



17^{ème} Colloque international de l'ADMEE Europe

Lisbonne - 18-20 novembre 2004

"L'évaluation des compétences : entre reconnaissance et validation des acquis de l'expérience "

L'évaluation des compétences linguistiques des adultes en français langue étrangère dans une perspective de multi-référentialisation

Marc DEMEUSE, Service de Pédagogie théorique et expérimentale, Université de Liège et IREDU, Université de Bourgogne et CNRS

Franck DESROCHES,

Alexandra CRENDAL,

Philippe OSTER,

François RENAUD, Direction des Relations internationales de l'Enseignement, Chambre de Commerce et d'Industrie de Paris (DRI-E/CCIP).

Et Xavier LEROUX.

Mots-clés : français langue étrangère, référentiel, compétences linguistiques, adultes non francophones

Résumé

L'évaluation des compétences linguistiques des adultes en français langue étrangère est l'objet de nombreux enjeux : reconnaissance et validation des acquis de l'expérience ou d'apprentissages formels non certifiés, notamment. Le Conseil de l'Europe, mais aussi les services canadiens de l'immigration (Citoyenneté et Immigration Canada) ou les autorités responsables de l'enseignement supérieur des pays francophones (comme la France ou la Communauté française de Belgique), définissent des standards destinés à permettre l'évaluation des compétences linguistiques de sujets adultes dans le domaine du français. Les objectifs de ces organismes peuvent être très variables et les enjeux, pour les candidats à l'évaluation, plus ou moins importants (depuis la reconnaissance d'un niveau de français dans la perspective d'une valorisation barémique jusqu'à la possibilité d'accéder à des études supérieures en France ou en Belgique, ou à être admis comme immigrant économique sur le territoire canadien).

Si, explicitement ou non, les différents organismes font recours à une approche par compétences, les référentiels peuvent se présenter de manière extrêmement variée : le Conseil de l'Europe, à travers son cadre commun de référence, et Citoyenneté et Immigration Canada, développent un document très complet et publié alors que les autorités académiques belges et françaises privilégient le travail de commissions chargées de créer ou d'agréer des outils (tests ou examens) d'évaluation.

D'autres acteurs, publics ou privés, développent par ailleurs, en rapport avec les référentiels mentionnés, des outils qu'ils soumettent à l'agrément des différentes autorités publiques. C'est dans ce cadre, celui de l'opérationnalisation à travers un outil d'évaluation particulier, dont l'objectif est de correspondre à un ensemble de référentiels, que cette communication se situe. Elle mobilise à la fois une équipe de conception et un chercheur indépendant chargé d'assurer le contrôle qualité de la production et du suivi des tests.

Après une mise en correspondance des différents référentiels externes et du référentiel propre de l'outil d'évaluation lui-même, les auteurs mettront en évidence les stratégies de développement des différentes formes équivalentes (formes parallèles) du test pour chacune des quatre composantes (expression et compréhension de l'oral et de l'écrit).

A travers l'analyse de la mise en œuvre d'un test particulier de « français général » pour adultes dont la langue maternelle n'est pas le français, l'exposé tentera de mettre en évidence la complexité d'une réponse spécifique à des prescrits distincts, sous contrainte d'une procédure de qualité totale en voie de certification.

1. Les enjeux de l'évaluation en français langue étrangère

1.1. Le contexte évolutif de l'évaluation en FLE : les composantes visibles du changement

Le domaine de l'évaluation en français langue étrangère s'est beaucoup développé ces quinze dernières années, et notamment en France. Plusieurs événements **en politiques linguistiques et éducatives** ont poussé ce domaine à évoluer : la convergence des politiques linguistiques européennes, la valorisation de la reconnaissance et des acquis de l'expérience et les recherches dans le domaine de l'évaluation en français langue étrangère, alliant désormais analyses qualitatives et quantitatives, pré-expérimentées avec les premiers référentiels en français langue étrangère comme le Français fondamental et reconduites dans les standards d'ALTE, de DIALANG et enfin du Cadre européen commun de référence pour les langues (CECR).

D'autres facteurs, **sociologiques**, ont aussi amené ces changements, et notamment la mondialisation des échanges et la mobilité accrue des étudiants, des professionnels, des personnes pour des raisons économiques, géopolitiques, éducatives. C'est dire combien les enjeux communicationnels sont devenus grands. Désormais, communiquer vite et efficacement est un enjeu majeur tant pour les individus que pour le monde de l'entreprise, la société, un état. Dans ce cadre-là, l'évaluation de la compétence en langue s'impose car elle permet au candidat et à l'institution de mesurer les réelles compétences en langue permettant une meilleure intégration. Tel est le cas pour les autorités canadiennes qui accordent un quart des points (24 / 100) à la maîtrise des langues française et anglaise à tout travailleur qualifié présentant une demande de visa.

Troisièmement, le **facteur temporel** s'impose et a ouvert de nouvelles perspectives à l'évaluation en langues. Pour reprendre l'expression d'Eric Delamotte (in Le commerce des langues, 1999), la « temporalité est renouvelée » en ce sens que la notion de temps influe sur la conception des outils et l'interprétation des résultats. Deux niveaux temporels peuvent être dégagés : l'un chronométrique, l'autre chronologique. Au niveau chronométrique, deux arguments de force ressortent :

◇ Pour le candidat, l'évaluation instantanée qu'offre un test de langue permet de valider, en deux heures, un niveau de compétence et d'obtenir, sous 48H00, des résultats pour présenter un dossier d'entrée à l'Université ou une demande de visa en vue d'études ou d'émigration future.

◇ Pour les institutions, l'exactitude du jugement et de la prise de décision sont de mise. Aussi, interpréter les résultats d'un diplôme obtenu il y a une dizaine d'années les intéressent de moins en moins. Elles se tournent petit à petit vers des instruments de mesure qui photographient la compétence linguistique du candidat au moment où il va être recruté par un office d'immigration ou admis à l'université.

Au niveau chronologique, les enjeux sont importants et ont amené des améliorations dans le domaine de la conception d'épreuves en français langue étrangère :

◇ La mise en place du portfolio européen et l'incitation à la « formation tout au long de la vie » amènent aujourd'hui les candidats à capitaliser toute expérience de l'utilisation des langues et à acter toute évaluation qui validerait un niveau de compétence. Or, si l'on veut mesurer la progression, il est important que le candidat passe une évaluation standardisée, étalonnée sur un cadre de référence comme le Cadre européen commun de référence pour les langues (CECR).

◇ Le même problème se pose pour des centres de formation qui utiliseraient l'évaluation d'un candidat comme un outil de positionnement. Les centres de formation en langues ont besoin de se référer à des outils qui mesureront toujours le niveau du candidat dans les quatre habiletés langagières et qui leur permettront de placer plus justement le candidat.

Le facteur temps traité ci-dessus montre combien **l'univers symbolique** de l'évaluation semble avoir changé de paradigme ces dix dernières années. En effet, la notion d'efficacité et d'équité apparaît de plus en plus dans les discours institutionnels. Dans le chapitre consacré à l'évaluation dans le Cadre européen commun de référence pour les langues (Conseil de l'Europe, 1998, p135), il est fait mention de ces aspects : « (...) *dans un programme de langue, la compétence de l'apprenant n'est pas la seule chose évaluée – par exemple, la rentabilité de certains matériels pédagogiques ou méthodes, le type et la qualité du discours effectivement produit, la satisfaction de l'enseignement et celle de l'apprenant, l'efficacité de l'enseignement, etc.* ». De plus, l'apprenant n'est plus seulement acteur de son apprentissage ; il est aussi de plus en plus maître de son parcours ; il capitalise ses savoirs et savoir-faire et apprend à être « *conscient de l'état présent de ces connaissances* » (recommandation du Conseil de l'Europe, p12).

Enfin, **au niveau de l'éthique**, il est intéressant de constater une avancée certaine. En effet, les institutions qui sélectionnent des candidats à l'immigration ou à l'entrée à l'université font de plus en plus appel à des opérateurs externes spécialistes de l'évaluation en langues et qui offrent des outils adaptés à la demande de telles institutions. Ainsi, le sélectionneur n'est plus partie prenante de l'évaluation linguistique en elle-même, mais sélectionne le candidat en fonction de seuils d'acceptabilité. Par exemple, suite à l'amendement d'une nouvelle loi sur l'immigration, les autorités canadiennes (Citoyenneté et Immigration Canada) ont décidé d'externaliser la passation de l'épreuve linguistique et de la confier à un organisme spécialiste de l'évaluation en langues.

Le contexte de l'évaluation en langues a donc bien évolué depuis les années 70 et de nouvelles voies s'ouvrent à la recherche interdisciplinaire. Cette évolution s'est faite de concert avec la mise en place de standards linguistiques qui ont été longuement éprouvés scientifiquement : les Standards linguistiques canadiens (Canada anglophone), les Niveaux de compétence en français langue seconde (Québec) et le Cadre européen commun de référence pour les langues (Europe). D'autres standards existent dans des domaines plus spécifiques : par exemple, les « Stanlag » pour l'OTAN.

1.2. De l'évaluation aux enjeux

Jean-Claude Mothe, Directeur du BELC, constatait il y a dix ans que l'évaluation en langues vivantes n'avait pas encore tout à fait droit de cité en France. Aujourd'hui, il en va différemment.

Dans le monde francophone, il y a pléthore de certifications. Elles sont actuellement recensées dans le cadre du projet CERTEL, à l'initiative de l'Agence Universitaire de la Francophonie (AUF). Le rapport devrait paraître en 2005.

En France, les certifications sont très nombreuses : certaines donnent un diplôme, d'autres une attestation. Elles peuvent être internes à une institution ou externes : elles sont alors proposées dans le monde entier, ce qui n'est pas sans poser un certain nombre de difficultés en termes d'homogénéisation des épreuves, de standardisation, d'étalonnage et de calibrage. C'est le cas notamment des tests de langue dont nous verrons plus loin ce que la dénomination recouvre du point de vue scientifique.

Les enjeux de ces certifications sont énormes, et à tout point de vue. La validation d'une compétence en français langue étrangère permettra au candidat d'émigrer, voire d'être sélectionné dans une université francophone. De tels enjeux amènent à sécuriser le dispositif de passation des épreuves et de délivrance des résultats, demande qui est d'ailleurs émise par les recruteurs et autorités.

Par ailleurs, d'autres instances comme le Conseil de l'Europe ont aujourd'hui ce souci de garantir des certifications de qualité. Le Cadre européen commun de référence pour les langues se réclame de cette approche qualité et propose, dans cette optique, un outil à partir duquel toutes les certifications pourront être reliées, suite à une validation interne et externe de l'outil. Ainsi, candidats et Institutions pourront comparer les niveaux dans toute l'Europe.

Les deux enjeux définis précédemment concernent le TEF. C'est donc dans ce cadre, celui de l'opérationnalisation à travers un outil d'évaluation particulier, dont l'objectif est de correspondre à un ensemble de référentiels, que cette communication se situe. Elle mobilise à la fois une équipe de conception et un chercheur indépendant chargé d'assurer le contrôle qualité de la production et du suivi des items.

1.3. La réponse d'un test de « français général » pour adultes non francophones

Le Test d'Évaluation de Français (TEF), créé en 1998, avait pour premier enjeu de répondre à la demande des écoles de la Chambre de Commerce et d'Industrie de Paris et des Établissements d'Enseignement à l'étranger. Ces organismes cherchaient à sélectionner, le plus justement possible, les étudiants étrangers sur la base, entre autres, des compétences linguistiques.

Dès sa création, les sept niveaux du TEF ont été indexés sur les six niveaux de compétence du Cadre européen commun de référence pour les langues. Pour l'élaboration des sept niveaux TEF, de nombreuses études ont été menées.

En juillet 2002, les autorités canadiennes adoptent une nouvelle loi sur l'immigration, et Citoyenneté et Immigration Canada référence le TEF comme test de français langue étrangère. Ce référencement donne lieu à de nouveaux travaux : une recherche est alors menée sur la mise en correspondance du cadre de référence du TEF avec les Standards linguistiques canadiens.

Cette communication tentera de montrer l'expérimentation qui a été menée par une équipe pluridisciplinaire.

2. Conception et construction du TEF

Avant d'aborder l'expérimentation, il est important de bien définir le champ d'action du Test d'Évaluation de Français d'un point de vue théorique et dans l'application de la théorie.

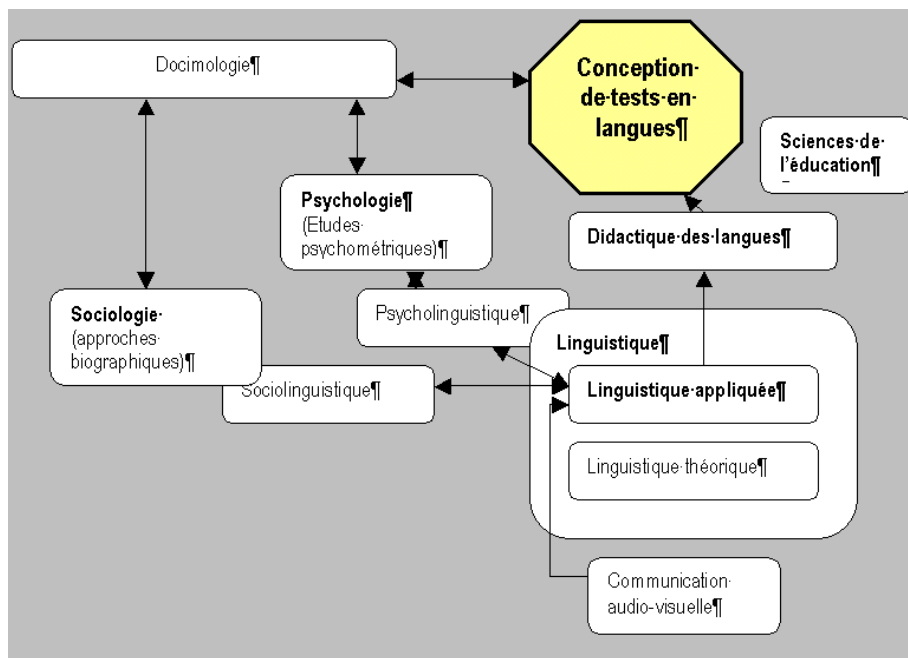
2.1. L'élaboration d'un test : un exercice à la « croisée des disciplines »

« Pour mériter le nom de test, un examen doit être standardisé, fidèle, valide et étalonné ». Cette définition donnée par Gilbert de Landsheere (Dictionnaire de l'évaluation et de la recherche en éducation, p 295) établit d'entrée des normes minimales de construction et de conception d'un test.

Le processus d'élaboration d'un test fait appel à plusieurs disciplines. C'est d'ailleurs ce qui en fait toute sa richesse comparé à une évaluation traditionnelle qui ne s'appuie que sur la pratique pédagogique du concepteur ou du formateur. Néanmoins, la mise en interaction de plusieurs champs disciplinaires peut rendre parfois difficile les études menées sur cet objet comme l'expérimentation sur la multiréférentialisation l'illustrera.

Comme le montre le schéma n°1, les tests sont à la croisée des disciplines. Du chercheur au praticien, en passant par l'évaluateur, les enjeux sont différents et les perceptions de l'outil également. Néanmoins, quelle que soit la position de l'acteur chercheur, évaluateur ou décideur, le cadre de référence du test et son indexation, voire multi-indexation sur des référentiels nationaux restent des points de référence.

Schéma 1 – La conception de tests et les disciplines théoriques et appliquées connexes



Source : Alexandra Crendal, d'après un schéma de Paul Rivenc sur la didactique des langues

2.2. Vers la mise en application : la construction du test

Les deux tableaux suivants, n°1 et n°2, sont un extrait de la table de spécification du TEF, test qui se compose de deux séries d'épreuves, les épreuves dites « obligatoires » (compréhension écrite, compréhension orale, lexique / structure) et les épreuves dites facultatives (expression écrite, expression orale).

Tableau 1 – *Epreuves obligatoires en 3 parties indissociables sous forme de questionnaire à choix multiple*

| | Durée | Structure | Score sur |
|----------------------|--------------|----------------------|-------------------|
| Compréhension écrite | 60 min | 50 questions | 300 points |
| Compréhension orale | 40 min | 60 questions | 360 points |
| Lexique / structure | 30 min | 40 questions | 240 points |
| TOTAL | 2H10 | 150 questions | 900 points |

Source : Le TEF, 6 années d'expertise, mars 2004, brochure éditée par la CCIP-DRH/E, p 7.

Tableau 2 – *Epreuves facultatives en deux parties indépendantes*

| | Durée | Structure | Score sur |
|------------------------|--------------|--------------------|------------------|
| Expression écrite (EE) | 60 min | 2 sujets à traiter | 450 points |
| Expression orale (EO) | 35 min | 2 sujets à traiter | 450 points |

Source : Le TEF, 6 années d'expertise, mars 2004, brochure éditée par la CCIP-DRH/E, p 7.

La conception des épreuves est centralisée et répond aux critères définis dans la table de spécification du TEF. Elle est réalisée par des concepteurs sélectionnés pour leurs compétences et régulièrement formés à l'utilisation des référentiels. Après évaluation et validation pédagogique, les questions à choix multiples sont pré-testées auprès de candidats. Enfin, des analyses sont menées sur les items à l'aide de la méthode de réponse à l'item (MRI) et de l'analyse de Rasch.

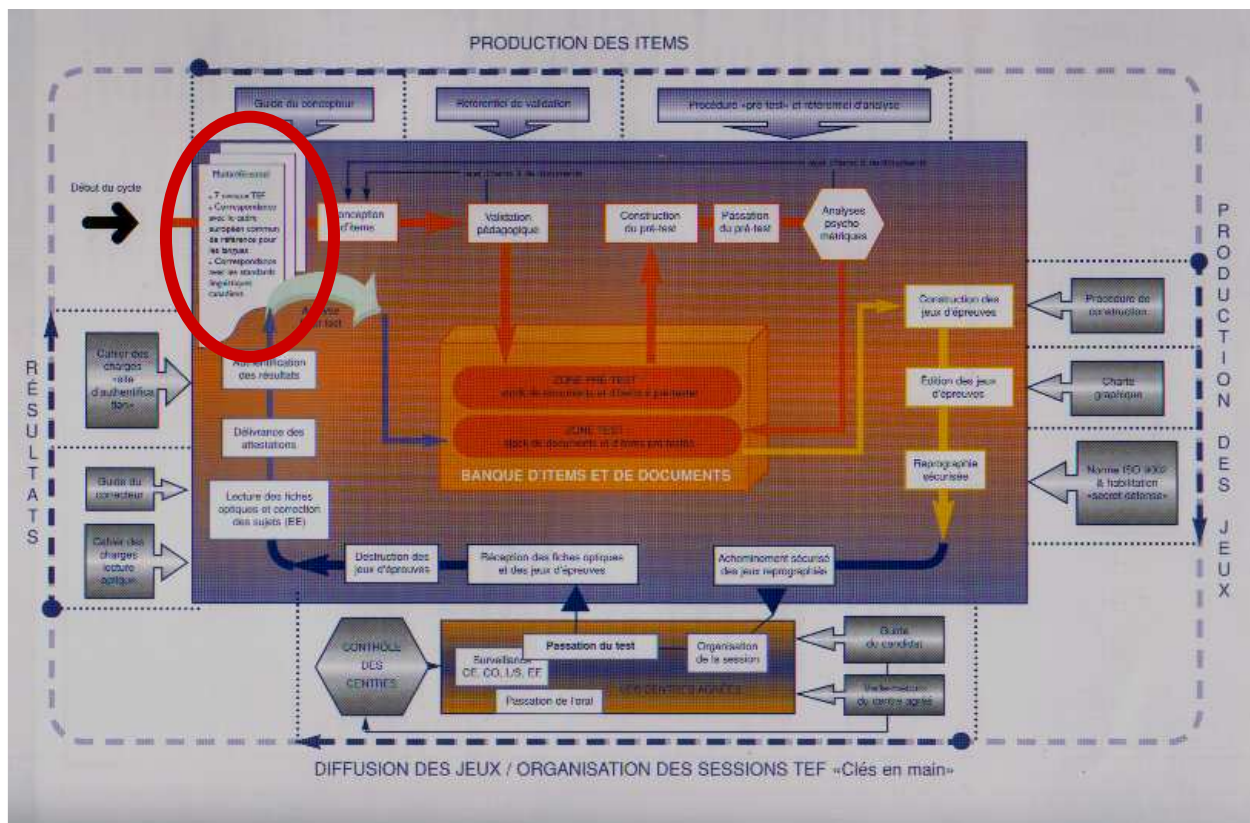
Les activités du TEF, au-delà de la conception, sont multiples. Deux grands processus interagissent : celui de la conception (production des items, production des jeux) et celui de la réalisation (diffusion des jeux, organisation des sessions TEF et communication des résultats). La qualité de ces deux processus est en permanence assurée. Le schéma n°2 présente ce cycle du Test d'Évaluation de Français (TEF) et la mesure de sa qualité.

3. La référentialisation : un référentiel ... des référentiels

3.1. Le rôle du référentiel dans le test

Le cadre de référence du TEF s'est d'abord appuyé sur le Cadre européen commun de référence pour les langues. Ces références ont permis d'élaborer les grilles de correction, nécessaires à l'évaluation des épreuves d'expression orale et écrite, et de proposer des commentaires individualisés sur l'attestation du candidat. Si le candidat ou un centre de formation le souhaite, il peut à partir des scores obtenus et des commentaires attribués en fonction du niveau atteint se référer au Cadre européen commun de référence pour les langues.

Schéma 2 – Cycle du Test d'Évaluation de Français (TEF) et mesure de sa qualité



Source : Le TEF, 6 années d'expertise, mars 2004, brochure éditée par la CCIP-DRIVE, p 16.

La référentialisation devient un enjeu majeur lorsque certains niveaux correspondent à la délivrance de points par des autorités telles que Citoyenneté et Immigration Canada, et que ces standards ne trouvent pas d'équivalent dans le système de notation proposé par le test. Cela devient encore plus impliquant lorsque ce même candidat obtient de 1 à 4 points par compétence s'il atteint au moins le niveau 3 du TEF ou le standard 4 des SLC, comme nous le montre le tableau n°3.

C'est dans ce cadre que s'est imposée la multiréférentialisation du TEF sur les Standards linguistiques canadiens.

Tableau 3 – Tableau d'équivalence des scores et niveaux TEF et de leurs mises en correspondance avec les SLC. Extrait du site Internet de Citoyenneté et Immigration Canada (juillet 2002).

| Niveau | Points (par compétence) | Résultats du test, par compétence | | | |
|------------------------------|--|--|--|--|--|
| | | Parler (expression orale) | Écouter (compréhension orale) | Lire (compréhension écrite) | Écrire (expression écrite) |
| Élevé (CLB/SLC 8-12) | 1 ^{ère} langue officielle : 4 | niveau 5 niveau 6 (349-450 points) | niveau 5 niveau 6 (280-360 points) | niveau 5 niveau 6 (233-300 points) | niveau 5 niveau 6 (349-450 points) |
| | 2 ^e langue officielle : 2 | | | | |
| Moyen (CLB/SLC 6-7) | 2 | niveau 4 (271-348 points) | niveau 4 (217-279 points) | niveau 4 (181-232 points) | niveau 4 (271-348 points) |
| Élémentaire (CLB/SLC 4-5) | 1 (max. de 2) | niveau 3 (181-270 points) | niveau 3 (145-216 points) | niveau 3 (121-180 points) | niveau 3 (181-270 points) |
| Néant (CLB/SLC 0-3) | 0 | niveau 0 niveau 1 niveau 2 (0-180 points) | niveau 0 niveau 1 niveau 2 (0-144 points) | niveau 0 niveau 1 niveau 2 (0-120 points) | niveau 0 niveau 1 niveau 2 (0-180 points) |

Source : <http://www.cic.gc.ca/francais/qualifie/qual-3-1.html>.

3.2. Une langue commune

La multiréférentialisation a donc des vertus à plusieurs égards. Elle permet une plus grande mobilité pour le candidat et facilite la traduction des niveaux pour les décideurs, les formateurs et les gestionnaires de programmes. Elle rend les résultats entre les individus plus équitables comme nous le montre ce passage extrait de l'introduction des Standards linguistiques canadiens.

« Sustrina, 25 ans, est arrivé au Canada d'Indonésie avec des habiletés limitées en français. Peu de temps après son arrivée, on évalue sa compétence linguistique et, sur la foi de ses résultats, il fut placé dans une classe de français de niveau 6. Plus tard cette même journée, Sustrina rencontra quelques uns de ses amis immigrants et, désireux de comparer son niveau d'habileté linguistique avec le leur, annonça qu'il se situait au niveau 6 en français langue seconde et demanda à chacun quel était son niveau. Une amie répondit qu'elle était classée débutant avancé et qu'elle ne comprenait pas ce qu'il voulait dire en parlant de niveau. Un autre répondit qu'il comprenait parfaitement ce que Sustrina voulait dire parce que, à son école de langue, il avait été classé au niveau 4. Il ajouta qu'on avait sans doute fait une erreur à l'École de Sustrina ou qu'il avait mal compris car, ils étaient bien d'accord là-dessus, les habiletés de Sustrina en français n'étaient certainement pas meilleures que celles de son amie, au contraire. »

C'est aussi avoir une langue commune qui permet une reconnaissance à l'échelle internationale d'une compétence en français langue seconde et étrangère.

3.3. Tableau synoptique de quatre référentiels en français langue étrangère

Le tableau n°4 propose de décrire, de façon succincte, quatre référentiels et de les comparer. Il s'agit du cadre de référence du TEF, du Cadre européen commun de référence pour les langues (CECR), des Niveaux de compétence en français langue seconde québécois (NCFLS) et des Standards linguistiques canadiens (SLC). D'autres référentiels auraient pu être mentionnés étant donné leur lien étroit avec le CECR (échelles d'ALTE ancrées sur le CECR et échelle Dialang, une application du CECR dans une version multilingue) ou avec le TEF (standards de l'ACTFL, American Council On the Teaching of Foreign Languages).

Tableau 4 – Tableau synoptique de quatre référentiels

| | TEF | CECR | SLC | NCFLS |
|-----------------------------|---|---|--|--|
| Description acronyme | Test d'Évaluation de Français | Cadre européen commun de référence pour les langues | Standards linguistiques canadiens | Niveaux de compétence en français langue seconde |
| Nombre de pages | 10 pages | 191 pages | 186 pages | 45 pages |
| Date d'édition | 1998 | 2001 | 2002 | 1998 |
| Editeur | Chambre de Commerce et d'Industrie de Paris Direction des Relations Internationales de l'Enseignement | Didier Conseil de l'Europe | Centre des niveaux compétence linguistique canadiens | Gouvernement du Québec |
| Mandataire | Chambre de Commerce et d'Industrie de Paris | Conseil de l'Europe Division des langues vivantes, Strasbourg | Citoyenneté et Immigration Canada | Groupe interministériel sur le cadre commun pour la francisation des immigrants adultes |
| Source | CECR ACTFL | Threshold Level Waystage Alte | Canadian Language Benchmarks (CLB) | Canadian Language Benchmarks (CLB) |
| Visées | <ul style="list-style-type: none"> • Evaluation | <ul style="list-style-type: none"> • Apprendre • Enseigner • Evaluer | <ul style="list-style-type: none"> • Normes nationales pour la formation des adultes en FLS • Conception et administration de programmes • Evaluation | <ul style="list-style-type: none"> • Arrimage des programmes • Base pour l'évaluation • Outil de communication • Outil de pilotage |
| Compétences | <ul style="list-style-type: none"> • Compréhension écrite • Compréhension orale • Lexique / structure • Expression orale • Expression écrite | <ul style="list-style-type: none"> • Comprendre • Parler • Écrire | <ul style="list-style-type: none"> • Lire • Écouter • Parler • Écrire | <ul style="list-style-type: none"> • