

Utilisation du CECRL pour développer des tests dans six langues pour les professeurs des écoles : avantages et limites

Vincent Folny, CIEP, France, folny@ciep.fr
Darina De Jaegher, Štátny pedagogický ústav Bratislava, Slovaquie, jaegher@statpedu.sk
Anna Butašová, Univerzita Komenského v Bratislave Pedagogická fakulta, Slovaquie
butasova@fedu.uniba.sk
Juraj Butaš, Katolícka univerzita Ružomberok Pedagogická fakulta, Slovaquie
jurbutas@gmail.com



Synergies Europe n° 6 - 2011 pp. 51-61

Dates de soumission/acceptation : 29 mars - 9 septembre 2011

Résumé : Depuis 2009, le Štátny pedagogický ústav collabore avec le CIEP (Centre international d'études pédagogiques) afin de développer des examens en six langues à destination des professeurs des écoles primaires. Cette collaboration entre deux institutions européennes est l'occasion d'un partage de savoirs et de savoir-faire mais aussi de mieux comprendre l'utilité des outils développés par le Conseil de l'Europe. Les décisions, les orientations prises au cours du projet sont une bonne illustration de la manière dont on utilise le CECRL en Europe. Cela révèle les besoins des développeurs de tests qui ne sont pas satisfaits aujourd'hui.

Mots-clés : CECRL, examen, construit, spécifications, CIEP

Abstract: Since 2009, the Štátny pedagogický ústav collaborates with the CIEP (Centre international d'études pédagogiques) in order to develop tests in six languages for primary school teachers. This collaboration between two European institutions is an opportunity for sharing knowledge and know-how but also to better understand the utility of tools developed by the Council of Europe. Decisions, directions taken during the project are a good illustration of how the CEFR is used in Europe. This reveals test developer needs that are currently wanting.

Keywords: CEFR, test, construct, specifications, CIEP

1. Le contexte d'émergence et la description du projet

1.1. Le contexte d'émergence

En 2009, le Štátny pedagogický ústav (Centre national pour l'éducation) à Bratislava (Slovaquie) a contacté le CIEP (Centre international d'études pédagogiques) afin d'envisager une collaboration pour développer des examens en six langues (anglais, allemand, espagnol, français, italien et russe). Il s'agissait de développer des examens pour évaluer le niveau de compétence linguistique d'élèves-enseignants destinés, à l'issue d'une formation de plusieurs années, à enseigner une langue étrangère.

Les examens ont pour objectif principal d'accompagner, à intervalles réguliers, l'évaluation des candidats au professorat (de leur entrée à leur sortie du dispositif de formation). Ils visent à évaluer la maîtrise d'une langue, indépendamment du type d'enseignement reçu. Ces examens ne permettent pas de vérifier la capacité à enseigner une langue mais bien à évaluer un niveau de maîtrise qui rend possible son enseignement. Les examens sont conçus pour six langues. Pour toutes les langues, quatre niveaux de compétence sont évalués, soit les niveaux A1, A2, B1 et B2 (Conseil de l'Europe, 2001). Chaque niveau du CECRL correspond à une année de formation, la formation pouvant durer 4 ans et moins si le futur enseignant a déjà un niveau supérieur au niveau A1 lorsqu'il commence sa formation. En fin de parcours, tous les futurs enseignants doivent avoir le niveau B2, quelle que soit la langue évaluée.

Le projet s'inscrit dans la tradition éducative slovaque dans laquelle l'apprentissage de six langues est proposé à l'ensemble des institutions scolaires (State Institute of Education, 2009). Dès le début du projet, l'accent a été mis sur la comparabilité des examens dans les six langues. Il s'agissait de concevoir des examens suffisamment comparables quant au format des épreuves, au mode d'administration, aux procédures et modalités de correction mais aussi quant aux contenus et niveaux visés dans chacune des langues. Dans ce contexte, l'utilisation du CECRL, et notamment des échelles développées par Brian North (2000), ont été d'une forte utilité. Bien évidemment, dès le début du projet, l'ensemble des acteurs était conscient que la comparabilité des examens ne pouvait être entièrement couverte par la seule utilisation du CECRL. Cette utilisation devait être accompagnée par la mise en place de procédures visant à encadrer le développement des examens.

