ASp
la revue du GERAS

39-40 | 2003
Varia

Un point sur l'approche qualitative de l'évaluation

Claire Bourguignon et Jean-Paul Narcy-Combes

Édition électronique

URL : <http://journals.openedition.org/asp/1377>

DOI : 10.4000/asp.1377

ISBN : 978-2-8218-0392-3

ISSN : 2108-6354

Éditeur

Groupe d'étude et de recherche en anglais de spécialité

Édition imprimée

Date de publication : 1 mai 2003

Pagination : 171-181

ISSN : 1246-8185

Référence électronique

Claire Bourguignon et Jean-Paul Narcy-Combes, « Un point sur l'approche qualitative de l'évaluation », *ASp* [En ligne], 39-40 | 2003, mis en ligne le 18 mai 2010, consulté le 19 avril 2019. URL : <http://journals.openedition.org/asp/1377> ; DOI : 10.4000/asp.1377

Ce document a été généré automatiquement le 19 avril 2019.

Tous droits réservés

Un point sur l'approche qualitative de l'évaluation

Claire Bourguignon et Jean-Paul Narcy-Combes

Introduction

- 1 Plusieurs années de travail en commun sur l'évaluation nous ont donné envie d'unir nos efforts pour approfondir la réflexion sur une approche qualitative de l'évaluation en langue.
- 2 L'ampleur de la tâche est telle que nous en sommes encore au stade du débroussaillage. Nous ne parvenons qu'à ouvrir des pistes, et nous restons dans le contexte de l'évaluation certificatrice. Nos propos ne sauraient s'appliquer aux autres types d'évaluation. En prenant la définition qu'en donne Chapelle : *a meaningful interpretation of observed behavior* (Bachman & Cohen 1998 : 33), nous pouvons dire que notre *construct* est maintenant mieux élaboré, il nous conduit non seulement à prendre en compte l'acte de parole, mais également ce qui a conditionné l'énonciation de cet acte. Voilà donc par où cet article commencera, ce qui permettra ensuite de prendre la mesure des apports et des limites de ce *construct* pour l'évaluation. Dans la mesure où il ne prend pas encore le contenu en compte, nous aborderons ensuite l'état de l'art sur les effets que jouent le contenu et le contexte dans les performances des apprenants. Nous pourrions alors déterminer provisoirement les caractéristiques de ce que nous appellerons une évaluation orientée-sujet. Au vu de ce qui intéresse le domaine LANSAD, nous tâcherons alors de déterminer en quoi tout ce qui précède affecte le CLES et ce que seraient alors les besoins en formation des enseignants sur ce point.

1. Notre problématique ou comment notre *construct* s'est mis en place

- 3 La journée de travail sur l'évaluation qu'Heather Hilton a organisée à Chambéry, en septembre 2002, était la deuxième rencontre consacrée à l'évaluation sous l'égide du GERAS. Elle s'inscrivait dans la continuité de la première pour approfondir la réflexion entamée au Colloque de Grenoble en 2001 qui a fait l'objet du numéro 34 de cette revue.
- 4 Or, la poursuite du troisième larron de Jean-Paul Narcy-Combes (2001) pouvait tout à fait se faire dans l'espace d'évaluation tel que Claire Bourguignon l'avait défini. Nous avons décidé, non seulement de travailler ensemble, mais de conjuguer nos réflexions dans une problématique commune.
- 5 Nous sommes donc partis de l'hypothèse que l'évaluation était un acte générateur d'un espace de communication, cet espace étant constitué de plusieurs paramètres : l'évalué, l'évaluateur, la langue et le contenu. (cf. schéma 1 en annexe).
- 6 C'est dans ce cadre qu'il nous est apparu que le type d'évaluation adoptée était fonction de la relation que ces différents paramètres entretenaient. À partir de là, nous avons dégagé deux paradigmes.

