



GLOTTOPOL

Revue de sociolinguistique en ligne

n° 23 – janvier 2014

*Inaccessibles, altérités, pluralités :
trois notions pour questionner les langues
et les cultures en éducation*

Numéro dirigé par Cécile Goï, Emmanuelle
Huver et Elatiana Razafimandimbimanana

SOMMAIRE

Cécile Goï, Emmanuelle Huver, Elatiana Razafimandimbimanana : *Les inaccessibles de l'altérité et de la pluralité linguistiques et culturelles. Enjeux et perspectives pour l'éducation.*

Cécile Goï : *Élèves allophones nouveaux arrivants et altérité en éducation : de l'inaccessible des pratiques des enseignants et des chercheurs à l'ineffable ontologique de l'être.*

Elatiana Razafimandimbimanana : *Quelque part entre des inaccessibles : une façon de conceptualiser la photographie et le sens en sociolinguistique.*

Emmanuelle Huver : *Les inaccessibles de l'évaluation en langue(s). Impensé ? Impasse ? Ferments ?*

Ann Feuteun : *Les langues de l'école pour négocier les « inaccessibles des langues » à l'école.*

Joanna Lorilleux : *L'expérience formative à l'épreuve de la catégorisation ?*

Vanessa Andreotti : *Strategic criticism and the question of (in)accessibility of the Other.*

Frédérique Lerbet-Sereni : *L'expérience de l'indicible/invisible : l'inaccessible comme mode d'accès renouvelé au connaître.*

Valérie Spaëth : *La question de l'Autre en didactique des langues.*

Véronique Castellotti : *L'hétérogénéité, fondement de l'éducation linguistique ? Vers des perspectives alterdidactiques.*

Compte-rendu

Fabienne Leconte : Debono Marc, 2013, *Langue et droit – Approche sociolinguistique, historique et épistémologique*, Editions modulaires européennes, collection Proximités, Bruxelles, 388 pages, ISBN : 978-2-8066-0770-6

GLOTTOPOL – n° 23 – janvier 2014

<http://glottopol.univ-rouen.fr>

LES INACCESSIBLES DE L'ÉVALUATION EN LANGUE(S) IMPENSÉ ? IMPASSE ? FERMENTS ?

Emmanuelle Huver

Université F. Rabelais (Tours), EA 4246 PREFics – Dynadiv

1. Généalogie d'un texte¹

Le programme Paraadiv, qui se nourrit de et s'inscrit dans les travaux et les réflexions menés au sein de l'EA 4246 PREFics-Dynadiv, part entre autres de l'idée que le chercheur doit « *intégrer dans ses interprétations le cheminement intellectuel et affectif qui est à l'origine de [celles-ci]* » (Kilani, 1992 : 16). Ce liminaire vise ainsi à expliciter les intentions, les intuitions et les convictions que j'identifie comme particulièrement saillantes dans le cheminement qui a mené à la rédaction de cet article. Si je l'ai qualifié de généalogique (et non d'historique), c'est bien justement pour insister sur le fait que les éléments présentés ici le sont au regard d'une logique rétrospective construite en fonction d'enjeux présents, et non comme une suite linéaire d'événements passés dont le déroulement ne dépendrait pas du point de vue du chercheur et des enjeux (présents) qu'il y perçoit / projette / présuppose, etc. De même, les auto-citations dans cet article ne sont pas à comprendre comme une stratégie auto-promotionnelle, mais comme la marque d'un travail de relecture et de réinterprétation de certains de mes travaux.

1.1. Paraadiv : de quelques principes visés

Cet article s'appuie tout d'abord sur les objectifs de Paraadiv et sur les démarches qui en découlent. Dans ce programme, il s'agit d'explorer les liens entre différentes dimensions qui nous semblent impliquées dans l'accueil des enfants nouveaux arrivants (désormais ENA²) à l'école française, en nous appuyant sur la pluralité des interprétations qui les construisent et les articulent (cf. introduction de ce numéro). Pour ce qui concerne plus particulièrement le présent propos, nous³ souhaitions, initialement, ne pas « plaquer » notre seule interprétation

¹ Cet article a bénéficié de la relecture vigilante d'E. Razafimandimbimanana et des deux évaluateurs du comité scientifique, que je remercie. Je reste bien sûr seule responsable des propos qui sont tenus.

² Je conserverai ici l'acronyme ENA, et non celui, plus récent, d'EANA – Elèves *Allophones* Nouvellement Arrivés (terminologie proposée dans les Instructions Officielles d'octobre 2012 relatives à ce public) – dans la mesure où, nous le verrons plus loin, le terme d'*allophonie* tend à mettre en transparence la pluralité du répertoire langagier de ces élèves.

³ Désormais et sauf mention contraire, *nous* renverra dans cet article à Cécile Goï et moi-même, Elatiana Razafimandimbimanana nous ayant rejointes ultérieurement et s'étant par ailleurs plus intéressée aux élèves « jeunes migrants plurilingues » qu'aux enseignants. Même si les propos développés ici s'appuient en grande

sur les séances d'évaluation que nous avons observées et filmées, mais, au contraire, écouter, entendre, comprendre (?) ce que les enseignants avaient à dire de ces séances, d'où la mise en œuvre d'une démarche d'auto / hétéro-confrontation sur la base d'extraits de séances d'évaluation diagnostique sélectionnés par nos soins. D'une certaine manière, et comme dans de nombreux protocoles de ce genre, notre objectif initial était peut-être, en allant chercher le point de vue des acteurs, d'avoir *accès* à ce qui reste inaccessible par l'observation seule, à savoir « *les interprétations que les individus eux-mêmes donnent de leurs comportements* » (Bigot et Cadet, 2011 : 28), les discours des enseignants « *constitu[a]nt [alors] une voie d'accès privilégiée à la part **inobservable** inhérente à toute activité* » (*op. cit.* : 13, c'est moi qui souligne⁴).

Toutefois, nous étions conjointement conscientes du caractère illusoire de cette démarche consistant à mettre en regard agir et discours sur l'agir afin de repousser (voire d'éliminer) la part inaccessible des pratiques et des discours d'autres que nous. Nous avons donc cherché – et cette intuition s'est consolidée au fil de nos réflexions au sein du PREFics-Dynadiv et de Paraadiv – à penser des modalités de recherche (i.e. tout à la fois des dispositifs, des postures et des écritures) qui ne visent pas tant à exposer et juxtaposer les discours des acteurs sur leur(s) agir(s) (en excluant les nôtres), qu'à croiser des interprétations (celles des acteurs *dont* les nôtres), afin de voir en quoi ces interprétations s'interrogeaient, se dérangeaient et/ou se confirmaient (nos interprétations étant de fait intégrées au réseau d'interprétations ainsi produites). Ce faisant, il s'est agi également de rendre visible (accessible ?) la part d'interprétation dont nous considérons qu'elle traverse irréductiblement toute recherche, non seulement pour revendiquer une forme de subjectivité dans la recherche en sciences humaines, mais aussi pour travailler à partir de celle-ci en la considérant comme un principe opératoire de construction des connaissances.

1.2. La thématique de l'évaluation et son traitement

Le programme Paraadiv présentait pour moi l'intérêt de faire se rencontrer deux thématiques qui m'occupaient depuis mes premiers travaux de recherche et que j'avais jusqu'ici traitées séparément, à savoir l'évaluation d'une part et les ENA d'autre part. En outre, cette articulation était peu (voire pas) thématisée dans la littérature de recherche existante – tout du moins en France –, alors qu'elle me semblait pouvoir permettre de traiter d'enjeux plus transversaux à la DDL (notamment les liens entre évaluation et altérité / diversité) et/ou à l'école française (notamment la reconnaissance et le traitement de l'altérité dans cette institution). Enfin, ces enjeux et les problématiques qui en découlent font l'objet d'une importante demande sociale, notamment depuis 2001 par les instructions officielles introduisant une évaluation dans les langues de scolarisation antérieure des élèves (cf. la récurrence de cette thématique dans les formations d'enseignants, les missions des CASNAV relatives à ce thème, etc.).

Mon intérêt pour les questions liées à l'évaluation renvoie également à la conviction de l'insuffisance d'une entrée dans ces questions par le seul angle méthodologique (Huver, 2012b). Ainsi, dans le cadre du programme Paraadiv, nous n'avons pas souhaité analyser des outils ou des pratiques pour les rattacher à tel ou tel courant méthodologique pré-identifié ou encore pour identifier quels outils ou pratiques seraient (ou non) conformes à *la* méthodologie d'enseignement du français présentée comme la plus opératoire et, à ce titre, à mettre en œuvre. En outre, contrairement à ce que souhaiteraient la plupart des enseignants, ainsi que l'institution scolaire, il ne s'agissait pas non plus d'élaborer des outils d'évaluation plus

partie sur les réflexions menées avec Cécile Goi, ils sont également réinterprétés au filtre de questionnements qui me sont plus personnels (cf. 1.2.). Il va donc sans dire que je garde seule la responsabilité des propos tenus ici.

⁴ Sauf mention contraire dans les citations ultérieures, les mises en gras sont de mon fait.

performants, plus efficaces ou plus opératoires que ceux existants, puisque nous partions justement de la préconception inverse selon laquelle les outils sont indissociables de leurs usages⁵. De ce point de vue, notre travail de recherche peut être considéré comme fondé sur un paradoxe, puisque nous avons choisi de travailler sur une thématique émanant d'une demande sociale forte et clairement identifiée tout en la traitant sous un angle non attendu, non souhaité, voire contesté, par les différents acteurs dont cette demande émane⁶.

Au-delà de Paraadiv et du public plus particulièrement concerné dans ce programme, travailler sur les questions d'évaluation en DDL renvoie enfin à une problématique de recherche plus personnelle et ancienne, à savoir la critique radicale de la domination d'une conception objectiviste de l'évaluation en DDL. Ce travail au sein de Paraadiv m'a ainsi semblé constituer une piste pertinente pour densifier cette position critique en en reconsidérant les fondements argumentatifs. Ce travail de densification conceptuelle visant à rendre visible le caractère construit de toute évaluation (y compris celles qui se présentent comme objectives et neutres) a en partie consisté à mettre en lumière les liens entre outils et visées politiques, voire idéologiques (sur le public des ENA plus précisément : Huver, 2012a). Cependant, je n'ai pas souhaité réduire la réflexion à cet angle d'argumentation. C'est pourquoi d'une part, j'ai choisi (et Cécile Goï avec moi) de m'intéresser aux liens entre outils et acteurs en travaillant sur les usages évaluatifs, c'est-à-dire sur les manières singulières, situées et parfois paradoxales ou contradictoires dont les acteurs s'approprient les outils et les démarches d'évaluation et leur donnent sens au regard de leurs contextes institutionnels d'intervention, de leurs parcours biographiques professionnels, de leurs imaginaires personnels et de notions didactiques circulantes⁷. D'autre part, j'avais le souhait de construire une cohérence vigilante constante entre questionnements didactiques et questionnements épistémologiques : en d'autres termes, je pense que cette interrogation radicale de l'objectivité en évaluation ne peut pas ne pas s'appuyer sur (et venir à l'appui d')une réflexion épistémologique plus large sur le statut de l'objectivité dans la construction des connaissances en sciences humaines en général, et en DDL en particulier.

C'est cet arrière-plan et ce cheminement qui ont progressivement donné corps à une réflexion sur les liens entre inaccessibles et évaluation, dont je m'attacherai dans cet article à thématiser les enjeux et les conséquences, pour le champ de l'évaluation mais aussi plus largement pour le champ de la DDL et de la recherche en DDL. Dans cette perspective, je présenterai d'abord brièvement les liens que j'identifie entre mesure, objectivité et conception méthodologiste de l'enseignement et de l'évaluation en langue. Puis, j'interrogerai une

⁵ Le terme *usage* renvoie au caractère conjoint et indissociable des pratiques et des représentations (Castellotti, 2012). Ceci implique d'interroger les conceptions de la didactique qui consistent essentiellement à élaborer des outils et des dispositifs, sans poser la question de leur mise en œuvre (i.e. postulant donc, au moins implicitement, que celle-ci sera conforme aux intentions des concepteurs, comme si ces intentions, et plus largement, le sens étaient transparents).

