

LE PROCESSUS D'ÉLABORATION DE TESTS

Rodica STANCIU-CAPOTĂ

«Le but général de toute forme d'évaluation en langue est d'avoir un échantillon des compétences langagières des candidats qui permette d'obtenir une représentation réaliste de leur niveau de capacité d'utiliser la langue en situation hors test.» [1 : 2] affirme M. Milanovic dans «Evaluation des compétences en langues et conception des tests», ouvrage qui vient à l'appui des enseignants/évaluateurs dont le dessein est d'appliquer le mieux possible les principes du CECRL, cadre qui se défend de préconiser une méthodologie, mais énonce quelques grands principes:

- place l'apprenant au centre du processus d'apprentissage/ enseignement ;
- donne des outils pour apprendre (savoir apprendre), dans le but d'autonomisation, d'individualisation du parcours, et d'auto évaluation ;
- affirme la perspective actionnelle ;
- conçoit l'apprenant comme acteur social.

En distinguant les différentes activités langagières et en proposant un choix de critères pertinents pour en évaluer les divers niveaux, le Cadre européen commun de référence pour les langues propose un découpage rationnel qui devrait aider les enseignants à mieux structurer l'apprentissage, à calibrer documents, tests et tâches d'évaluation, et donc à mieux répondre aux objectifs des programmes. En déclinant les composantes de la compétence communicative langagière (compétence linguistique, sociolinguistique et pragmatique). Le CERCLE permet aux enseignants d'avoir une idée claire de ce qu'ils sont censés enseigner/évaluer.

L'évaluation de la capacité à utiliser la langue en situation est la plus grande nouveauté que le cadre apporte. Car, ce qui prédominait jusqu'ici en classe de langue, c'était l'évaluation du savoir, centrée sur les contenus, et portant sur ce qui a été enseigné. L'évaluation de la performance, c'est-à-dire ce qu'un sujet peut effectivement faire, pratiqué par des tâches communicatives et langagières, permet dorénavant à chacun de se situer. Cette performance se décompose en compétence linguistique (grammaticale, lexicale, phonologique), sociolinguistique et pragmatique (interactionnelle et fonctionnelle).

Le CECRL propose un inventaire très précis des composantes de chacune des compétences (linguistique, sociolinguistique, pragmatique) que l'apprenant doit maîtriser à chaque niveau, des situations de communication, des tâches et des objectifs de communication. En tant qu'outil de référence, le CECRL a été conçu pour être utilisable dans des contextes divers : formation continue ou initiale, apprenant adulte ou enfant, milieu professionnel ou scolaire, etc. L'approche actionnelle qu'il préconise ne fait que reprendre les principes de l'approche communicative avec une insistance particulière sur la réalisation de tâches : le niveau de compétence d'un apprenant y est défini en fonction du nombre de tâches qu'il peut réaliser. La maîtrise d'une langue passe par une combinaison de savoirs et de savoir-faire plaçant les apprenants à la fois en position d'« apprendre » et d'« apprendre à », moyennant des mises en situation et un accompagnement adéquat.

Le CECRL met en lumière la complexité de l'acte de communication et donne, tant pour l'enseignement que pour l'évaluation, des entrées qui ne sont pas purement linguistiques mais aussi socioculturelles et pragmatiques. Si son but est de présenter une échelle de référence commune, ce but n'est pas de prôner la mise en place d'un test unique ou d'une seule certification. Mais le fait que les systèmes d'évaluation s'appuient sur le cadre pour l'élaboration des épreuves et des niveaux permet d'apporter une reconnaissance européenne sinon internationale des acquis des apprenants.

Dans ce contexte, l'élaboration des tests d'évaluation devient une des priorités de l'enseignant/évaluateur à l'aide duquel le CECRL, présente, dans le chapitre 4, un modèle de compétences langagières, important « parce qu'il fournit une base utile à la définition du camp de compétences à évaluer. » [1 : 2] et définit en détail les caractéristiques du contenu d'un test.

