Tucson en Arizona la troisième conférence sur le développement et l'évaluation de la compétence interculturelle². Pour définir cette compétence, les organisateurs reproduisent celle proposée par Byram lui-même :

1. La forme féminine des marques linguistiques sera privilégiée dans cet article en réaction à l'utilisation systématique des formes masculines linguistiques dans la plupart des recherches francophones.
2. http://cercll.arizona.edu/doku.php/development/conferences/2012_icc

«Intercultural competence is [the ability] “to see relationships between different cultures - both internal and external to a society - and to mediate, that is interpret each in terms of the other, either for themselves or for other people”». «It also encompasses the ability to critically or analytically understand that one’s “own and other cultures” perspective is culturally determined rather than natural.»
Byram M. (2000). *Assessing intercultural competence in language teaching*. *Sprogforum*, 18 (6), 8-13.

Mais cette définition, qui a tant d’influence dans les mondes de la recherche globaux, pose problème, comme l’a démontré une minorité de chercheurs (Zarate et Gohard-Radenkovic 2004 ; Dervin 2010). Dans cet article, je souhaite réinterroger ces critiques et questionner ainsi l’évaluation éventuelle de la compétence dans le contexte de la didactique des langues. Pour commencer, afin de pouvoir envisager cette dernière, il faudrait savoir ce qui (peut) se cache(r) derrière ce *Sampo* didactique et pédagogique.

Déconstruire l’objet magique : ce que la compétence interculturelle ne peut pas/plus être...

Je vais étudier dans un autre pays et j’ai besoin de compétences interculturelles.

Ce fragment, entendu de la bouche d’une jeune étudiante finlandaise de français, reflète assez bien la *doxa* actuelle sur le déplacement, la mobilité et la rencontre de l’autre : si je bouge, je dois apprendre à être «interculturel». Mais que met-on derrière ce mot magique ? Les définitions sont pléthoriques et souvent contradictoires, même si elles semblent partager un certain nombre d’a priori (cf. par exemple la liste d’outils proposés par l’*Intercultural Communication Institute* à Portland aux États-Unis³ ; Lavanchy et al., 2011). Je propose d’examiner dans ce qui suit ce que la compétence interculturelle *n’est pas/plus*.

Premièrement, cette compétence n’est ni culturelle ni – pire ! – culturaliste. En d’autres termes, c’est l’*inter-* de la notion *interculturel* qui devrait intéresser ceux qui veulent la pratiquer plus que l’obsédant concept fatigué de culture (Breidenbach et Nyíri 2009 : 10). Il faudrait notamment dépasser l’idée souvent présente dans ces définitions que la compétence interculturelle permet « *de vivre heureux et de travailler avec succès au contact d’une autre culture* » (définition glanée sur internet). C’est une vérité de La Palice de dire que ce ne sont pas des cultures que nous rencontrons mais des individus complexes (Dervin 2012) et pourtant...

De même, pour Adrian Holliday (2011 : 39), l’approche descriptive des cultures souvent adoptée pour identifier les éléments à acquérir pour

3. <http://www.intercultural.org/tools.php>

devenir «interculturel» est problématique. Il explique que cette approche est avant tout idéologique car elle présente un système stable (Dahlen 1997) qui contribue à «assigner à résidence» (Maffesoli 2010 : 81) en imposant représentations et stéréotypes sur l'autre et le soi. Ainsi, Breidenbach et Nyíri (2009 : 50) dénoncent la survalorisation de l'idée de Confucianisme pour parler des Chinois et expliquent de façon ironique que : «*In any case using "Confucianism" to explain the values of 1.4 billion Chinese is surely as absurd as trying to derive the behavior of contemporary Europeans from the Bible or from Plato's Republic.*» De nombreux enseignants de langues utilisent d'ailleurs ces «lois» pour travailler avec certains apprenants, en France comme ailleurs, et leur imposent quelquefois des modèles d'apprentissage, d'habitudes, de traditions qu'ils considèrent «normaux» pour eux (Debono 2011). Mais comme le rappelle clairement Lemke (2008 : 20) : «*Normality is always a mystification of normativity, a social lie that succeeds in part by introducing simplistic, low-dimensional category grids for pigeon-holing us.*»

Continuons à dire ce que le *Sampo* interculturel n'est pas. Trois expressions clés serviront à démanteler les mythes de cette légende : le différentialisme à outrances, l'individualisme et la stabilité. Le premier élément souligne la tendance à vouloir travailler avant tout sur les différences culturelles comme base de la compétence (Abdallah-Pretceille 1986). Bien sûr, il y a des différences entre les apprenants «étrangers» mais ces différences peuvent être individuelles, professionnelles, générationnelles, émotionnelles, interactives, etc. De surcroît, on a tendance à l'oublier, les individus partagent souvent des caractéristiques qui vont bien au-delà des frontières imaginées des États-Nations ou des néo-frontières (Est/Ouest). Il est d'ailleurs intéressant de noter que ce dernier argument est souvent écarté par de nombreux chercheurs qui ont tenté de conceptualiser la compétence interculturelle en didactique comme ailleurs. Ainsi, dans son modèle *the Developmental Model of Intercultural Sensitivity*, le culturaliste américain Milton Bennett (2004) propose une étape intermédiaire «négative» dans le développement de la compétence interculturelle qu'il a intitulée la «minimisation des différences»... Cela reflète assez bien le fétichisme de la différence dite culturelle, dont on devrait s'éloigner.

Le deuxième élément qui ne peut constituer le graal que nous cherchons est celui de l'individualisme. En effet, la plupart des modèles de compétence interculturelle se concentrent sur l'individu «à évaluer» mais écartent systématiquement la relation qu'il est amené à construire avec un autre lors d'une rencontre. Pourtant, il est difficile de ne pas être d'accord avec M. Maffesoli (2010 : 80) lorsqu'il écrit «[...] il n'y a pas de moi sans l'autre, les autres : ma loi c'est l'autre. Je n'existe que dans et par l'esprit de l'autre, que dans et par le regard de l'autre». Pour finir, et en liaison avec ce qui précède, la compétence interculturelle ne peut être stable : d'abord parce qu'elle dépend d'une

interaction avec une autre personne dans des contextes précis et ensuite car l'action n'est pas toujours en phase avec les discours et vice versa. Prenons l'exemple de Raymond Apthorpe (2005) à Taiwan. L'anthropologue avait vécu avec des villageois qui avaient affirmé à plusieurs reprises qu'ils ne mangeaient pas de buffle car l'animal avait une dimension sacrée. Un jour, on lui sert un ragoût de l'animal. Il interroge alors le cuisinier :

Apthorpe : So this is Buffalo stew
 Cuisinier : Yes it is
 Apthorpe : I thought I had learned from you that you do not eat buffalo meat
 Cuisinier : Yes we don't eat buffalo
 Apthorpe : But this is buffalo stew
 Cuisinier : Yes but it is not our buffalo, it was our neighbour's

Il en va de même pour les discours sur la compétence interculturelle : ce n'est pas parce qu'un apprenant de langues dit avoir appris à éviter les stéréotypes que dans la «réalité» il ne réduit pas l'autre ou lui-même à des alibis culturels (Abdallah-Preteille 1986). La démarche angéliste qui consiste à vouloir faire croire à ce dernier phénomène a été remise en question il y a bien longtemps, notamment en psychologie sociale (Howarth 2002). Mais nous vivons dans des sociétés de la performance et de la visibilité (Aubert et Haroche 2011) où l'illusion de l'efficacité et de la prospérité est bien ancrée – d'où sans aucun doute par exemple l'omniprésence de l'objectif *savoir éviter les conflits* en éducation interculturelle.

R *e-fabriquer le Sampo et évaluer ses effets ?*

Des changements dans la façon de conceptualiser la compétence interculturelle et de la mettre en place dans la didactique du FLE paraissent donc indispensables. Ce renouvellement doit aller, il me semble, dans le sens de ce qu'Alban Bensa (2010 : 36-37) affirme à propos des méthodes anthropologiques : «il faut s'affranchir de l'idée absurde d'une adhésion pleine et entière des acteurs à leur propre monde, sans que jamais leur perplexité, leur questionnement, leur éloignement relatif par rapport à ce qu'ils vivent, ne soient examinés». À partir de là, en suivant le conseil de Noiriel à propos du racisme (2007 : 23), il nous faut trouver des instruments pour aider les apprenants à «penser et agir pour eux-mêmes» plutôt que de leur dire «ce qu'ils doivent penser ou faire». Une approche critique de l'évaluation peut aider à emprunter cette direction et éviter, par exemple comme on le fait dans le monde de l'entreprise, de juger de ces compétences à partir d'un simple test sur les connaissances dites culturelles, par le biais d'un QCM sur ordinateur.

Disons dès le début que l'évaluation sommative surtout quantitative n'a pas vraiment sa place ici car il semble difficile d'envisager de donner une note à un individu (ou plutôt à deux individus car on n'évaluera pas seulement un interlocuteur, cf. *supra*) dans le cadre de la compétence interculturelle. Le formatif, par le biais de l'auto-évaluation « contrôlée » (c'est-à-dire apprise - cette forme d'évaluation n'étant pas « naturelle »), de la co-évaluation et de l'évaluation en groupes, complémentée par des évaluations sommatives des savoir-faire qui mènent les apprenants à identifier ce qu'ils savent faire et à se fixer des objectifs, représente l'option la plus forte pour l'interculturel : son apprentissage étant « à vie » et sans fin, ce n'est pas une note uniquement qui doit sanctionner le processus d'apprentissage mais des discussions autour des aspects éthiques, idéologiques, relationnels et représentationnels des rencontres interculturelles (Abdallah-Preteuille et Porcher 1998 ; Barbot et Dervin 2012 ; Holliday 2010).

Ainsi, avant/pendant et après avoir visionné un film ou un documentaire, après avoir lu un roman ou avoir vécu une expérience qu'ils considèrent comme interculturelle, les apprenants s'interrogeront sur l'« ensemble élaboré de stratégies qui se déroulent dans un contexte d'énonciation précis » entre les acteurs (Chauvier 2011 : 18) : identifications multiples, mises en scène, représentations, manipulations, alibis, etc. (Dervin 2012). En bref, les anomalies et les dissonances qui peuplent « la vie politique inhérente à toute rencontre » (Chauvier, *ibid.*). Les apports de la pragmatique discursive (énonciation, dialogisme, analyse non-culturaliste des conversations) sont inestimables pour ce faire (Dervin 2012).

Les apprenants peuvent examiner des situations en s'inspirant de l'exemple de Bensa (2010 : 70) : « Quand il arrive que des individus se revendiquent de leur "culture", c'est qu'ils s'adressent à d'autres individus dont ils imaginent qu'ils sont très différents d'eux, détenteurs eux-mêmes d'une autre "culture". La culture n'est pas un cadre de référence en soi qui s'impose aux personnes d'un même groupe, c'est toujours d'abord le regard de l'autre sur l'autre ». De plus, « l'enveloppe » temporelle et spatiale (Maffesoli, 2010 : 28) dans laquelle les interactions se déroulent, souvent caractérisée de « contexte », sera également discutée. Pour finir, on entraînera les apprenants à identifier et à analyser dans cette enveloppe, les aspects ludiques, oniriques et imaginaires (Maffesoli, *ibid.* : 48) constitutifs des discours et des actions en société. Breidenbach et Nyíri (2009) ont problématisé l'ensemble des éléments présentés ici dans leur ouvrage intitulé *Seeing culture everywhere*. Leurs propositions sont tout à fait transférables dans le contexte de la didactique des langues à tous les niveaux. Leur point central consiste d'ailleurs aussi à développer une compétence interculturelle, c'est-à-dire « *the critical ability to question the implicit and explicit assumptions behind cultural claims and the power dynamics that they may be concealing* » (*ibid.* : 340). À partir de là, ils proposent les questionnements suivants,

qui peuvent être utilisés comme base pour l'auto-évaluation/évaluation. Face à ce qui est présenté comme une vérité culturelle, les apprenants se demanderont :

	QUESTIONS
Discours sur la culture	<ul style="list-style-type: none"> - Quelles affirmations implicites et explicites sur la culture, à propos de quels groupes ? - Quelles frontières sont posées entre les groupes ? - Les différences sont-elles internes ou externes ? (<i>ibid.</i> : 343, ma traduction)
Droit à la parole	<ul style="list-style-type: none"> - Qui dit quoi sur la culture ? Pourquoi ? - Au nom de qui parlent-ils de façon explicite ou implicite ? - Pourquoi ont-ils le droit de parler ? Quelles voix sont silencieuses ? (<i>ibid.</i> : 344)

On pourra se rapporter également au texte classique de l'anthropologue J. Clifford (1986/2011), intitulé *Vérité partielles, vérités partiales*, ou aux propositions récentes de la sociolinguiste Ingrid Piller (2011) qui vont dans le même sens.

Il y a quelques années, j'ai participé au projet ECC-Lang que le ministère finlandais de l'Éducation avait mis en place. Dans le cadre de ce projet, un certain nombre d'enseignants-chercheurs impliqués dans l'enseignement du français à l'université, ont tenté de créer des « grilles » pour évaluer les compétences interculturelles. Nous nous sommes vite rendu compte que l'évaluation par « notes » uniquement des compétences interculturelles était une impasse car, mis à part la définition de compétences d'analyse, prenant en compte tous les éléments mentionnés plus haut, l'approche par connaissances sur la France et les espaces francophones était peu intéressante car elle nécessitait un positionnement de la part des enseignants (choix des connaissances) sur lequel nous n'arrivions pas à nous entendre – l'idée d'un niveau seuil étant inenvisageable. Nous avons donc « réduit » ces compétences à des compétences universitaires par niveau (1^{re} année, Licence, Master), adaptées donc au contexte qui nous concernait. Les propositions des paragraphes précédents n'avaient pas séduit les collègues car elles auraient requis de leur part un apprentissage de ces éléments avant de les proposer aux étudiants. La démarche interculturelle proposée ici est interdisciplinaire car il faut se familiariser avec l'anthropologie, la sociologie, la psychologie sociale, etc. Dans les formations d'enseignants en Finlande, ces domaines sont peu explorés car on a tendance à se concentrer sur les apports des sciences de l'éducation uniquement – même pour l'interculturel...

C onclusion

C'est une forme de violence totalitaire que de vouloir le bien des autres
 à la place des autres.
 (Maffesoli 2010 : 21.)

La citation qui clôt cet article surprendra sans aucun doute par son côté politiquement incorrect. Pourtant, la traque du *Sampo* interculturel que nous venons d'effectuer a fait émerger des signes de « violence totalitaire » dont l'enseignant de langues pourrait se démarquer. Ainsi vouloir imposer un modèle de compétence interculturelle (néo)culturaliste pour rendre les rencontres « heureuses » et « efficaces » est une illusion dangereuse qu'il faut apprendre à dépasser. Il en va de même si l'on essaie de « vendre » aux apprenants l'idée que tout, dans l'interculturel, peut s'expliquer, se calculer, être contrôlé par des recettes magiques – ou *miracles* pour filer la métaphore du *Sampo* – et que l'on peut donc évaluer formellement la compétence qui en découle.

La compétence interculturelle peut avoir des effets qui ne sont ni calculés ni calculables – car nous avons affaire à des interactions humaines donc non contrôlables (actions/discours) et complexes (« l'enveloppe »). C'est pourquoi une approche critique, socioconstructiviste et intersubjective de la compétence et de son évaluation est nécessaire.

Il manque, pour finir, une caractéristique souvent négligée dans l'enseignement des langues mais qui a depuis longtemps été travaillée par exemple dans le domaine de l'éducation multiculturelle critique (s'inspirant des travaux de la pédagogie critique) : le combat pour la justice sociale (ex : Banks 2011). Il ne semble plus utile en effet d'enseigner une compétence interculturelle « molle » et élitiste (par exemple pour rencontrer une « bourgeoisie parisienne imaginaire » ou « se débarrasser des stéréotypes »). Il faudrait ainsi passer à une étape supérieure et amener les apprenants à prendre des risques, en dénonçant par exemple toute situation d'inégalité (pour soi et les autres), de racisme (avec ou « sans races »), de sexisme, et toute forme de préjugé, de discrimination, etc., interpersonnelles ou institutionnelles (Räsänen 2009 : 37). Pour parodier Bourdieu, il faudrait donc faire de la compétence interculturelle un vrai « sport de combat ».

Bibliographie

- ABDALLAH-PRETCEILLE M. et PORCHER L. (1998), *Éthique de la diversité et éducation*, Paris, PUF, L'Éducateur.
- ABDALLAH-PRETCEILLE M. (1986), *Vers une pédagogie interculturelle*, Paris, PUF.
- APTHORPE R. (2005), "It's (the) culture, stupid!" Why "adding culture" is unlikely to make any serious difference to international developmentalism, *The Asia Pacific Journal of Anthropology*, Vol. 6 (2), p. 130-141.
- AUBERT N. et HAROCHE C. (2011), *Les tyrannies de la visibilité*, Toulouse, Erès.
- BANKS J.A. et MCGEE BANKS C. (2008), *Multicultural education: issues and perspectives*, New York, Wiley.
- BARBOT M.-J. et DERVIN F. (éds.) (2011), *Rencontres interculturelles et formation, Éducation permanente*, n° 186.
- BENNETT M.J. (2004), *Becoming interculturally competent*, dans Wurzel J.S. (éd.), *Toward multiculturalism: a reader in multicultural education*, Newton, MA, Intercultural Resource Corporation, p. 62-77.
- BENSA A. (2010), *Après Levi-Strauss*, Paris, Textuel.
- BREIDENBACH J. et NYIRI P. (2009), *Seeing culture everywhere*, Washington, University of Washington Press.
- BYRAM M. (1997), *Assessing Intercultural Communicative Competence*, Clevedon, Multilingual Matters.
- CHAUVIER E. (2011), *Anthropologie de l'ordinaire*, Toulouse, Anacharsis.
- CLIFFORD J. (1986/2011), *Vérités partielles, vérités partiales*, *Journal des anthropologues*, n° 126-127, p. 385-433.
- DAHLEN T. (1997), *Amongst the interculturalists*, Stockholm, Stockholm University Press.
- DEBONO M. (2011), *Pour une pédagogie du conflit en Chine ?*, *Synergies Chine*, n° 6, p. 127-140.
- DERVIN F. (2010), *Assessing intercultural competence in language learning and teaching : A critical review of current efforts in higher education*, dans Dervin, F. et Suomela-Salmi, E. (éds.), *New Approaches to Assessment in Higher Education*, Bern, Peter Lang, p. 157-174.
- DERVIN F. (2012), *Impostures interculturelles*, Paris, L'Harmattan.
- HOLLIDAY A. (2011), *Intercultural communication and ideology*, London, Sage.
- HOWARTH C. (2002), "Identity in whose eyes ? : the role of representations in identity construction", *Journal for the Theory of Social Behaviour*, 32 (2), p. 145-162.
- LAVANCHY A., GAJARDO A. et DERVIN F. (éds.) (2011), *Anthropologies de l'interculturalité*, Paris, L'Harmattan.
- LEMKE J. (2008), *Identity, development and desire : critical questions*, dans Caldas-Coulthard C. et Ledema R. (éds.), *Identity trouble*, London, Palgrave MacMillan.
- MAFFESOLI M. (2010), *Le temps revient, formes élémentaires de la postmodernité*, Paris, DDB.
- NOIRIEL G. (2007), *Racisme : la responsabilité des élites*, Paris, Textuel.
- PILLER I. (2011), *Intercultural Communication. A Critical Introduction*, Edinburgh, EUP.
- RÄSÄNEN R. (2009), *Teachers' intercultural competence and education for global responsibility*, dans Talib M.-T., Loima J., Paavola H., Patrikainen S. (éds.), *Dialogs on Diversity and Global Education*, Bern, Peter Lang, p. 29-49.
- ZARATE G. et GOHARD-RADENKOVIC A. (2004), *La reconnaissance des compétences interculturelles : de la grille à la carte*, Paris, Didier.

Dans les valises du plurilinguisme : quelle conception des normes et des erreurs ?

MARTINE MARQUILLÓ LARRUY
UNIVERSITÉ DE POITIERS, FRANCE

À Bernard Py,
dont le regard perspicace et le sourire bienveillant vont
nous manquer.

Le désir de plurilinguisme européen vient bousculer un certain nombre de conceptions bien ancrées dans les esprits. Le singulier de la norme est lui aussi en passe de devenir obsolète tant la diversité de la variation linguistique se fait présente, explicite... et légitime. Comment ce bouillonnement de diversité(s) s'accommode-t-il d'une évaluation de plus en plus techniquement rigoureuse voire pointilleuse ? Comment les formateurs peuvent-ils s'y retrouver entre l'élasticité d'une conception plurilingue de l'apprentissage des langues et la certitude quelque peu rigide d'un étalonnage plus quantitatif que qualitatif que la pratique vient le plus souvent interroger ? Cette contribution propose donc de réfléchir à l'articulation entre plurilinguisme, normes et évaluation.

Les valises du plurilinguisme

LE PLURLINGUISME EN EUROPE : LUXE OU NÉCESSITÉ ?

D'un point de vue institutionnel, le plurilinguisme européen est en marche. Le rêve exprimé par Victor Hugo en 1883 dans le codicille de son testament¹ est aujourd'hui une réalité, si on considère l'existence de la bibliothèque *Europeana*, l'utilisation des quatre langues co-officielles (catalan, valencien, galicien et Euskera) pour le débat des motions des différents groupes parlementaires du sénat espagnol², ou encore le fait que le plurilinguisme constitue un des emblèmes de l'identité européenne.

1. « Je lègue tous mes manuscrits, et tout ce qui sera trouvé, écrit ou dessiné par moi, à la Bibliothèque Nationale de Paris qui sera un jour la Bibliothèque des États-Unis d'Europe »

2. Les questions concernant l'État demeurant, elles, en espagnol.

À ceux qui trouvent que la traduction coûte cher à l'Europe, il est répondu que « certaines valeurs valent plus que leur simple coût », que ces mesures induisent le respect de l'autre et permettent de vivre ensemble et/ou que le coût du multilinguisme pour l'ensemble des institutions européennes (Parlement, Commission, Conseil économique et social) représente à peine le prix d'un café pour chaque Européen (2,20 € par citoyen et par an). Nécessaire ici, et peut-être considérée comme un luxe, là, il reste que la promotion du plurilinguisme est omniprésente.

Si le plurilinguisme finit par agacer, tant il est partout, il n'est pas sorti de rien. En effet, le souci majeur, au début des années 1950, était de construire chez les Européens d'autres représentations que celles d'un nationalisme étroit et chauvin qui avait conduit aux atrocités de la Deuxième Guerre mondiale. Dans le cours qu'il professe en 1944-1945 au Collège de France, Lucien Febvre, cofondateur avec Marc Bloch de *l'École des Annales*, lance un appel à constituer « la nation européenne dont les nations d'aujourd'hui seraient les provinces », (Febvre et Crouzet 2012 : 317). C'est dans ce contexte que l'on peut comprendre la rédaction d'un manuel d'histoire de la civilisation française destiné à de jeunes enfants qui mettait l'accent non sur la singularité de la nation mais sur la richesse d'un héritage fait d'emprunts et d'acclimations diverses³.

À l'origine de cette nouvelle doxa contemporaine du plurilinguisme, il y a donc un idéal humaniste d'équité, de reconnaissance de la légitimité, de la diversité et surtout la prise en considération que l'Europe contemporaine est le fruit de croisements, d'intenses brassages faits de guerres et de conquêtes certes, mais aussi de l'appropriation d'héritages culturels intriqués très divers, en un mot de *métissages* (cf. les exemples donnés par Laplantine et Nous 1997). Au-delà des revendications nationalistes, le nouveau paradigme plurilingue nous invite donc à **renouer** avec une histoire que la création des États-Nations au XIX^e siècle nous avait fait oublier (voir Baggioni 1997 ou encore Martinell Gifre et E. Erlendsdóttir 2005).

LE PLURILINGUISME : L'ENVERS DU DÉCOR

Les preuves du volontarisme institutionnel sont légion tant au niveau européen qu'au sein des institutions francophones. Entre autres exemples possibles, on peut citer le slogan de l'Organisation internationale de la francophonie (« la voix de la diversité ») ou encore la quatrième priorité de la politique de la langue de la Délégation générale à la langue française et aux langues de France (« favoriser la diversité linguistique »). Cependant, cette promotion institutionnelle du plurilinguisme n'est pas exempte de contradictions et de quelques couacs...

On peut ainsi mettre en regard la résolution du Parlement européen du 24 mars 2009 (*Le multilinguisme, un atout pour l'Europe et un engagement commun*, dont l'article 6 élargit le plurilinguisme aux domaines

3. Cet ouvrage, *Nous sommes des sang-mêlés*, n'a vu le jour que très récemment (Febvre et Crouzet 2012).

culturels et linguistiques) et l'obligation *implicite* mais *avérée* de soumettre un projet en langue anglaise et ce, en dépit de ce qui est énoncé dans les règles de fonctionnement (voir, par exemple, le site de COST, *European coopération in science and technology* ; cf. également Maurer 2011, ou, sur les problématiques liées à l'anglais langue véhiculaire de l'enseignement supérieur, Truchot 2010).

Au-delà de ces dysfonctionnements liés à l'irrésistible poussée de l'anglais, d'autres éléments entachent et obèrent les principes humanistes que j'évoquais ci-dessus et doivent nous garder d'une représentation trop idéalisée ou angélique, sinon du plurilinguisme lui-même qui se trouve ainsi instrumentalisé, du moins des institutions qui le promeuvent. En 1999, Francis Gillery réalise un documentaire, *Le cartable de Big Brother*, diffusé à une heure tardive sur France 3, mais que l'on peut aisément regarder encore aujourd'hui sur la toile, qui décorquait, en un peu moins d'une heure, les négociations, tractations et autres démarches de lobbying menées par de grands groupes industriels (et en particulier *l'European Round table*) auprès de la Commission européenne pour déconstruire les services publics de l'éducation et s'emparer du juteux terrain de la formation à travers le développement et le contrôle de ce que l'on appelait alors les *autoroutes de l'information*. Ce que ce film dénonçait comme avenir probable est devenu notre quotidien. On peut par ailleurs constater la place centrale que joue l'évaluation dans ces processus :

« Le risque, je pense, c'est d'arriver à établir un standard, "undisputed standard", pour juger l'ensemble des citoyens de l'Europe, pour juger de leurs connaissances et de leurs compétences de base. On a évacué l'histoire parce que c'était trop délicat, la géographie l'est peut-être, et encore. Et on va prendre un certain nombre de choses comme les maths qui doivent être indiscutables [...].

Et puis, en plus de cela, ces connaissances, ces compétences de base, on va les inscrire dans une carte. Donc, c'est une initiative dont il est dit clairement qu'elle est transnationale. [...] Donc, vous voulez chercher un emploi, on va vous demander votre système d'accréditation puisque c'est cela qui fera foi. Parce que les diplômes papier, ils ont beau être irremplaçables, [...] ils sont tellement marqués d'un côté national de votre culture nationale, que ça doit être anti-européen. L'informatique, ça garde la trace de tout. Donc, vous avez les moyens de savoir si une personne a passé pour la 25^e fois son test. [...] C'est pire que le fichage. Enfin, c'est un fichage comme personne n'avait pensé pouvoir le faire. » Extraits de l'interview d'un ingénieur expert de la Commission européenne, Gillery : *Le cartable de Big Brother*.

En témoigne, en France, le *Livret personnel de compétences* qui atteste l'acquisition des connaissances et compétences du socle commun, de l'école élémentaire à la fin de la scolarité obligatoire et qui est utilisé à l'école primaire depuis 2008, depuis 2009 au collège et qui suscite des actions en justice de parents militants.

Dans une conférence relativement récente (2009), Jean-Paul Bronckart se demande si « la notion de *compétence* est pertinente en éducation ». Il démêle dans un rappel historique très éclairant l'écheveau complexe des ancrages de cette notion. Il insiste sur ses deux origines

distinctes qui conduisent à des statuts et des évolutions différentes voire *contradictoires*, et ce quand bien même ces deux conceptions interfèrent. Il distingue ainsi, d'une part, la compétence en linguistique et en didactique des langues (de Chomsky à la reconfiguration de la notion opérée par Hymes : cf. à ce propos Coste, de Pietro et Moore : 2012) et, d'autre part, la compétence issue du champ du travail et de la formation des adultes. Dans le champ de l'éducation – formation il constate que la notion de compétence est revendiquée par des groupes sociaux qui ne partagent pas les mêmes conceptions de l'éducation. Il oppose ainsi une conception *généreuse*, pour laquelle il prend précisément comme exemple la définition de la *compétence plurilingue et pluriculturelle* dans le CECR et la met en contraste avec les discours du patronat qui :

« considèr[ent] que la dérégulation économique (c'est-à-dire la destruction des mécanismes de contrôle de l'État et de contreponds des syndicats) doit se prolonger par une dérégulation éducative, c'est-à-dire par l'intégration des systèmes de formation dans une logique de marché. Et ces discours proposent aussi que les finalités de l'enseignement soient redéfinies en termes de compétences, cette fois en tant que capacités pratiques à exercer un métier sans tenir le moindre compte des finalités générales de socialisation et d'acculturation jusqu'à présent assignés à l'enseignement obligatoire » (Bronckart 2009 : 6).

Dans un ouvrage qui a suscité bien des débats, Bruno Maurer (2011) propose une lecture critique très détaillée de publications émanant du Conseil de l'Europe qui accompagnent et explicitent la notion de *compétence plurilingue* dans une **visée politique**. Il montre, par exemple, comment une notion telle que la *compétence partielle* peut être **détournée** au profit d'une économie peu soucieuse d'éthique. Il reprend à ce sujet les propos de Frath :

« Ainsi, un secrétaire devrait avoir, mettons, un niveau C1 en compréhension et production orale, pour répondre au téléphone, C2 en production écrite si elle doit rédiger des lettres elle-même, et seulement B1 si elle doit les recopier. [...] L'avantage pour l'employeur est évident : plus besoin d'investir dans la formation maison des jeunes diplômés, qui pourraient être licenciés aisément, puisqu'on trouvera toujours des demandeurs d'emplois aux compétences limitées, et donc cartifs de cette industrie » (Frath 2002, cité par Maurer 2011 : 83).

Les réflexions qui précèdent montrent d'une part la complexité des facteurs intriqués dans le plurilinguisme et d'autre part la nécessité de prendre en compte ce que Lise Dubois (2006) nomme *la reconfiguration des valeurs attribuées aux ressources linguistiques*. La nouvelle donne de l'économie mondiale dans laquelle les services, l'information et la communication jouent un rôle majeur invitent à suivre attentivement les travaux de chercheurs comme Alexandre Duchêne, Monica Heller et Cécile Canut (cf. notamment le n° 136 de *Langage et société*, dirigé par C. Canut et A. Duchêne, qui a pour thème : « Appropriation politique et économie des langues ») qui observent l'exploitation des ressources langagières et des locuteurs par l'économie néolibérale, *la diversité linguistique devenant de plus en plus un facteur de rentabilité* (Duchêne 2011 : 82).

Sans écarter la critique socio-politique que l'on peut faire de la notion de **compétence partielle**, Marinette Matthey (2005) montre toutefois le risque qu'il y aurait à vouloir jeter le bébé avec l'eau du bain : selon elle, considérer que les compétences partielles équivalent à des *objectifs au rabais*, c'est faire le lit des tenants du *tout ou rien* et contribuer ainsi au « tout-à-l'anglais ». M. Matthey note également la diversité des interprétations dont cette notion fait l'objet et prend acte de son intérêt. Celle-ci met en effet en avant la dimension fonctionnelle des langues et correspond à la prise en compte des besoins langagiers diversifiés des locuteurs. Cependant, elle n'en pointe pas moins :

« [qu']en finir avec le purisme et le perfectionnisme dans l'apprentissage des langues est une très bonne chose, mais qu[e cela] se heurte malheureusement aux pratiques séculaires d'évaluations normatives des productions linguistiques en situation scolaire. Il y a un risque sérieux, que l'école mette de moins en moins l'accent sur les aspects formels des langues dans l'enseignement, tout en le maintenant dans l'évaluation, faute de ne pas savoir évaluer ce qu'elle enseigne. [...] Il est donc urgent que des linguistes et des didacticiens tentent d'élaborer de nouveaux objets d'évaluation, compatibles avec la notion de *compétences partielles* » (Matthey 2005 : 153).

Dans ce contexte relativement nouveau et délicat, la question de l'évaluation des compétences partielles suppose de repérer et d'explicitier les effets du plurilinguisme et de la variation sur certaines notions traditionnellement associées à l'évaluation, et notamment sur les conceptions circulantes de la norme (ou des normes) et sur l'évolution du statut attribué aux erreurs.

Plurilinguisme et didactique des langues : *carambolage de représentations et apprivoisements en cours...*

Si le secteur économique s'est rapidement emparé, et avec profit, de la nouvelle idéologie plurilingue, son appropriation en didactique des langues, parce qu'elle implique un changement dans les *représentations* et les *pratiques*, requiert un temps plus long. Des déplacements disciplinaires l'ont accompagnée et ont ouvert de nouveaux espaces de réflexion.

Même si certains chercheurs comme L. Dabène avaient souligné la **nécessité** d'une politique volontariste à l'égard du plurilinguisme dès le milieu des années 1980 (Dabène 1984), il n'est pas faux de considérer que ce domaine est encore aujourd'hui en pleine émergence, tant sur le plan de la recherche que dans ses implications didactiques. Ces dernières, mal cernées, entrent en contradiction avec nombre de représentations traditionnelles dont on a du mal à se défaire. Dans un article

consacré aux nouvelles approches théoriques du langage et du plurilinguisme, G. Lüdi (2011 : 51) évoque la critique faite par Larsen-Freeman (2006) du modèle dominant de l'acquisition des langues secondes, selon lequel, de manière traditionnelle, les langues cibles natives seraient stables et homogènes et l'acquisition consisterait en un rapprochement de la langue cible par étapes clairement distinguables et linéaires.

Alors que l'on sort à peine d'une représentation négative des effets du bilinguisme, la référence au locuteur natif comme finalité à atteindre ou la référence à une norme unique ou encore la représentation du bilinguisme comme une compétence équilibrée dans les deux langues font partie de ces lieux communs que le plurilinguisme vient bousculer. Pour B. Py (2005), après une première adhésion de surface, relevant presque du « politiquement correct », le plurilinguisme entre en collision avec des représentations issues d'une tradition scolaire plus habituée à supprimer les variations qu'à composer avec elles et qui renvoie à la représentation figée d'une langue standardisée et fixée une fois pour toutes. Dans ce cadre, le rapport à la norme est envisagé sous une forme dichotomique en termes de « juste » ou « faux ». À ceci s'ajoute la méconnaissance de fonctionnements langagiers tels que la créativité ou l'approximation.

Par ailleurs, D. Coste a expliqué à plusieurs reprises (2006, 2008) que la *compétence plurilingue* n'était originellement qu'un « pari, voire [un] coup de force didactique » (Coste 2008 : 93), qui a été nourri et « épaissi » par la créativité et la ténacité conceptuelle de quelques chercheurs : les concepts ne sont pas nouveaux, dit-il, mais ils sont « retravaillés, précisés, remis en interrelation au fil des chapitres sous l'angle des **contacts de langues** et de la constitution d'un **répertoire plurilingue** » (la mise en gras est mienne, Coste 2006 : 247). Ainsi, ce qui fonde le concept de *répertoire plurilingue*, c'est la notion de *variation*, qu'il s'agisse du degré de compétence ou de la variation que constitue la diversité des langues, tout autant que la dimension dynamique et évolutive des langues qui suivent les fluctuations de la vie des individus.

Ce changement de perspective, cette *révolution* au sens premier du terme, a conduit à reconsidérer l'ensemble des phénomènes qui entrent en jeu dans la mouvance des langues en contact, tels que les *alternances de codes*, les *interférences*, les *emprunts*, les *calques* qui ont quitté – du moins dans le domaine de la recherche – le camp de la stigmatisation pour faire l'objet de descriptions fines qui rendent compte des **fonctionnalités** que jouent dans les échanges langagiers ces marques de passages translinguistiques. Ainsi, dès 1987, Georges Lüdi, et avec lui toute l'équipe de chercheurs de Bâle-Neuchâtel, à travers la valeur d'hypéronyme de l'expression *marques transcodiques*, choisissait de catégoriser de manière aussi neutre que possible ce qui auparavant était considéré comme des insuffisances langagières, des écarts à la norme, soit des erreurs.

Les deux notions cardinales dans le domaine des contacts de langues, celles de *fossilisation* et de *transfert*, ont fait récemment l'objet d'une attention renouvelée. La notion de *fossilisation*, entendue classiquement comme un arrêt ou un effet de plateau dans l'évolution de l'apprentissage, notion centrale dans la théorie du fondateur du concept d'*interlangue*, Selinker (1972), est reconsidérée et nuancée par Stéphanie Costa-Galligani dès 2001 :

« Compte tenu des différentes propriétés inhérentes à l'apprentissage d'une langue, il nous paraît difficilement concevable d'envisager la fossilisation en termes d'arrêt du processus d'acquisition. D'autant plus que les pratiques langagières sont contraintes d'évoluer au contact d'autres modèles de langue chez l'apprenant. Il nous paraît plus pertinent de réserver les termes de *relâchement* ou de *démobilisation* pour qualifier le stade de l'acquisition où "l'impulsion à apprendre" en langue cible diminue. » (Costa-Galligani 2001 : 61)

L'idée à retenir, concernant les questions d'évaluation, est bien celle de **l'ébranlement du caractère définitif de la fossilisation**. Par exemple, les travaux de Boettinger, Park et Timmis (2010), par une démarche originale de co-observation/supervision/autonomie appuyée sur des stratégies métacognitives montrent des défossilisations de leurs propres pratiques langagières.

En outre, on note aujourd'hui – sous l'influence d'un domaine qui prend de plus en plus de consistance, celui qui gravite autour des questionnements de l'apprentissage d'une L3 – un véritable renversement de perspective qui consiste à considérer que le bilinguisme, tout comme le monolinguisme, ne sont que des cas particuliers du plurilinguisme. Sans que cela disqualifie les travaux centrés sur les relations L1/ L2, la prise en considération d'une L3 introduit une complexité qui correspond à un changement d'échelle, comme le montre Ulrike Jessner (2008 : 271) en contrastant les possibilités d'acquisition associées à deux langues *versus* trois langues. Lorsque deux langues sont en présence, soit elles sont acquises simultanément, soit successivement, alors que l'ajout d'une troisième peut entraîner au moins les cas suivants :

1. L1 \Rightarrow L2 \Rightarrow L3. Les trois langues sont apprises consécutivement ;
2. Lx/Ly/Lz. Les trois langues sont apprises simultanément
3. Lx/Ly \Rightarrow L3. Lx et Ly sont apprises simultanément et avant l'apprentissage de la L3
4. Lx \Rightarrow Ly/Lz. La L2 et la L3 sont apprises simultanément après l'apprentissage de la L1
(d'après Jessner 2008 : 271)

Outre le fait que ceci rend d'autant plus délicate la manière de nommer et de qualifier ces différentes langues, ces distinctions ne sont pas superflues car elles ont des incidences sur les perméabilités translinguistiques que l'on regroupait auparavant sous le nom de *transferts* et qui sont fédérées aujourd'hui sous la bannière des *Crosslinguistic*

influence (CLI) dans la littérature consacrée à l'acquisition de trois langues au moins (voir de Angelis et Dewaele 2011, ainsi que les travaux pionniers de Jasone Cenoz (1997, 2004) sur l'apprentissage de l'anglais au Pays basque, ou, pour la France, Trévisiol et Rast 2004, ainsi que Bono 2008).

Ainsi, une de nos étudiantes, chinoise et enseignante de français expérimentée, a rédigé, pour les besoins d'une mise en situation où l'on observait des processus rédactionnels en langue étrangère, un petit texte argumentatif en italien, qui est la dernière langue qu'elle a apprise.

cela pro ragazzo
avere i ragazzi
jeter ragazza
fe ragazze dare misna m
volere -> vuoi possono
attencion
zion
posso
parents
posso
case
possono andare
hanno la timida
sono
fanno
En multi paese
alalla scuola. In Europa, molti pensano che
i ragazzi e le ragazze sono uguali.

Les observatrices qui devaient rendre compte de ses manières de faire ont relevé une stratégie récurrente : sur la partie supérieure de sa feuille, la scriptrice avait délimité un espace sur lequel elle notait les mots qui lui faisaient défaut en italien. Elles ont également remarqué que les verbalisations se faisaient en langue espagnole, ce qui signifie que c'est l'avant-dernière langue apprise qui lui a servi le plus souvent de tremplin pour rédiger son texte en italien. Cependant, l'observation attentive de son brouillon permet de voir que l'espagnol n'est pas la seule langue d'appui : on y trouve des mots français : *jeter*, *parents*, *cela*, des mots suivis de leur équivalent en espagnol : *penser/pensar* ; des mots, enfin, qui traduisent un entre-deux linguistique. Le mot *attencion* porte des traces particulièrement emblématiques de ces zigzags de la pensée entre plusieurs langues : la double consonne renvoie à la graphie du français mais le préfixe final est celui qui lui correspond en espagnol bien qu'il se trouve ici dépourvu d'un accent. À cette notation s'ajoute au-dessus une variante (peut-être phonétique ?) *zion* qui montre que la graphie de ce mot a suscité une intense réflexion métalinguistique de son auteur d'autant que, dans le texte rédigé au-dessous, on observe de plus une rature sur ce qui correspondrait à la graphie en langue française *atten /t/cion*, langue qui, depuis deux ans, est celle de son environnement immédiat.

Cet effet de loupe sur un extrait de texte montre combien le plurilinguisme, pour peu que l'on veuille bien ne pas l'occulter, renvoie à un renouvellement de la conception de l'erreur et de la norme qui se

trouvent forcément convoquées lors de l'évaluation. En matière d'interprétation d'erreurs, au-delà de la conception qui considère que les erreurs sont des indices d'une étape de l'apprentissage, il est nécessaire aujourd'hui d'insister sur l'idée de la relativité de l'erreur : il n'y a d'erreur que par rapport à une norme ou plutôt des normes spécifiques qui ont été préalablement définies en fonction de la situation de production et du contexte d'apprentissage. Ce qui pose la délicate question de la norme ou plutôt des normes de référence.

L'énaurme : *quelles normes, entre purisme et variation, pour la didactique du français langue étrangère dans un contexte plurilingue ?*

Le mauvais jeu de mots de l'intitulé de ce paragraphe a pour seul but d'attirer l'attention sur la difficulté de la question posée au regard de ce trait si particulier à la francité, qui fait dire à J.-M. Klinkenberg que les Français ont « une hypertrophie de la glande grammaticale » et qui explique la parution régulière d'ouvrages consacrés quasi exclusivement à la question du purisme (par exemple, Leeman-Bouix 1994, Meschonnic 1997, Paveau et Rosier 2008...) ou aux modalités d'essentialisation de la langue (Merlin-Kajman 2003, Canut 2010). Françoise Gadet avance l'idée que, si la France a conservé cette *position ultra-dominante* dans la détermination de la norme et si elle se comporte comme si elle était *la propriétaire exclusive* de la langue, c'est parce que, contrairement au destin des autres langues européennes diffusées dans le nouveau monde, *toutes, sauf le français, auraient plus de locuteurs hors du berceau* (Gadet 2008 : 7).

L'enseignant de français, sans doute plus que les enseignants des autres langues, doit composer aujourd'hui avec cette tension entre une forte inclination *lettrée* pour la *belle langue* et le foisonnement de la pluralité des français qui, selon le contexte, permettra d'accepter *j/le kiffe grave...* ou à tout le moins, en explicitera le sens, pour ne pas obérer la compétence de compréhension de l'apprenant. Depuis les travaux fondateurs des sociolinguistes et des créolistes, la variation a fait beaucoup de chemin, à tel point que Gilles Siouffi (2010) s'interroge sur la *normativité des variantes* et que Mario Barra Jover (2010) propose dans la revue *Langue Française* un titre proche de l'oxymore : *formaliser la variation...* Sans rentrer dans la subtilité des réflexions de ces deux linguistes, il reste qu'il n'est plus guère possible aujourd'hui de se référer à une norme unique et de ne pas s'intéresser, au moins

en tant que repère, aux incidences que peuvent avoir sur la norme (et donc sur l'évaluation) les cinq grands types de *variations* qui font aujourd'hui consensus (Gadet 2000, 2003) :

- la variation *géographique* ou *diatopique*, qui documente les particularités lexicales de différents espaces géographiques. Ces variantes, taxées de régionalismes, ont pu faire l'objet de connotations péjoratives et étaient donc à exclure de la langue standard (*barrer*, régionalisme dans l'ouest de la France pour *fermer une porte à clé*, est d'utilisation générale au Québec) ;
- la variation *diachronique* ou *historique* qui se réfère à l'évolution de la langue, dont l'orthographe porte un témoignage privilégié ;
- la variation *diastratique* qui est corrélée à des facteurs sociaux (quat'[/[esplik]...) et qu'il est parfois délicat de dissocier de...
- la variation *diaphasique* qui renvoie à la *situation*. Qualifiée par le passé de « niveaux de langue », cette dénomination est abandonnée au profit de *registre* ou *style*.
- et, enfin, la variation qui renvoie à la *distinction du canal oral/écrit*. Cette dernière variation se trouve investie d'une manière particulière par les nouveaux médias ou ce que Sandrine Wachs et Corinne Weber nomment *les nouveaux univers langagiers* (Facebook, blogs, forums de discussion, tweeter, etc.). Le brouillage de frontières entre écrit et oral s'accompagne de normes qui échappent aux carcans traditionnels et marquent la nécessité pour la didactique d'aujourd'hui de « tendre vers une culture des normes, [ce] qui est alors une manière de revisiter l'ensemble de la problématique de la variation, en mettant en avant les ressources plurielles et singulières, qui sont autant de traces d'appartenances et d'identités ». (Wachs et Weber 2011 : 124.)

L'ensemble de ces paramètres variationnels a un impact sur l'élaboration des normes : la variation géographique a conduit à l'élaboration de normes locales que l'on a qualifiées de *normes endogènes* par contraste avec la référence aux normes de l'Hexagone que l'on qualifie alors de *normes exogènes* (voir Bavoux et al. 2008). Certes, *norme* est un mot ambigu et Alain Rey a dès 1972 proposé de distinguer une **norme objective** qui renvoie aux régularités du système de la langue, un **usage** qui réfère à un stade intermédiaire entre le système et la *parole* et une **norme prescriptive** qui s'appuie sur des évaluations juridiques (il faut dire/ne pas dire).

En 1997, dans l'ouvrage dirigé par Marie-Louise Moreau, outre la distinction faite ci-dessus entre *norme endogène* et *norme exogène*, sont distinguées **cinq types de normes** : (1) des *normes de fonctionnement* qui correspondent aux pratiques linguistiques des membres d'une communauté ; (2) des *normes descriptives* qui décrivent et rendent explicites les régularités observées mais qui sont moins nombreuses que les premières car toutes les parties de la langue n'ont pas été également analysées et décrites ; (3) des *normes prescriptives* qui hiérarchisent les normes à imiter. Les deux derniers types renvoient aux représentations :

(4) les *normes évaluatives* attribuent des valeurs esthétiques ou morales (une langue vulgaire, élégante...); (5) les *normes fantasmées*, éloignées des pratiques réelles, peuvent s'articuler aux quatre autres. La norme est ici considérée comme un idéal inatteignable et vis à vis de laquelle tout le monde serait pris en défaut.

D'un point de vue pédagogique, ces distinctions sont très utiles : lorsqu'on refuse une production telle que « je vais au coiffeur » on applique une norme prescriptive ; si au contraire, on l'accepte cela signifie que l'on prend comme référence une norme de fonctionnement. Si l'on veut faire un travail lexical pour la création d'une poésie à la manière d'Henri Michaux, on admettra des formes comme *flaquer* (sauter dans des flaques) ou *arriérer* (sur le modèle d'avancer) et dans ce cas on situera la créativité et la stylistique de l'écart dans le cadre du système de la langue...

La principale difficulté de l'enseignant de français, cependant, est de faire tenir tout à la fois la variation et ce que, faute de mieux, on pourrait nommer le *cœur de la langue*, c'est-à-dire cette langue commune qui permettra à l'étudiant allophone de se faire comprendre dans l'ensemble de la francophonie. Donner les moyens d'apprécier la variété langagière n'est donc pas pour autant dissociable de l'enseignement d'une variété que l'on pourrait qualifier de « transvariationnelle » sur le mode des *scripta transdialectales* du Moyen Âge. J.-P. Jaffré montre à propos des variations orthographiques combien elles dépendent – pour être comprises – de la référence à une norme orthographique permanente. Il donne l'exemple de *L'omnivore*, titre d'un ouvrage sur la diététique. Ici le simple ajout d'une lettre à un mot, le « h » initial à *omnivore*, crée un effet de sens, qui n'a pas d'équivalent à l'oral mais, nous dit-il, « ce jeu sémantique serait impossible sans le détour par une norme orthographique stable » (Jaffré 2010 : 321).

Au terme de ce parcours, l'effet de loupe sur ces quelques implications réelles des composantes du plurilinguisme et de la variation, tant du point de vue de l'apprenant que du point de vue de la langue qui fait l'objet de l'apprentissage, avait pour but de mettre en évidence des éléments qui avaient été peu mis en relief jusqu'ici. Le souhait est qu'ils permettent à l'enseignant une interaction plus éclairée et sereine au sein de sa classe ainsi que l'élaboration de démarches d'évaluation – sommative et formative – tenant compte de ces phénomènes.

Bibliographie

- BAGGIONI D. (1997), *Langues et nations en Europe*, Paris : Payot et Rivages.
- BARRA-JOVER M. (2010), « Le » français ou ce qui arrive lorsqu'un état de choses est observé comme une entité », *Langue française*, Les français : formaliser la variation, n° 168, p. 3-18.
- BAVOUX CL., LAMBERT-FELIX P. et WHARTON S. (2008), *Normes endogènes et plurilinguisme*, Paris-Lyon : ENS éditions.
- BERTRAND O. et SCHAFFNER I. dir. (2011), *Variétés, variations et formes du français*, Paris : éditions de l'École Polytechnique.
- BOETTINGER F., PARK J. et TIMMIS I. (2010), « Self-directed noticing for defossilisation : Three case studies », *International Journal of English Studies*, n° 10-1, p. 43-64.
- CANUT C. (2010), *Une langue sans qualité*, Limoges : Lambert-Lucas
- BONO M. (2008), « Influences interlinguistiques dans l'apprentissage d'une L3 : quand les langues secondes l'emportent sur la langue première », dans Moore D. et Castellotti éd.s. : *La compétence plurilingue : regards francophones*, Berne : Peter Lang, p. 147-166.
- CENOZ J. (2003), « Facteurs déterminant l'acquisition d'une L3 : Âge, développement cognitif et milieu », *Acquisition et interaction en langue étrangère AILE*, n° 18, p. 38-51.
- CENOZ J. (1997), « L'acquisition de la troisième langue : bilinguisme et plurilinguisme au pays basque », *Acquisition et interaction en langue étrangère AILE*, n° 10, p. 159-175.
- COSTA-GALLIGANI (2001), « Compétences, production et représentations métalinguistiques chez des migrants espagnols de longue date », *Langue française*, n° 131, p. 52-65.
- COSTE D. (2006), « Postface », dans Moore D., *op. cit.*, p. 245-255.
- COSTE D. (2008), « Éducation plurilingue et langue de scolarisation », *Les cahiers de l'ACEDLE* n° 5, p. 91-107.
- COSTE D. et al. (2012), « Hymes et le palimpseste de la compétence de communication Tours, détours et retours en didactique des langues », *Langage et société* 1/2012, n° 139, p. 103-123.
- DE ANGELIS G. et DEWAELE J.-M. (2011), *New Trends in Crosslinguistic Influence and Multilingualism Research*, Bristol/Buffalo/Toronto : Multilingual Matters.
- DUCHENE A. (2011), « Néolibéralisme, inégalités sociales et plurilinguisme : l'exploitation des ressources langagières et des locuteurs », *Langage et société*, 2011/2 n° 136, p. 81-108.
- GADET F. (2003), « Le style comme perspective sur la dynamique des langues » Introduction, *Langage et société*, 2004/3 n° 109, p. 1-8.
- GADET F. (2008), « Préface. Du bon usage du purisme quand on est linguiste », dans Paveau M.-A et Rosier L., *op. cit.*, p. 5-9.
- HELLER M. (2008), « Repenser le plurilinguisme : langue, postnationalisme et la nouvelle économie mondialisée », *Diversité urbaine*, www.erudit.org, p. 163-176.
- JAFFRE J.-P. (2010), « De la variation en orthographe », *Études de linguistique appliquée* n° 159, juillet-septembre, p. 309-323.
- LAPLANTINE F. et NOUSS A. (1997), *Le Métissage*, Paris, Flammarion.
- LÜDI G. (2004), « Pour une linguistique de la compétence du locuteur plurilingue », *Revue française de linguistique appliquée*, Vol. IX, p. 125-135.
- MARQUILLO LARRU M. (2003), *L'Interprétation de l'erreur*, Paris : Clé international.
- MARTINELL GIFRE E. et ERLENDSDÓTTIR E. (2005), *La conciencia lingüística europea : nuevas aportaciones de impresiones de viajeros*, Barcelona : Promociones y Publicaciones Universitarias.

- MATHEY M. (2005), « Plurilinguisme, compétences partielles et éveil aux langues. De la sociolinguistique à la didactique des langues », dans Pouliot M., Boulea E. et Bronckart J.-P., dir. *Repenser l'enseignement des langues : comment identifier et exploiter les compétences*, Lille : Presses du Septentrion, p. 139-159.
- MAURER B. (2011), *Enseignement des langues et construction européenne. Le plurilinguisme, nouvelle idéologie dominante*. Paris : Éditions Archives Contemporaines
- MERLIN-KAJMAN (2003), *La langue est-elle fasciste ? Langue, pouvoir, enseignement*, Paris : Seuil.
- MOREAU M.-L. (1197), *Sociolinguistique : les concepts de base*, Sprimont : Mardaga.
- PAVEAU M.-A. et ROSIER L. (2008), *La langue française. Passions et Polémiques*, Paris : Vuibert.
- PY B. (2005), « Représentations langagières et malentendus dans l'enseignement plurilingue », *Éducation et sociétés plurilingues* n° 19, décembre 2005, p. 73-84.
- SELINKER L. (1972), « Interlanguage », *International Review of applied linguistics*, vol. X-3, p. 209-231.
- SIOUFFI G. (2011), « Les variantes ont-elles une normativité ? », dans Bertrand O. et Schaffner I. dir., *Variétés, variations et formes du français*, Paris : éditions de l'École Polytechnique p. 13-30.
- TREVISOL P. et RAST R. (2004), « Présentation. L'acquisition d'une L3 », *Acquisition et interaction en langue étrangère*, AILE, n° 24.
- SPAETH V. dir. (2010), « Le français au contact des langues », *Langue française* n° 167.
- TRUCHOT C. (2010), « L'enseignement supérieur en anglais véhiculaire : la qualité en question », http://www.diploweb.com/spip.php?page=imprimeretid_artic... Consulté de 8 novembre 2011.
- WACHS S. et WEBER C. (2011), « ça l'fé grave ou le statut des nouveaux univers langagiers en didactique du français et des langues » dans Bertrand O. et Schaffner I. dir., *Variétés, variations et formes du français*, Paris : éditions de l'École Polytechnique p. 111-126.

E nseignement bilingue et évaluation : réflexions sur la conception de tests de niveaux

LAURENT GAJO, JEAN-MARC LUSCHER
ET CECILIA SERRA
UNIVERSITÉ DE GENÈVE, SUISSE

I ntroduction

L'évaluation en didactique des langues étrangères s'est donné de nouvelles marques suite à la publication du CECR en 2001. Ce document de référence a en effet permis de regarder de manière transversale, comparable et différenciée les compétences disponibles à un moment donné au sein du répertoire pluriel d'un locuteur-apprenant. Élaboré autour de la notion de compétence plurilingue (et pluriculturelle), les instruments d'évaluation du CECR ne permettent toutefois pas de saisir le plurilinguisme en tant que tel. Cette compétence prétendument unique, articulée, ne se laisse ainsi saisir que par une évaluation cumulée de ses composantes. Une telle limite ne saurait convenir aux approches relevant de la didactique du plurilinguisme et travaillant justement sur les articulations entre langues. Plus encore, elle peine à fournir les outils nécessaires à l'évaluation d'un enseignement de type bilingue, où les langues interviennent dans le cadre de disciplines non linguistiques (DNL).

Dans cet article, après avoir posé les enjeux particuliers de la didactique du plurilinguisme et, en particulier, de l'enseignement bilingue, nous reviendrons sur des outils d'évaluation comme le CECR ou le CARAP et nous nous arrêterons sur les défis que représente l'élaboration d'un test pour une filière bilingue. Nous baserons une partie de notre propos sur un projet mené actuellement à l'École internationale de Genève. En conclusion, nous reviendrons sur des notions comme celles de *répertoire pluriel* et de *ressource*.

Didactique du plurilinguisme, enseignement bilingue et évaluation

La didactique du plurilinguisme (DPL) représente un paradigme nouveau, résultant de certaines évolutions en didactique des langues, mais se distinguant de celle-ci. Candelier (2008) parle d'approches plurielles, par opposition à des approches singulières, focalisées sur une langue en particulier. Il classe parmi les approches plurielles la pédagogie interculturelle, l'éveil aux langues, la didactique intégrée et la didactique des langues voisines (intercompréhension entre langues proches). Il n'y range pas nécessairement l'enseignement bilingue, ce que fait de son côté Gajo (2009), qui considère que celui-ci réalise pleinement les enjeux de la didactique du plurilinguisme en travaillant, de manière multi-intégrée, entre les langues et entre les langues et les disciplines.

La DPL comporte une dimension sociale (voir l'ouvrage en hommage à Louise Dabène coordonné par Billiez 1998), contextualisée et dépasse les seuls objets linguistiques. De ce point de vue, elle correspond bien à la perspective actionnelle postulée par le CECR, les langues intervenant dans des tâches concrètes d'interaction avec l'autre, d'accès à des contenus divers, dans des environnements plus ou moins familiers mais porteurs de sens. Ainsi, par exemple, les cultures éducatives et les disciplines scolaires constituent aussi bien des objets que des conditions d'apprentissage. Évaluer les compétences développées dans le cadre d'approches relevant de la DPL suppose de considérer la langue comme en contact et en contexte : la langue avec d'autres langues, la langue dans des contenus particuliers, la langue dans des tâches données. Il n'est dès lors pas sûr que la notion de compétence, largement débattue en didactique (voir notamment Dolz et Ollagnier 2002 ; Pekarek Doehler 2009), convienne ici. Nous y préférons celle de ressource, sur laquelle nous reviendrons en conclusion.

Pour illustrer les enjeux nouveaux posés par la DPL en termes d'évaluation, nous évoquons brièvement chacune des approches (voir Gajo, à paraître) et nous nous arrêtons plus longuement sur la dernière, l'enseignement bilingue :

- *enseigner vers la L2* : il s'agit d'une didactique de la sensibilisation, de l'éveil, dans laquelle nous rangeons certains aspects de la pédagogie interculturelle et de l'éveil aux langues ; le travail envisagé porte davantage sur les représentations que sur les pratiques et engendre un retour métalinguistique sur la L1 (et la culture 1) ; l'évaluation devrait dès lors porter sur la perception même de la diversité, l'identification des langues et des cultures ;

- *enseigner avec la L2* : il s'agit d'une didactique de la comparaison, de la contrastivité, dans laquelle nous classons en priorité la didactique des langues voisines, même si l'on pourrait y placer aussi, dans une certaine mesure, l'éveil aux langues ; l'évaluation doit porter ici sur le développement d'une « méta-grammaire », de règles de passage, sur une perception des niveaux de structuration linguistique et les spécificités des quatre compétences ;
- *enseigner la L2 comme une des L* : il s'agit ici de la didactique intégrée, qui consiste à enseigner une langue en tenant compte des autres langues de l'école et même du curriculum expérientiel des élèves (voir Beacco et al., 2010) ; les ressources travaillées s'apparenteront à celles de la catégorie précédente, mais demanderont avant tout une réflexion en amont de la part des enseignants ;
- *enseigner en L2* : il s'agit de l'enseignement bilingue, qui consiste à proposer une partie au moins du programme scolaire dans une L2 et qui touche les langues dans les disciplines, ou les disciplines par les langues.

L'enseignement bilingue se situe à l'articulation entre langues et disciplines et suppose donc un processus d'intégration que l'on retrouve dans les expressions européennes *CLIL* (Content and language integrated learning ; voir notamment Coyle, Hood et Marsh 2010) et *EMILE* (Enseignement d'une matière par l'intégration d'une langue étrangère ; voir notamment Baetens Beardsmore 1999). Tout en étant conçues comme un terrain favorisant l'usage et l'apprentissage des langues, les DNL peuvent bénéficier de la diversité du travail bilingue. L'enseignement bilingue implique donc une prise au sérieux de la dimension verbale de l'enseignement et de toutes les DNL, qu'il convient dès lors de rebaptiser *DdNL* (disciplines dites non linguistiques ; Gajo 2009). Le développement des savoirs en L2 se tisse ainsi dans une dynamique complexe traversée par :

- le lien entre contenu et langue ;
- les références croisées entre L1 et L2, la première fournissant un substrat de représentations au façonnage des savoirs ;
- l'articulation entre savoirs communs et savoirs spécialisés ;
- l'inscription des savoirs dans des formes discursives culturellement situées (types de textes, d'énoncés, de supports, etc.), plus ou moins transférables de L1 à L2, plus ou moins spécifiques de l'école ou de telle ou telle DdNL.

En bref, évaluer les productions des élèves en enseignement bilingue exige une prise en compte de l'*effet langue* dans le contenu, ou de l'*effet contenu* dans le discours. L'élaboration des savoirs s'organise en effet dans une intrication entre langue(s), contenu(s) et discours qu'il s'agit aussi d'évaluer comme telle.

N^otre terrain : l'École Internationale de Genève

Fondée en 1924, l'École Internationale de Genève est aujourd'hui organisée en huit écoles (du pré-primaire au secondaire supérieur) réparties sur trois campus, et encadre environ 4 300 élèves. La population des élèves est hétérogène à plus d'un titre : on recense 81 langues premières, des cultures diversifiées et stratifiées, du fait de la composition familiale et/ou du parcours international de l'élève, et une durée de fréquentation scolaire fluctuante due à la mobilité des familles. Pour donner un exemple, le cas le plus complexe est celui des élèves qui fréquentent l'école pour une durée de trois ans et qui arrivent sans connaître ni l'anglais ni le français, les deux langues d'enseignement. En ce laps de temps, ils doivent être mis à niveau dans au moins une des deux langues tout en poursuivant leur cursus scolaire.

Un autre facteur d'hétérogénéité, souvent sous-estimé, vient du corps enseignant qui était encore dernièrement recruté sur la base de l'appartenance linguistique, en pays anglophones ou francophones. De ce point de vue, deux (macro) traditions didactiques et pédagogiques se confrontent et doivent pouvoir converger lorsque ces mêmes enseignants sont appelés à collaborer dans le cadre de l'enseignement bilingue.

L'École Internationale offre, par ailleurs, divers types d'enseignement en deux langues. On y trouve des filières en anglais parallèles à celles en français, des programmes « duels » où chaque langue encadre la moitié du cursus et où l'élève a le choix d'étudier une discipline dans l'une ou l'autre langue, et les enseignements bilingues, qui prennent de plus en plus pied, aussi dans la tentative d'établir à long terme un modèle plus gérable, cohérent et efficace d'enseignement.

Plusieurs expériences d'enseignement bilingue ont maintenant cours à l'École Internationale de Genève, allant du modèle CLIL/EMILE à environ 50 % dans les deux langues sur l'année scolaire, au récent modèle d'origine canadienne de double immersion totale, en alternance sur deux semestres, que nous sommes en train d'évaluer et dont nous parlons ici.

La particularité de l'organisation de ce programme bilingue sur le site de la Châtaigneraie est en effet que la langue « une », c'est-à-dire la langue de scolarisation, n'est pas la même durant toute l'année : durant le premier semestre, les cours sont donnés en français alors que, durant le deuxième semestre, ils sont systématiquement donnés en anglais. Les cours de langue, en français et en anglais, sont quant à eux dispensés durant toute l'année scolaire. Une continuité qui vise à faciliter le relais d'une langue à l'autre et ensuite à maintenir les acquis, lorsque l'enseignement va basculer vers l'autre langue.

Les élèves relèvent donc *de facto* de trois, voire quatre catégories du point de vue de leur accès aux savoirs enseignés :

- 1) élèves dont la langue 1 est le français
- 2) élèves dont la langue 1 est l'anglais
- 3) élèves dont la langue 1 est autre

Au sein de cette dernière catégorie, on peut théoriquement distinguer les élèves ayant de meilleures connaissances en français de ceux pour lesquels l'anglais est la L2 la mieux connue. De fait, une grande partie des élèves de la catégorie 3 relèvent du deuxième cas : L1 autre/L2 anglais /L3 français. L'École étant établie en Suisse romande, les élèves seraient donc en milieu homoglotte la première partie de l'année (enseignement en français) et alloglotte la seconde. Mais dans les faits, certaines familles installées à Genève depuis peu et se considérant « en transit » vivent dans l'espace virtuel de la « Genève internationale » et utilisent quasi exclusivement l'anglais dans leur vie professionnelle comme dans leur vie privée : la langue ordinaire de communication des enfants est ainsi l'anglais et il n'est pas exagéré d'avancer que c'est lors de l'enseignement en anglais qu'ils sont en milieu homoglotte. On aura compris que certains élèves sont placés en situation d'alloglossie quelles que soient les langues de scolarisation et de socialisation. La notion de répertoire pluriel prend ainsi tout son sens et il est difficile de dresser un tableau clair des différentes situations linguistiques individuelles.

Il est fondamental de cerner les enjeux des caractéristiques de ce public, qui rendent difficilement opérants les outils habituels de la didactique des langues et les cadres de référence actuels.

É valuer l'enseignement/apprentissage bilingue : utilité et limites du CARAP et du CECR

Nous discuterons ici rapidement de deux instruments de référence du Conseil de l'Europe – le CARAP et le CECR.

LE CARAP

Le CARAP (Cadre de référence pour les approches plurielles des langues et des cultures <http://carap.ecml.at>) paraît de prime abord adapté à notre problématique : l'enseignement bilingue en contexte pluri-lingue de l'École Internationale relève certainement d'une approche plurielle. Pourtant, il ne nous semble pas être directement utilisable pour notre étude, ni dans ses objets ni dans ses objectifs. Il concerne en effet explicitement quatre approches plurielles et aucune ne cerne notre objet d'études :

- *La didactique intégrée* : elle ne recouvre que très imparfaitement notre domaine de recherche, dans la mesure où les élèves de nos différents dispositifs n'ont pas forcément l'occasion de mettre en lien les différentes langues de l'École ou les enseignements de L2 et en L2 ;
- *L'approche interculturelle* : l'interculturalité est un fait à l'École Internationale, mais ne constitue ni une matière, ni même une préoccupation ; elle semble au contraire être perçue, voire conçue comme « allant de soi » ;
- *L'éveil aux langues* : cette dimension existe de fait de par la prise en compte de deux langues d'enseignement et la présence d'une hétérogénéité linguistique constitutive, mais elle n'intervient pas en tant qu'approche dans le programme qui nous occupe ici ;
- *L'intercompréhension entre les langues parentes* : cette démarche n'est pas non plus prise en compte de manière explicite, notamment en raison d'une focalisation didactique première sur le couple « français-anglais », langues pourtant largement apparentées.

Certes, les instruments proposés par le CARAP concernent le développement de la compétence plurilingue et en cela il nous intéresse, c'est pourquoi nous abordons encore deux de ses aspects : les descripteurs et le concept de « ressource ».

Les **descripteurs** du CARAP sont répartis en trois catégories : les savoirs (« K », *knowledge*), les savoir-être (« A », *attitudes*) et les savoir-faire (« S », *skills*). Par ailleurs, dans les documents disponibles en ligne, ils sont ordonnés sous forme de tableaux dont les colonnes représentent les années scolaires : il serait ainsi possible de proposer à l'École Internationale de compléter ses cursus par des cours de sensibilisation aux langues et à leur emploi.

Les **savoirs** sont d'ordre métalinguistique. Exemple :

K6 : Savoir qu'il existe entre les langues des ressemblances et des différences

Toutes les connaissances viennent de l'extérieur de l'individu, elles lui parviennent soit par transmission explicite par un enseignement, soit par la compréhension de l'environnement et l'interprétation du contexte. On peut donc s'attendre que, même en l'absence d'enseignement spécifique, les élèves de l'École Internationale, dont la plupart pratiquent plus de trois langues, possèdent quelques connaissances sur les langues.

Les **savoir-être** sont, au moins pour certains d'entre eux, largement présents à l'École Internationale, par exemple :

A2 : Sensibilité à l'existence d'autres langues/cultures/personnes // à l'existence de diversité des langues/cultures/personnes

Il est ici décrit la raison même de l'existence de l'institution et la condition pour y inscrire son enfant...

Les **savoir-faire** réfèrent pour l'essentiel aux capacités de *percevoir*, *d'observer* et *d'analyser* les savoirs. Ils leur sont ainsi indissociables.

Ces descripteurs représentent ce qui est nommé dans le CARAP des « ressources ». Cette notion, nouvelle, nous paraît cerner heureusement un concept nécessaire dans le processus d'enseignement/apprentissage des langues et jusqu'ici imparfaitement désigné tantôt par « compétences », « capacités », « dispositions », « performances » entre autres. Il s'agit de ce que l'individu apprenant/utilisateur de la langue doit ou peut mobiliser pour effectuer la tâche requise. Bien que mises en œuvre par un individu dans une circonstance donnée, les ressources peuvent être (au moins partiellement) identifiées, listées et faire l'objet d'une prise de conscience. Cette façon de cerner ce phénomène par cette appellation nous semble intéressante et à même d'être exploitée au-delà du CARAP, par exemple pour préciser certains aspects du CECR. Mais voyons maintenant comment nous comptons adopter/adapter le CECR, les hypothèses qui sous-tendent notre démarche et les moyens mis en œuvre pour les vérifier/infirmer.

LE CECR DANS UN PROGRAMME BILINGUE

Le CECR est en principe strictement fondé et adapté pour l'enseignement/apprentissage de la L2. Il affirme l'intérêt et la richesse des compétences plurilingues, prône la connaissance de plusieurs langues étrangères et promeut l'intercompréhension, mais ne se penche pas sur la langue première, ni du point de vue acquisitionnel, ni pour l'évaluation des compétences langagières. Il ne paraît donc pas adéquat pour décrire les savoirs, les savoir-être et les savoir-faire d'élèves engagés dans un programme tel que celui de l'École Internationale. Nous avons pourtant décidé de l'exploiter en nous fondant sur les trois hypothèses suivantes :

Première hypothèse (H1) :

L'utilisation de la langue d'apprentissage à des fins utiles, pour suivre l'enseignement d'une autre discipline et apprendre/acquérir de nouvelles connaissances, favorise l'apprentissage de cette langue.

Il s'agit de l'application de la perspective actionnelle au domaine scolaire, prônée par le CLIL et c'est l'hypothèse fondatrice de notre démarche auprès de l'École Internationale de Genève, nous l'avons déjà évoqué plus haut.

La deuxième hypothèse repose sur le postulat suivant : les savoirs déclaratifs sont propres à chaque langue (moyennant les potentialités de transferts entre langues proches), mais les savoir-faire sont transposables d'une langue à l'autre : c'est l'aspect transversal des activités langagières. Autrement dit, lorsqu'un apprenant est en mesure de développer une courte argumentation monologique en L1, il a la capacité de le faire également en L2, à condition de posséder les connaissances nécessaires dans cette langue. Notre hypothèse (qui est également prise en compte dans un enseignement CLIL) est que le mouvement inverse est également possible.

Deuxième hypothèse (H2) :

L'apprenant/utilisateur d'une langue peut s'approprier un savoir-faire d'abord en L2, avant de le transposer dans sa L1.

De là va découler la troisième hypothèse.

Troisième hypothèse (H3), découlant de H1 + H2 :

Les descripteurs du CECR sont exploitables pour l'évaluation de capacités en L1.

Cette dernière hypothèse est pour nous à la fois la plus importante et la plus problématique. Elle est indispensable à notre démarche, mais nécessite quelques précisions méthodologiques.

Premièrement, comme évoqué dans son sous-titre, nous fondons H3 sur les deux autres hypothèses, liées à l'enseignement CLIL : ce type d'enseignement tisse un lien fort entre L1 et L2. Le fait que les connaissances scolaires, traditionnellement transmises en L1, soient transmises ici en L2 donne un sens à l'apprentissage de L2 (H1) et entraîne que les stratégies et les savoir-faire qu'elles impliquent nécessairement sont aussi appris/acquis en L2 (H2).

Deuxièmement, il est raisonnable de distinguer les savoir-faire décrits dans le CECR pour les 2 niveaux A de ceux qui correspondent aux 4 niveaux B/C. Les premiers réfèrent, au moins pour les activités orales, à des capacités qui seraient, en L1, de l'ordre des acquisitions préscolaires, celles de l'enfant de 4 à 5 ans. Les seconds, en revanche, ne peuvent exister sans enseignement explicite, probablement accompagnés d'observations et d'inférences complexes de la part des utilisateurs :

B1 : Peut écrire des textes articulés simplement sur une gamme de sujets variés dans son domaine en liant une série d'éléments discrets en une séquence linéaire. (Production écrite générale, CECR 2005, p. 51)

B2 : Peut développer méthodiquement une argumentation en mettant en évidence les points significatifs et les éléments pertinents. (Monologue suivi, CECR 2005, p. 50)

Ces deux exemples montrent que les descripteurs de ces niveaux peuvent référer à des « ressources », au sens du CARAP et que, comme tels, ils correspondent à des activités ou à des tâches développées dans le cadre scolaire, également en L1.

Nous ne proposons pas pour autant de conserver tels quels les quatre derniers niveaux du CECR et d'attribuer les niveaux B1 à C2 aux natifs, sans autre précaution. Cela serait inadéquat pour au moins deux raisons. La première est que l'attribution à un niveau résulte de la conjonction entre les savoir-faire, actualisés en LE, et les savoirs déclaratifs dans cette langue (essentiellement le lexique et la grammaire). Or, dans la LM, des formes grammaticales peuvent être connues et maîtrisées même si elles ne sont pas ou sont peu utilisées. Et la deuxième est que certaines activités décrites dans les niveaux supérieurs – C1 et surtout C2 – restent largement inaccessibles à des natifs n'ayant pas suivi d'enseignement spécifique.

L'entrée que nous proposons pour regrouper les ressources pertinentes en L1 repose sur la typologie textuelle et, en particulier, sur la distribution en « description », « explication » (relations de cause-conséquence), « narration » et « argumentation ». *A priori*, ces quatre catégories regroupent tous les types de textes, en réception, comme en production. Elles sont présentées ici dans l'ordre de difficulté de leur compréhension/conception. Il est entendu que chaque type de texte peut être conçu de manière plus ou moins simple ou complexe. La capacité de mobiliser des ressources pour décrire intervient avant celle qui est nécessaire pour raconter ou pour argumenter, mais lorsque l'utilisateur est capable de raconter ou d'argumenter, il est non seulement toujours capable de décrire, mais il le fait de manière plus complexe.

Le test que nous avons conçu fait systématiquement intervenir, de manière contrôlée et isolée, un moment de description (d'images), un moment de justification et un moment d'argumentation, avec une alternance entre la langue de scolarisation (L1, ici le français) et une langue enseignée/d'enseignement (L2, ici l'anglais). Nous avons choisi un entretien court avec alternance des deux langues pour mesurer, d'une part, les capacités de chaque élève de l'échantillon à produire des énoncés dans les deux langues (donc toujours au moins une L2) et vérifier, d'autre part, la souplesse de passer d'une langue à l'autre qui caractérise les possesseurs d'un répertoire pluriel.

Évaluer l'alternance de codes représente un vrai défi pour l'enseignement bilingue, mais touche bien le cœur du fonctionnement d'un répertoire pluriel. Comme le souligne un texte récent du Conseil de l'Europe, « pour qu'elle soit intégrée comment un élément à évaluer, l'alternance codique doit impérativement être réalisée avec un minimum de maîtrise – dans une moindre mesure pour les évaluations informelles, bien plus pour les évaluations standardisées où les enjeux sont élevés » (Lenz et Berthele 2010 : 25). Ce même texte suggère par ailleurs l'intérêt d'une meilleure exploitation de la compétence de médiation prévue par le CECR. Celle-ci est particulièrement pertinente dans les classes bilingues, où il s'agit de développer des stratégies à l'articulation des deux langues d'enseignement et, en outre, au croisement des langues et des disciplines.

E *njeux de l'élaboration d'un test en classe bilingue*

L'évaluation d'un enseignement/apprentissage bilingue soulève souvent la question de savoir si on évalue la langue ou le contenu, ou si on a des instruments de mesure suffisants pour départager l'impact de l'un sur l'autre (cf. Short 1993 ; Coyle, Hood et Marsh 2010). Une

réponse satisfaisante consiste à mettre les élèves dans une situation dans laquelle ils doivent avoir recours à des typologies discursives, comme décrire, raconter, argumenter, etc., pour accomplir des activités que l'enseignement aura par ailleurs entraînées, sur la base d'un contenu déjà connu. Ce faisant, la production de l'élève s'échelonne sur un axe de complexité, allant du plus simple au plus élaboré, à la mesure de ses capacités langagières du moment (Serra 2007). Et puisque chaque type textuel s'appuie sur des formes transversales du langage, la composante linguistique pourra être analysée avec discernement comme une étape ou l'évolution d'un « savoir-dire ». Autre aspect non négligeable, tout élève sera en mesure de s'exprimer, même si sa maîtrise linguistique est basique, car chaque type textuel peut être réduit à une phrase simple.

Dans le cas qui nous occupe, nous avons fait passer le test à un échantillon de 12 élèves de 5^e et 12 élèves de 6^e année primaires, sélectionnés prioritairement sur la base des critères linguistiques mentionnés ci-dessus (L1 = français, anglais ou autre) et en tenant compte également de leurs résultats scolaires. L'échantillon comprenait donc pour chacune des catégories linguistiques 4 élèves : un plus faible, un plus avancé et deux moyens. Ce critère externe nous a été utile pour confronter notre évaluation à celle de l'école et, par exemple, pour questionner une mauvaise production et la situer par rapport au comportement observé régulièrement en classe. Le test a été passé à la fin du premier semestre durant lequel l'enseignement avait été donné, comme dit précédemment, en français.

Le contenu du test portait sur des projets d'enseignement que les élèves avaient poursuivis au début ou au milieu de la période scolaire considérée, l'un sur l'énergie, l'autre sur les biomes (ou écosystèmes). Au début du test, l'élève devait répondre à des questions de « mise en route », sur son âge, ses langues et son trajet, le matin même, pour l'école, le tout donnant lieu à un discours de type informel. L'élève avait ensuite une série de petites cartes représentant quelques éléments du contenu. Dans le cas de l'énergie, par exemple, des photos de diverses sources d'énergie (éclair, panneaux solaires, moulin à eau, etc.) ou de consommation d'énergie (ampoules, voitures, etc.). L'élève devait les décrire et ensuite les classer selon ses critères, qu'il devait par la suite justifier. Ces deux étapes faisaient donc appel à la description, à la catégorisation et à la justification, autant d'activités cognitives familières que le développement langagier en L2 pouvait ou non amplifier.

La présence d'élèves de langue 1 s'est avérée de grande utilité, car elle a montré que la capacité de justifier ses choix et d'argumenter était en réalité beaucoup plus liée que nous ne le pensions à la capacité cognitive d'organiser la pensée qu'aux moyens linguistiques à disposition ; ainsi, quelques élèves de langue 2 avaient une meilleure production que ceux de langue 1, même si « objectivement » ces derniers auraient dû avoir de meilleurs résultats.

Après ces étapes, l'enquêtrice lisait aux élèves un court texte (environ 7 lignes) dans lequel est exprimée une opinion sur le thème concerné. Dans le cas de l'énergie, par exemple, le texte incitait à un usage écologique quotidien tout en soulignant l'importance de la production nucléaire. Les élèves n'avaient pas le texte sous les yeux et devaient seulement écouter. À la fin de la lecture, l'enquêtrice donnait la consigne de prendre position sur l'opinion exprimée par le texte et de changer, pour ce faire, de langue, c'est-à-dire de passer – de manière inattendue – du français à l'anglais. Une activité complexe qui mettait en lien direct la compréhension et la production orale, ainsi que la langue 1 et la langue 2.

Un premier résultat est que tous les élèves plurilingues (3^e catégorie) ont changé de langue sans aucune hésitation, démontrant par là leur grande flexibilité langagière, alors que quelques élèves de la 2^e catégorie (anglais en L1) ont brièvement hésité, même si ce changement leur permettait de s'exprimer dans leur langue forte : il leur était apparemment plus accessible de rester en situation de test, donc dans la langue d'apprentissage.

À la fin de l'année scolaire, le même test doit être passé aux mêmes élèves sur d'autres contenus et en suivant le rythme du programme : la langue du test sera à ce moment l'anglais, avec un retour au français en fin d'entretien. La comparaison des deux étapes doit permettre d'évaluer l'efficacité de ce type particulier d'enseignement bilingue avec, à l'appui, des tests ciblés sur les compétences disciplinaires.

Conclusion

Évaluer les acquis spécifiques de l'enseignement bilingue reste un défi majeur pour la didactique, qui peine à sortir de paradigmes « singuliers ». L'*effet langue* dans le contenu se laisse approcher, notamment, par l'analyse de compétences discursives impliquant des processus cognitifs comme la catégorisation et présentant une certaine transversalité d'une langue à l'autre. De son côté, l'*effet contenu* dans le discours demande encore l'identification d'observables à systématiser pour l'évaluation. Par ailleurs, le travail bilingue se laisse appréhender par un recours délibéré à l'alternance codique, dont la didactisation demeure un enjeu.

Pour l'élève, le répertoire pluriel et les contenus disciplinaires sont tout à la fois des objets à développer et des ressources potentielles en lien avec des tâches didactiques particulières. Le défi épistémologique consiste alors à articuler la notion de ressource, fortement contextualisée, aux besoins d'une perspective d'évaluation tendant vers une forme de standardisation.