1.2. La description du projet

La Cellule Qualité et Expertises (CQE) du Département Evaluation et Certifications (DEC) au CIEP a été identifiée pour mettre en œuvre la demande de collaboration faite par le Štátny pedagogický ústav. La CQE a souhaité doter le centre slovaque d'outils lui permettant d'améliorer la qualité de ses examens. En aucun cas, il ne s'agissait pour la CQE de développer elle-même les examens mais bien d'accompagner le Štátny pedagogický ústav dans la conception intellectuelle du projet, la mise en place de savoir-faire issus de la littérature scientifique ou de l'expertise professionnelle. Pour les deux institutions, ce projet s'inscrivait dans une perspective de renforcement de la diversité linguistique au sein d'institutions éducatives au cœur de l'espace européen.

La CQE a, dans un premier temps, opéré un audit prospectif. Il était important d'évaluer les procédures devant être mises en place et de réaliser un inventaire des compétences disponibles. Lors de cet audit, des informations ont été prises sur :

- la population cible, sa définition, ses caractéristiques ;
- le but et le contexte d'utilisation des examens ;
- les moyens humains mis à disposition pour développer les épreuves ;
- les actions envisagées pour la mise en relation avec le CECRL ;
- le ou les cadres conceptuels envisagés pour développer les examens ;
- les compétences et contenus à évaluer ;
- les procédures à mettre en place pour assurer la qualité de la mesure ;
- le type d'analyses psychométriques faites sur les items, les tâches.

Bien que le projet se situe dans un contexte scolaire, les procédures d'audit appliquées ont été similaires à celles décrites dans les *Procedures for auditing* développées par les membres de l'association ALTE (Bolton, 2010).

À l'issue de l'audit prospectif, il a été décidé de travailler sur le construit des examens afin, non seulement, de définir les compétences évaluées, mais aussi, de s'assurer que l'ensemble des acteurs du dispositif avait bien la même interprétation de la nature du projet. Nous souhaitions vérifier que le projet était compris à des niveaux de hiérarchie différents et entre les acteurs des différentes langues. Lorsque des définitions ont été données aux construits et que ces dernières ont été validées par l'ensemble des responsables pédagogiques, les spécifications ont alors été mises au point. Là encore, elles ont été validées par l'ensemble des responsables pédagogiques. Dans ce projet, les construits et les spécifications des examens étaient identiques quelle que soit la langue évaluée.

Après cette étape conceptuelle, la CQE est intervenue à plusieurs reprises en Slovaquie. Elle a accompagné le processus de développement des examens mis en œuvre par le Štátny pedagogický ústav. À ce jour, les examens pour les niveaux A1, A2 et B1 ont déjà été développés et administrés, ceux du niveau B2 le seront prochainement.

2. Utilisation du CECRL dans le cadre du développement des examens

2.1. Description des construits et du cadre conceptuel

Le construit est une traduction de l'anglais « *construct* ». Il peut aussi être traduit dans la littérature francophone par « construit théorique », « construit hypothétique » ou encore « concept hypothétique ». La définition du « construit » utilisée dans cet article est empruntée à M. Laurier (2005) :

« [...] On peut définir un construit comme un ensemble de variables qui représente le réseau d'habiletés, d'attitudes et de connaissances composant une compétence, de même que les interactions entre celles-ci et les autres compétences. » (Laurier, M. 2005 :66)

Présenter l'ensemble des construits du test ainsi que le cadre conceptuel (Butaš, Butašová et Folny, 2011) serait sans doute inapproprié dans cet article, c'est pourquoi pour des raisons d'économie, il n'en sera donné qu'un résumé. Ce résumé a pour objectif de mieux faire comprendre la nature des examens, de présenter les procédures d'utilisation du CECRL dans ce contexte, et enfin, faire percevoir la différence entre construits et échelles de compétence.

Le cadre conceptuel choisi pour ces examens s'inspire de la description de la compétence en langue, inscrite dans une logique communicative, telle qu'elle est proposée par Bachman (1990), Bachman et Palmer (1996, 2010), mais aussi par Canale et Swain (1980), et Canale (1983). Pour ce qui est des échelles de niveaux, les travaux de Brian North (2000), utilisés dans le CECRL, servent de référence. Pour chacun des niveaux, les compétences de compréhension orale, compréhension écrite, de structures de la langue et d'expression orale sont évaluées.

Construit de la compréhension de l'oral

La compréhension de l'oral est la compétence qui permet à un individu en position de « spectateur », de comprendre des contenus délivrés dans une situation de communication (plus ou moins prévisible) par un ou plusieurs interlocuteurs ayant différents statuts, que les discours soient spontanés ou bien encore écrits et prononcés à voix haute. La compréhension porte aussi bien sur des thèmes concrets qu'abstraites. Cette activité langagière appartient au domaine personnel, public, éducationnel ou professionnel (Conseil de l'Europe, 2001). La compétence de compréhension orale se singularise par la nécessité de l'auditeur de comprendre les contours globaux et les informations essentielles des discours et par une moindre capacité à analyser en détail le discours des autres (Song, 2008).

Pour l'actualisation du construit dans les tâches du test, les échelles et les sous-échelles du CECRL du niveau A1 au niveau B2 qui ont été sélectionnées sont :

- compréhension générale de l'oral ;
- comprendre une interaction entre locuteurs natifs ;
- comprendre en tant qu'auditeur ;
- comprendre des émissions de radio et des enregistrements.

Construit de la compréhension de l'écrit

La compréhension de l'écrit est la compétence qui permet à un individu de comprendre des contenus linguistiques exprimés dans un ensemble d'écrits diversifiés et adaptés à la situation de communication dans laquelle ils s'inscrivent. Ces écrits peuvent être composés par un ou des scripteurs ayant un statut variable (frère, amie, collègue, supérieure hiérarchique,...), que le discours soit synchrone ou bien encore différé. Les écrits peuvent être concrets ou abstraits et appartenir au domaine personnel, public, éducationnel ou professionnel.

La compétence de compréhension écrite se singularise par la nécessité du lecteur de comprendre non seulement les contours et les informations essentielles des discours, mais aussi par une capacité à analyser les textes dans le détail (Song, 2008).

Pour l'actualisation du construit dans les tâches du test, les échelles et les sous-échelles du CECRL du niveau A1 au niveau B2 qui ont été sélectionnées sont :

- compréhension générale de l'écrit ;
- lire des instructions ;
- lire pour s'orienter ;
- lire pour s'informer.

Construit des structures de la langue

Ce que l'on entend par « structures de la langue » concerne à la fois la connaissance de la langue et la capacité à utiliser ses structures dans des contextes langagiers particuliers représentatifs de familles de situations. Les structures de la langue peuvent être des structures orales ou écrites. Toutefois, leur évaluation passe systématiquement par l'écrit. Les structures pouvant être évaluées sont celles qui sont comprises dans le périmètre de la « norme » de la langue. Autrement dit, il s'agit d'évaluer les structures qui permettent

de faire passer un message dans un maximum de situations. Sont privilégiées les structures qui sont comprises et utilisées par la majorité des locuteurs / scripteurs de la langue.

Les structures de la langue revêtent un caractère discret, autrement dit, il s'agit bien de segments isolés de la langue apparentés le plus souvent à des connaissances. Le candidat doit démontrer sa capacité à utiliser ces connaissances en fonction d'un contexte aussi minimal soit-il.

L'évaluation de connaissances discrètes a pour objectif d'assurer une couverture de contenu plus équilibrée en prenant en compte différents aspects de la compétence, et pas uniquement des composantes ayant un niveau d'intégration plus élevé (comme pour l'épreuve d'expression orale où les tâches proposées demandent un niveau d'intégration plus élevé des connaissances, attitudes et habiletés). Il s'agit donc d'évaluer la compétence du candidat à un niveau de granulation plus fin, notamment pour évaluer la précision de sa connaissance (voire de la maîtrise) de la langue.

Pour l'actualisation de ce construit dans les tâches du test, les éléments suivants sont pris en considération :

- connaissances morphosyntaxiques ;
- connaissances lexicales.

Sont également pris en compte les descriptions de niveau de référence (DNR) pour les langues bénéficiant déjà de référentiels¹ adossés au CECRL.

Construit de l'interaction orale

L'interaction orale est la résultante de la performance de deux ou plusieurs individus ayant des statuts symétriques ou asymétriques qui « co-construisent » un discours ou, dit plus simplement, construisent un discours se développant au fil des échanges verbaux. L'interaction orale a pour singularité d'être au carrefour de la compétence en langue de deux ou plusieurs individus qui interagissent de manière peu prévisible au cours de leurs échanges. En aucun cas, l'interaction ne modifie la compétence en langue de chacun des acteurs participants aux échanges verbaux (Bachman, 2007). Autrement dit, même s'il y a un processus d'apprentissage de la langue au cours de l'interaction orale (aussi minimal soit-il), en situation d'examen, on n'évalue pas ce processus. Surtout, il n'est pas suffisamment significatif pour que la performance du candidat varie significativement. En revanche, le niveau de compétence du candidat évalué a bien un impact sur la performance finale, sur le résultat de l'interaction. L'interaction entre pairs est évaluée avec des futurs professeurs de niveaux très similaires même si la recherche n'a pas réussi à démontrer qu'une différence de niveau impactait la performance des candidats (Csépes, 2009 ; Iwashita, 1999 ; Nakatsuhara, 2004).

Pour l'actualisation de ce construit dans les tâches du test, les échelles et les sous-échelles du CECRL du niveau A1 au niveau B2 qui ont été sélectionnées sont :

- interaction orale générale ;
- discussions informelles ;
- échange d'information ;
- interviewé et être interviewé.

2.2. Synthèse de l'utilisation du CECRL pour le développement des examens

A la lecture des construits, on comprend que dans ce projet, le CECRL est utilisé essentiellement pour les descripteurs et les échelles de niveaux développées par Brian North (2000). Le choix d'évaluer la compréhension orale, la compréhension écrite et des structures de la langue est autant dû au contexte d'utilisation des examens qu'à une « tradition » dans l'évaluation en langue (Lado, 1961 ; Carroll, 1961, 1972). En revanche, l'introduction de l'évaluation de l'interaction orale a été rendue possible, en grande partie, grâce au CECRL. En effet, la place attribuée à l'interaction orale dans le CECRL a donné une légitimité certaine à l'évaluation d'une compétence jusque là peu évaluée dans les systèmes scolaire et universitaire slovaques. Cette introduction de l'évaluation de l'interaction orale était d'autant plus nécessaire que cette compétence est essentielle pour un professeur enseignant une langue.

Au-delà de l'utilisation des descripteurs de niveau, les « domaines » évoqués dans le CECRL et les « sous-échelles » de compétence ont eux-aussi été utilisés. Le choix des sous-échelles a été guidé par la volonté de mettre le contenu des examens en adéquation, autant que possible, avec les activités langagières des futurs professeurs des écoles. Les domaines identifiés par le CECRL, soit « personnel », « public », « éducationnel » et « professionnel » ont été conservés. Une utilisation prolongée de ces catégories (notamment au CIEP) a déjà démontré qu'elles sont opérantes. Elles permettent de distinguer des catégories de contenu de manière fiable. Elles facilitent le travail réalisé sur la couverture de contenu surtout pour distinguer le contenu des examens de niveau différent. Enfin, elles ouvrent la possibilité d'une meilleure qualité des inférences faites à partir des résultats sur la compétence du candidat. Ici, le terme « inférence » renvoie au sens qu'il prend traditionnellement dans la linguistique appliquée. Il s'agit du processus d'interprétation amenant à généraliser les résultats à un test pour décrire la compétence d'un candidat ou son positionnement sur un niveau de compétence. La qualité du processus inférentiel s'évalue en vérifiant la validité des liens entre les construits du test, les spécifications, la performance du candidat, les procédures permettant de calculer le score, la fidélité du test, la généralisation des résultats à plusieurs versions d'un examen ou encore à l'extrapolation de la compétence observée à des situations / tâches non présentes dans le test. Pour plus de précisions concernant cette notion, nous renvoyons le lecteur vers les ouvrages de Bachman et Palmer (2010), Chalhoub-Deville, Chapelle et Duff (2006) et Ercikan et Roth (2009).