Dans notre démarche d'adaptation des principes du CECRL à l'enseignement du français économique, l'élaboration des tests d'évaluation

– que ce soit une évaluation diagnostique, continue ou sommative – s'avère une des tâches des plus astucieuses de l'enseignant qui devient de la sorte évaluateur non seulement des performances de ses étudiants, mais aussi de ses propres performances professionnelles.

Le processus d'élaboration des tests est un processus cyclique et réitéré, puisqu'il renvoie toujours à ce qu'on a enseigné « pour remédier les problèmes d'acquisition par une réévaluation permanente » [1 : 17]. Nous y ajouterons qu'il s'agit aussi d'une réévaluation des méthodes et des stratégies didactiques utilisées, des étapes parcourues et aussi des tests antérieurs.

Pour élaborer un test, quels que soient le moment de l'évaluation ou le type de compétence évaluée, le concepteur de tests doit parcourir selon M. Milanovic six étapes :

Percevoir le besoin d'un test

Dans cette première étape du processus d'élaboration de tests, on remarque le fait que deux sont les situations dans lesquelles le test est obligatoire au moins du point de vue du moment de l'évaluation:

- **L'évaluation initiale** (ou diagnostique) qui est un préalable indispensable à l'enseignant qui doit *placer* l'apprenant dans un groupe de niveau qui corresponde à ses acquis antérieurs et *établir* à partir des acquis de l'apprenant une progression ;
- **L'évaluation sommative** (ou finale), qui *rend compte* d'un parcours, d'une démarche didactique et qui *place* l'apprenant sur une échelle de valeurs au niveau de son groupe ou au niveau d'un cadre objectif, tel le CECRL

L'évaluation formative (de parcours) est un type d'évaluation dont chaque enseignant / évaluateur établit le moment et l'objectif (des raisons didactiques : évaluer une étape, une compétence, vérifier des acquis, régler le rythme d'enseignement, etc.)

Planifier le test

Pendant cette étape de l'élaboration des tests, on « recueille des données sur les demandes précises des candidats » [1 : 5], on *collecte* des informations et on *planifie* en fait le test, du point de vue du besoin déjà établi et du moment où on passera le test, des contraintes didactiques et institutionnelles.

Concevoir le test

C'est le moment où on *définit* le type de test le plus approprié au niveau des apprenants, à leurs besoins et au but de l'enseignant, mais aussi le moment où on *établit* des spécifications (forme du test, contenu, longueur des items, sujets choisis, etc.). C'est aussi le moment où on cherche l'adéquation du type de test au type de compétence qu'on veut évaluer. On établit de même les « variables et contraintes » [1 : 5] qui affectent la forme et le contenu du test.

Elaborer le test

Cette quatrième étape suppose une *expérimentation* apte à donner une image précise de l'appropriation du test à l'objectif proposé, son adéquation et sa validité. C'est une phase qui permet l'élaboration des barèmes ou des échelles de notation, mais aussi une étape durant laquelle on peut faire des modifications à tous les éléments du test.

Passer le test

Cette étape d'*opérationnalisation* est une phase dans laquelle le test est définitif, de même que la notation, et il est administré aux apprenants.

Contrôler les résultats du test

C'est à ce moment qu'on *corrige* et *vérifie* les résultats de la passation du test, *analysant* à la fois les résultats des apprenants évalués, mais aussi les caractéristiques du test.

A ce que l'on peut voir, élaborer des tests, surtout dans le domaine qui nous intéresse le plus, le français économique, n'est pas chose facile. D'autant plus qu'au moment de l'élaboration des tests d'évaluation il faut faire attention à ce que le test soit *valide*, *fidèle*, *faisable* et ait de l'*impact* sur les acteurs impliqués dans l'acte d'évaluation des compétences.