Bibliographie

- BAETENS BEARDSMORE H. (1999), «Consolidating experience in plurilingual education», dans Marsh D. et Marsland B. (Eds.), *CLIL Initiatives for the Millennium* (p. 24-30). Université de Jyväskylä, Continuing Education Centre.
- BEACCO J.-C. et al. (2010), *Guide pour le développement et la mise en œuvre de curriculums pour une éducation plurilingue et interculturelle*, Conseil de l'Europe.
- BILLIEZ J. (éd.) (1998), *De la didactique des langues à la didactique du plurilinguisme. Hommage à Louise Dabène*. Grenoble, CDL-LIDILEM.
- CANDELIER M. (2008), «Approches plurielles, didactiques du plurilinguisme: le même et l'autre», *Cahiers de l'ACEDLE* 5/1, p. 65-90.
- COYLE D., HOOD Ph. et MARSH D. (2010), *CLIL : Content and Language Integrated Learning*, Cambridge University Press.
- DOLZ J. et OLLAGNIER E. (éds) (2002), *L'énigme de la compétence en éducation*, Bruxelles, De Boeck Université.
- GAJO L. (2009a), «De la DNL à la DdNL : principes de classe et formation des enseignants», *Les Langues modernes* 4-2009, p. 15-24.
- GAJO L. (à paraître), «Plurilinguisme : enjeux sociaux et ressources éducatives», *Revue japonaise de didactique du français*.
- LENZ P. et BERTHELE R. (2010), *Prise en compte des compétences plurilingue et interculturelle dans l'évaluation*, Conseil de l'Europe.
- PEKAREK DOEHLER S. (2009), «Démystifier les compétences : vers une pratique écologique d'évaluation», dans Galatanu O., Pierrard M. et Van Raemdonck D. (éds), *Construction du sens et acquisition de la signification linguistique dans l'interaction*. Berne : Peter Lang, p. 19-38.
- SERRA C. (2007), «Assessing CLIL at primary school : A longitudinal study» *International Journal of Bilingual Education and Bilingualism*, 10, 5, p. 580-602.
- SHORT D. (1993), «Assessing integrated language and content instruction» *TESOL Quarterly*, 27, 4, p. 627-56.

D *'une histoire à
ses perspectives :
évaluer après
le CECR ?*

DANIEL COSTE

Petit retour à Rüşchlikon

DANIEL COSTE

ÉCOLE NORMALE SUPÉRIEURE DE LYON, FRANCE

Les deux Rüşchlikon

On ne sait peut-être pas que le nom de Rüşchlikon, village suisse à proximité de Zürich, est doublement associé aux travaux du Conseil de l'Europe dans le domaine des langues vivantes. D'une part, c'est là, en 1971, au cours d'une rencontre organisée par la Suisse, que ce qui allait devenir la série des projets « langues vivantes » étalés sur un quart de siècle (1973-1998) et pilotés continûment par John Trim a été préfigurée ; le volume *Systèmes d'apprentissage des langues vivantes par les adultes* paraît en 1973, rassemblant, entre autres, des contributions de John Trim, David Wilkins, René Richterich, Jan van Ek. L'ambition est alors de définir des objectifs par unités capitalisables, susceptibles de donner lieu à évaluation et à validation plus clairement spécifiées, plus accessibles et plus étagées que les diplômes traditionnels, à critères souvent opaques et demandant de longues années de préparation, et de plus en plus considérés comme inadaptés à la formation continue des adultes. C'est alors que la notion de « besoins langagiers » est mobilisée et que le dispositif de caractérisation notionnelle-fonctionnelle des objectifs d'apprentissage est mis en place. Il sera à la base des « niveaux-seuils » dont les premières réalisations sont publiées peu après : *The Threshold Level* (Van Ek 1975) et *Un niveau-seuil* (Coste et al., 1976), accompagnant le mouvement des approches dites communicatives. Les niveaux-seuils sont alors conçus comme des socles sur lesquels bâtir d'autres niveaux.

Le second Rüşchlikon, vingt ans plus tard, est celui, mieux connu, qui marque en 1991 les origines de deux instruments appelés à un certain avenir : le *Cadre européen commun de référence pour les langues* (dorénavant CECR ; Conseil de l'Europe 2001) et le *Portfolio européen des langues* (PEL).

Mon propos est ici de revenir sur certains aspects de ce second Rüşchlikon, mais il faut d'abord souligner que les questions relatives à l'évaluation ont été débattues au cours des années 1970 et 1980 et ne se focalisaient pas uniquement sur les gradations possibles à l'intérieur d'un système d'unités capitalisables et sur la définition de niveaux

inférieurs ou supérieurs au Niveau-seuil, tels que, pour l'anglais et respectivement, *Waystage* et *Vantage*.

On ne saurait en effet oublier que tout un courant, interne aux travaux du Conseil, défend alors le principe d'une centration sur l'apprenant, sur sa capacité d'autonomisation progressive et sur ses possibilités d'auto-évaluation. Les chercheurs du CRAPEL de Nancy, tels Michel Kuhn et Henri Holec, ne sont guère favorables à une multiplication et à une hiérarchie d'un ensemble de niveaux gradués mobilisant les mêmes catégories et types de descripteurs que les niveaux-seuils. À cet égard, l'étude de Henri Holec, *Autonomie et apprentissage des langues étrangères* (1979), marque une étape importante.

L'évaluation est l'objet majeur du second Rüsçhlikon. Même si ce n'est pas le seul. Le titre du symposium est clair : « Transparence et cohérence dans l'apprentissage des langues vivantes en Europe : objectifs, évaluation, certification »¹. Sont intéressées et présentes des institutions qui délivrent des certificats et diplômes, tels le Local Examinations Syndicate de l'université de Cambridge, le Goethe Institut, les Universités populaires, l'Alliance française, les Universités pour étrangers de Sienne et de Pérouge. Le constat général est que les certifications existantes, très nombreuses (et singulièrement pour l'anglais) sont difficiles à situer les unes par rapport aux autres et que les usagers, quels qu'ils soient (apprenants, enseignants, administrateurs, employeurs), peinent à s'y retrouver. Les organisateurs suisses, au nombre desquels les Eurocentres (fondation créée par la chaîne de magasins Migros), ont des propositions à mettre en débat : projet d'un cadre commun de référence dont des échelles de niveaux commencent à être élaborées à partir de recherches empiriques menées en Suisse par Brian North (Eurocentres) et Günther Schneider (Université de Fribourg) et projet d'un portfolio des langues, porté par Rolf Schärer, alors directeur général des Eurocentres.

Parmi les participants français ou de pays francophones, on relève en particulier les noms de Louis Porcher (Université Paris 3), René Richterrich (Université de Lausanne), Annie Monnerie (CIEP), Simonne Lieutaud (Alliance française). Alors professeur à l'Université de Genève, j'ai à assurer l'exposé introductif à la session des travaux consacrés aux objectifs au-delà du niveau-seuil. L'intitulé de cet exposé est « Remarques sur la conception des objectifs et des modalités d'évaluation au-delà du niveau-seuil ».

Quelles options au-delà du niveau-seuil ?

C'est la position que je présentais alors sur laquelle je voudrais revenir ici pour la commenter rétrospectivement². Non pour revendiquer un

1. Le rapport du Symposium de Rüsçhlikon de 1991 a été produit par Brian North en 1993 et n'a pas donné lieu à publication spécifique. On peut toutefois en consulter la version scannée, contenant le texte des différentes interventions, à l'adresse figurant dans les références finales de la présente contribution.

2. Pour plus de commodité dans la distinction des voix énonciatives, je me permettrai de rappeler et résumer les analyses que je faisais dans cet exposé en référant à «DC91».

point de vue personnel, mais pour pointer les options qui se présentaient et dont, pour l'essentiel, seule une a été retenue au bout du compte, celle que l'on connaît avec les niveaux de compétence étagés de A1 à C2. Pour autant, certaines des orientations que je vais rappeler ci-dessous n'ont pas été totalement perdues de vue³.

Le point de départ de l'analyse, commandant l'essentiel de l'argumentation, consiste à rappeler que l'ensemble des niveaux-seuils qui avaient été élaborés pour différentes langues, tout en présentant certaines variations dans le dispositif de catégorisation, répondait à un modèle central commun et que la question était désormais de savoir s'il convenait de poser une différence sensible entre l'avant et l'après niveau-seuil

Au-delà du niveau-seuil, la question est bien de savoir si le modèle d'ensemble pour la construction des objectifs d'apprentissage reste le même ou si, une fois franchi le seuil, on s'engage dans un espace où les règles du jeu ne sont plus tout à fait les mêmes.

Et il s'agit de montrer

- qu'il est peut-être moins important de délimiter des niveaux de compétence générale ou partielle soigneusement étagés les uns par rapport aux autres que de s'interroger sur les types de savoirs et de savoir-faire à propos et à l'intérieur desquels il y a lieu de distinguer des degrés de maîtrise ;
- que l'enjeu devient dès lors de caractériser la multiplicité des objectifs post-niveau-seuil, dans la perspective d'une plus grande transparence et d'une plus grande cohérence, comme autant de modules ou de combinaisons variables de modules appartenant, eux, à des séries en nombre plus limité.
- que, si les objectifs poursuivis peuvent être extrêmement diversifiés au-delà du niveau-seuil, il n'en demeure pas moins que des modes d'évaluation et de certification devraient aussi exister, qui permettent d'apprécier, de manière légère, tant des capacités largement transversales mises en œuvre à des fins multiples que des capacités plus partielles ou ponctuelles.

Dans l'exposé de 1991, c'est à une mise en regard de différentes positions qu'il était procédé.

Pour DC91, des orientations comme celle retenue par Van Ek (avec qui Trim a travaillé à la mise au point d'un Threshold Level actualisé pour l'anglais, le *Threshold Level 90*) postulent que, moyennant divers ajustements, des objectifs au-delà du niveau-seuil sont à penser, pour plusieurs niveaux successifs, selon des catégories et avec des mises en relation de ces catégories tout à fait similaires à celles que déploient déjà les différents niveaux-seuils. On s'inscrit alors dans une continuité de niveaux possibles, caractérisables les uns par rapport aux autres sur le mode de l'extension et de la complexification. Et les modalités d'évaluation peuvent de même relever de cette continuité graduée.

La position de tenants, tel Brian North, d'un étagement de savoir-faire observables pour différentes capacités, est, elle, considérée par DC91 comme à la fois plus analytique et plus radicale. Plus analytique en cela que, dans le principe, l'éclatement et la subdivision des capacités introduisent des distinctions multiples (à l'intérieur de la compréhension écrite par exemple, on peut distinguer entre des niveaux pour la

3. Des développements relatifs à une approche modulaire de la construction curriculaire et de l'évaluation sont esquissés dans le chapitre 8 du CECR et le *Guide pour le développement et la mise en œuvre de curriculums pour une éducation plurilingue et interculturelle* (Beacco et al., 2010) comporte des propositions compatibles avec ce type d'orientation.

lecture du journal et des niveaux pour la lecture littéraire ou encore celle des modes d'emploi techniques), plus radicale en ce sens que, du niveau 1 au niveau n , on postule que la capacité considérée reste la même, que c'est bien « la même chose » que l'on apprécie de manière continue et à coup de reformulations successives de descripteurs composant traits communs et traits distinctifs entre le niveau n et, d'une part, le niveau $n-1$, d'autre part, le niveau $n+1$.

Dans un cas (Van Ek) comme dans l'autre (North) toutefois, c'est bien une continuité dans le progrès global ou partiel des apprenants que l'on cherche à situer par divers modes d'échelonnement et par une évaluation des performances en fonction de cet échelonnement.

Certes, constate DC91, tant pour les apprenants que pour les enseignants, les responsables d'institutions de formation, les acteurs de secteurs professionnels, il est précieux de disposer d'échelles, de grilles ou de cartes sur lesquelles rapporter une appréciation des connaissances, des performances, de l'expérience langagière. Par ailleurs, la transparence et la cohérence voudraient en effet qu'on parvienne à la mise au point de standards et de normes de référence qui permettraient d'étalonner, pour mieux les situer les uns par rapport aux autres, ces instruments de mesure existants que sont, pour plusieurs langues, les grandes certifications internationales reconnues.

Il convient de prendre garde toutefois à ce que le souci louable de comparaison, d'harmonisation éventuelle, voire d'évolution des pratiques d'évaluation dans le sens d'une meilleure prise en compte de la capacité à communiquer, ne conduise pas à des simplifications ou à des distorsions abusives. Et DC91 soulève plusieurs questions :

- *Doit-on estimer que les connaissances et les capacités en langue sont du même ordre et descriptibles avec les mêmes instruments d'analyse à quelque niveau de développement qu'on se situe ? L'acquisition se présente-t-elle comme un continuum caractérisable de bout en bout selon les mêmes découpages catégoriels et les mêmes paramètres ?*
- *Dans la mesure où la multiplication des niveaux distingués pour une capacité donnée induit l'idée d'une continuité graduable, faut-il considérer que des capacités diverses (par exemple, l'expression écrite et la compréhension orale) sont étagées de la même manière et avec le même nombre de niveaux sur une même grille récapitulative ?*
- *Est-ce qu'une multiplication des degrés dans l'accomplissement d'une tâche communicative (par exemple, la gestion d'une interaction) n'aboutit pas à des distinctions et à des raffinements, en particulier pour les niveaux les plus élevés, qui font qu'on s'interroge sur l'économie de l'ensemble (le passage d'un niveau 7 à un niveau 8 n'est-il pas beaucoup plus coûteux pour l'apprenant que le passage d'un niveau 2 à un niveau 3) ?*
- *Est-ce que, pour les degrés les plus élevés d'une échelle, la spécification de ce qu'un apprenant est à même de faire dans la langue étrangère est possible indépendamment d'une indication sur les objets de discours ? La capacité à argumenter peut-elle s'apprécier hors articulation avec une topique déterminée ?*
- *La multiplication des niveaux distingués et le recours aux mêmes modes de description des objectifs entraînent des modalités d'évaluation elles-mêmes homologues tout au long du cursus. Est-ce là encore la procédure la plus économique et la plus efficace, si transparente et cohérente paraît-elle ?*

– Un étagement en niveaux jusqu'aux étapes les plus élevées, surtout si l'on entend arriver à une grille récapitulative, ne privilégie-t-il pas de fait la verticalité par rapport à l'horizontalité, la progression à l'intérieur des capacités initialement distinguées plus que la diversification de ces capacités, l'introduction de nouvelles, la recomposition autre des capacités premières ?

– La référence à des échelles d'appréciation ne conduit-elle pas, qu'on le veuille ou non, à valoriser les parcours continus plus que les réorientations, les changements de cap, la poursuite d'objectifs pour partie déconnectés par rapport à ceux qu'on avait d'abord visés ? Si louable que soit la volonté de transparence et de cohérence pour toutes les parties concernées, elle ne devrait pas avoir pour effet de donner une prime à la norme et à l'orthodoxie dans les parcours d'apprentissage.

D es modules de différents types ?

Ces interrogations de 1991 restent sans doute ouvertes, ce qui au demeurant ne saurait constituer une preuve de leur pertinence. DC91 les poursuit toutefois en posant que, au-delà du niveau-seuil, bon nombre d'apprenants entendent développer plus avant une compétence partielle (ou privilégiée), telle que lire des textes spécialisés ou écouter des programmes culturels ou prendre part à un échange commercial dans une branche professionnelle. « Il ne s'agit pas alors de se limiter à cette capacité particulière, mais de considérer que, quant au reste, le niveau-seuil atteint est suffisant. » D'où l'idée que poursuivre des validations et *a fortiori* des certifications couvrant – en principe – tout l'éventail des capacités ne serait pas vraiment le lot commun.

L'enjeu est aussi, pour l'avenir, que des compétences partielles, « déséquilibrées » si l'on veut, deviennent des objectifs tout aussi valides, reconnus, valorisés, que des objectifs généraux et globaux, conçus en termes de changement de niveau dans toutes les capacités.

Et d'où aussi, pour DC91, le choix d'une approche modulaire plutôt que rigoureusement stratifiée. À condition bien entendu que ces modules, validables comme tels, puissent aussi entrer dans des combinaisons plus générales, si le parcours personnel de l'acteur social considéré le conduit, dans son apprentissage « tout au long de la vie », à souhaiter cette demande d'équivalence ou cette validation des acquis de l'expérience en vue d'un objectif ultérieur moins « partiel ». De tels dispositifs existent dans d'autres domaines (notamment, en France, pour ce qui touche aux diplômes et certifications dans différents secteurs d'activité professionnelle⁴) et on ne voit pas pourquoi l'apprentissage des langues en serait privé.

Là où le texte de 1991 s'engage de façon plus iconoclaste au regard de la dominante d'une perspective communicationnelle (ou même « actionnelle »⁵), c'est en considérant que les dits modules post niveau-seuil

4. Voir à ce propos le Portail de la Validation des Acquis de l'Expérience : www.vae.gouv.fr

5. Adjectif ici anachronique, puisque la « perspective actionnelle » et la catégorisation de l'apprenant comme acteur social n'apparaissent que plus tard, au cours de la rédaction du CECR.

peuvent fort bien ne pas avoir des objectifs directement communicationnels, mais relever d'autres apports susceptibles d'intéresser les apprenants. Quatre grands types de modules sont ainsi distingués :

- Le premier est évidemment celui des connaissances et capacités langagières au sens où l'entend le modèle des niveaux-seuils, au sens aussi où l'entendent la plupart des échelles de compétence.[...]
- Le second est celui des savoirs sur la langue et ses usages, acquis en tant que tels et indépendamment d'une visée ou d'une mise en œuvre communicative immédiate. Ces savoirs «méta» peuvent être d'ordres divers et plus ou moins explicitement articulables à l'emploi effectif de la langue considérée. Mais ce qui importe ici, c'est que des apprenants soient susceptibles d'avoir pour objectif l'acquisition de telle ou telle zone de ces savoirs et que la connaissance acquise soit susceptible d'évaluation en termes de «savoir sur» plus que d'«usage de».
- Le troisième est celui de savoirs autres que métalangagiers, qu'il s'agisse de connaissances anthropologiques, géographiques ou autres touchant aux locuteurs de la langue qu'on apprend, ou qu'il s'agisse de savoirs d'ordre scientifique ou littéraire relevant de disciplines constituées (physique des solides ou sociocritique de la littérature). [...] Le progrès en langue, dans une visée donnée, peut se trouver très directement dépendant aussi de l'acquisition de savoirs nouveaux autres que linguistiques [...]
- Le quatrième pourrait se ramener à ce qu'on désigne génériquement sous l'appellation de «stratégies d'apprentissage» et de «stratégies de communication». Tout ce qui relève de l'apprendre à apprendre rentre dans ce chapitre [...].

Dans ses commentaires, DC91 insiste sur le peu d'originalité de ces distinctions de secteurs, souligne qu'il ne s'agit pas de négliger l'importance du premier, mais estime que cette classification sommaire élargit l'offre «horizontale» de formation au lieu de la canaliser seulement dans une continuité «verticale».

Rétrospectivement...

Rétrospectivement il y a lieu de noter que, bien évidemment, la définition des contenus de chacun de ces types de modules et les modalités d'évaluation correspondantes diffèrent les unes des autres selon le secteur concerné : des modules portant sur les stratégies ne se confondent pas, à ces égards, avec ceux portant sur des savoirs. On ne se trouve plus dans une logique de constance et de hiérarchisation continue de descripteurs comme c'est le cas dans ce qui aboutira aux échelles de niveaux du CECR.

La place faite aux stratégies est significative de l'ambiance didactique des années 1990 et ne serait probablement pas envisagée de la même manière aujourd'hui, mais rien n'empêche d'imaginer des modules consacrés aux stratégies d'intercompréhension ou encore aux stratégies de compréhension orale ou au déchiffrement critique des médias et des discours diffusés sur la toile, etc. Quant à ce qui relève des modules

turnés plus vers une réflexion plus formelle, ils ne se limiteraient pas à un travail de grammaire explicite ; on peut penser à des observations d'ordre sociolinguistique, à des formes de « didactique intégrée » (au sens comparatiste de l'expression) ou à des activités du type de ce qui a été largement développé par l'Éveil aux langues, ou encore à des analyses entrant dans le champ de l'interculturel. Et pour le secteur des modules de construction de savoirs « autres », il est clair que l'on touche là au CLIL/EMILE, c'est-à-dire à l'enseignement bilingue (ou bi-/plurilingue), avec ou sans alternance des langues.

Autrement dit, le dispositif suggéré en 1991 pourrait être repensé, plus de vingt ans plus tard, en relation notamment à ce qu'on désigne – avec des variantes quant aux contenus et démarches didactiques ainsi regroupés, sous l'appellation « approches plurielles » (éveil aux langues, intercompréhension, didactique intégrée, enseignement bilingue, pédagogie interculturelle) (Candelier 2008 ; Cavalli 2008 ; Candelier et De Pietro 2011). C'est ainsi esquisser une autre manière de penser des parcours d'apprentissage et des types d'évaluation diversifiés. Les orientations du *Guide pour le développement et la mise en œuvre de curriculums pour une éducation plurilingue et interculturelle* (Beacco et al., 2010) vont à quelques égards dans le même sens. L'enjeu, y compris pour les curriculums scolaires, revient à poser des changements internes de dynamique au cours des trajectoires d'apprentissage. Et on voit bien que quasiment toutes les approches plurielles soulèvent des questionnements majeurs et stimulants quant aux modalités d'évaluation pouvant y être appliquées (voir Huver et Springer, 2011)⁶.

En tout état de cause, qu'il s'agisse de publics scolaires ou d'adultes et que l'on ait ou non recours à des formes explicites de modularisation selon des secteurs similaires à ceux proposés en 1991, un des bénéfices secondaires mais non négligeables des options ainsi ouvertes est bien de développer chez les apprenants (et les enseignants ?) concernés à la fois : – une culture d'apprentissage (et d'enseignement ?) : comprendre qu'on peut apprendre de diverses manières et que cette diversité assure une certaine économie curriculaire (économie pouvant être prise ici aux deux sens courants du terme : organisation et gain) ; – et une culture de l'évaluation : l'apprentissage est multidimensionnel et à cette multidimensionnalité correspondent des modes pluriels d'évaluation.

Enfin, il est permis de penser que ces ruptures apparentes de continuité, ces possibilités de changer d'objets et de manières de travailler et d'évaluer les résultats sont de nature à prévenir « l'effet plafond » (*ceiling effect*), la saturation et le piétinement sans progrès que l'on connaît dans bien des cursus de langue quand tout continue selon une routine dans le format d'exercice et d'évaluation, sans la stimulation de démarches et d'expériences d'apprentissage nouvelles.

Il ne vaut pas la peine de résumer et commenter plus avant le texte de 1991. On y trouve aussi l'amorce de ce qui deviendra, dans le CECR,

6. Un dispositif d'évaluation à base modulaire et, qui plus est, avec des modules de types et de formats différents autorise des options et des souplesses dans les parcours proposés aux apprenants, mais s'avère évidemment plus complexe s'il s'agit de penser en termes de niveaux étagés du genre A1, A2, etc. Les voies à explorer relèvent d'une dynamique « d'évaluation alternative » (Huver et Springer 2011).

l'activité de médiation et, en outre, des prises de position par rapport aux questionnements initiaux de l'intervention, en particulier celle-ci :

Pour un ensemble d'objectifs de nature ou de composition modulaire, correspondant non seulement à des visées globales, mais aussi à l'acquisition de savoirs et de savoir-faire partiels, complémentaires, relais ou très spécifiques, on n'a pas à penser l'évaluation uniquement en termes de performances communicatives simulées ou authentiques. Selon les modules et les secteurs concernés, d'autres critères et d'autres modes d'évaluation peuvent être mis en œuvre, relevant du contrôle de connaissances académiques ou de la résolution de problèmes autres que langagiers. Il est clair que l'on est aujourd'hui moins bien armé pour aborder ces autres objets et modes d'évaluation de manière satisfaisante qu'on ne l'est pour ce qui touche aux habiletés langagières; mais il y a justement à travailler dans ce sens.

Et encore :

Dans une conception modulaire forte, il faut poser que les modules sont relativement autonomes les uns par rapport aux autres et s'organisent selon leurs propres modes, non selon des catégories ou des critères qui devraient nécessairement être les mêmes pour tous. À considérer les quatre secteurs que l'on distinguait plus haut [...], il est clair que des degrés peuvent être distingués à l'intérieur de chacun des ordres de maîtrise propres à ces secteurs ; mais que, d'un secteur à l'autre et parfois d'un module à l'autre, les degrés de maîtrise distingués n'ont pas à répondre aux mêmes critères d'étagement.

Si j'ai rappelé certains aspects du «second Rüsçhlikon» et caractérisé l'analyse que j'y avais alors esquissée, ce n'est pas pour marquer, après coup, une différence ou une prise de distance à l'égard des choix finalement retenus pour le CECR. Membre du groupe de rédaction, j'ai toujours pleinement assumé les décisions finalement arrêtées, quitte à souligner à diverses reprises (notamment Coste 2006 ; Coste 2007) les risques de dérive que présentaient certains des usages institutionnels ou commerciaux des niveaux et de leurs échelles.

Le succès, aussi spectaculaire qu'inattendu et ambigu du CECR, passé très vite du statut premier de référence au rôle de mètre étalon (tant soit peu élastique), montre assez qu'une attente et une demande existaient pour un instrument de cette nature. Quels que soient les us et abus auxquels il s'est prêté, mais aussi grâce aux débats qu'il a suscités et aux critiques, souvent vives, qu'il a rencontrées, il a, en un peu plus d'une décennie, marqué profondément ce que Francis Goullier nomme l'espace éducatif européen. L'outil est riche d'un fort potentiel et autorise des mises en œuvre beaucoup plus fines que celles auxquelles, dans l'ensemble et jusqu'à présent, il a été soumis.

Les perspectives modulaires que, sans prétention à l'originalité, j'avais tracées à Rüsçhlikon n'ont peut-être pas perdu toute pertinence, mais étaient, de par la complexité probable de leur mise en œuvre, taxables d'irréalisme. Les rementionner ici, dans un recueil riche d'analyses et de propositions novatrices et alors que les questions relatives à l'évaluation sont plus que jamais d'actualité (voir Huver et Springer 2011), c'est, par contraste, marquer aussi combien le paysage s'est transformé.

Bibliographie

- BEACCO J.-C., BYRAM M., CAVALLI M., COSTE D., EGLI CUENAT M., GOULLIER F. et PANTHIER J., (2010), *Guide pour le développement et la mise en œuvre de curriculums pour une éducation plurilingue et interculturelle*. Strasbourg, Conseil de l'Europe, 2010.
- CANDELIER M. (2008), « Approches plurielles, didactiques du plurilinguisme : le même et l'autre », *Cahiers de l'ACEDLE*, 2008, p. 65-90.
http://acedle.org/IMG/pdf/Candelier_Cah5-1.pdf
- CANDELIER M. et DE PIETRO J.-F. (2011), « Les approches plurielles : cadre conceptuel et méthodologie d'élaboration du Cadre de référence pour les approches plurielles », dans : Blanchet P. et Chardenet P., *Guide pour la recherche en didactique des langues et des cultures. Approches contextualisées*, Paris, Éditions des Archives contemporaines, p. 261-273.
- CAVALLI (2008), « Didactiques intégrées et approches plurielles », *Babylonia 1*, 2008, p. 15-19.
- CONSEIL DE L'EUROPE, Conseil de la Coopération Culturelle, (1991-1993), *Transparence et cohérence dans l'apprentissage des langues vivantes en Europe : objectifs, évaluation, certification*, Rapport du Symposium tenu à Rüschtikon, Suisse, 10-16 novembre 1991.
http://www.coe.int/t/dg4/linguistic/Ruschlikon1991_fr.pdf.
- CONSEIL DE L'EUROPE (2001), *Cadre européen commun de référence pour les langues. Apprendre, enseigner, évaluer*, Strasbourg, division des Politiques linguistiques et Paris, Didier.
- COSTE D. (1991), « Remarques sur la conception des objectifs et des modalités d'évaluation au-delà du niveau-seuil », dans Conseil de l'Europe, *Transparence et cohérence dans l'apprentissage des langues vivantes en Europe : objectifs, évaluation, certification*, Rapport du Symposium tenu à Rüschtikon, Suisse, 10-16 novembre 1991, p. 86-95.
- COSTE D. (2006), « Le Cadre européen commun de référence pour les langues : traditions, traductions, translations. Retour subjectif sur un parcours », *Synergies Europe*, n° 1.
- COSTE D. (2007), « Le Cadre européen commun de référence pour les langues : contextualisation et/ou standardisation ? » Colloque international de la FIPF, *Le Cadre européen, une référence mondiale ?*, Sèvres, 19-21 juin 2007.
- COSTE D. (à paraître), « "Approches plurielles" et/ou "Approche plurielle" ? Entre diversité du singulier et singularité du pluriel »
- COSTE D., MOORE D. et ZARATE G. (1997), *Compétence plurilingue et pluriculturelle*, Strasbourg : Conseil de l'Europe, 1997 (Repris dans *Le français dans le monde. Recherches et applications*, « L'apprentissage des langues dans le cadre européen », juillet 1998).
- HOLEC H. (1979), *Autonomie et apprentissage des langues étrangères*, Conseil de l'Europe, Paris, Hatier.
- HUVER E. et SPRINGER C. (2011), *L'évaluation en langues*, Paris, Didier, coll. Langues et didactique.
- TRIM J.-L. M., RICHTERICH R., VAN EK J. et WILKINS D. (1973), *Systèmes d'apprentissage des langues vivantes par les adultes*, Strasbourg, Conseil de l'Europe.

V *aria*

AMINATA DIOP

L *'originalité des CALF comme instance de diffusion du français langue seconde/ étrangère au Tchad*

AMINATA DIOP
UNIVERSITÉ PARIS 3, FRANCE

Au Tchad, le français est l'une des langues officielles du pays mais elle demeure étrangère à d'importants segments de la population, dont ceux exerçant le pouvoir politique et militaire. La situation singulière de la francophonie tchadienne, pour partie produit de son histoire, tout particulièrement de l'école coloniale française, explique la création des centres d'apprentissage de la langue française (CALF) comme moyen de pallier les lacunes d'un système éducatif défaillant. Ces centres visent à alphabétiser en français des adultes qui ont précédemment bénéficié d'une scolarité en arabe littéraire ou qui ont eu une scolarité insuffisante en français. Cette instance d'intervention didactique sous la tutelle du ministère tchadien de l'Éducation et bénéficiant de l'appui de la Coopération française tend maintenant, sous la contrainte des baisses de crédits, à prendre son autonomie. À partir de l'analyse de données bibliographiques et d'enquêtes réalisées auprès d'enseignants et d'apprenants, cet article se propose de mettre en lumière l'originalité de l'institution CALF dans le monde francophone qui tient à son insertion dans un contexte particulier, à son double parrainage français et tchadien – certes évolutif – et à la nature de son impact sur la formation linguistique des adultes au Tchad.

G *éopolitique et langue française au Tchad*

À partir de 1900, avec la colonisation française, la langue de l'Hexagone s'implante et se développe lentement mais aussi de manière différenciée dans l'espace tchadien. À l'est et au nord, zone du Sahel

et du Sahara où vivaient des populations musulmanes, la France a dû faire face à une forte résistance, d'abord militaire puis idéologique avec le refus de l'école (Khayar 1976). L'attention coloniale se concentra sur le sud et l'ouest du pays, « le Tchad utile », zone plus peuplée, plus fertile et où vivaient des peuples animistes : la culture du coton se développe et alimente les industries textiles de métropole ; parallèlement, selon le modèle en cours en AEF, l'école publique française y obtient l'adhésion des populations mais elle ne prend véritablement son essor qu'après la conférence de Brazzaville en 1944. Ainsi au moment de l'Indépendance, les élites tchadiennes francophones capables d'assumer des responsabilités dans le gouvernement et l'administration sont très peu nombreuses et proviennent essentiellement du sud du pays (Gardinier 1998).

Depuis 1960, les régimes autocratiques et dictatoriaux se succèdent, au rythme des coups d'État et des rébellions venues notamment des frontières du Soudan et de la Libye. La France soutiendra, de manière indéfectible, ces différents régimes, tant sur le plan militaire que civil. À partir de la guerre de 1978, le pouvoir sera détenu par des clans et ethnies arabophones du nord-est du pays. L'arabe littéraire devient langue officielle en 1982, au même titre que le français. Au quotidien, les populations tchadiennes – 108 communautés linguistiques sont recensées – communiquent, selon les régions, à l'aide des langues véhiculaires, notamment l'arabe tchadien, qui prend de plus en plus d'ampleur, en particulier dans les villes (Jullien de Pommerol 1997). Jusqu'à nos jours, l'administration et l'éducation restent toujours largement dominées par la pratique du français, les clivages identitaires sous-jacents. L'école qui devrait contribuer à la construction d'une nation unifiée et aux valeurs communes est toujours en grande difficulté, ce qui se traduit notamment par une faible rétention (50 %) et une faible acquisition¹ des élèves au cours du cycle primaire (Banque Mondiale 2007). Quand, de surcroît, beaucoup d'enfants, en particulier parmi les populations arabo-musulmanes, ne fréquentent pas l'école ou n'achèvent pas le cycle primaire, l'analphabétisme des adultes est très élevé (81 %), grevant par là même les potentialités de développement du pays. Ainsi, le CALF s'insère dans un contexte où il existe de réels besoins de formation pour les adultes en particulier arabophones, tant en alphabétisation qu'en FLE.