1.1. Le paradigme de la connaissance

- 7 Le premier paradigme, celui de la connaissance implique une centration sur l'objet langue que l'on peut représenter à l'aide du schéma 2 situé en annexe.
- 8 Dans ce cadre, se met en place un espace d'évaluation que nous avons qualifié précédemment d'« espace d'évaluation orienté-objet »¹. Il s'agit pour l'évaluateur de vérifier la connaissance que l'évalué a de la langue, objet qui lui est extérieur et qui constitue le contenu même de l'évaluation. Ainsi, dans un processus de transmission unidirectionnelle, de l'évalué vers l'évaluateur, l'évaluation se fait sur le mode vrai / faux dans une logique qui s'apparente plus à la logique mathématique qu'à une logique communicationnelle. En effet, l'évaluateur se focalisera sur l'erreur par rapport à des attendus identiques pour tous les évalués. Nous nous apercevons donc que l'évalué, en tant que sujet, est évacué de ce type d'évaluation.
- 9 Pour reprendre les propos de B. Rey (2000), nous pouvons dire que, dans ce cas, il s'agit d'évaluer un « objet mort » selon le principe qu'« évaluer un savoir, c'est évaluer un objet mort qui n'est qu'un assemblage réglé et organisé de signes. Si un savoir a un sens, cela signifie qu'il produit du sens »
- 10 Notre objectif étant de réfléchir à un type d'évaluation portant sur l'aptitude à communiquer de l'évalué, en d'autres termes, à une évaluation de la compétence de communication de l'évalué en tant que sujet communicant, il apparaît clairement que l'évaluation orientée-objet ne correspond pas à notre *construct* de départ.
- 11 Ceci nous amène au deuxième paradigme que nous avons retenu, celui de la compétence avec une centration sur le sujet.

1.2. Le paradigme de la compétence

- 12 Une fois établi que nous nous situons sur le paradigme de la compétence du sujet, il est nécessaire de définir ce que nous entendons par ce concept. Il ne s'agit pas, pour reprendre les termes de P. Charaudeau de la « compétence d'un sujet idéal »,
- il ne s'agira pas non plus de la compétence de communication d'un seul sujet produisant le message, puisque l'acte de langage [ne se réduit pas] à la seule transmission de l'intention du sujet communiquant ; il ne s'agira pas d'avantage de la seule compétence du sujet-interprétant, puisqu'interpréter, c'est aussi faire des hypothèses sur le processus de transmission d'une intention ; il s'agira plutôt de la compétence d'un sujet que nous appellerons sujet analysant qui intègre les activités supposées du sujet produisant et du sujet interprétant le langage. (1983 : 84)
- 13 Nous nous intéressons au sujet, à la fois sujet collectif et sujet singulier, c'est-à-dire celui qui, tout en utilisant la langue pour dire quelque chose à quelqu'un dans une situation donnée par rapport à un contenu partagé, « se dira » en faisant sienne cette langue par rapport à sa subjectivité.
- 14 Nous sommes conscients du problème que cela peut poser au niveau de l'évaluation car, comme le souligne Bachman : « *Even if it were possible to examine our subjects' brains directly, we would see little that would help us determine their level of language ability* » (1995 : 43).
- 15 C'est la raison pour laquelle il faut réfléchir aux critères de performance que nous devons retenir comme reflétant un degré de compétence. Pour reprendre les propos de Bachman à nouveau, cela signifie que :
- We must therefore decide what specific procedures, or operations, we will follow to elicit the kind of performance that will indicate the degree to which the given construct is present in the individual. (*ibid.*)
- 16 Cela va se traduire par la mise en place d'un espace d'évaluation que nous avons appelé précédemment un « espace d'évaluation orienté-sujet » (*ibid.*), représenté sur le schéma 3 en annexe.
- 17 L'évaluation portera dès lors sur la manière dont le sujet-évalué mobilise ses connaissances linguistiques en situation de communication par rapport à un contenu. Ceci se fera dans une relation interactionnelle avec un évaluateur, lui aussi, considéré en tant que sujet, c'est à dire n'évaluant pas selon des critères purement linguistiques préalablement définis, qui ne nécessitent aucune implication de sa part. L'utilisation de la langue se fera par rapport à un contenu lié à la situation de communication dans un espace d'évaluation qui sera un véritable espace de communication construit autour de la négociation de sens par rapport à une situation (interlocuteurs, contenu, contexte) selon différents degrés d'imprévisibilité.
- 18 Nous voyons que nous nous trouvons, dans ce cadre, face à un type d'évaluation qui n'est plus quantitative autour de critères vrai / faux mais qualitative autour de critères liés à l'appropriété, l'efficacité, la pertinence et la cohérence de l'acte de parole de l'évalué.

1.3. La prise en compte de l'acte de parole en situation d'évaluation

- 19 Que signifie prendre en compte l'acte de parole en situation d'évaluation ?
- 20 Tout d'abord, cela signifie définir la « communication » non pas comme une simple transmission d'informations mais bien comme un acte impliquant la relation à l'Autre

dans une dialectique de différenciation et de mise en commun par rapport à une situation donnée, prenant en compte le contexte de communication et le contenu de cette communication.

- 21 Cela revient à reconnaître que « le problème principal de la communication, n'est pas celui du réel mais celui des partenaires. » (Lamizet 1992). Ceci veut dire que l'évaluation doit dépasser la seule prise en compte de l'aspect locutoire du langage, c'est-à-dire le niveau de la signification.
- 22 Dépasser le niveau de la signification signifie s'intéresser au sens négocié par le sujet, c'est-à-dire prendre en compte la dimension illocutoire du langage où dire une parole possède une certaine valeur, et la dimension perlocutoire, où dire signifie obtenir certains effets².
- 23 Bien évidemment la mise en place d'un espace d'évaluation orienté-sujet, en prenant en compte l'acte de parole, dont les caractéristiques seront données dans la quatrième partie de ce travail, a aussi ses limites.

2. Apports et limites de notre *construct*

- 24 En nous intéressant tant aux effets illocutoires qu'aux effets perlocutoires, nous souhaitons avoir une vision complexe de l'emploi de la langue dans un environnement social et professionnel réaliste, ce qu'une approche quantitative ne peut apporter. Shohamy montre combien, malgré leur rigueur, les résultats du TOEFL ne peuvent compenser les méfaits de la pauvreté du *construct* de référence de ce test : « *Post-test analyses of TOEFL scores, no matter how rigorous they are cannot make up for a limited representation of the language construct of the test* » (2000 : 536).
- 25 Néanmoins, élaborer des tests, sans aucun réductionnisme, selon le *construct* réaliste auquel nous sommes parvenus, entraîne des conséquences qu'il conviendra de traiter :
  - La prise en compte du contenu et la focalisation sur le fond causent une baisse apparente de la performance linguistique. Les tests qualitatifs ne deviendront acceptables pour tous qu'après un changement des représentations sur cette performance, en particulier sur l'importance de sa 'correction' linguistique, et sur les attentes qui en découlent.
  - La réaction des utilisateurs face à un test varie en fonction de leur *construct* personnel plus ou moins explicité. Des réactions négatives peuvent donc biaiser les résultats de personnes qui passeraient les tests en rejetant encore le *construct* qui les sous-tend.Malgré cela, il est clair que la pertinence du qualitatif se trouve dans :
  - la cohérence de ces tests par rapport aux visées du candidat,
  - leur efficacité face à ses attentes,
  - l'adéquation (le fait d'être approprié) au contexte.
- 26 De plus, comme un test qualitatif prend l'interaction en compte, il permet de mesurer la réaction du candidat à l'imprévisible, ce qui est très utile. Cohen (Bachman & Cohen 1998) rappelle cependant qu'il convient de faire attention aux entraînements, *preplanned strategies*, que les apprenants intériorisent si le test est trop prévisible dans les situations qu'il propose, et donc paradoxalement dans l'imprévisible qu'il met en place !
- 27 Puisque la mesure est adaptable, on donne un profil qui décrit l'apprenant et non un niveau qui le classe, elle est cohérente avec les notions de variabilité de l'interlangue suivant les contextes d'emploi (cf Tarone in Bachman & Cohen 1998), qu'elle permet de

vérifier. La mesure sera toujours positive, la sanction ne sera pas un échec, ce qui favorise la motivation et élimine les phénomènes de dévalorisation.