Dans cette démarche assez laborieuse d'élaboration de tests d'évaluation, M. Milanovic souligne que l'enseignant/évaluateur devra, après avoir établi le type de compétence qu'il veut évaluer (se) répondre à des questions du genre :

- ❑ Quelles sont les variables professionnelles à prendre en considération lors de la planification du test (les situations de communication réelles/ professionnelles dans lesquelles on utilisera la langue, le contexte socioprofessionnel, le niveau souhaité pour ces situations, etc.) ?
- ❑ Quelles sont les variables matérielles (le nombre de personnes évaluées, la durée du test, les moyens techniques à utiliser, le nombre de salles, etc.) ?
- ❑ Quelles sont les caractéristiques du contenu du test (le but des tâches, type de tâches, types d'exercices pour chaque tâche, types de textes/supports, etc.) ?
- ❑ Quelles sont les caractéristiques techniques (durée du test, nombre de parties, nombre d'items par tâche, note par item, note finale, critères de correction pour les types de production, etc.) ?
- ❑ Quelles sont les procédures (date et lieu de l'examen, nombre d'heures d'enseignement/ apprentissage pour la préparation du test, etc.) ?
- ❑ Quel est le but du test ?
- ❑ De quelle manière on fera l'analyse du test et à quoi serviront les résultats ?

Des réponses à ces questions dépend à notre avis la réussite d'une évaluation par test, d'une évaluation des compétences de compréhension, production et interaction en langues.

La production de tests

Dans le processus d'élaboration des tests, l'étape de la production est une des plus importantes, supposant pour leur concepteur une prise en charge de bon nombre de variables qui mènent à la réalisation d'un test « au niveau et au contenu exigé et couvrant ce que l'on veut évaluer » [1 : 20]. M. Milanovic met en évidence dans son guide l'obligation du concepteur d'assurer un équilibre entre le niveau de *difficulté* du test, le *contenu*, la *progression* des tâches et des types d'items qui leur correspondent [1 : 20].

Le choix des tâches, qui constituent le squelette du test, doit être très bien justifié par l'objectif proposé : l'évaluation d'une certaine compétence. Car, comme on le verra par la suite, les différents types de compétences demandent différents types de tâches spécifiques, des tâches dont les items doivent mettre en évidence les compétences de l'apprenant.

De manière générale, les tâches (décomposées dans le test en items) sont formées de plusieurs éléments :

- a) **La consigne** (« instructions données pour réagir à un support donné » [3 : 27]. Elle doit être claire, cohérente, comprendre toutes les informations et les contraintes nécessaires. Si l'apprenant évalué

est mis devant une consigne nouvelle, elle sera accompagnée par un exemple. La langue utilisée dans la consigne sera simple et inférieure au niveau de la langue évaluée, le vocabulaire appartiendra aux ressources de l'apprenant. La consigne, selon Christine Tagliante, doit aussi : « être univoque, formulée de façon à ce que la performance soit observable, comporter des indications de durée, indiquer éventuellement le seuil d'acceptabilité » [4 : 92].

- b) **Le support textuel** (textes, dialogues dont le rôle est de servir de déclencheur d'une réaction). Pour l'enseignant d'une langue de spécialité, le français économique dans notre cas, le choix du support est extrêmement important, vu que ce support doit contribuer par son contenu même au développement d'une compétence socioculturelle, d'un savoir être et non seulement d'un savoir faire. En fonction du type de compétence qu'on veut évaluer, le CECRL propose dans le chapitre 4 des genres et des types de textes/supports desquels on pourrait choisir quelques uns pour l'évaluation de certaines compétences [3 : 76] :

I. supports oraux

- les annonces publiques et les instructions
- les discours, les exposés
- les commentaires sportifs
- les informations radio et télévisées
- les débats publics
- les conversations téléphoniques
- les entretiens d'embauche
- etc.

II. supports écrits

- les articles de journaux ou de magazines
- les modes d'emploi, les étiquettes
- les dépliants
- les formulaires et les questionnaires
- les lettres d'affaires et professionnelles
- les e-mails, les messages
- les notes de services, les compte-rendus et les rapports
- etc.

- c) **La réponse de l'évalué** (par choix ou par production, en fonction du type d'item).

d) **La grille de correction** (qui facilite l'évaluation et définit les critères à évaluer en fonction de la compétence évaluée). Elle doit donner suffisamment d'informations sur les réponses acceptées, sur les éléments qui seront ou non pris en considération, en fonction du type de tâche, réduire le plus possible le degré de subjectivité du correcteur surtout s'il s'agit d'une production, soit-elle orale ou écrite.

e) **Le barème de notation** (qui attribue le nombre de points par tâche/par item et limite la subjectivité). Dans l'élaboration du barème de notation on doit tenir compte du fait que les méthodes de notation diffèrent en fonction de la compétence évaluée ou de sa complexité. A propos de cela, Christine Tagliante disait dans l'ouvrage cité précédemment : « Evaluer une compétence complexe, c'est prendre en compte non seulement les divers savoirs et savoir-faire dont elle est constituée, mais également tout ce qui relève de la sensibilité, de l'imagination, de l'opinion personnelle et de l'affectivité de l'individu. En un mot, de sa personnalité. Dès que l'on quitte l'évaluation des connaissances élémentaires, la mesure se transforme en appréciation, les barèmes étant là pour canaliser la subjectivité de l'évaluateur et garantir à l'évalué la validité de son évaluation » [4 : 87].

Les types de tests d'évaluation

Comme les objectifs et la méthodologie d'enseignement du français économique visent le développement des compétences communicationnelles voire actionnelles et non plus la manipulation de structures grammaticales ou lexicales telles quelles, « les situations d'évaluation doivent mesurer des habilités langagières en situations de communication aussi authentiques que possibles » [5 : 64]. De là, une conclusion : le type de test à choisir lors d'une évaluation en FOS n'est pas chose facile. Le choix dépend d'une multitude de facteurs qui déterminent ou orientent le concepteur de tests dans sa démarche qui consiste en premier lieu à choisir le type de test qui lui permettra d'évaluer le mieux un ou plusieurs types de compétences.

Partant de la taxonomie de Benjamin Bloom (mémorisation, compréhension, application, analyse, synthèse, esprit critique), Christine Tagliante présente dans son ouvrage « L'évaluation et le Cadre Européen Commun » les types « classiques » de tests, soulignant le fait que ceux-ci ne diffèrent pas des tests à appliquer

en FLE, respectivement en FOS, en évaluation initiale, formative ou sommative. Ce qui fait la différence c'est la notion de *tâche* à évaluer et non plus d'*acquis* classique. On évalue ainsi ce que l'apprenant *sait faire* et non seulement ce que l'apprenant *sait*.

Selon C. Tagliante, toute activité d'élaboration de tests devrait commencer par la rédaction d'une fiche descriptive [4 : 90] qui pourrait rendre l'activité d'évaluation plus cohérente et plus structurée et qui permettrait la réalisation de tests d'évaluation appropriés à l'objectif d'évaluation proposé.

Les dix rubriques de cette fiche seraient :

- a. Le descripteur évalué
- b. Les actes de langage ou les actes de parole
- c. L'objectif à maîtriser
- d. Les savoirs requis
- e. Le type de tâche
- f. Le public
- g. Le niveau taxonomique
- h. Le support
- i. Le domaine
- j. Le sous-domaine

Ce type de fiche serait un outil très nécessaire dans la phase de la conception des tests et elle faciliterait le travail du concepteur de tests qui doit choisir entre plusieurs types en fonction de leurs caractéristiques.