L'utilisation du CECRL s'est avéré être moins pertinente ou indispensable pour la rédaction des construits, des spécifications et l'économie générale du projet.

Pour ce qui est des construits, si les définitions données dans le CECRL (Conseil de l'Europe, 2001) ne sont en rien en désaccord avec celles proposées dans le projet slovaque, en revanche, elles se sont révélées être insuffisantes. Il ne faut pas négliger le fait que le CECRL a pour vocation de proposer des construits et des descripteurs de performance « universels ». Certes, si cela n'invalide en rien le CECRL, cela situe pourtant la difficulté à l'utiliser pour des examens ayant une population et des utilisations bien définies et limitées. Par ailleurs, le CECRL est un outil de vulgarisation scientifique, une compilation de savoirs et de savoir-faire parue à un instant t. Or, cette vulgarisation date au moins des années 1990 (Alderson, 2007 ; North et Schneider, 1998). Depuis 10 ans, la recherche a avancé et il est, maintenant, indispensable d'accompagner l'utilisation du CECRL

par de nouveaux travaux et, bien sûr, par le retour sur les pratiques professionnelles. Alors que les échelles de compétences sont et resteront encore longtemps valides et opérantes (parce que le travail effectué par Brian North est par nature universel et les échelles du CECRL sont utilisées fréquemment pour fixer des normes), les définitions données aux construits et aux différentes composantes de la compétence en langue évoluent. Pour ne prendre qu'un exemple, le construit de l'interaction orale est de plus en plus souvent envisagé sous l'angle d'une interaction entre pairs et non plus uniquement entre un examinateur et un candidat. Cet aspect de l'interaction n'est pas significativement traité dans le CECRL. En revanche, de récents travaux permettent de mieux appréhender ce construit et le fonctionnement des tâches permettant de l'évaluer et de le mesurer (Ducasse, 2010 ; Taylor et Wigglesworth, 2009).

Le cadre, au point 9.2.1 (Conseil de l'Europe, 2001 : 136), consacre peu de place aux spécifications des examens en renvoyant le lecteur vers plusieurs chapitres. Il va de soi que le cadre n'a pas pour ambition d'être un manuel pour la conception des examens. Toutefois, on peut remarquer que même les manuels conçus pour la conception d'examens ne sont guère plus précis. Le très utile et pertinent *Manual for Language Test Development and Examining* (Conseil de l'Europe, 2011 : 23-25) n'est pas prolixe sur ce point. Il semble que le travail sur les spécifications des examens reste souvent dans le giron des compétences des organismes professionnels. Il est pourtant important que l'ensemble des acteurs du monde éducatif comprenne le lien entre les construits, les spécifications du test, l'interprétation des résultats et les conséquences liées à leurs utilisations. Ce point crucial mériterait donc d'être traité plus en profondeur. Depuis quelques années, les critiques portées sur le CECRL sont récurrentes. Cependant, si certains griefs sont fondés et vrais, ils n'en sont pas pour autant pertinents. On a ainsi reproché au CECRL de proposer des spécifications insuffisamment précises pour faciliter la mise en relation des tâches avec les descripteurs du cadre (Weir, 2007). Toutefois, comme le rappelle Bachman (2002), la difficulté d'une tâche est fonction de ses caractéristiques mais aussi de celles des candidats et de l'interaction entre les deux. Il est donc peu probable que l'on puisse trouver des spécifications universelles correspondant exactement au niveau du CECRL. En revanche, il serait utile d'expliquer comment on s'assure de la validité des spécifications au regard, à la fois, des échelles du CECRL, des construits et de ou des utilisations de l'examen. Il deviendrait alors plus aisé de ne pas rédiger des examens essentiellement focalisés sur les niveaux du CECRL mais aussi de prendre en considération l'utilisation des résultats.

Pour ce qui est de l'économie générale du projet, le CECRL (2001) ne constitue qu'une composante des compétences ayant permis de réaliser le projet. Si, et c'est sans doute normal, le CECRL ne constitue pas un apport exclusif dans la gestion d'un projet, nous recommandons la lecture du *Manual for Language Test Development and Examining* (Conseil de l'Europe, 2011). Cette publication (récemment mise à jour) est fortement utile à la conduite d'un projet. De ce point de vue, elle complète utilement le CECRL.