3. Contenu et contexte revisités

- 28 Pour comprendre la teneur de cette partie, il est préférable de revenir à cette définition de l'activité langagière qui reviendrait à :

Accomplir des tâches sociales relevant d'un bagage de savoirs (didactique des disciplines) en négociant du sens suivant les conventions d'une communauté donnée (pragmatique) en utilisant la langue (une des langues) de cette communauté (linguistique). (Narcy-Combes 2001 : 7)

- 29 Tenir compte de cette définition nous impose de considérer le contexte plus attentivement, en particulier au vu des résultats des recherches sur les effets de ce contexte et du contenu sur la production des apprenants.

3.1. Prise en compte du contexte

Dans la mesure où les interactants construisent leur espace d'évaluation :

- Il est important de noter comment les candidats voient les indices de contextualisation (*contextualization clues*), car c'est la finesse de cette perception qui leur permet de gérer les situations de manière adaptée.
  - Il importe également de vérifier s'il y a méprise au niveau du domaine de référence abordé (désaccord/*mismatch* entre les interlocuteurs).
- 30 Ces compétences sont très liées à la connaissance que le candidat a du domaine ou des domaines auxquels le contexte renvoie (Douglas 1998 : 151).
- 31 Dans la mesure où notre intérêt se porte sur le sens négocié en commun dans une interaction dont l'espace est construit lors du test, nous sommes amenés à prendre en compte tout autant le locutoire (signifiant), que l'illocutoire (la fonction) et le perlocutoire (l'effet), ce que des tests quantitatifs ne permettent pas.
- 32 Rappelons enfin que, dans des situations professionnelles, la prévisibilité est grande, mais c'est la réaction à l'imprévisible (incident, accident) qui est cruciale. Dans certains contextes professionnels on peut donc comprendre que des attentes élevées se fassent sentir sur ce point.
- 33 Il y aurait deux seuils pour que le candidat puisse réagir de façon appropriée au contexte: un seuil de spécificité contextuelle (authenticité de la situation) et un seuil de développement de l'interlangue. Il importe alors de les connaître et de les mesurer.

3.2. Effet du contenu

- 34 Dans l'état actuel de notre réflexion, nous ne pouvons qu'apporter des éléments qui confirment que le contenu, que véhicule la langue de la spécialité, devrait être pris en compte de façon plus rigoureuse. Comment le faire reste à trouver...
- 35 Il semble que :
- Les succès scolaires (*academic skills*) sont translangues.
  - Les compétences cognitives affectent plus le discours et l'écrit que la compétence grammaticale.

- Plus les exigences au niveau du contenu sont élevées, moins la forme pourra être parfaite.
- Une bonne maîtrise du contenu va de pair avec une maîtrise des structures rhétoriques des réponses.
- Si le développement linguistique du candidat est peu avancé, l'organisation sera quasi conceptuelle et rendra l'interaction difficile.
- Si on ne mesure que le contenu, on valorise les candidats avancés dans le domaine car la connaissance du domaine augmente le contrôle du candidat.

36 Il est donc évident que l'incidence du contenu sur la performance est loin d'être négligeable, et que des épreuves certificatrices ne sauraient être créées sans y réfléchir.