⁶ Ce qui pose la question des articulations à construire entre recherche et intervention.

⁷ C'est pourquoi nous avons dans un premier temps analysé différents outils d'évaluation diagnostique utilisés avec des ENA, sous l'angle notamment de l'image des ENA que ces outils produisaient et dont ils étaient eux-mêmes le produit. Nous avons ensuite proposé à certains de nos étudiants de Master, locuteurs des langues des évaluations et scolarisés dans des espaces géographiques ou institutionnels dans lesquels ces langues étaient langues de scolarisation, de passer ces évaluations et de nous indiquer les difficultés, erreurs, écarts ressentis. Nous avons enfin observé et filmé une séance d'évaluation diagnostique. Nous en avons sélectionné certains extraits qui nous semblaient particulièrement saillants pour notre problématique et les avons montrés à des enseignants intervenant auprès d'ENA, qui les ont commentés. C'est sur l'ensemble de ce matériau que nous nous sommes appuyées au sein de Paraadiv, c'est aussi ce matériau que je mobilise dans cet article ainsi que, plus largement, l'ensemble de mes expériences d'enseignante, de formatrice et de chercheur en matière d'évaluation en langues. Pour plus de détails sur la démarche adoptée dans le projet (construction et interprétation des observables) et sur les conceptions épistémologiques qui les fondent, cf. Goï ici même, ainsi que Goï et Huver (2013a et 2013c).

conception de l'évaluation qui envisage les inaccessibles comme un phénomène à éradiquer et donc comme une impasse. Enfin je proposerai une lecture de la notion d'inaccessible comme ferment⁸ possible d'une conception subjective / interprétative de l'évaluation.

2. Domination d'un paradigme objectiviste – les inaccessibles : des impensés

Le champ de l'évaluation en langue est largement dominé par un paradigme objectiviste (Huver et Springer, 2011), ce qui se manifeste par une centration sur les questions de *mesure* (et de modalités de son optimisation) d'une part et de *traçabilité* d'autre part (celle-ci étant envisagée comme une modalité d'optimisation de celle-là). Ce constat n'est pas propre, et de loin, au public des ENA, mais il est nécessaire de s'y arrêter un instant dans la mesure où il influe largement sur les représentations circulantes (véhiculées par les concepteurs de tests, les chercheurs dans le domaine de l'évaluation en langue, les organismes certificateurs notamment) et donc sur les pratiques d'évaluation, y compris auprès du public qui nous intéresse. Par ailleurs, c'est cette réflexion plus globale qui m'a conduite à réfléchir à la question de ce que nous avons nommé ici inaccessible (mais qui pourrait aussi se décliner comme le non objectivable, le non mesurable et/ou le non traçable).

2.1. Objectivité et mesure : la quantification comme repoussoir de la subjectivité

Les discours (scientifiques et/ou d'expertise) dans le champ de l'évaluation en langue sont actuellement fortement marqués d'une exigence d'objectivité : celle-ci est historiquement située, mais elle se trouve ré-activée dans la situation politico-économique contemporaine.

La question de l'objectivité est en effet une problématique centrale dans les recherches sur l'évaluation et elle constitue même un fil rouge de celles-ci. Qu'elle soit envisagée comme une nécessité, un horizon d'attente ou une illusion, la plupart des recherches portant, de près ou de loin, sur les questions d'évaluation, ont dû, et doivent encore, se situer par rapport à elle. La problématique de l'objectivité est à ce point structurante du champ qu'on peut avancer l'idée (que je ne pourrai développer plus avant ici) que l'histoire des travaux portant sur l'évaluation peut être revisitée à l'aune des diverses réponses et modalités de positionnement au regard de cette problématique.

Pour ce qui concerne plus particulièrement le champ de la DDL, on observe un recentrage massif autour de la problématique de l'objectivité conçue comme un critère de qualité des situations d'évaluation et des outils de jugement⁹. Il faut souligner que cela n'a pas toujours été le cas, puisque Bolton (1987) a même avancé la thèse inverse, en opposant, dans les approches communicatives (et, plus largement les méthodologies non centrées exclusivement

⁸ Les sens figurés de *ferment* renvoient en effet d'une part à l'origine et d'autre part au bouillonnement, à l'effervescence. C'est bien ainsi que je considère la question des inaccessibles pour l'évaluation : l'origine d'une réflexion sur la subjectivité en évaluation, mais aussi le bouillonnement que cette réflexion est susceptible d'induire. Plus largement, Bachelard (1947) a consacré de très belles pages à la rêverie autour du mou, de la pâte et de la fermentation, qui imprègnent le sens que j'accorde au terme *ferment*. Enfin, à l'heure où je finalise cet article, je relis par hasard De Robillard (2007 : 178) et retrouve ce terme, dont j'avais oublié qu'il l'avait lui aussi mobilisé, pour sa puissance évocatrice (« le ferment fait « pousser » (...) la matière par multiplication et transformation ») d'une part, et pour son potentiel corollaire de critique de la notion d'outil d'autre part, autre aspect qui, on le verra, sera central dans cet article : « Si je propose la métaphore du ferment, c'est pour trouver un pendant à celle d'outil, technolinguistique. L'outil travail *sur* un autre objet, et cette relation n'est pas réciproque, le ferment travaille *dans, avec*. » (*ibid.*).

⁹ Les termes hypéronymiques *situation d'évaluation* et *outil de jugement* sont empruntés à la littérature scientifique canadienne francophone (Scallon, Lussier, etc.) : le premier englobe toutes les sortes d'activités mises en œuvres à des fins d'évaluation (activités, exercices, tests, etc.) et le second toutes les sortes d'outils permettant d'interpréter les productions résultant des activités (barème, grilles de critères, check lists, etc.), sans présumer de leur orientation méthodologique et/ou épistémologique.

sur la compétence linguistique au sens strict), les critères d'objectivité et de validité (1) et en plaidant pour la préséance de ce dernier (2).

(1) Essayer d'élaborer des tests communicatifs pose aux auteurs de ces tests des problèmes, avant tout celui du conflit entre objectivité de l'interprétation et fiabilité du test d'un côté et validité de l'autre. [...] Si l'on va dans le sens de l'objectivité de dépouillement et de fiabilité la plus grande possible, les tests exigent une directivité sans équivoque des réponses que l'on demande, mais si l'on va dans le sens de l'objectif "compétence communicative interactive", les tests doivent laisser la plus grande place possible à des réponses individuelles créatives. Mais [...] on diminue la fiabilité de la mesure. La docimologie classique demandait aux tests d'être objectifs, fiables et valides. Mais dans les tests communicatifs, ces trois critères s'opposent. » (Bolton, 1987 : 4-5)

(2) « Un test fiable mais non valide est sans valeur. » (Bolton S., 1987 : 10)

Si la problématique de l'objectivité en évaluation est historiquement récurrente, le regain d'intérêt dont elle bénéficie actuellement tient à mon sens *entre autres*¹⁰ à un contexte socio-économique qui en permet la réactivation (Huver, à paraître 2014). En effet, actuellement, l'évaluation est envisagée non plus seulement dans sa dimension didactique de construction et de certification des apprentissages, mais aussi comme un instrument de régulation des politiques publiques, des politiques linguistiques éducatives ou encore des politiques migratoires (Huver et Springer, 2011). Prenons l'exemple du DELF scolaire, proposé aux ENA sur la base du volontariat : désormais, il ne constitue plus seulement une reconnaissance symbolique des acquis par l'institution scolaire, mais devient de plus en plus visiblement un instrument d'évaluation de l'efficacité des dispositifs dans lesquels les ENA sont scolarisés. Dans certains cas, le DELF constitue également un argument dans le cadre des procédures de régularisation : l'obtention du diplôme est alors présentée comme attestant d'un processus d'intégration effectif et en cours, ce qui témoigne d'une certaine conception des liens entre langue (ici langue française) et insertion, la maîtrise de la langue étant conçue à la fois comme la preuve de l'intégration et sa condition nécessaire et préalable (Huver, 2012a).

Cette intrication entre les dimensions didactiques, politico-économiques et idéologiques¹¹ de l'évaluation joue un rôle décisif dans la prégnance de la revendication d'objectivité telle qu'elle se manifeste dans nombre de discours et de travaux sur l'évaluation en DDL. En effet, la présentation de ces enjeux comme particulièrement « forts »¹² vient régulièrement à l'appui de la demande d'outils évaluatifs objectifs, de manière à garantir des jugements eux-mêmes objectifs.

« Lorsque des décisions prises sur la base d'informations collectées par le biais d'une évaluation entraînent des conséquences à fort enjeu, cette évaluation doit être réalisée par des spécialistes à partir d'instruments standardisés. » (Lenz et Berthele, 2010 : 25)

¹⁰ Cette analyse essentiellement socio-économique n'est néanmoins pas suffisante à elle seule et il conviendrait d'approfondir une analyse plus épistémologique et historicisante des discours scientifiques notamment : il y a ainsi bien évidemment un enjeu pour les chercheurs à techniciser l'évaluation, puisqu'ils en seraient ainsi, de fait, les seuls experts.

¹¹ Au sens de « système d'idées, philosophie du monde et de la vie » (Petit Robert 2012 : 1272).

¹² Les niveaux d'intervention conçus comme relevant d'« enjeux forts » seraient ici à questionner. On pourrait en effet juger que l'éducation, la formation et l'apprentissage constituent également des « enjeux forts », par exemple de démocratisation de l'éducation ou de formation tout au long de la vie, dont on a largement souligné l'importance dans une société de la connaissance et qui supposeraient la diffusion d'usages évaluatifs plus orientés vers une visée formative que certificative (Huver et Springer, 2011 : 282).

Dès lors que l'objectivité constitue un critère central d'évaluation de la qualité des évaluations en langues (évaluations certificatives notamment), il convient de la garantir, ce qui suppose de concevoir et de mettre en œuvre des outils impartiaux et transparents (la focalisation sur les outils, l'impartialité et la transparence étant conçues comme autant de gages d'objectivité). La psychométrie est alors posée comme la discipline scientifique de référence la plus adaptée à cet impératif d'objectivité et de transparence dans la mesure où « *en plus de permettre une quantification, [elle] donne une description plus objective et s'avère plus économique en temps et en argent qu'une évaluation « subjective » consciencieuse conçue pour fournir des informations comparables* » (Bernier et Pietrulewicz, 1997 : 14. Cf. également à ce propos Tardieu (2013) et Riba (2013), qui reprennent – sans l'interroger – cette association psychométrie – objectivité – transparence).

Cette inscription forte de l'évaluation en langue dans la psychométrie justifie et a pour conséquence le recours à des techniques fondées sur l'usage des mathématiques et des statistiques à des fins de standardisation et d'étalonnage, réservant ainsi la conception des outils d'évaluation à des experts spécialistes de ce domaine (Huver et Springer, 2011 : 30-31).