Des outils d'évaluation « ouverts » aux outils « fermés », du QCM à la production orale ou écrite ou à l'expression orale en continu, un large éventail de types d'activités d'évaluation [20] est à la disposition du concepteur de tests en français sur objectifs spécifiques, un éventail qui rend compte aussi de la complexité de sa tâche de faire correspondre un type de test, d'activité, d'exercice, d'item à chaque moment de l'évaluation, à chaque type de compétence évaluée. Pour ce qui est de l'évaluation en français économique, nous avons constaté une multitude d'outils d'évaluation dont il pourrait se servir dans sa démarche et que nous avons essayé par la suite de classer en trois grands groupes.

Les types d'outils d'évaluation en fonction de la réponse attendue

Les réponses à donner aux différents types d'items peuvent être fermées (la réponse est sous la forme d'une croix dans une case, par exemple, et permet une évaluation tout à fait objective de

l'acquisition des connaissances), ouvertes (on applique un barème critérié, l'appréciation entrant en jeu à la place de la comptabilisation des réponses justes ou erronées) ou semi-ouvertes (qui combinent les deux types antérieurs). A partir de cela on peut distinguer des outils d'évaluation :

- à réponses fermées : les QCM, l'appariement, la discrimination auditive, le test de classement, l'exercice à trous, etc. ;
- à réponse ouvertes : la production écrite guidée, le résumé, l'analyse, la synthèse, le commentaire, le jeu de rôles, a production orale en continu, etc. ;
- à réponses semi-ouvertes : le test de closure, l'exercice de transformation, le QROC (question-naire à réponses ouvertes courtes), etc.

Les types d'outils d'évaluation en fonction de la compétence évaluée

La compétence de communication se réalise à travers la réception, la production et l'interaction. De là, sa déclinaison en compréhension orale et écrite, expression orale et écrite et interaction que l'enseignant doit évaluer. On a recensé pour l'évaluation de :

- *la compréhension orale*: les exercices d'écoute (dont les avantages consistent dans le fait qu'ils « sont mesurables (il n'y a pas d'ambiguïté dans la réponse) et adaptables (ils peuvent être utilisés à tous les niveaux et concernent la compréhension globale aussi bien que détaillée) » [6 : 31], le QCM, l'appariement, le QRO, le texte à trous, le test de closure, les exercices de transformation, etc.
- *la compréhension écrite* : le QCM, l'appariement, le classement, l'exercice à trous, le QROC, le commentaire, la synthèse, l'analyse, etc.
- *l'expression orale* : le commentaire, le jeu de rôles, l'exposé, la description, l'argumentation, etc.
- *l'expression écrite* : les exercices à trous, les exercices de transformation, la production guidée, le résumé, l'analyse, la synthèse, etc.
- *l'expression orale en continu* : le résumé, le commentaire, la synthèse, la dissertation, etc.

A ce que l'on peut observer, certains types d'exercices peuvent être utilisés dans l'évaluation de plusieurs compétences, leur élaboration devant tenir compte du type de compétence évaluée.

Les types d'outils d'évaluation en fonction de l'étape du processus d'enseignement/apprentissage

Comme « le moment de l'évaluation est un moment formateur non seulement pour l'apprenant, mais aussi pour l'enseignant qui – grâce à ces manières de faire – apprend à se taire pour – tout en restant la référence absolue de la classe – donner à celui qui apprend la place la plus active possible » [6 : 65], l'établissement du moment de l'évaluation dépend de l'objectif que l'enseignant/évaluateur s'est proposé.