4. L'évaluation orientée-sujet

4.1. Ses caractéristiques

- 37 Pour que le système d'évaluation retenu soit valide, il faut qu'il y ait adéquation avec le *construct* de départ. Ceci permet de définir cinq caractéristiques principales.
- 38 Tout d'abord, l'évaluation devra porter sur trois axes. Le premier axe sera un axe linguistique, ce dernier n'étant pas conçu en termes de connaissances accumulées mais en relation avec la théorie de l'interlangue. C'est à dire qu'il s'agira d'évaluer jusqu'à quel point le sujet-évalué se détache de L1 en situation de communication. Le deuxième axe sera un axe pragmatique en liaison avec l'acte de parole, la situation d'évaluation et le contexte. Dans ce cadre intervient le concept d'appropriété et de pertinence. Le troisième axe portera sur le contenu. Dans ce cadre, l'objectif sera de voir jusqu'à quel point la maîtrise du contenu intervient dans la « réussite » (concept sur lequel nous allons revenir dans le cinquième point) en situation d'évaluation. En d'autres termes, il s'agira d'évaluer le degré d'imprévisibilité auquel le sujet-évalué peut s'adapter.
- 39 Une fois déterminés les axes sur lesquels porte l'évaluation, il convient de définir un support approprié.
- 40 Nous avons, à plusieurs reprises, souligné l'importance de concevoir la communication comme un acte global, ce qu'une atomisation des différentes composantes soumises à évaluation ne permet pas de mesurer. Ainsi, communiquer, c'est à la fois comprendre, produire mais aussi interagir. C'est la raison pour laquelle nous proposons un support d'évaluation composé de tâches distinctes, mais reliées entre elles, en vue d'une résolution de problème. Un support type « scénario » nous paraît donc aller dans ce sens. Il s'agit en fait d'une situation où le sujet-évalué se projette dans le rôle d'un personnage ayant à prendre une décision. Pour ce faire, il doit entreprendre de réaliser un certain nombre de tâches thématiquement reliées entre elles qui s'intègrent dans une chaîne de traitement de l'information allant de son recueil, en passant par son tri, son classement, sa hiérarchisation, tout ceci devant mener à sa restitution et à sa présentation dans des formats imposés par le contexte, le tout conduisant à une prise de décision. Nous voyons, dès lors, que le scénario relève d'une approche holistique de la communication. Les compétences du sujet-évalué, dans ce cadre, dépassent donc largement la compétence linguistique. Elles font appel à un processus inférentiel nécessitant la prise en compte des stimuli ostensifs, dont l'objectif est de diriger l'attention du sujet-évalué qui aura à les repérer en fonction d'une intentionnalité qui se dégagera des tâches qu'il aura à effectuer. La compétence discursive apparaîtra dans les stratégies qu'il utilisera dans sa

prise de décision, quant à la compétence situationnelle, elle sera jugée en fonction du respect du format que le contexte détermine.

- 41 Ainsi, un tel support permettra d'évaluer simultanément les trois axes que nous avons définis précédemment en adéquation avec le *construct* de départ que nous nous étions fixé.
- 42 Un dernier point fondamental vient s'ajouter à la mise en place d'un scénario. Il s'agit de la présence d'un évaluateur interlocuteur dans ce cadre. Son rôle sera double. Ce sera un pourvoyeur supplémentaire d'informations que le sujet-évalué devra recueillir dans une interaction basée essentiellement sur le mode questions / réponses, ceci dans le respect d'une des caractéristiques de la communication qui consiste à transmettre et à recevoir de l'information. La communication n'étant pas réduite à cette composante, il mettra en place avec le sujet-évalué une relation interactive sur la base de la négociation de sens, à charge pour le sujet-évalué d'argumenter pour convaincre son interlocuteur du bien fondé de la décision qu'il a prise.
- 43 Ceci nous amène à la troisième caractéristique de l'évaluation orientée-sujet, à savoir le rôle joué par l'évalué et l'évaluateur.
- 44 La résolution de problème induite par le scénario oblige l'évalué à être actif. Il ne subit pas un test, il s'engage dans une action en respectant la situation présentée dans le dit scénario. Ainsi, il mobilisera un savoir et un savoir-faire qui seront révélateurs d'un savoir-action.
- 45 Dans le même temps, l'évaluateur n'est plus un simple vérificateur de connaissances contraint à confronter les réponses de l'évalué à des attendus préalablement définis. Il est lui aussi actif dans une démarche interactive devant permettre à l'évalué de montrer son savoir-action.
- 46 La quatrième caractéristique de l'évaluation orientée-sujet découle des trois précédentes. L'évaluation ainsi conçue revient à mettre en place un réel espace de communication, seule condition, nous semble-t-il, pour que l'on puisse juger de l'aptitude à communiquer de l'évalué.
- 47 Le dernier problème qui se pose concerne les critères d'évaluation. Il n'est pas besoin de préciser qu'une évaluation quantitative, qu'elle soit chiffrée ou sur le mode binaire est inconcevable dans ce cadre puisque la compétence est intrinsèque au sujet-évalué. Nous devons envisager de mesurer la performance selon des critères révélateurs de la compétence de communication que nous souhaitons évaluer. Ces critères permettront ainsi de prendre en compte l'observable, selon deux axes cette fois. Le premier, l'axe linguistique, permettra de juger de la qualité de la langue utilisée (ceci n'étant pas remis en cause mais jugé insuffisant dans le cadre de l'évaluation de l'aptitude à communiquer). Le deuxième axe, que nous appellerons l'axe pragmatique, portera sur les deux derniers axes présentés au début de cette partie et permettra d'évaluer ce que nous avons appelé précédemment les compétences discursives et situationnelles. Nous ne nous trouvons plus là face à une évaluation quantitative mais qualitative, qui donnera l'obtention non d'une note mais d'un niveau de communication.
- 48 Bien évidemment, le choix de ces critères aura fait l'objet d'une réflexion en amont en liaison avec le *construct* que nous avons défini. Se pose alors la question de l'objectivité de ce type d'évaluation, concept pour le moins important car il semble être un des critères de qualité retenu par la plupart des tests de langue existant sur le marché.

4.2. L'évaluation orientée-sujet est-elle objective ?

- 49 Une autre manière de poser la question, au vu de l'intitulé de ce travail, est la suivante : l'évaluation qualitative est-elle plus subjective que l'évaluation quantitative ? La première réponse à cette question viendra des propos suivants tenus par Bachman :

Language tests are subjective in nearly all aspects. Test developers make subjective decisions in the design of tests and test writers make subjective decisions in producing test items and, with the exception of objectively scored tests, scorers make subjective judgements in scoring tests. Language tests are based on either course syllabi or theories of language proficiency, both of them represent subjective judgements of the individuals who devise them, and tests are therefore neither infallible nor exhaustive definitions of the abilities we may want to test. For these reasons, a test score should be interpreted as a reflection of a given test developer's selection of abilities included in a given syllabus. (1995 : 37)

- 50 Nous ajouterons à ce qui vient d'être dit que si la validité d'un test repose sur l'adéquation entre son objectif de départ et la pertinence des résultats obtenus en fonction des objectifs fixés, sa fiabilité dépendra de l'exactitude de la mesure obtenue. Cependant, pour que la mesure soit exacte, il faut que ce que l'on souhaite mesurer relève de critères d'exactitude, ce qui n'est pas le cas de la communication langagière.
- 51 Ainsi, l'évaluation ne passera pas par le choix d'une réponse juste incompatible avec l'acte de communication mais par ce qu'Austin appelle la « réussite » d'un acte de parole, terme que nous avons utilisé précédemment. Cette réussite est liée à la pertinence des énoncés qui varie selon les interlocuteurs et la situation. Ainsi la réussite sera mesurée en fonction des critères de performance reflétant d'un côté la compétence linguistique et d'un autre la compétence pragmatique.
- 52 De ce fait, si aucun système d'évaluation ne peut être considéré comme objectif, l'important, en revanche, c'est le degré de pertinence lié aux critères choisis en fonction de l'objectif annoncé. Le choix que nous avons fait impose une logique différente de celle qui régit les tests de langue habituellement. En effet, comme le souligne Michel Perrin dans le rapport final du DCL,

À la correction en creux qui obéit à la logique du stylo rouge soustracteur – démarche somme toute négative – on opposera l'appréciation positive résultant de la dynamique du vert : tout ce qui « passe » de manière pertinente, à un niveau donné de complexité, d'approfondissement, de nuance, tant pour la réception des messages que leur production. Bref avec ce qu'ils savent faire sur le plan linguistique, il s'agit de jauger ce que les candidats arrivent à faire sur le plan pragmatique. Et de traduire ce jugement essentiellement qualitatif en notations arithmétiques, compatibles avec les exigences d'objectivité de tout système de certification. (Baïssus 1995 : 29)

- 53 Ceci étant, comme le dit très clairement J.M. Baïssus : « Une révision déchirante s'impose pour les enseignants de langue traditionnels : il faut s'accommoder de l'imperfection pourvu que la tâche (professionnelle) soit effectivement accomplie dans une langue étrangère » (1995 : 15).

Tel est le point que nous aborderons dans la dernière partie de ce travail.

5. Conséquences sur le CLES

- 54 Par rapport à ce qui vient d'être vu, on peut avancer que la liaison avec le savoir disciplinaire facilite la création de tests, car il y a dans ce cas réduction de l'imprévisibilité en situation d'évaluation : ce qui peut être intéressant au niveau d'une évaluation institutionnelle où il est parfois difficile de créer les épreuves complexes dans leur organisation qui seules permettent de mesurer les réactions à l'imprévisible. La composante avec contenu relèverait donc préférablement de l'institution du candidat (Douglas 1998 : 150-51). On peut également imaginer un réseau d'institutions similaires qui en travaillant ensemble verraient le travail s'alléger et la créativité des concepteurs de tests augmenter.
- 55 Il est par contre difficile de concevoir que des épreuves complexes, mettant en place des situations où les candidats gèrent de l'imprévisible, puissent être gérées localement de façon fiable et rigoureuse. Un organisme universitaire semblable à UCLES (*University of Cambridge Local Examinations Syndicate*) se révélerait fort utile au-delà des bouleversements culturels qu'impose sa mise en place.

6. Besoin en formation des examinateurs

- 56 Les examinateurs, dans le contexte français, ne sont pas préparés à fonctionner dans des systèmes certificateurs tels que le suggérerait une approche qualitative. Il leur faudrait réfléchir en termes de :
- Responsabilité épistémologique telle que Kelly la définissait en 1955, comme le confirment ces lignes de Shohamy :
Test-users also have a responsibility. Test-users need to be cognizant of the properties of the instruments they employ and ensure appropriate interpretation and use of the instruments they employ. (2000 : 537)
- 57 Ceci implique la connaissance des divers *constructs* qui définissent l'évaluation en langues et une réflexion sur leur pertinence en ce qui concerne le contexte et les besoins des candidats.
- Compétence dans le contenu. Si celui-ci est d'un niveau élevé, sa connaissance est importante pour l'évaluateur dont la sévérité linguistique peut être augmentée à cause de son manque de compréhension.
Raters or judges of context-based tests may need to have some sense of the content accuracy of responses, and may rate context-based test responses more conservatively as a function of not knowing how accurate the responses are. (Douglas 1998 : 153)
  - Analyse rhétorique interlinguale pour éclaircir certaines évaluations où les attentes des examinateurs peuvent être injustifiées au vu des résultats des études sur l'interlangue :
Rhetorical/grammatical interlanguage analysis may be necessary to disambiguate gross subjective ratings on context-based tests. (Douglas 1998 : 154)

Conclusion

- 58 Au-delà de tout positionnement sur l'évolution du CLES actuel, nous sommes dans une situation où seul un test correspond à presque tous les critères que nous avons définis, le

DCL. Il pourrait paraître opportun qu'un institut national en assure la mise en place, en particulier au niveau de la création des scénarios, comme on pourrait espérer qu'un tel institut puisse piloter le CLES, dont les orientations sont moins professionnelles.

- 59 Le travail est considérable pour mettre en place une évaluation certificatrice rigoureuse et efficace dans ce pays. La création d'un groupe de recherche émanant du GERAS avec pour objectif de contribuer à mettre en place un tel institut de recherche et de formation nous semble un objectif à moyen terme tout à fait cohérent qui assurerait la pérennité de la réflexion au-delà des inévitables changements gouvernementaux. Le rôle d'un tel institut ne serait pas d'élaborer des lois, mais de créer, et d'expérimenter des tests à la demande des organismes intéressés. Le déterminer pourrait être un des objectifs de notre communauté.
- 60 Austin, J.L. 1970. *Quand dire, c'est faire*. Paris : Éditions du Seuil.
- 61 Bachman, L.F. 1995. *Fundamental Considerations in Language Testing*. Oxford : OUP.
- 62 Baïssus, J.-M. 1995. *Rapport final du DCL*. Ministère de l'Éducation.
- 63 Bourguignon, C. 2001. « La communication langagière : évaluation, médiation, remédiation », Colloque du GERAS, Université de Bordeaux 2.
- 64 Bygate, M., P. Skehan and M. Swain (dir.). 2001. *Researching Pedagogic Tasks, Second Language Learning, Teaching and Testing*. Londres : Longman.
- 65 Chareaudeau, P. 1983. *Langage et discours, Éléments de sémiolinguistique, Théorie et pratique*. Paris : Hachette.
- 66 Dortier, J.-F. 1999. « Pensée et langage : les limites de la traduction automatique ». In *Le Cerveau et la Pensée*. Auxerre : Sciences Humaines Éditions.
- 67 Douglas, Dan. 1998. « Testing methods in context-based second language research ». In Bachman L.F. et A. Cohen (dir.), *Interfaces Between Second Language Acquisition and Language Testing Research*, Cambridge : Cambridge University Press, 141-155.
- 68 Duverger, J. (dir.). 2000. *Le français dans le monde*, numéro spécial, janvier : Actualités de l'enseignement bilingue. Paris : Hachette.
- 69 Ellis, R. 1997. *SLA Research and Language Teaching*. Oxford : Oxford University Press.
- 70 Halliday, A. 1999. « Evaluating the discourse: the role of applied linguistics in the management of evaluation and innovation ». In Rea-Dickens P. et K.P. Germaine (dir.) *Managing Evaluation and Innovation in Language teaching*. Londres : Longman.
- 71 Kelly, G. 1955. *The Psychology of Personal Constructs*. New York : Norton.
- 72 Lamizet, B. 1992. *Les lieux de la communication*. Liège : Pierre Mardaga Éditeur.
- 73 Mager, R.F. 1996. *Comment mesurer les résultats de l'enseignement*. Paris : Bordas.
- 74 Meyer, F. (dir.). 2001. *Language for Special Purposes: Perspectives for the New Millenium*. Tübingen: GNV.
- 75 Narcy-Combes, J.-P. 2001. « Sens, forme, ne manque-t-il pas un troisième larron ? ». *ASp* 34, 3-13.
- 76 Rey, B. 2000. « Un apprentissage du sens est-il possible ? ». In J.-M. Barbier et O. Galatanu (dir.), *Signification, sens, formation*. Paris : PUF, 107-126.
- 77 Shohamy, E. 2000. « The relationship between language testing and second language acquisition revisited ». *System* 28, 541-553.

- 78 UPLEGESS (ed). 1988. *L'évaluation des compétences en fin de cursus*. Compiègne : Université de Technologie.

ANNEXES

Fig. 1 Espace d'évaluation

Fig. 2 Espace d'évaluation orienté-objet

Fig. 3 Espace d'évaluation orienté-sujet

NOTES

1. « La communication langagière : évaluation, médiation, re-médiation », Colloque du GERAS, Bordeaux, mars 2001.
2. Nous avons repris les définitions d'aspect locutoire, illocutoire et perlocutoire données par Austin (1970).

RÉSUMÉS

Pour comprendre les points forts d'une approche qualitative de l'évaluation en langue étrangère, il convient de bien comprendre ses fondements théoriques et les apports récents des recherches en acquisition des langues qui la justifient. L'objectif de cet article est de permettre aux enseignants du secteur LANSAD de faire le point et de mieux cerner les implications que les projets tels que celui du CLES peuvent avoir pour eux.

In order to understand the strong points of qualitative foreign language evaluation it seems preferable to have a clear view of its theoretical background and of the recent results of SLA research which justify it. The aim of this article is to enable the teachers of foreign languages to students of other disciplines to review the problem and to figure out what projects such as CLES may imply for the profession.

INDEX

Mots-clés : approche qualitative, cadre théorique de référence, contenu, contexte, évaluation, objectivité, orientation sujet, tâche, validité

Keywords : construct, content, context, evaluatee-centredness, objectivity, qualitative approach, task-based approach, validity

AUTEURS

CLAIRE BOURGUIGNON

Claire Bourguignon est maître de conférences HDR à l'IUFM de l'Académie de Rouen. Son domaine de recherche est la didactique des langues, particulièrement la relation enseignement, formation, évaluation, appliquée à l'anglais de spécialité et l'anglais général.

claire.bourguignon@rouen.iufm.fr

JEAN-PAUL NARCY-COMBES

Jean-Paul Narcy-Combes est professeur à l'Université de La Rochelle et appartient au CERC (Equipe d'accueil de l'Université de Nantes). Il participe également aux travaux en anglais de spécialité (tâches, évaluation). Ses recherches portent sur les stratégies d'apprentissage et la théorisation de la pratique, l'épistémologie et la méthodologie de la recherche, en lien avec les TIC. jean-paul.narcy-combes@wanadoo.fr