3. Emergence de besoins d'amélioration des outils existant et besoin de nouveaux outils

Souvent, lorsqu'on utilise le CECRL pour concevoir un examen, surgissent des tensions liées au besoin de développer un outil qui soit à la fois universel et adapté à une population et un usage particuliers. Or, il n'y a pas d'examens complètement universaux ou d'examens

utilisés pour une population complètement homogène, avec des utilisations ciblées et stables. Le positionnement des examens n'est en rien dichotomique car il correspond plutôt à un positionnement sur un continuum.

Ici, le choix opéré par le Štátny pedagogický ústav reflète le fait que les futurs professeurs des écoles sont en situation d'apprentissage de la langue et qu'ils deviennent à l'issue des formations, eux-mêmes, des enseignants. Le contenu et les domaines choisis pour la rédaction des tâches et des items reflètent cette double orientation. Dans ce contexte, il n'est pas possible de concevoir des épreuves uniquement orientées sur le domaine éducationnel (les futurs professeurs étant eux-mêmes en situation d'apprentissage) ou uniquement sur les niveaux du CECRL (les participants aux formations devant enseigner les langues dans un cadre professionnel). Le choix des compétences de production et de réception, orale et écrite correspond à l'utilisation par les professeurs de la langue tant au niveau professionnel que personnel. Ainsi l'évaluation des compétences à l'oral (compréhension orale et interaction orale) prédomine-t-elle sur celle de l'écrit. L'oral est un élément essentiel pour l'enseignement d'une langue à des jeunes enfants, mais aussi, pour qu'un professeur reste au contact d'une culture et qu'il entretienne, diversifie ou améliore ses compétences en langue. L'évaluation des structures de la langue, elle, a pour objectif de vérifier la précision des connaissances linguistiques (aspect important de la compétence professionnelle d'un professeur). Enfin, l'évaluation de la compréhension écrite a pour objectif de s'assurer que le futur professeur pourra lire, soit du matériel pédagogique qui lui sera destiné, soit des documents lui permettant de rester au contact de la / les culture(s) correspondant à la langue apprise. *In fine*, les spécifications du test reflètent le contexte particulier slovaque avec deux éléments de contexte majeurs : l'utilisation de la langue à des fins d'enseignement et la pratique régulière de cette langue hors contexte scolaire.

La comparabilité des épreuves rédigées en 6 langues demande également de résoudre la tension entre le besoin d'universalité (d'homogénéité des épreuves) et la prise en compte de réelles différences d'ordre linguistique, sociolinguistique et pragmatique. Si, étant donné son caractère universel, le CECRL ne propose pas de descriptions pour chacune des langues, les DNR ont permis d'offrir des compléments intéressants au cadre. Portant, l'utilisation de ces DNR dans le cadre du développement d'examens en plusieurs langues pose problème. Dans ce projet, nous ne disposons pas de DNR pour la langue russe. Par ailleurs, les DNR pour l'anglais, l'allemand, le français, l'espagnol et l'italien ont été développés de manière différentes (soit à partir de l'analyse de corpus de documents, de corpus de copies ou d'avis d'experts,...). Les DNR ayant été développés pour une utilisation par langue, il est difficile de les utiliser dans des projets « multilingues ». Le « *A Core Inventory for General English* » (North, Ortega et Sheehan, 2011) pourrait, s'il était développé pour plusieurs langues, représenté un complément fondamental au CECRL. En effet, cet outil regroupe toutes les notions / fonctions essentielles permettant de décrire les niveaux du CECRL en anglais. Dans ce travail, les auteurs ont vérifié que la communauté professionnelle reconnaît bien ces notions comme étant représentatives de chacun des niveaux. Ce travail, s'il était généralisé aux autres langues, permettrait de disposer de référentiels développés à partir d'une même méthodologie pour plusieurs langues.