« Our goal at ETS and the purpose of standardized tests are the same — to provide fair, valid and reliable assessments that produce meaningful results. Standardized testing, if done carefully and with a high degree of quality assurance, can eliminate bias and prevent unfair advantages by testing the same or similar information under the same testing conditions. Standardized tests allow the comparison of test takers from different areas of the state, the country and the world. »
http://www.ets.org/understanding_testing/purpose.¹³

« As leaders in the field of language assessment, all our exams are specifically designed to be fair to all test takers, whatever their age, gender, nationality, first language, culture or ethnic background. Quality and fairness are at the centre of everything we do:

- *Quality – this means validity, reliability, impact and practicality.*
- *Validity – the extent to which exam marks can be considered a true reflection of underlying ability.*
- *Reliability – the extent to which test results are consistent, accurate and, therefore, dependable.*
- *Impact – the positive effect of a test on candidates and other users, including society as a whole.*
- *Practicality – the extent to which a test is workable in terms of the resources needed. »* (<http://www.cambridgeesol.org/what-we-do/research-validation.html>)¹⁴.

¹³ « L'objectif d'ETS et des tests standardisés est le même : fournir des évaluations justes, valides et fiables, qui produisent des résultats significatifs. Les tests standardisés, s'ils sont exécutés soigneusement et avec l'assurance d'un haut degré de qualité, peuvent éliminer les biais et prévenir des injustices en testant les mêmes informations (ou des informations similaires) dans les mêmes conditions. Les tests standardisés permettent de comparer les personnes testées de différentes régions, de différents pays ou même du monde entier. » (ma traduction)

¹⁴ « En tant que leaders dans le champ de l'évaluation en langue, tous nos examens sont construits pour être justes envers tous les candidats, quel que soit leur âge, le genre, leur nationalité, leur langue première, leur culture ou leur origine ethnique. La qualité et l'équité sont le centre de tout ce que nous faisons :

- Qualité – cela signifie validité, fiabilité, impact et opérationnalité.
- Validité – la mesure dans laquelle les traces dans les examens peuvent être considérées comme un reflet véritable des capacités sous-jacentes.
- Fiabilité – la mesure dans laquelle les résultats des tests sont cohérents, précis et donc, fiables.
- Impact – l'effet positif d'un test sur les candidats et sur les autres utilisateurs, dont la société dans son ensemble.

Le recours préconisé, assumé et explicite à la psychométrie en DDL contribue ainsi à faire basculer l'évaluation du côté de la mesure, puisque, justement, la psychométrie est la science de la mesure¹⁵. Se crée alors un axe paradigmatique fort associant objectivité, mesure et démarches quantitatives (paradigme objectiviste), posé comme opposé à la subjectivité, à l'interprétation et aux démarches qualitatives (paradigme subjectiviste).

Par ailleurs, la subjectivité est, dans les discours circulants sur l'évaluation, régulièrement associée à l'opacité et donc à la partialité et à l'injustice, et apparaît donc comme indéfendable dans le cadre des « enjeux forts » précédemment évoqués. L'objectivité est à l'inverse résolument associée à la transparence ; la mesure et la quantification (« la réalité incontestable des chiffres »...) ont alors une fonction de repoussoir de la subjectivité et de l'opacité. L'appui sur le recueil de traces, qui ont un statut de preuve, participe également de cette exigence d'objectivité et de transparence : se dessine ainsi un fil rouge reliant objectivité, mesure, transparence et traces.

2.2. De l'objectivité à la traçabilité : le fil rouge de la transparence

Dans le paradigme objectiviste, la notion de trace :

- *est avancée pour garantir l'objectivité de l'évaluation* : la trace a alors une fonction de *preuve*, puisqu'elle est sélectionnée et/ou présentée pour établir l'existence indubitable du phénomène qui fait l'objet de l'évaluation ;
- *est considérée comme objective* dans la mesure où elle est conçue comme pré-existant et/ou pouvant exister en dehors des représentations, des intentions et des pré-supposés de la personne qui les sélectionne, i.e comme absolues et nécessaires (*vs* contingentes et relatives).

Cette suprématie des traces en évaluation est par exemple mise en évidence dans cette page de garde d'un portfolio d'Education Physique et Sportive (1) ou encore dans le témoignage suivant (2), qui traite des portfolios de compétences professionnelles.

• Opérationnalité - la mesure dans laquelle un test est faisable en termes de ressources nécessaires » (Ma traduction)

¹⁵ initialement et étymologiquement de l'intelligence

Mon portfolio _____ Nom / Prénom : _____

Comment utiliser ce « PORTFOLIO » ?

À partir d'ici tu peux "naviguer" selon ton envie et te rendre indifféremment dans les cahiers suivants :

CAHIER A
À la découverte de ton corps

A1 Qui es-tu ? p. 2

A2 Ton corps, ton squelette, tes muscles p. 7

A3 Tu prends soin de ton corps p. 17 (Alimentation, étirements)

CAHIER B
À la découverte de mes apprentissages

B1 L'apprentissage et la fiche d'analyse p. 2

B2 Mes fiches d'analyse et mes apprentissages classés par activité physique et sportive p. 5

B3 Synthèse de mes apprentissages p. 14

CAHIER C
Preuves et Traces

C1 École primaire p. 3

C2 Cycle d'Orientation p. 4

C3 Écoles supérieures p. 5

C4 Activités Physiques et Sportives extrascolaires p. 6

Tu ne sais pas que choisir, alors suis-moi, allons au cahier A
À la découverte de ton corps

Portfolio EP 2000 Présentation page 4 Version juillet 2001

(1) http://www.edu.ge.ch/co/renard/humanitaire/portfolio/00_Cahier_Presentation.pdf

(2) « Ainsi les IUFM créeront-ils un logiciel d'«e-portfolio» où chaque compétence (du référentiel appauvri, bien sûr) sera évaluée à partir d'un «dépôt» ou d'une «soumission» informatique du stagiaire. L'obsession du contrôle s'installera : chaque formateur d'IUFM devra se demander «Quel type de trace vais-je demander aux stagiaires pour prouver objectivement l'atteinte de la compétence ?». «Kèltipdetrass ?» sera le mot d'ordre obsédant envahissant la préparation de ses cours. «Latrass ! latrass ! d'abord latrass ! Le contenu - éventuel - ensuite ! ». Il faudra que cela rentre effectivement dans l'e-portfolio. » <http://www.r-lecole.fr/AncienSite/iufm/portf.html>¹⁶

Si cet extrait et l'article dans lequel il s'insère prête par ailleurs à discussion (notamment autour des conceptions de l'éducation et de la formation qu'il contribue à véhiculer), il est néanmoins intéressant de constater qu'il critique l'outil portfolio en dénonçant une « obsession » du contrôle, associée ici aux traces et à l'objectivité. Or, cet impératif de traçage et de traçabilité à des fins de contrôle, ainsi que la croyance dans l'objectivité de la trace dont celui-ci relève, s'inscrivent dans un mouvement sociétal plus large qui déborde amplement le champ de la DDL. Très répandu dans le domaine de l'agro-alimentaire, on le retrouve également à différents niveaux de la gestion industrielle et dans le secteur du contrôle qualité ; les liens avec les problématiques liées à l'évaluation sont, dans ce dernier cas, évidents. Viruégas (2005 : 13) évoque ainsi les liens historiques entre traçabilité et évaluation (ou plutôt mesure), en faisant remonter le concept de traçabilité « au début des années 60 dans des manuels militaires américains de définition des bonnes pratiques de mesure ».

On peut par ailleurs observer une terminologie transversale à ces divers domaines d'intervention (de l'enseignement des langues au secteur industriel en passant par le secteur des services) autour des termes *transparence*, *planification*, *performance*, *précision* : ces termes et la manière dont ils sont investis relèvent clairement du paradigme objectiviste

¹⁶ Je remercie Marc Debono de m'avoir fait part de cet exemple.

évoqué ci-dessus, mais contribuent également à renforcer sa présence et sa prégnance, comme l'illustrent ces extraits :

« La demande d'information et de **transparence** des consommateurs et des utilisateurs devient de plus en plus forte vis-à-vis de la chaîne de distribution et de la chaîne d'approvisionnement : le consommateur souhaite **une information précise**. Cette attente traduit sa préoccupation, elle est relayée par les politiques qui ont transformé la traçabilité des produits en obligation. (...) **La traçabilité est une nécessité** dans le cadre des actions :

- de maîtrise de la qualité : la mise en œuvre d'un système de traçabilité (...) **garantit** (...) **la véracité** des informations présentes sur les produits ;
- de **maîtrise** de la logistique (...);
- de respect de la **règlementation**. » (Faraggi, 2006 : 1-3)

« We adhere to the following maxims of test impact :

- Maxim 1: **PLAN** – Use a **rational** and **explicit** approach to test development
- Maxim 2: **SUPPORT** – Support **stakeholders** in the testing process
- Maxim 3: **COMMUNICATE** – provide **comprehensive, useful and transparent** information
- Maxim 4: **MONITOR** and **EVALUATE** – collect **all** relevant data and analyse as required. » (<http://www.cambridgeesol.org/assets/pdf/general/pogp.pdf>)¹⁷

Dès lors que la garantie d'objectivité et de véracité passe par un recueil de traces, que faire avec, sans, malgré ce qui ne laisse pas de trace, i.e. la non trace, le non traçable ? Poser cette question suppose préalablement de reconnaître l'existence d'éléments inaccessibles et donc une part d'opacité constitutive au sein de l'évaluation. En d'autres termes, si on se place du point de vue objectiviste, on pourrait estimer que la réflexion se trouve alors dans une impasse, puisque *le fait même* de poser la question de ce qui est / reste / semble inaccessible parce qu'altéritaire mine les fondements du paradigme épistémologique dans lequel ce point de vue s'inscrit.

C'est sous cet angle que je souhaite, dans ce qui suit, revisiter les travaux sur l'évaluation développés dans Paraadiv, qui sont d'abord passés par une réflexion sur les outils, dont la prégnance de la demande peut s'interpréter comme autant de stratégies pour sortir de l'impasse en s'attaquant aux inaccessibles au moyen de procédures technicistes.

3. Le technicisme pour sortir l'évaluation de l'impasse des inaccessibles ?

La demande d'outils livrés « clé en mains », qui me semble très forte en DDL en général (cf. le nombre important d'articles scientifiques consistant en la création, l'amélioration ou l'évaluation d'outils, à mettre en lien avec la manière dont est conçue l'intervention en DDL), renvoie plus largement à la question des « bonnes pratiques ». Elle relève, à mon sens, d'une forme d'illusion de neutralité techniciste, qui *in fine* se nourrit de et nourrit le paradigme objectiviste précédemment évoqué (je n'ai ici pas les moyens de développer ces aspects et notamment la question des relations entre technicité et positivisme – cf., à cet effet, les travaux de Merleau-Ponty, Simondon, Ellul, etc.). Cette réflexion présente des enjeux

¹⁷ « Nous adhérons aux maximes suivantes concernant l'impact des tests :

- Maxime 1 : Planifier. Utiliser une approche rationnelle et explicite pour l'élaboration des tests ;
- Maxime 2 : Soutenir. Soutenir les parties prenantes dans le processus d'évaluation (de test) ;
- Maxime 3 : Communiquer. Fournir des informations complètes, utiles et transparentes ;
- Maxime 4 : Contrôler et évaluer. Collecter toutes les données pertinentes et les analyser selon les procédures requises » (ma traduction).

particulièrement visibles et importants dans le cadre de travaux sur l'évaluation, dans la mesure où les questions d'objectivité y prennent une saillance toute particulière.

3.1. Neutralité, transparence, exhaustivité : la technique contre les inaccessibles

De manière assez large en DDL, et en particulier dans le champ de l'évaluation en langues pour des publics migrants, la demande sociale se concentre très fortement sur la conception d'outils (certaines des citations *supra* sont très parlantes, de ce point de vue).