Histoire de l'institution CALF et du parrainage franco-tchadien

Depuis leur création, l'histoire des CALF peut être découpée en trois phases caractérisées par des enjeux et des modalités spécifiques de

1. « 47 % seulement des sortants d'un cycle primaire complet acquièrent suffisamment de connaissances de base pour rester alphabète à l'âge adulte » (Banque Mondiale 2007).

partage des rôles entre la France et le Tchad. Une expérience annonciatrice débuta en 1975 avec la mise en place d'un laboratoire de langue à l'université de N'Djamena par la Coopération française. Les enseignants expatriés avaient alors participé à la formation en français d'adultes ayant poursuivi des études supérieures en arabe en Égypte, au Liban, en Libye... Les résultats pédagogiques furent au rendez-vous mais la France, n'ayant pas dans ses priorités le recyclage de diplômés formés dans les pays arabes, stoppa l'expérience en 1977. Ce n'est que quelques années plus tard, forcée par les événements, que la France révisa sa stratégie.

À l'époque du CAVIFAT ou centre audiovisuel pour les fonctionnaires arabophones tchadiens (1983-1990) qui correspond au régime d'His-sène Habré, la France et le Tchad semblent avoir conclu un pacte gagnant-gagnant. Les cadres arabophones, formés dans les pays arabes et proches du nouveau pouvoir, disposeront des armes linguistiques (le français) pour gérer une administration et profiter d'un État qui reste francophone nonobstant la promulgation de l'arabe comme langue officielle. La France, en formant ces nouvelles élites, espère, par les réseaux d'entente mutuelle tissés, maintenir son influence sur les décisions de l'État et obtenir des marchés pour ses sociétés industrielles : exploitation du pétrole, infrastructures à construire ou reconstruire après la guerre civile, etc. Plus concrètement, le Tchad fournit les locaux du CAVIFAT à N'Djamena tandis que la France assure le fonctionnement et met à disposition un important arsenal d'assistants techniques et de ressources pédagogiques, sa tâche étant aussi de former les futurs enseignants tchadiens de FLE.

À partir de 1991, durant le régime d'Idriss Déby, le CAVIFAT rebaptisé CALF s'ouvre non seulement aux fonctionnaires tchadiens mais à tous les publics adultes et étend son implantation géographique pour constituer un réseau de cinq centres positionnés dans les plus grandes villes du pays : le CALF le plus ancien et le plus important en termes d'effectifs d'apprenants (60 %) est situé dans la capitale multilingue, N'Djamena. Les quatre autres appartiennent à des régions à dominante arabophone (Abéché et Adré) et francophone (Moundou et Sarh). Tout en contribuant au bilinguisme officiel, le CALF, se voulant donc plus démocratique, doit aider à compenser l'analphabétisme mais aussi faciliter l'intégration des adultes arabophones dans une société qui reste toujours francophone dans ses émanations administratives, économiques formelles et diplomatiques. La France dispose, quant à elle, d'un outil pour étendre le tissu de la francophonie à toutes les couches de la population adulte. Durant cette période, l'État tchadien met à disposition des locaux et paye le salaire des enseignants détachés du ministère de l'Éducation (instituteurs, professeurs de français, formateurs) auprès des CALF. La France supporte les différents investissements, le fonctionnement et les appuis pédagogiques au travers de projets pluriannuels. Au sein du conseil

d'administration (CA) des CALF, le Tchad est en position forte face à la France pour prendre les décisions stratégiques. Cependant, avec son poids financier et son dispositif d'assistance technique, la Coopération française garde la maîtrise des réalisations et des orientations.

À partir de 2005, faute d'un réel engagement financier du Tchad devenu pétrolier, la France, ne souhaitant pas financer indéfiniment les CALF, met fin à son intervention sous forme de projets : sans être désavouée par son État-partenaire, elle s'engage dans un processus de pérennisation et d'amélioration de l'image des CALF. En 2006, doté de nouveaux statuts associatifs, le réseau des CALF s'ouvre potentiellement à des partenariats multiples : selon les principes du libéralisme mondialisé, la Coopération française et l'État tchadien ne doivent plus être les seuls bénéficiaires et contributeurs des centres qui sont invités à chercher des marchés de formation auprès de la société civile, des entreprises privées, etc., pour augmenter leurs capacités notamment financières. De nouvelles modalités de pilotage, de coordination et de gestion sont également définies. Au sein du CA de l'association, la Coopération française, maintenant positionnée en tant qu'observatrice, semble avoir transféré son pouvoir décisionnel aux acteurs directs des CALF (le coordinateur, les enseignants et les apprenants) qui feront face aux instances politiques tchadiennes (ministères de l'Éducation, de la Culture et de l'Économie). En matière de coordination, la France, au travers de ses expatriés dont le nombre décroît progressivement, abandonne officiellement sa position de substitution pour un appui technique à des responsables enseignants tchadiens. La notoriété du centre de N'Djamena, matérialisée par une forte croissance des effectifs d'apprenants (quadruplement de 2002 à 2007 pour atteindre 3150), a créé des conditions propices à l'atteinte de l'autonomie financière de ce CALF vis-à-vis des budgets de fonctionnement français. Cette période d'austérité n'empêchera néanmoins pas la France de construire, en 2008, un nouveau CALF à N'Djamena, près du Centre culturel français, qui permettra, au-delà de doubler la capacité d'accueil en apprenants du centre leader et de promouvoir l'image des CALF, de proclamer sa propre présence, celles de la Francophonie et de la Langue française au Tchad.

Impact sur la formation des adultes à N'Djamena

Les apprenants du CALF de N'Djamena appartiennent aux populations arabo-musulmanes ayant l'arabe tchadien en partage et dont le pluri-linguisme – atout d'apprentissage – reste limité. En dehors des élèves,

étudiants et sans profession, ces adultes, très majoritairement des hommes, sont des commerçants et artisans, des enseignants et professeurs ou des fonctionnaires de l'administration, de la gestion, du contrôle, de la sécurité et de l'armée. Avec une telle diversité sociolinguistique, il est difficile d'associer une motivation et une stratégie d'apprentissage à chaque profil d'apprenant. Néanmoins, pour la majorité, l'apprentissage du français est considéré comme un tremplin socio-économique ou professionnel (accéder à de meilleurs postes pour les fonctionnaires, développer leur clientèle francophone pour les commerçants, étudier en France...) ou comme une manière d'éviter la marginalisation ou l'indexation dans la société tchadienne où cette langue étrangère revêt une importance cruciale dans la vie quotidienne et active. Depuis 2005, les apprenants sont de plus en plus nombreux à poursuivre leur parcours de formation sur plusieurs sessions trimestrielles de cours : progressant dans l'échelle des niveaux de compétences, ceux-ci ambitionnent ainsi d'obtenir au moins le DELF ; les autres moins endurants ne souhaitant qu'avoir des bases pour parler, lire et écrire.

L'analyse des résultats au DELF et au DALF sur la période 2002-2007 montre que le CALF de N'Djamena forme avec efficacité des apprenants capables d'un usage «élémentaire» de la langue française (A1-A2). Sa capacité à former efficacement des apprenants, aptes à un usage «indépendant» du français (B1-B2) et à une «maîtrise» du français (C1-C2), est, en revanche, beaucoup moins évidente et va décroissante avec la difficulté du diplôme. En cause, notamment, les limites en compétences sociolinguistiques et interculturelles des enseignants. La qualité de l'apprentissage semble cependant s'être maintenue au cours du temps, sans pour autant s'améliorer. Mais tous les apprenants ne sortent pas du CALF en ayant obtenu un diplôme officiel reconnu internationalement. Les apprenants inscrits dans les classes d'alphabétisation et de post-alphabétisation (43 % des effectifs) seront nombreux à abandonner dès leur première session trimestrielle de cours ou sortiront, au bout de deux ou trois sessions, en sachant juste parler, lire et écrire.

Afin d'appréhender la manière dont le CALF contribue à une meilleure intégration des apprenants au monde francophone tchadien, une enquête spécifique auprès des anciens apprenants de N'Djamena a été réalisée en 2007. Il apparaît que la valorisation des compétences langagières en français sous forme d'ascension dans l'échelle sociale, hiérarchique et/ou des revenus est très variable et est fonction, par ordre d'importance, du niveau d'études de l'apprenant, de son positionnement social initial, des réseaux d'influence mobilisables mais aussi de sa francophilie et des opportunités d'emplois en milieu francophone. Pour mieux connaître ces impacts, une typologie a été construite sur la base du niveau d'études – conditionnant le rythme d'apprentissage et le niveau maximum de compétences langagières en français

accessible – et du secteur d'activités (privés, élèves-étudiants, fonctionnaires). Sept trajectoires professionnelles génériques ont été définies et sont présentées ci-après² :

– Les « Commerçants ou artisans analphabètes » n'ont pas été à l'école ou n'ont pas achevé leur cycle primaire, ils sortent du CALF en sachant lire, écrire et parler le français à un niveau débutant à intermédiaire, ce qui leur permet de gagner en autonomie et de progresser dans l'échelle des revenus sans pouvoir cependant changer de groupe social.

– Les « Débrouillards lettrés » ont achevé au moins un cycle d'études secondaire. Entrant au CALF avec un niveau de français intermédiaire, ils en sortent avec un niveau avancé. Marchands ambulants et/ou travailleurs temporaires, avec leur potentiel intellectuel, leur dynamisme et leurs compétences en français, ils se créent des opportunités qui peuvent leur permettre de gravir l'échelle des revenus ou d'obtenir un emploi plus durable.

– Les « Élèves francophones » s'inscrivent au CALF pour combler leurs lacunes en français. De culture arabo-musulmane, leurs parents francophiles, conscients des difficultés de l'école tchadienne, les ont incités à améliorer leurs compétences langagières, afin de mieux s'intégrer dans le monde francophone tout en s'ouvrant de meilleures opportunités d'emploi. Sortant du CALF avec un niveau avancé, l'élève francophone a une réelle perspective d'ascension sociale notamment dans le privé.

– Les « Étudiants arabophones » ont effectué l'ensemble de leur cursus dans des institutions strictement arabophones. Entrant au CALF à un niveau débutant, ces adultes en sortent au niveau intermédiaire et valorisent ensuite facilement leurs connaissances en français en gravisant les échelons de la hiérarchie et des revenus associés. Leur connexion au monde francophone est facilitée, mais leur intégration y est beaucoup moins réelle dans la mesure où ils évoluent professionnellement dans des milieux d'abord arabophones.

– Les « Enseignants et communicateurs », après une Licence littéraire en arabe, sont affectés au ministère de l'Éducation nationale tchadienne ou à celui de la Communication. Fortement motivés, ils sortent du CALF avec un niveau avancé en français. Ces fonctionnaires valorisent facilement leurs nouvelles compétences en diversifiant leurs revenus, mais leur ascension hiérarchique dépend plus étroitement de la mise en action de solides réseaux socio-ethniques d'influence. Leur ouverture au monde francophone est réalisée notamment par l'accès aux médias.

– Les « Scientifiques et techniques » sont des fonctionnaires ayant obtenu un diplôme universitaire Bac +2 à +4 dans les pays arabes. Leur motivation à apprendre le français est modérée. Ceux qui n'abandonnent pas sortent du CALF avec un niveau intermédiaire à avancé et peuvent le cas échéant, sur la base de leurs compétences scientifiques de base, valoriser financièrement ces nouvelles connaissances au gré des projets de coopération notamment, ce qui conditionnera aussi leur plus grande ouverture et intégration au monde francophone.

2. Le niveau en français des apprenants correspond ici à celui de leur classe de langue au sein du CALF, niveau apprécié par un test d'évaluation interne à l'institution.

– Les «Élites intellectuelles», de niveau d'études d'au minimum Bac +5, ont un très bon rythme d'apprentissage d'autant plus que la plupart a été sensibilisée dans d'autres systèmes alphabétiques que l'arabe littéraire (anglais, russe, etc.). Cible initiale du CAVIFAT puis du CALF, ces fonctionnaires valorisent très bien leurs apprentissages du français (DELFI ou DALF) en termes d'ascension sociale (progression dans la hiérarchie, augmentation de revenus) et s'intègrent effectivement dans le monde francophone. En effet, occupant les postes clefs, ces élites devenues «Décideurs» seront inévitablement en contact avec les coopérations extérieures, bilatérales ou multilatérales, et, en premier lieu, la France.

Ces sept trajectoires peuvent être considérées comme de probables scénarios du devenir des apprenants actuels. Mais du fait de l'inexistence d'une base de données des anciens apprenants, il n'a pas été possible de quantifier les différentes trajectoires professionnelles identifiées. Globalement, on peut cependant estimer que 17 % des apprenants du CALF de N'Djamena (ceux qui suivent au moins trois sessions trimestrielles de cours) pourront améliorer leurs compétences langagières en français et espérer les valoriser dans leur quotidien, sous une forme ou sous une autre.

Une efficacité à renforcer

Le CALF constitue un «appareil» formatif institutionnel, en ce sens que son objectif de transmission de la langue française est aussi animé par des motifs culturels et politiques. Sa capacité à alphabétiser et à augmenter les compétences linguistiques en français des adultes arabophones compense les lacunes du système éducatif tchadien et celles de la formation tout au long de la vie. Il facilite l'intégration des arabophones dans le tissu économique et dans la fonction publique francophone. Son dynamisme se mesure donc à sa capacité à s'adapter au contexte national et à proposer des activités pédagogiques répondant aux besoins de la société et des publics spécifiques.

Si cette instance didactique, qui suit les évolutions scientifiques en matière d'innovations pédagogiques (des méthodes audiovisuelles aux méthodes communicatives)³, semble avoir bien géré son succès et son développement, il n'en demeure pas moins qu'il lui faudra trouver les moyens d'être encore plus efficace afin de contribuer davantage au bilinguisme arabe-français et ne pas se limiter à demeurer un réseau de centres en capacité d'alphabétiser des adultes en quantité et de les former à un usage simplement élémentaire du français. La réalisation de ce saut pédagogique qualitatif devrait passer par une attention particulière portée à la formation des enseignants.

3. Le «paquet pédagogique» de la méthode Panorama constitue aujourd'hui le support privilégié d'enseignement-apprentissage au CALF.

Bibliographie

- BANQUE MONDIALE (2007), *Le système éducatif tchadien. Éléments de diagnostic pour une politique éducative nouvelle et une meilleure efficacité de la dépense publique*, Document de travail n° 110, série : le développement humain en Afrique, 199 p.
- GARDINIER D. (1988), « L'enseignement colonial français au Tchad (1900-1960) », in *L'Afrique et l'Asie modernes*, n° 161, p. 59-71.
- JULLIEN DE POMMEROL P. (1997), *L'arabe tchadien. Émergence d'une langue véhiculaire*, Karthala, Paris, 174 p.
- KHAYAR I. (1976), *Le refus de l'école. Contribution à l'étude des problèmes de l'éducation chez les musulmans du Ouaddai (Tchad)*, Maisonneuve, Paris, 140 p.

Recherches et applications

La revue de référence
de la recherche
universitaire
francophone

Abonnement 1 an

**2 Recherches
et Applications**

À paraître
janvier et juillet,
de l'année en cours

29€

Soit 20% d'économie

REPLIR À CÔTÉ



BULLETIN D'ABONNEMENT

ABONNEMENT 1 AN 29 € (soit 20% de réduction)

**2 Recherches & Applications à paraître
(janvier et juillet de l'année en cours)**

Je m'abonne

NOM :

.....

PRÉNOM :

.....

ADRESSE :

.....

.....

CODE POSTAL :

.....

VILLE :

.....

PAYS : TÉL :

.....

COURRIEL :

.....

Je règle

CHÈQUE BANCAIRE À L'ORDRE DE SEJER

VIREMENT BANCAIRE

Préciser les noms et adresse de l'abonné ainsi que le numéro de facture si vous l'avez.

Crédit Lyonnais **30002-00797-0000401153D** clé **08**

IBAN **FR36 3000 2007 9700 0040 1153**

BIC/SWIFT: **CRLYFRPP D08**

VIREMENT POSTAL

Préciser les noms et adresse de l'abonné ainsi que le numéro de facture si vous l'avez.

CCP 30041-00001-2697014T020 clé **79**

MANDAT

À l'ordre de : **SEJER** (merci de joindre la photocopie du mandat)

CARTE BANCAIRE (VISA, EUROCARD-MASTERCARD)

N° de carte

Date de validité

Signature

Cryptogramme

(Les trois derniers chiffres
au dos de votre carte.)

BULLETIN À RENVoyer À :

LE FRANÇAIS DANS LE MONDE - 9 BIS, RUE ABEL HOVELACQUE

7503 PARIS - FRANCE - TÉL. : 01.72.36.30.67

N° d'éditeur : 00000000 – CGI – Janvier 2013
Imprimé en France par La Nouvelle Imprimerie Laballery
58500 Clamecy – N° d'impression : 000000

La Nouvelle Imprimerie Laballery est titulaire de la marque Imprim'Vert®