Dans le projet slovaque, l'utilisation des échelles du CECRL et la rédaction des construits ont permis de concevoir des examens en six langues et de trouver des éléments de comparaison pour le contenu des épreuves. Il est ainsi possible de vérifier (et de

démontrer) que, au niveau du contenu, les épreuves sont homogènes, quelle que soit la langue évaluée, ce qui n'échappe ni aux experts ni à un public moins averti.

Pourtant, il aurait été intéressant de disposer d'échelles à la fois à vocation universelle et ciblant plus précisément l'utilisation de la langue dans un contexte d'enseignement. Le CECRL gagnerait sans doute à proposer des échelles visant des familles de compétences correspondant à différentes utilisations de la langue et des populations plus homogènes. On pourrait avoir des échelles plus spécifiques et en harmonie avec les échelles déjà développées par B. North (2000). Il serait ainsi intéressant de développer sur tous les niveaux ou un nombre de niveaux plus restreints des échelles pour :

- les enfants ;
- les émigrés ;
- les étudiants ;
- les professeurs de langue ;
- les publics des professionnels ;
- les publics peu ou pas instruits ;
- ...

L'objectif de la réalisation de ces échelles serait de doter les développeurs d'examens d'outils à la fois plus adaptés à une situation spécifique et suffisamment universalistes pour permettre de respecter l'adossement aux niveaux du CECRL. Si le choix de concevoir ou pas des examens spécifiques appartient aux développeurs, il est raisonnable de penser que cela inciterait et faciliterait le développement d'outils plus en accord avec les inférences faites à partir des résultats et leur interprétation (Bachman, 2007, 2010).

Dans le projet slovaque, des échelles plus précises faites pour les professeurs des écoles primaires auraient pu servir à concevoir toutes ou plusieurs tâches dans les épreuves. Elles auraient pu ne pas être utilisées pour les bas niveaux (où l'objectif est de faire acquérir une maîtrise générale de la langue) mais, en revanche, avoir un poids plus important pour le niveau B2 (là où on évalue la maîtrise de la langue dans une logique de certification professionnelle). *In fine*, le CECRL aurait ainsi servi pour développer des examens plus en accord avec leurs utilisations et aurait sans doute apporté un surcroît de légitimité à la démarche sans tomber dans le « tout spécifique » ou le « tout universaliste ».

Conclusion

L'utilisation du CECRL avec pour objectif de développer des examens pour les futurs professeurs des écoles primaires a été fructueuse. Elle a permis d'inscrire le projet dans une vision de la compétence en langue uniforme et de pouvoir adosser les épreuves à un même référentiel. Toutefois, l'adossement au CECRL ne constitue qu'un aspect lié au développement des épreuves. Il n'autorise pas l'économie de la rédaction des construits, de réfléchir à la meilleure adéquation possible entre les construits et les spécifications ou encore de mettre en place des procédures pour que la conception de l'examen mais aussi sa gestion opérationnelle soient les plus efficaces possible. Le CECRL, malgré toutes les précautions qui doivent être prises, se révèle être un excellent outil pour donner cohésion et liant au développement des tests et certifications mais aussi une réelle visibilité auprès du grand public. Mais ce, bien sûr, à condition de ne pas oublier comment seront utilisées les inférences faites à partir des résultats et de rester attentif à l'avancée des connaissances.