Les outils produits pour l'évaluation des ENA recouvrent des formes et des fonctions si diverses qu'il peut sembler exagéré d'affirmer qu'ils relèvent majoritairement du paradigme objectiviste précédemment évoqué. En effet, tous – loin de là – ne prennent pas la forme d'un *test* de langue, prototypique de l'évaluation – mesure (ainsi, c'est le DELF scolaire et non le TCF qui est proposé aux ENA souhaitant faire reconnaître leurs acquis en français). De même, les évaluations en langues dites « d'origine », si elles prennent des formes généralement fermées (pour des raisons de facilité d'utilisation), ne sont pas étalonnées sur des niveaux au moyen de techniques issues de la psychométrie.

... Tout du moins, l'héritage n'est-il pas explicité, conscientisé, accepté. L'objectivisme ne se réduit pas à ces dimensions formelles. On peut en effet considérer qu'au-delà de cette diversité d'outils, il existe un allant de soi commun relevant de ce que l'on pourrait appeler une « illusion de transparence », qui recoupe certaines des valeurs au fondement du paradigme objectiviste. Dans ce paradigme en effet, l'opacité, l'immatériel, le non-traçable, lorsqu'ils sont même pensés, sont considérés comme nuisibles à la qualité de l'évaluation, générant de l'arbitraire, voire du conflit, quant aux jugements produits. Par conséquent, ils constituent des éléments qu'il convient d'endiguer et d'éradiquer. Pour ce faire, il s'agit de mettre en place des *outils* appropriés, qui sont donc de fait considérés à la fois comme objectifs et garants de l'objectivité¹⁸. Leur opacité éventuelle est dans ce cadre essentiellement d'ordre technique et peut à terme être supprimée dès lors que l'on se dote des moyens (i.e. des outils) adéquats.

Par exemple, Viruéga (2005 : 18-20) considère l'incertitude comme un élément inhérent au secteur de la gestion industrielle, qui remet en question les systèmes de traçabilité standards. Il indique également que « *la traçabilité est un système dynamique composé d'une partie immatérielle* [et qu']il est donc inutile de réduire un système de traçabilité à un système d'étiquetage ou à une fiche suiveuse » (*op. cit.* : 73-74). On pourrait donc penser qu'il reconnaît, au moins partiellement, une part d'opacité et de subjectivité au sein de ces systèmes. Cependant, conjointement, il estime que la traçabilité constitue un outil de *maitrise* de l'incertitude : les démarches qu'il propose restent ainsi finalement toutes d'ordre technique, puisqu'il s'agit de mettre en place des normes ISO, des démarches d'étalonnage, de codification, etc., bref, des outils pour passer « *de la qualification à la quantification* » (*op. cit.* : 185). Dans ce cadre, ni le statut épistémologique de la trace comme signe univoque, ni celui de l'outil comme moyen objectif de produire des traces univoques ne sont interrogés. Par ailleurs, il est intéressant de noter que les outils visent des formes d'exhaustivité, i.e. visent à repousser les limites de ce qui est inaccessible : Viruéga (2005 : 184) parle ainsi significativement de « *méthodologies couvrant toutes les dimensions nécessaires* ».

Ce détour par le secteur de la gestion industrielle peut certes sembler caricatural au lecteur. On retrouve pourtant des conceptions similaires dans le champ de l'évaluation en langues, où le traçage est perçu comme gage de rentabilité et de fiabilité. En effet, là aussi, objectivité et inaccessibles sont non seulement posés comme antinomiques mais aussi associés à des connotations valorisantes du côté de l'objectivité et dévalorisantes pour la subjectivité (et les

¹⁸ Et ce, à l'instar des traces (cf. *supra*), et sans doute justement parce qu'ils ont pour fonction de faire produire des traces (situations d'évaluation) et de classer celles-ci (outils de jugement).

inaccessibles qui la fondent). Il n'est alors guère étonnant que l'évaluation (et/ou la réflexion sur l'évaluation) prenne des formes assez technicisées, la technique étant envisagée comme un moyen de produire de la transparence et donc comme un moyen d'endiguer la subjectivité et l'opacité de l'activité évaluative (i.e. les inaccessibles qu'elles génèrent et dont elles découlent)¹⁹. De ce fait, le « bon » outil est, comme dans le domaine de la gestion industrielle, pensé sur le mode de l'exhaustivité, de la transparence et de la rentabilité :

- exhaustivité, puisque les outils visent à rendre accessible tout ce qui, sans eux, resterait inaccessible. C'est ainsi que l'on peut interpréter les outils qui démultiplient les activités et/ou les critères, de manière à appréhender l'entièreté de l'objet évalué ;
- transparence, puisque l'accessibilité des données évaluées est conçue comme dépendante de la qualité de l'outil en lui-même et pour lui-même, en dehors de la diversité de ses usages ;
- rentabilité, puisque la transparence, l'objectivité, la fiabilité etc. constituent autant d'arguments d'excellence du produit (en l'occurrence de la certification en langue), alors que des formes d'évaluation perçues comme plus subjectives (lettres de recommandation, portfolios, etc.) sont plus facilement suspectées de « parti-pris ».

Or, on peut concevoir les outils de manière radicalement différente : non pas comme une technique, i.e. une pratique codifiée indépendante des acteurs qui la conçoivent et la mettent en œuvre, mais comme inséparables de leurs utilisateurs et des modalités d'utilisations (i.e. des usages) et donc comme constitutivement producteurs et produits de représentations – diverses et en tension.

3.2. Les outils et leurs usages : produits et producteurs de représentations

L'évaluation renvoyant, étymologiquement et notionnellement, à la question des valeurs, nous sommes parties, dans le cadre de Paraadiv, de l'idée que les outils d'évaluation, plus que les outils d'enseignement, peuvent être envisagés comme des « terrains » permettant d'explorer les valeurs, les attentes, et les normes des acteurs (enseignant, institution, etc.), notamment, pour la DDL, les rapports à la norme et à la diversité / altérité linguistiques. C'est pourquoi, dans une première phase, nous avons travaillé à analyser les outils d'évaluation mis à disposition des enseignants dans le cadre des évaluations diagnostiques des ENA à leur arrivée dans le système scolaire français, non pas sous l'angle des méthodologies de référence ou encore de la qualité²⁰ des outils, mais plutôt sous l'angle des représentations sous-jacentes à ces outils et à la manière dont ceux-ci contribuent, en retour, à construire des représentations de la langue, de la norme et de l'altérité/pluralité linguistique et culturelle (Huver, 2009 ; Goï et Huver, 2010).

Il s'est ainsi agi de repérer, dans les activités et référentiels d'évaluation, les « empreintes » (ou en tout cas, ce que nous identifions comme telles – cf. Goï et Huver, 2013a et c) des représentations de leurs concepteurs concernant notamment le public visé et la part d'altérité dont celui-ci serait porteur. Il en est ressorti que les outils d'évaluation projettent des images particulières des enfants évalués, qui, selon les cas, peuvent apparaître comme sans langue, (non) francophones ou encore allophones, mais plus rarement comme plurilingues (ce que la terminologie adoptée par l'Education Nationale dans les Instructions Officielles d'octobre 2012 – enfants *allophones* nouvellement arrivés – tend à confirmer).

¹⁹ Pour une perspective généalogique sur ces questions, on peut lire avec profit *Surveiller et punir* de M. Foucault, qui montre de manière pertinente l'arc existant entre examen, trace et technologie et la manière dont cette association s'est construite.

²⁰ *Qualité* étant ici à opposer à *pertinence*, la qualité pouvant être a-contextuelle, alors que la pertinence est nécessairement située.

Des enfants sans langue

Certains concepteurs proposent de recourir à une épreuve d'intelligence non verbale afin d'« évaluer les capacités intellectuelles de sujets peu scolarisés ou peu familiers avec la culture française » (<http://www.cnam.fr/kp/france/profs/test/samuel/garde.htm>). Cet exemple illustre bien le fait que la culture et la formation professionnelles jouent un rôle indéniable dans la manière d'appréhender le public et les modalités de son évaluation : l'outil évoqué ici a en effet été créé dans le cadre d'un ouvrage coordonné par Ginette Francequin, enseignante chercheuse en psychologie clinique et sociale et dont certain.e.s des auteur.e.s exercent dans en tant que conseillers d'orientation psychologues (cf. également *infra*). Par ailleurs (mais les deux aspects ne sont pas indépendants), les tests se fondent sur l'idée qu'il est possible d'accéder à l'intelligence, aux compétences et aux stratégies cognitives (l'influence de la psychométrie est ici manifeste), et ce, sans mobilisation de ressources langagières. L'enfant est alors considéré, d'une certaine manière, comme « sans langue », puisque la part centrale de la langue / des langues dans le développement cognitif et culturel est ici complètement occultée et/ou impensée et/ou universalisée.

Des enfants non francophones

La plupart des évaluations diagnostiques en français portent sur des connaissances : la langue est ainsi envisagée comme une somme d'éléments discrets (les connaissances) que l'on peut évaluer indépendamment. En outre, ces connaissances sont essentiellement d'ordre linguistique (*vs* langagier, discursif, sociolinguistique, etc.), puisque les éléments discrets évalués renvoient aux catégories classiques de la linguistique structuraliste (morphologie, lexique, syntaxe, etc.). Enfin, la langue de référence (i.e. la norme linguistique) est ici une langue écrite, homogénéisée, franco-centrée, voire littéraire : ainsi, dans un exercice de conjugaison du *Kaléidoscope polyphonique* (Broussy *et al.*, 1998), les enfants doivent conjuguer les verbes de textes de J. Giraudoux et de J. Verne.

Certaines activités sont certes plus communicatives et intègrent par exemple des critères comme l'intelligibilité du message²¹. Toutefois, là encore, la norme de référence reste largement franco-centrée et homogénéisée. Enfin, sauf exception (par exemple Rameau, 2004), aucun des outils d'évaluation ne considère l'enfant au regard non de la seule (non) maîtrise d'un code, mais de la diversité des ressources (dont possiblement des ressources non langagières et/ou non francophones) qu'il mobilise à des fins d'insertion linguistique et sociale. Dans ces outils, l'enfant apparaît donc essentiellement comme non francophone, puisqu'il est évalué « en creux » au regard de la distance entre ses savoirs sur *le* français et ce que sait (devrait savoir) le locuteur natif, l'élève idéalisé.

Des enfants allophones

Le recours à des évaluations dans d'autres langues que le français (i.e. dans la / une des langue(s) de référence des élèves) contribue à construire l'enfant, à ses propres yeux comme aux yeux des acteurs de l'établissement d'accueil (enseignants de FLS, enseignants de la classe « ordinaire »...), non pas – ou pas seulement – comme *non* francophone, mais comme allophone (locuteur – d'au moins – une autre langue), non pas comme *incompétent*, mais comme doté de compétences construites avant, ailleurs, autrement.

²¹ Reste à savoir comment ce terme est interprété par les évaluateurs. Le terme *intelligibilité* renvoie en effet aux deux sens suivants :

- « *ce qui peut être perçu par l'ouïe* », (Petit Robert, 2012 : 1350) auquel cas on s'intéressera à la qualité de la réalisation phonologique du message ;
- « *ce qui est aisé à comprendre* », (*ibid.*) auquel cas on s'intéressera à une forme de performativité et d'efficacité communicative plus globale.

Nous verrons plus loin que cette question de l'interprétation, bien que centrale, est souvent largement évacuée.

Nous verrons que la mise en œuvre de ces évaluations n'est pas sans poser question, ce qui nous a d'ailleurs incitées à travailler la question des outils en relation avec leurs usages (cf. *infra*). On peut toutefois d'ores et déjà souligner le fait que la conception de ces épreuves n'est pas sans rapport avec la manière dont l'école – et la société – française pense et considère l'altérité. En effet, ces épreuves sont des épreuves françaises rédigées en français qui sont *ensuite* traduites dans les langues de migration les plus fréquentes, en pensant cette traduction comme une simple transposition (i.e. comme un acte technique et neutre) : la lisibilité et la faisabilité des activités s'en trouve, du point de vue des évalués, largement entravées (pour des exemples : cf. Huver, 2009 et Goï et Huver, 2010).

L'introduction d'évaluations dans « la » langue de scolarisation antérieure des élèves est ainsi le produit d'une évolution des représentations sur ces élèves (déplacement des perceptions de la non francophonie à l'allophonie), et, conjointement, elle participe à diffuser cette évolution. Ces exemples illustrent l'intérêt qu'il y a à mettre en œuvre des évaluations qui permettent de faire apparaître les enfants sous un angle autre que celui de leur incompétence en français (au sens de français de France, écrit voire littéraire). Mais ils illustrent en même temps les difficultés de l'institution scolaire française à penser l'altérité de manière décentrée et polycentrée (ou le souhait plus ou moins volontaire de ne pas le faire). L'altérité est ainsi majoritairement pensée à partir d'une identité posée comme universelle, donc sous l'angle de la (réification de la) *différence* (vs de l'écart, si on reprend la distinction de F. Jullien, 2008). Il est ainsi significatif que cette tentative de prise en compte de l'altérité aboutisse dans certains cas à des formes d'assignation culturaliste à la différence. C'est par exemple le cas dans une activité de compréhension écrite du *Kaléidoscope polyphonique*, dans laquelle l'enfant doit répondre aux questions après qu'on lui a retiré le texte, au motif que « *les élèves qui arrivent ont souvent suivi des scolarités dans des systèmes où la pédagogie est souvent basée sur une mémorisation intensive et néglige l'aspect analytique et critique* » (Broussy *et al.*, 1998 : s.p.).

3.3. De quelques conséquences de ce constat

Si des travaux existent sur les liens entre outils d'enseignement et représentations (cf. notamment les travaux de G. Zarate sur les représentations de l'étranger, qui remontent au début des années 90), cette analyse n'est pas / peu menée (à ma connaissance en tout cas) pour les outils d'évaluation. Or, les quelques exemples avancés ci-dessus montrent que les activités d'évaluation ne sont pas exemptes de représentations de la langue, de l'école, de l'altérité, etc. On peut même dire que l'outil d'évaluation, en ce qu'il pose par définition et par étymologie la question des valeurs, constitue un *prisme* qui permet d'appréhender certaines compétences des ENA tout en renvoyant une image particulière, *parmi d'autres possibles*. Cependant, les conséquences de ce constat ne sont que peu tirées, alors qu'elles ne sont pas tout à fait du même ordre pour l'évaluation que pour l'enseignement, dans la mesure où la problématique de l'objectivité est particulièrement structurante dans le champ des recherches en évaluation.

Reconnaitre le fait que toute situation d'évaluation et/ou que tout outil de jugement est empreint de représentations, et que toute acte évaluatif comporte une part d'interprétation revient à dire que tout outil d'évaluation est subjectif, même ceux dont les modalités de passation sont les plus standardisées, étalonnées, objectivées. De ce point de vue, on pourrait être tenté d'affirmer – dans un premier temps en tout cas – que la question des inaccessibles signe l'acte de décès de la transparence et de l'objectivité en évaluation. Face à ce constat, deux voies peuvent être explorées. C'est par exemple ce qu'indique Ljalikova (2010 : 72-74)

« Il apparaît qu'aux paramètres quantitatifs et qualitatifs, bref mesurables et observables, s'ajoutent aussi des paramètres non quantifiables ou relatifs à un contexte et

à un paradigme particulier. Par exemple, il arrive que les mêmes crus, conservés dans la même cave, n'aient pas le même goût. À chaque fois, proposer du vin et ouvrir une bouteille comporte un peu de risque. (...) Le problème humain est de reconnaître l'expérience de l'immense. Il nous semble que de façon générale, il est possible de repérer deux grandes tendances (...). Certains essaient de réduire les choses inconnues aux dénominateurs connus. Ainsi, nous assistons à la multiplication des critères de référence qui témoigne de la volonté de définir le qualitatif par le quantitatif et le quantifiable. Les autres reconnaissent l'immense et l'imprévisible. »

En d'autres termes, une de ces voies consisterait à tenter d'éradiquer les inaccessibles au bénéfice d'une objectivité retrouvée, l'autre à faire avec ces inaccessibles pour construire, sur cette base, une autre conceptualisation de l'évaluation.

La voie de l'« éradication » peut être illustrée au moyen de la métaphore du *bullet time*, développée dans l'introduction de ce numéro. On pourrait *multiplier* les activités afin d'accéder à l'ensemble des informations jugées nécessaires et de saisir l'enfant dans *toutes* ses dimensions et dans *toute* sa complexité. Certains des premiers textes rédigés dans le cadre de ce qui allait devenir Paraadiv portent d'ailleurs la trace de cette aspiration à accéder de manière exhaustive – par des outils, autre fait significatif – à ce qui resterait, avec d'autres pratiques et d'autres outils, inaccessible à l'évaluateur.

*« Il nous semble que se dessine ici une perspective de recherche-intervention intéressante visant à imaginer des dispositifs – et des **activités** – permettant, **sinon d'évaluer, au moins d'avoir accès à la diversité** potentielle des pratiques littéraires des enfants et de leurs modalités de transmission /appropriation : hors de l'école (pratiques et représentations littéraires familiales), mais aussi hors de la classe (avec les groupes de pairs, dans la cour de récréation par exemple). » (Huver, 2009 : 162).*

Si l'on caricature ce raisonnement, il faudrait construire des activités évaluant toutes les dimensions linguistiques, mais aussi toutes les dimensions communicatives du français (ou plutôt des français, de toutes les variantes les français). Il faudrait également que les élèves puissent être évalués dans toutes leurs langues de scolarisation antérieures (en partant des « cultures » éducatives locales dans toutes leurs variétés). Il faudrait aussi inclure des éléments d'information sur leurs pratiques langagières en dehors de l'école, puisque celles-ci constituent des points d'appui potentiels pour les apprentissages et la socialisation scolaires. Peut-être également des tests permettant de diagnostiquer certains troubles (dyslexie, etc.) afin de distinguer les difficultés qui relèveraient du pathologique de celles qui relèveraient d'une simple méconnaissance de la langue. La liste est potentiellement infinie et donne le vertige²². En d'autres termes, la « solution » qui consisterait en une multiplication / juxtaposition la plus exhaustive possible des informations relatives aux élèves par le biais d'outils toujours plus nombreux et/ou plus perfectionnés s'avère au final être également une impasse...

Impasse tant que l'on raisonne dans le même cadre, sans en percevoir et en admettre les limites, pour penser qu'elles peuvent être dépassées, voire transgressées. Ainsi, plutôt que de considérer les inaccessibles en évaluation comme des éléments à éradiquer dans la mesure où ils nuisent à la qualité (i.e. à l'objectivité) du processus, on peut choisir de « sortir du cadre » en reconnaissant le caractère constitutivement subjectif et opaque de l'évaluation (au lieu de lutter contre). De ce point de vue, la question des inaccessibles (même si, initialement, elle n'était pas explicitement formulée en ces termes) a agi comme un ferment, dans la mesure où elle nous a amenés à déplacer notre réflexion des outils aux postures c'est-à-dire à nous

²² D'ailleurs, le CECR (2001 : 145) souligne que « *quelle que soit l'approche adoptée, toute pratique de l'évaluation doit **réduire le nombre de catégories** possibles à un nombre manipulable. L'expérience montre qu'au-delà de quatre ou cinq catégories on est cognitivement saturé et que sept catégories constituent un seuil psychologique à ne pas dépasser. Il faut donc **faire des choix**.* »

intéresser à la manière dont les enseignants travaillent avec ces outils et à ce qu'ils ont à dire de ces usages au regard de la manière dont ils perçoivent et construisent l'altérité de leurs élèves.

4. Des outils aux postures : les inaccessibles comme ferments

4.1. Des outils à leur(s) interprétation(s)

Dans le champ de l'évaluation, il est d'usage de distinguer évaluation quantitative et évaluation qualitative, le caractère qualitatif étant essentiellement défini « en creux », c'est-à-dire comme du non quantitatif, assimilé au non chiffré / chiffrable et au non mesuré / mesurable. Ainsi, prototypiquement, une évaluation quantitative serait une évaluation qui donne lieu à une note (chiffrée), alors qu'une évaluation qualitative donnerait lieu à une appréciation (commentaire, icône de type smiley ou couleur, etc.)²³.

Or, cette opposition ainsi posée dans l'absolu est, de ce fait même, simpliste et clivante. Prenons l'exemple des *Portfolios Européens des Langues* (désormais PEL), souvent présentés comme emblématiques d'une évaluation qualitative et formative : une rapide incursion dans les textes institutionnels visant au « cadrage » de cet outil, dans les textes de recherche ou encore dans des études de cas permet de mettre à jour le fait que les PEL relèvent, *à la fois et plus ou moins*, des logiques quantitative et qualitative.

À la fois, dans la mesure où les PEL sont, dès leur conception, traversés par des logiques contradictoires qui co-existent : ils sont ainsi présentés comme des outils d'évaluation qualitative, tout en empruntant abondamment à une terminologie renvoyant à cet idéal de transparence, de fiabilité et d'exhaustivité caractéristique du paradigme objectiviste

« *Le Portfolio européen des langues vise à documenter la capacité langagière plurilingue et les expériences dans d'autres langues de son détenteur de manière complète, concrète, transparente et fiable.* » (Schneider et Lenz, 2000 : 4)

« *On s'attend à ce que les Portfolios – selon leur contenu effectif – appuient et valorisent les points suivants : des formes d'évaluation plus exhaustives et plus complètes (notamment en évaluation formative), l'auto-évaluation, (...).* » (Schneider et Lenz, 2000 : 7)

Plus ou moins, dans la mesure où ces logiques antagonistes co-existantes sont activées à des degrés et selon des modalités variables par les acteurs, en fonction, entre autres, de leurs représentations de l'évaluation, de leurs perceptions des enjeux et de la manière dont ils articulent ces dimensions, ce qu'illustre notamment l'exemple évoqué *supra* du e-portfolio à destination des enseignants ou encore les études de cas d'utilisation du PEL dans différents pays membres du Conseil de l'Europe produites par D. Little (Little coord., s.d.).

« *Lorsque nous avons reçu le « Portfolio européen des langues » dans notre établissement, nous avons tous été très séduits par ce qui nous est apparu comme un instrument, prêt à l'usage, d'auto évaluation – celle-ci étant certainement l'une des premières étapes dans le sens de l'autonomie des élèves. Nous avons donc présenté le Portfolio en classe, indiqué aux élèves la manière de le remplir... et nous avons attendu que le miracle se produise ! Un mois plus tard, au moment de vérifier les premiers résultats, avec nos « partenaires », dans le cadre des sessions organisées en classe de seconde, nous avons constaté que rien n'avait changé. Certes, les élèves semblaient intéressés par l'idée du Portfolio ; mais il était évident que, sans un engagement encore*

²³ Ce qui renvoie, de manière plus générale, aux modalités de construction de l'opposition qualitatif / quantitatif. Cf. à ce propos Castellotti (2012).

plus marqué de la part des professeurs, le PEL allait rester à l'état de « manuel scolaire » - un de plus ! -, que les élèves ne daigneraient pas ouvrir par eux-mêmes. » (Little coord., s.d. : 19)

Ce détour par l'exemple des PEL illustre à mon sens les faits suivants, qui ont constitué un fil rouge des travaux sur l'évaluation menés dans le cadre de Paraadiv :

- aucun outil ne relève en lui-même de l'une ou l'autre sorte d'évaluation, ni par principe, ni exclusivement ;
- penser les outils (i.e. ses caractéristiques formelles) comme le principal critère discriminatif entre évaluation qualitative et quantitative relève d'une conception techniciste de la didactique en général et de l'évaluation en particulier, en ce que « *les principes de la rationalité technique sont les mêmes que ceux de la pensée scientifique, parmi lesquels le principe d'économie et de simplicité, qui conduisent à poser sans cesse la question de l'optimisation : comment obtenir le résultat cherché au meilleur coût.* » (Blay, 2006 : 777²⁴) ;
- penser les outils dans une perspective non techniciste suppose de ne pas les séparer des représentations, des interprétations, des imaginaires et des parcours biographiques et professionnels des acteurs. Le caractère qualitatif de l'évaluation ne tient alors pas tant à l'outil lui-même qu'aux conceptions de l'enseignant concernant l'auto-évaluation²⁵, l'évaluation qualitative, etc.