Ainsi, pour *l'évaluation initiale* (dont le but est de faire le point sur les connaissances acquises précédemment, de faire des pronostiques sur les progrès et les orientations probables, de placer l'apprenant dans un groupe plus ou moins homogène du point de vue des compétences) on fait passer aux apprenants des tests d'entrée, de placement ou des relevés de performances, à l'aide d'une diversité d'exercices qui permettent de vérifier chacune des compétences langagières. On a constaté qu'il n'y a pas de types d'outils d'évaluation privilégiés. Pour *l'évaluation en cours de formation* (dont le but est de diriger l'apprenant, de mettre en évidence ses difficultés et de guider l'enseignant par un retour sur ses objectifs) on emploie d'habitude des exercices autocorrectifs, des grilles d'auto évaluation, des exercices spécifiques pour la compétence évaluée. Ces types de tests qui accompagnent l'enseignement/apprentissage sont plus ciblés sur des objectifs concrets, comprenant parfois des exercices alternatifs sur une même compétence (grammaticale, lexicale, phonique, etc.). Quant à *l'évaluation finale* ou *sommative*, qui constitue le moment de bilan d'un projet, d'une séquence didactique, se finalisant par une note/un qualificatif plus ou moins officiel, elle privilégie tous les types d'outils d'évaluation, choisis en fonction de la compétence évaluée et du niveau de langue.

En conclusion, les types de tests utilisés pour ces trois types d'évaluation ne diffèrent pas essentiellement en fonction du moment de l'évaluation, la différence étant seulement dans l'objectif déclaré de chaque type d'évaluation (classer/ évaluer/noter)

RÉFÉRENCES ET NOTES POUR L'ÉVALUATION ET L'ÉLABORATION DE TESTS

1. Allal L. – *Vers une pratique de l'évaluation formative. Matériel de formation continue des enseignants*, De Boeck Université, 1999
2. Allal L., Bain D., Perrenoud Ph. – *Évaluation formative et didactique du français*
3. Castellotti V. Py B. (coord.) – *La notion de compétence ENS Editions*, Lyon, 2002
4. Collectif, *Enseigner le FLE. Pratiques de classe*, Éd. Belin, Paris, 2005
5. Cuq J.-P., Gruca I. – *Cours de didactique du français langue étrangère et seconde PU de Grenoble*, 2005
6. Dialogos no. 4 – *Évaluation et didactique des langues*, Editura ASE, Bucuresti, 2001
7. Hadji Ch. – *L'évaluation démystifiée*, ESE, Paris, 1999
8. Laurier M. D. – *Les principes de la mesure et de l'évaluation des apprentissages*, Gaëtan Morin Editeur, 2005
9. Lussier Denise – *Évaluer les apprentissages dans une approche communicative*, éd. Hachette, Paris, 1992
10. Milanovic M. sous la direction de, *Évaluation de compétences en langues et conception de tests*, Division des Politiques Linguistiques, Strasbourg, 2002
11. Monnerie-Goarin A., Lescure R. (coord.) – *Evaluation et certification en langue étrangère*, no. spécial du FDLM, 1993
12. Morissette D., *Guide pratique de l'évaluation sommative. Gestion des épreuves et des examens*, de Boeck, 2002
13. Noël- Jothy Fr., Sampsonis B. – *Certification et outils d'évaluation en FLE*, Hachette, Paris, 2006
14. Oprea C.-L. – *Strategii didactice interactive*, EDP RA, Bucuresti, 2007
15. Petitjean B. (coord.) – *L'évaluation*, no 44, Pratiques Hachette, Paris
16. Porcher L. – *L'enseignement des langues étrangères*, Hachette, Paris, 2004
17. Tagliante, Christine – *L'évaluation et le Cadre Européen Commun*, Éd. CLÉ International, Paris, 2005
18. *Un Cadre Européen Commun de Référence pour les Langues: apprendre, enseigner, évaluer*, Division des langues vivantes, Strasbourg, 2002
19. Veltcheff, Caroline, Hilton, Stanley – *L'évaluation en FLE*, Éd. Hachette, Paris, 2003
20. QCM, QCM à trois ou plusieurs choix, repérage, texte lacunaire, appariement, test de closure (reconstruction), QROC, classement, production guidée avec consignes, analyse de texte ou d'image, résumé, synthèse, tableau de production à double entrée, etc.