Bibliographie

- Bachman, L. F. 1990. *Fundamental considerations in language testing*. Oxford : Oxford University Press.
- Bachman, L. F. Palmer, A. S. 1996. *Language testing in practice*, Oxford : Oxford University Press.
- Bachman, L. 2002. Some reflections on task-based language performance assessment. *Language Testing*, n° 19. 453-476.
- Bachman, L. F. 2007. What is the construct? In J. Fox, M. Weshe, D. Bayliss, L. Cheng, C. Turner, C. Doe (Eds.), *Language testing reconsidered*. Ottawa : Ottawa University Press, pp. 80-111.
- Bachman, L. F., et Palmer, A. S. 2010. *Language testing in practice: developing language assessments and justifying their use in the real world*. Oxford : Oxford University Press.
- Bolton, S. 2010. Auditing Cambridge ESOL's main suite and BEC examinations. *Research Notes*, n° 39, pp. 31-33, http://www.cambridgeesol.org/rs_notes/offprints/pdfs/RN39p31-33.pdf [Consulté le 1 avril 2011].
- Butaš, J., Butašová, A., Folny, V. 2011. *Tvorba testov pre meranie úrovne A1, A2, B1 a B2 podľa Spoločného európskeho referenčného rámca pre jazyky : vzdelávanie učiteľov základných škôl v oblasti cudzích jazykov vsúvislosti s Konceptiou vyučovania cudzích jazykov na základných a stredných školách*. Bratislava : VERBUM
- Canale, M. 1983. From communicative competence to communicative language pedagogy. In : J. C. Richards, R. W. Schmidt (Eds). *Language and Communication*. London : Longman.
- Canale, M., et Swain, M. 1980. Theoretical bases of communicative approach to second language teaching and testing. *Applied Linguistics*, n° 1, pp. 1-47.
- Carroll, J. B. 1961. Fundamental considerations in testing English language proficiency of foreign students. In: H. B. Allen, R. N. Campbell (Eds). 1972. *Teaching English as a Second Language* (2nd ed.). New York : McGraw-Hill, pp. 313-321.
- Chalbound-Deville, M. et al. (Eds). 2006. *Inference and generalizability in applied linguistics : multiple perspectives*, Amsterdam & Philadelphia : Benjamins.
- Csépes, I. 2009. *Measuring oral proficiency through paired-task performance*. Language testing and evaluation. Vol.14. Frankfurt am Main : Peter Lang.
- Conseil de l'Europe. 2001. *Cadre européen commun de référence pour les langues - Apprendre, enseigner, évaluer*. Paris : Didier. (3e édition revue et corrigée), http://www.coe.int/T/DG4/Linguistic/Source/Framework_FR.pdf [Consulté le 1/04/11].
- Conseil de l'Europe. 2011. *Manual for Language Test Development and Examining*. Language Policy Division, http://www.coe.int/t/dg4/linguistic/ManualLangageTest-ALte2011_EN.pdf [Consulté le 9 septembre 2011].
- Ducasse, A. 2010. *Interaction in paired oral proficiency assessment in Spanish*. *Language testing and evaluation*. vol. 20. Frankfurt am Main : Peter Lang.
- Ercikan, K., Roth, W-M. (Eds.) 2009. *Generalizing from Educational Research: Beyond Qualitative and Quantitative Polarization*. New York : Routledge.
- Iwashita, N. 1999. The validity of the paired interview format in oral performance assessment. Melbourne papers in *Language Testing*, n° 8, pp. 51-66.

Lado, R. 1961. *Language testing*. New York : McGraw-Hill.

Laurier, M. et al. 2005. *Les principes de la mesure et de l'évaluation des apprentissages*. Montréal : Gaétan Morin éditeur.

Nakatsuhara, F. 2004. *An investigation into conversational styles in speaking tests*. Unpublished masters thesis. University of Essex.

North, B. 2000. *The Development of a common framework scale of language proficiency*. New York : Peter Lang.

North, B., Ortega, A., Sheehan, S. 2011. *A Core Inventory for General English*. British Council / EAQUALS, <http://www.teachingenglish.org.uk/sites/teacheng/files/Z243%20E&E%20EQUALS%20BROCHURErevised6.pdf> [Consulté le 9 septembre 2011].

North, B., Schneider, G. 1998. Scaling descriptors for language proficiency scales. *Language Testing*, n° 15, pp. 217-262.

State Institute of Education. 2009. *Conception of teaching foreign languages at primary and secondary schools: increasing the effectiveness of teaching foreign languages*. Bratislava.

Song, M. Y. 2008. Do divisible subskills exist in second language (L2) comprehension? A structural equation modeling approach. *Language Testing* n° 25, pp. 435-464.

Taylor, L., Wigglesworth, G. (Eds.) 2009. Special Issue: Pair Work in L2 Assessment Context. *Language testing*, n° 26.

Weir, C. J. 2007. Limitations of the Common European Framework for developing comparable examinations and tests. *Language Testing*, n° 22, pp. 281-300.

Notes

¹ La liste est consultable à http://www.coe.int/t/dg4/linguistic/dnr_FR.asp