Dans le cadre de Paraadiv, il nous a semblé de ce fait impératif de ne pas nous en tenir à une analyse « désincarnée », c'est-à-dire vidée d'une part de la présence et de l'action des enseignants et d'autre part de leurs interprétations de ces actions (les leurs et celles de leurs collègues). C'est pourquoi nous avons souhaité prolonger cette analyse des outils d'évaluation diagnostique par une phase d'observations (filmées) des pratiques d'enseignants mettant en œuvre ces outils et d'entretiens d'auto-hétéro-confrontations (Goï et Huver, 2012).

Lors des séances d'observation, nous avons pu constater cette importante diversité des pratiques. En effet, bien que les outils d'évaluation soient communs, les enseignants / évaluateurs se positionnaient spatialement de manière très différente (en face ou à côté des élèves), géraient l'espace de manière très différente (instaurant une distance ou une proximité physique entre l'élève et eux, laissant l'élève gérer ou non cette proximité, etc.), entraient dans l'évaluation de manière très différente (par les exercices, en commençant par ce qui leur semblait le plus simple ou le plus connu des élèves, par une discussion qui a pu, à certains moments, prendre la forme d'un interrogatoire (sur les connaissances en français ou le parcours biographique des élèves, etc.).

4.2. Cultures professionnelles ? Imaginaires ?²⁶

Une première interprétation centrée sur les cultures professionnelles

Ce constat fait, il restait à essayer de « comprendre » cette pluralité, que nous avons dans un premier temps attribuée à la diversité des trajectoires et des cultures professionnelles des

²⁴ L'auteur soulignant par ailleurs que cette conception installant science et technique dans une relation dépendance réciproque est en fait très récente (trois ou quatre siècles), donc historiquement située.

²⁵ Dans Huver et Springer (2011 : 226), j'ai ainsi évoqué le fait que les usages des PEL s'inscrivaient dans un continuum allant de l'autocontrôle à l'« auto-questionnement » (Jorro, 2000 ; Vial, 2001) et que ces usages dépendaient non seulement des outils, mais aussi, et surtout, des représentations des acteurs sur l'évaluation, le plurilinguisme, l'enseignement/apprentissage, les langues, le rôle et le statut de l'enseignant, etc.

²⁶ Cette partie constitue une version revisitée de Goï et Huver (2013a et 2013c). Sur la pluralité des interprétations d'une même notion selon qu'elles sont passées au filtre de telle ou telle épistémé : cf. également Vial (2001).

acteurs. Les sélections des extraits qui ont servi de base à l'auto / hétéro-confrontation portent d'ailleurs l'empreinte de cette intuition première, puisque nous avons choisi des extraits « mettant en scène » ces thématiques d'une manière qui nous a semblé à la fois particulièrement saillante et contrastée. En effet, lors des prises de vue durant la passation des évaluations diagnostiques, nous avons été frappées par la différence de pratiques et de postures entre deux enseignants, Caroline et Daniel. Ainsi, si Daniel adopte une attitude que nous considérons comme très aidante, rassurante et ouverte, Caroline adopte quant à elle une attitude de retrait face à une élève nettement en difficulté. Par ailleurs, nous avons essentiellement retenu de leur biographie personnelle et professionnelle (du moins de ce qui nous en était accessible²⁷) que Caroline est enseignante certifiée de Lettres depuis 10 ans et enseignante de FLS en collège depuis 3 ou 4 ans et que Daniel est professeur des écoles depuis une trentaine d'années et enseignant de FLS en collège depuis une quinzaine d'années.

Le choix des extraits soumis à l'auto/hétéro-confrontation résultait essentiellement de ce constat initial : nous souhaitions en effet voir quelles catégories interprétatives les témoins allaient mobiliser pour nous expliquer et pour s'expliquer la situation telle qu'ils la percevaient. Or, bien qu'il soit également possible d'y repérer des convergences, les différents enseignants ayant visionné ces extraits (dont Caroline et Daniel) ont unanimement perçu les pratiques d'évaluation présentées sous l'angle de leur contraste. L'un comme l'autre ont expliqué celui-ci de manière assez causaliste en mobilisant le critère de la formation et de la culture professionnelles et en avançant le jugement selon lequel les enseignants du premier degré étaient plus aptes à pratiquer ces évaluations, comme l'illustrent les trois extraits suivants, tirés des entretiens avec Caroline et avec Daniel.

(1) *Caroline : Mais je trouve que là-dessus, mes collègues du premier degré, ils ont une vision beaucoup plus large, mais moi, je suis pas prof de maths, je suis prof de lettres, j'ai vraiment un cursus de lettres, et j'ai pas la vision globale que peuvent avoir Daniel ou Gabrielle²⁸. Et ça c'est vraiment quelque chose qui nous manque, parce que c'est vraiment au niveau de la formation que ça joue (...). Je trouve que la formation est plus intéressante par rapport à ce genre de public que celle que moi j'ai eue. Et je trouve ça... dur (...) Un professeur des écoles, il va cibler directement la chose, alors que moi, je sais pas faire ça, parce que, parce que j'ai pas eu la formation.*

(2) *Daniel : Béatrice, même Gabrielle, on va tenir le même discours. (...)*

[CG] : donc Béatrice, Gabrielle, toi, vous êtes tous les trois

D : on est INSTIT, on est instits, oui, on a cette approche là.

[CG] : Parce que Béatrice, elle est arrivée après Caroline ou Frédérique, elle est moins ancienne dans la fonction.

D : c'est vrai, c'est vrai, tu as raison. Mais dans une évaluation, Béatrice et moi, on les sent pareil les évaluations ; nos comptes rendus, ils sont identiques, on va chercher les mêmes choses.

(3) *Daniel : j'ai l'impression que l'instit... nous instit, on va essayer de valoriser. Les personnes qui ont une formation plus CAPES, plus prof, ils vont plus tac, tac, tac, et pourtant on a les mêmes lignes à lire, le même item. Je pense que pour les évaluations, c'est pareil, on a un fonctionnement, une approche un peu...*

On voit ici que la catégorie « culture professionnelle » constitue un filtre puissant d'interprétation, puisqu'il est avancé fermement et à plusieurs reprises, et qu'il résiste à

²⁷ Sachant que cette « accessibilité » est variable selon les deux enseignants (que nous connaissions plus ou moins bien) et selon les chercheuses (qui ont des liens d'amitié plus ou moins anciens et approfondis avec l'un des deux enseignants).

²⁸ Béatrice, Gabrielle et Frédérique sont des collègues également impliqués dans ce dispositif d'accueil des ENA. Béatrice et Gabrielle sont professeurs des écoles ; Frédérique est certifiée, enseignante de français.

d'autres propositions d'interprétation (comme l'ancienneté, dans l'extrait ci-dessus). Cette unanimité était, pour nous chercheuses, *a priori* bien confortable et rassurante, puisqu'elle convergait avec nos propres interprétations initiales... Pourtant, c'est aussi cette forte convergence des interprétations qui nous a amenées à prendre conscience du fait que le paramètre de la culture professionnelle était mobilisé comme un allant de soi, une évidence si partagée entre nous (enseignants et chercheuses) qu'elle nous (chercheuses notamment) empêchait peut-être de percevoir / construire d'autres interprétations qui nous seraient jusqu'alors restées « inaccessibles », forêt cachée par l'arbre que constitue cette convergence issue d'un « entre soi ».

Complexification : la question des imaginaires évaluatifs

Il ne s'est alors pas tant agi d'écarter cette première interprétation que d'essayer d'en envisager d'autres possibles. Cette relecture des observations et des entretiens nous a ainsi amenées à mobiliser la notion d'imaginaire, en ce que celle-ci nous a semblée pertinente d'une part pour rendre compte de manière intégrée des pratiques, des discours sur les pratiques et des représentations des enseignants et, d'autre part, pour nous distancier d'une conception normative et rationaliste de la notion de représentation.

Il nous est en effet apparu que les discours de Caroline et de Daniel étaient travaillés par des imaginaires très différents, qui nous semblaient tout aussi pertinents que la « culture professionnelle » pour pouvoir rendre compte de ce contraste fort que nous percevions entre leurs pratiques respectives (en tout cas, telles qu'elles nous ont été données à voir lorsque nous les avons observées). Caroline considère ainsi l'évaluation sous l'angle de ce que Jorro (2000) nomme l'imaginaire de la maîtrise, qui se caractérise par l'insistance sur les outils, l'objectivité, l'exactitude et le prescriptif : elle se plonge en effet dans une prise de notes abondante, destinée à argumenter au mieux le diagnostic effectué. En outre, lors de l'entretien, elle justifie ses pratiques par le discours suivant, qui concentre en une formule non seulement le souci des élèves et la centration sur les outils, mais aussi la conception objectivisante de l'évaluation et la nécessité de penser sa pratique dans des cadres prescrits (« il faut »).

Caroline : C'est toujours compliqué, moi je me dis qu'il ne faut pas les mettre en difficulté avec quelque chose d'immédiatement trop dur, et en même temps, il ne faut pas non plus cibler trop bas, parce qu'il faut arriver à viser juste, pour pouvoir orienter le mieux possible.

Par contraste, les pratiques de Daniel sont pétries de ce que Jorro (2002) nomme l'imaginaire de la construction. Ainsi, lors des évaluations, il ne s'attache pas tant aux résultats des activités qu'à une observation attentive des élèves en train de mener ces activités. Par ailleurs, il est vigilant à des configurations sinon inaccessibles du moins inattendues, puisque, par exemple, au lieu d'inférer un lien de causalité entre nationalité et langue de scolarisation, il demande à deux élèves ukrainiennes où elles sont allées à l'école : celles-ci révèlent alors qu'elles ont été majoritairement scolarisées au Portugal et qu'elles préfèrent donc passer les évaluations en langues d'origines en portugais (et non en russe ou en ukrainien comme on aurait pu s'y attendre). Enfin, lors de son entretien, il accepte, voire revendique, le flou et l'imprévu dans son activité évaluative et, plus largement, dans son activité professionnelle.

Daniel : Que ce soit utile, que ce soit un outil utile, tout en sachant qu'il [l'outil d'évaluation] n'est pas VRRAI. Il te donne des pistes, mais il te donne pas tout, après une semaine de travail avec l'élève tu vas te dire, ah ben il a des qualités que j'avais pas vues. L'élève, il va prendre sa place et il va parler un petit peu tiens ce jour là, j'aurais pas pensé qu'il pouvait savoir tout ça.

Primat des tensions et de l'hétérogénéité

Ces inscriptions imaginaires ne sont pour autant pas dénuées de tensions (pas plus que ne l'étaient les inscriptions en termes de culture professionnelle). En effet, Daniel, bien qu'il semble vigilant à la part d'inattendu et d'inaccessible lié à la diversité linguistique²⁹ et qu'il revendique une part d'intuition et d'approximation dans son activité d'évaluateur, est dans le même temps très dubitatif, voire critique, sur le fait que Frédérique mène une partie de son évaluation en anglais avec un jeune adolescent kosovar.

Daniel : Par contre là, j'ai pas trop compris : il était anglais le petit ?

EH : non, il était albanais

D : alors pourquoi tout s'est fait en anglais ? il parlait anglais ?

CG : il parlait anglais, oui, et il parlait pas français

D : il parlait bien anglais. Mais c'est pareil, on le voit pas au travail (Daniel fait défiler le film et le regarde : gestuelle dubitative – soupir + haussement des épaules)

(...) Tu vois, mais là, c'est tout en anglais

CG : c'est les mathématiques là je crois

D : mmmh ouais, mais ça c'est si il avance pas, mais il faut lui laisser un temps pour qu'il fasse quelque chose je crois

EH : parce que là, je crois, elle veut pas faire un exercice, elle veut, elle lui demande comment c'était les mathématiques au Kosovo.

De même, même s'il revendique à certains moments le flou et le subjectif dans l'évaluation, il balance à d'autres moments vers des tendances à la maîtrise et à l'objectivité.

Daniel : j'ai l'impression que j'en fais trop, qu'il faut encore plus euh (mouvement des mains – apaisement, recul) plus lâcher, mais être LA, mais PLUS le laisser euh, je sais pas pourquoi

CG : c'est marrant, parce que Caroline, elle disait l'inverse

D : ah bon

CG : ouais, elle avait un regard sur elle, où elle trouvait qu'elle en faisait pas assez, que c'était plus accompagnant, la manière dont toi tu t'y prenais. C'est drôle de voir ce double regard euh

D : ben je crois qu'il ya besoin de les accompagner de toute façon, MAIS peut-être un peu plus de recul, en tout cas, moi c'est ça que je vois

EH : tu veux dire que sinon c'est trop subjectif dans l'évaluation

D : OUI, oui

EH : c'est biaisé

D : oui, ça peut biaiser un peu, j'crois

On voit bien là que ce qui traverse les discours et les pratiques des enseignants est plus de l'ordre de la tension, de l'hétérogène, d'une recherche d'équilibre sans cesse déséquilibrée, remise en question, remise sur le métier, que d'une appartenance définitive, homogène et figée à telle ou telle catégorie (que celle-ci soit celle des cultures professionnelles ou des imaginaires). Par conséquent, il semble plus fructueux de travailler à rendre compte de ces

²⁹ cf. sa demande aux élèves ukrainiennes évoquée *supra* ou encore le fait que lors de l'entretien, il affirme regarder « s'ils ont fait un peu d'anglais, s'ils comprennent à peu près ce qu'ils lisent dans leur langue », s'il y a « des trucs qu'il[s] transfère[nt] de [leur] langue ».

tensions qu'à analyser les pratiques et les discours des acteurs pour les typologiser (cf. par exemple pour Paraadiv Goï et Huver 2013b et, plus largement, Castellotti ici-même).

4.3. Evaluation et inaccessibles

Cette réflexion sur les articulations entre évaluation et inaccessibles de l'altérité a opéré par allers-retours et mises en lien entre deux niveaux (au moins) d'inaccessibles. D'une part, il s'est agi d'explorer la part d'inaccessible inhérente à tout outil et à toute démarche d'évaluation. Celle-ci reste le plus souvent impensée ou, lorsqu'elle est pensée, est évacuée par le biais de procédures technicisantes (issues des techniques de la mesure et des statistiques) appuyées sur et renvoyant à des discours visant à promouvoir l'objectivité, la transparence, la fiabilité, elles-mêmes garanties par ces procédures technicisantes. Par conséquent, reconnaître une part d'inaccessible et donc d'opaque dans l'évaluation (même la plus standardisée et « objective ») questionne de manière radicale le statut de l'objectivité en évaluation.

Les outils occupent ici une place centrale, en ce qu'ils sont conçus et présentés comme les garants de cette objectivité tant recherchée, comme s'ils pouvaient à *eux seuls* (i.e. de par leurs caractéristiques formelles et le traitement de celles-ci) la *garantir*. Les outils d'évaluation sont en effet pensés comme des moyens d'accéder directement au « réel », voire à *la vérité* (la (vraie) compétence, le (vrai) niveau, etc.). Ils sont donc pensés comme passifs, transparents, i.e. d'une part comme s'ils n'étaient pas transformés par les usages des évaluateurs et des évalués et, d'autre part, comme s'ils ne transformaient pas cette « réalité ». Travailler sur / avec cette opacité des outils suppose donc de poser la question de ce à quoi ils permettent (ou non) d'accéder, ce qui rejoint la question de l'objectivité précédemment posée. Cela suppose également de s'intéresser aux usages évaluatifs, à la manière, nécessairement singulière et située, dont les acteurs (ici notamment les évaluateurs) s'approprient les outils, les mettent en œuvre et leur donnent sens. Se dessine alors un autre niveau d'« inaccessible », celui de « l'au-delà de l'observable », des rapports à-, des représentations, des imaginaires (en l'occurrence relatifs à l'évaluation).

Ces deux « niveaux » – si tant est qu'on puisse les distinguer – renvoient à une problématique commune qui est celle de l'interprétation ou, autrement dit, du statut et de la légitimité de ce sur quoi on s'appuie pour interpréter. Dans le champ de l'évaluation en langues, la tendance majoritaire est de chercher des traces et de leur conférer le statut de preuve de la compétence ou de l'incompétence du sujet évalué. De même, les outils sont assimilés à des techniques, à des méthodologies, que Gadamer dénonce comme un piège dès lors qu'elles s'inspirent, souvent sans le savoir, des sciences exactes, la compréhension étant dès lors assurée par une réflexion non pas tant méthodologique que philosophique³⁰. Poser la question de ce qui est inaccessible à l'évaluateur et de la manière de prendre en compte et de travailler avec ces inaccessibles³¹ fait basculer le propos dans un autre paradigme épistémologique, ce qui se manifeste entre autres par un tout autre champ sémantique et terminologique. Par exemple, il n'est plus tant question de preuves que d'indices³², voire de « mises en altérité » (Huver, 2013), en ce que la convocation des traces permettrait essentiellement aux évaluateurs de discuter entre eux de leurs évaluations et de leur caractère

³⁰ Cf. également, à ce propos, les travaux de D. de Robillard.

³¹ Rappelons encore que ces inaccessibles ne constituent bien évidemment pas un « donné », mais sont relatifs à l'évaluateur (qui peut percevoir et comprendre ou non ou partiellement les phénomènes que l'on cherche à évaluer), à l'évalué (qui peut cacher ou non ou partiellement ces phénomènes), à l'outil (qui permet d'appréhender ou non ou partiellement ces phénomènes), etc.

³² Au sens que lui donne C. Ginzburg (1989 : 143, citant S. Freud) de « traits sous-estimés ou dont on ne tient pas compte, (...) de rebut de l'observation », permettant la mise en œuvre d'une « méthode d'interprétation basée sur les écarts, sur les faits marginaux, considérés comme révélateurs » (*op. cit.* : 146).

potentiellement divergent. Les traces ne sont dès lors plus séparables des interprétations qui en sont faites (et donc des raisons pour lesquelles elles sont reconnues comme des traces). En outre, les traces n'occupent plus la totalité de l'espace évaluatif, puisqu'il s'agit également de traiter avec le non perceptible, le non observable, le non tracé, le non traçable (par exemple, l'intuition, l'impression d'ensemble, l'histoire de la relation entre évalué et évaluateur, des informations connues de l'évaluateur mais non directement visibles dans le « corpus » que représenterait la copie ou l'ensemble des observables recueillis, voire ce que l'évaluateur ignore³³). Enfin, l'évaluateur n'est plus un exécutant qui met en application un protocole ou un outil – censément neutre et objectif ou objectivé –, mais un acteur avec ses inscriptions institutionnelles, son parcours, ses représentations, son imaginaire qui inter-agit par l'intermédiaire d'un outil (lui-même inséparable de certains enjeux, certaines inscriptions institutionnelles, certaines représentations, etc.) : une partie centrale de sa professionnalité consiste par conséquent à développer un point de vue critique et réflexif vis-à-vis des outils qu'il met en œuvre et de sa propre posture d'évaluateur.

5. Ouverture – Travailler avec les inaccessibles : s'attendre aux inattendus

Ce positionnement fondamentalement réflexif et critique s'appuie sur et participe à la construction de rapports à la diversité, à l'altérité, au flou, à l'inaccessible qui traversent les *postures* (Herreros 2002, Vial 2001) professionnelles des acteurs en présence. Ceci n'est pas sans me rappeler la proposition de V. Castellotti

« d'introduire le risque et l'insécurité comme facteurs centraux de l'apprentissage, la compétence fondamentale à construire étant la capacité à accepter et à gérer l'insécurité et à se préparer constamment à l'imprévisible, cet imprévisible pouvant également résider dans la ressemblance inattendue. » (Castellotti et Huver, 2012)

Au-delà de l'évaluation en langues elle-même, il me semble que les réflexions développées ici résonnent avec des questions et des problématiques qui interrogent plus globalement la recherche en DDL et la formation des enseignants de langue³⁴. En effet, que peut signifier « former à l'évaluation » (et former tout court, d'ailleurs), dès lors que l'on souhaite laisser une place, voire revendiquer une place, pour les inaccessibles, ainsi que pour la subjectivité et l'opacité qui en découlent ? Comment et pourquoi penser des outils (d'évaluation, de formation) si on soutient par ailleurs que la perspective de la réception prime sur celle de la production si bien que ces outils seront nécessairement interprétés et mis en œuvre autrement que dans la perspective dans laquelle ils ont été initialement pensés ?

Une première réponse réside évidemment dans la mise en œuvre de démarches de formation visant prioritairement à favoriser la construction de cette posture réflexive et critique évoquée plus haut. Cependant, faire une place aux inaccessibles en formation (et notamment dans des formations à l'évaluation) suppose également d'amener à conscientiser le fait que l'altérité et l'opacité – et donc le flou, le malentendu, l'incompréhension³⁵ – sont constitutifs de toute relation, aussi bien évaluative que formative. Par conséquent, de même qu'il s'agit de former à la réflexivité et à l'esprit critique par la réflexivité et une exigence de

³³ Comme par exemple les « connaissances ignorées des élèves » (Penloup, 2008). Il faut également souligner que ces éléments constituent des potentialités, des possibilités, mobilisables de manière plus ou moins faisable et pertinente selon les fonctions et les formes de l'évaluation.

³⁴ Et plus globalement sans doute la recherche en sciences humaines et la formation professionnelle. Les particularisations ici opérées ne préjugent pas d'une quelconque spécificité de ces champs mais marquent simplement les domaines d'intérêts et de réflexion dans lesquels je m'inscris plus particulièrement.

³⁵ Autant de termes qu'il conviendrait donc de réhabiliter. Cf., dans une perspective similaire, Castellotti (à paraître et ici-même) à propos du terme *hétérogénéité*.

critique argumentée, il s'agit ici de développer des modalités formatives travaillant à partir de la pluralité des interprétations, des opacités, des malentendus, des incompréhensions de manière à rendre les acteurs en formation vigilants à ces phénomènes et en capacité de travailler avec ceux-ci.

Du point de vue de la recherche enfin, c'est bien cette conscience du fait que certains éléments nous restent nécessairement inaccessibles qui a constitué le ferment de notre réflexion dans le programme Paraadiv. Elle a par exemple motivé une réflexion qui vise à remettre en question une approche uniquement méthodologique de l'évaluation pour nous intéresser aux usages évaluatifs. La pluralité de ces usages évaluatifs constatée, c'est aussi une interrogation sur ce que nous ne percevons pas, sur ce qui nous demeurerait caché (ou *a contrario* sur ce qui était trop ostensiblement évident), qui nous a amenées à relativiser une interprétation trop centrée sur des formes de catégorisation pour privilégier les tensions, l'hétérogénéité et la manière dont les enseignants font avec, sans, malgré elles. La condition et la conséquence de cette conscience vigilante réside dans la construction d'une posture de recherche qui tente d'articuler – tant dans la réflexion que dans l'écriture – : vigilance à l'inattendu et à l'altérité et conscientisation / explicitation de nos filtres interprétatifs, de nos convictions, de nos anticipations à des fins de construction / déconstruction / reconstruction / confrontation de nos interprétations.

Comme le remarque fort justement Gadamer (1953/1996) :

« ce n'est pas ce qui ne fait pas problème et ce qui ne vient pas satisfaire les attentes de notre recherche qu'il faut encourager. Il faut plutôt découvrir – et contre nous-mêmes – d'où peuvent provenir les heurts. »

En d'autres termes, que ce soit dans le champ de l'évaluation, de la formation ou de la recherche, apprendre à travailler avec et par les inaccessibles, c'est peut-être et avant tout apprendre à penser contre soi-même... et le donner à voir (?).

Bibliographie

NB : les références aux textes publiés dans ce numéro ne sont pas reprises dans la présente bibliographie.

- BACHELARD G., 1947, *La Terre et les rêveries de la volonté : essai sur l'imagination des forces*, José Corti, Paris.
- BERNIER J.-J. et PIETRULEWICZ B., 1997, *La psychométrie. Traité de mesure appliquée*, Gaëtan Morin Editeur, Montréal.
- BIGOT V. et CADET L. (dir.), 2011, *Discours d'enseignants sur leur action en classe. Enjeux théoriques et enjeux de formation*, Riveneuve, Paris.
- BLAY M. (dir.), 2006, *Dictionnaire des concepts philosophiques*, Larousse / CNRS éditions, Paris.
- BOLTON S., 1987, *Évaluation de la compétence communicative en langue étrangère*, Hatier, Paris.
- BROUSSY M., CHARPENTIER M., FRANCEQUIN G. et al., 1998, *Kaléidoscope polyphonique, Tome 1 – Quelques outils pour accueillir et positionner les adolescents étrangers*.
- CASTELLOTTI V., à paraître, « S'approprier une langue : une confrontation à l'altérité et une mobilisation de l'hétérogénéité », *Hommages à R. Delamotte-Legrand*.
- CASTELLOTTI V., 2012, « Recherches qualitatives : épistémologie, écriture, interprétation en didactique des langues », dans C. Goï (dir.), *Quelles recherches qualitatives en sciences humaines ? Approches interdisciplinaires de la diversité*, L'Harmattan, Paris.

- CASTELLOTTI V. et HUVER E., 2012, « Mobilités et circulations académiques : dynamiques, catégorisations, évaluations – ou : "Bougez, il en restera toujours quelque chose" », dans *Le discours et la langue*. Tome 3.2., pp.117-132.
- FARAGGI B., 2006, *Traçabilité. Réglementation, normes, technologies, mise en œuvre*, Dunod, Paris.
- FOUCAULT M., 1975, *Surveiller et Punir*, Gallimard, Paris.
- GADAMER H.-G., 1953/1996, *La philosophie herméneutique*, PUF, Paris.
- GINZBURG C. (1989, version originale : 1986), *Mythes, emblèmes, traces. Morphologie et histoire*, Flammarion, Paris.
- GOÏ C. et HUVER E., 2013a, « Diversités et réflexivités : Quelles empreintes pour quelles interprétations ? », dans Béziat J. (dir.), *Analyse de pratiques et réflexivité en formation en recherche et en contextes socio-éducatifs*, L'Harmattan, Paris, pp.177-193.
- GOÏ C. et HUVER E., 2013b, « Accueil des élèves migrants à l'école française : postures, représentations, pratiques ségrégatives et/ou inclusives ? », dans M.-M. Bertucci (coord.), *Glottopol*, 21, *Lieux de ségrégation sociale et urbaine : tensions linguistiques et didactiques ?*, http://glottopol.univ-rouen.fr/telecharger/numero_21/gpl21_08goi_huver.pdf.
- GOÏ C. et HUVER E., 2013c, « Des enseignants aux chercheurs en didactique des langues. Postures épistémologiques entre identités, cultures et imaginaires professionnels », dans Galli H., Verney-Carron N. et al. (dir.), *La didactique au prisme de l'épistémologie. Une approche plurielle*, Editions universitaires de Dijon, Dijon, pp.191-204.
- GOÏ C. et HUVER E., 2012, « Ecole et construction identitaire : travailler avec l'altérité des élèves nouveaux arrivants », dans Sauvage J. et Demougin F. (dirs.), *La construction identitaire à l'école. Perspectives linguistiques et plurielles*, L'Harmattan, Paris, pp.241-247.
- GOÏ C. et HUVER E., 2010, « Evaluer – accueillir – insérer : quel(s) prisme(s) pour quelle(s) projection(s) ? », dans *Cahiers de Sociolinguistique*, 15, Rennes : PUR, pp. 97-108.
- HERREROS G., 2002, *Pour une sociologie d'intervention*, Erès, Ramonville Saint Agne.
- HUVER E., à paraître 2014, « CECR et évaluation : interprétations plurielles et logiques contradictoires », dans *Cahiers du GEPE*.
- HUVER E., 2013, « La diversité au centre de projets à construire – Réflexions autour d'un projet de recherche en didactique des langues : le projet DIFFODIA », dans Castellotti V. (dir.), *Le(s) français dans la mondialisation*, Bruxelles, EME, pp. 349-365.
- HUVER E., 2012a, « L'évaluation des enfants nouveaux arrivants : problèmes spécifiques ou problématiques transversales ? », dans *VEI Diversité*, La pression évaluatrice – Quelle place pour les plus faibles ?, 169, pp. 175-181. *Accessible en ligne* : <http://www.cndp.fr/entrepot/ville-ecole-integration/realites-et-pratiques/premiers-pas-vers-la-comprehension-du-discours-pedagogique/reperes.html>.
- HUVER E., 2012b, « Evaluation, changement, représentations – Vers une interrogation de quelques allants de soi en didactique des langues », dans *Cahiers de l'ASDIFLE*, 23, pp.13-24.
- HUVER E., 2009, « Evaluation des enfants nouvellement arrivés en France : évaluer des enfants non francophones, allophones et/ou plurilingues ? », dans *Cahiers de l'ASDIFLE*, 20, pp. 149-168.
- HUVER E. et SPRINGER C., 2011, *L'évaluation en langues*, Didier, Paris.
- JORRO A., 2000, *L'enseignant et l'évaluation. Des gestes évaluatifs en question*, De Boeck, Bruxelles.

- JULLIEN F., 2008, *De l'universel, de l'uniforme, du commun et du dialogue entre les cultures*, Fayard, Paris.
- KILANI M., 1992, « Découverte et invention de l'autre dans le discours anthropologique », dans *Cahiers de l'ILSL*, Langue, littérature et altérité, pp. 3-16.
- LJALIKOVA A., 2010, « Le qualitatif et le quantitatif dans l'évaluation universitaire », dans *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*, F. Dervin et E. Suomela-Salmi (Eds.), Peter Lang, Bern, pp.57-78.
- LENZ P. et BERTHELE R., 2010, « Prise en compte des compétences plurilingue et interculturelle dans l'évaluation », *Guide pour le développement et la mise en œuvre de curriculums pour une éducation plurilingue et interculturelle – Etude satellite n° 2*, Conseil de l'Europe, Strasbourg.
http://www.coe.int/t/DG4/linguistic/Source/Source2010_ForumGeneva/Assessment2010_Lenz_FR.pdf, (consulté le 9 décembre 2012).
- LITTLE D. coord. (s.d.). *Exploitation du Portfolio Européen des Langues : neuf exemples*, Conseil de l'Europe, Strasbourg.
<http://www.coe.int/T/DG4/Portfolio/documents/Exploitation%20du%20PEL.pdf>.
- PENLOUP M.-C. (dir.), 2008, *Les connaissances ignorées : approche pluridisciplinaire de ce que savent les élèves*, INRP, Lyon.
- RAMEAU C., 2004, *Référentiel de compétences, Dispositif de « Soutien en FLS »*,
<http://refef.crifpe.ca/document/memoire/RAMEAU.pdf>, (consulté le 9 décembre 2012).
- RIBA P., 2013, dans « Les descripteurs de compétence en langue : rappels et mise en perspective », *Recherches et applications – Le français dans le monde*, 53, L'évaluation en didactique des langues et des cultures : continuités, tensions, ruptures, Huver E. et Ljalikova A. (coord.), pp.56-66.
- ROBILLARD de D., 2007, « La linguistique autrement : altérité, expérientiation, réflexivité, constructivisme, multiversalité : en attendant que le Titanic ne coule pas », dans Blanchet P., Calvet L.-J., De Robillard D., *Un siècle après le Cours de Saussure : la linguistique en question*, Paris, L'Harmattan, pp.81-228.
- SCHNEIDER G. et LENZ P., 2000, *Portfolio européen des langues. Guide à l'usage des concepteurs*, Conseil de l'Europe, Strasbourg.
http://www.coe.int/t/dg4/education/elp/elp-reg/Source/Publications/Developers_guide_FR.pdf (consulté le 9 décembre 2012).
- TARDIEU C., 2013, « Les tests de langues : une nécessité? », dans Huver E. et Ljalikova A. (coord.), *Recherches et applications – Le français dans le monde*, 53, L'évaluation en didactique des langues et des cultures : continuités, tensions, ruptures, pp.67-79.
- VIAL M., 2001, *Se former pour évaluer, se donner une problématique et élaborer des concepts*, De Boeck, Bruxelles.
- VIRUEGA J.-L., 2005, *Traçabilité. Outils, méthodes et pratiques*, Editions d'Organisation, Paris.

GLOTTOPOL

Revue de sociolinguistique en ligne

Comité de rédaction : Michaël Abecassis, Salih Akin, Sophie Babault, Claude Caitucoli, Véronique Castellotti, Régine Delamotte-Legrand, Robert Fournier, Stéphanie Galligani, Emmanuelle Huver, Normand Labrie, Foued Laroussi, Benoit Leblanc, Fabienne Leconte, Gudrun Ledegen, Danièle Moore, Clara Mortamet, Alioune Ndao, Isabelle Pierozak, Gisèle Prignitz, Georges-Elia Sarfati.

Conseiller scientifique : Jean-Baptiste Marcellesi.

Rédacteur en chef : Clara Mortamet.

Comité scientifique : Claudine Bavoux, Michel Beniamino, Jacqueline Billiez, Philippe Blanchet, Pierre Bouchard, Ahmed Boukous, Pierre Dumont, Jean-Michel Eloy, Françoise Gadet, Marie-Christine Hazaël-Massieux, Monica Heller, Caroline Juilliard, Jean-Marie Klinkenberg, Jean Le Du, Marinette Matthey, Jacques Maurais, Marie-Louise Moreau, Robert Nicolai, Lambert Félix Prudent, Ambroise Queffélec, Didier de Robillard, Paul Siblot, Claude Truchot, Daniel Véronique.

Comité de lecture pour ce numéro : N. Auger (Université de Montpellier) – Philippe Blanchet (Université Rennes 2) – Didier de Robillard (Université F. Rabelais, Tours) – Muriel Molinié (Université de Cergy-Pontoise) – Isabelle Pierozak (Université F. Rabelais, Tours) – Marie-Madeleine Bertucci (Université de Cergy-Pontoise) – Silvia Lucchini (Université Catholique de Louvain) – Marie-Claude Penloup (Université de Rouen) – Gabrielle Budach (Université de Southampton).

Laboratoire Dysola – Université de Rouen
<http://glottopol.univ-rouen.fr>

ISSN : 1769-7425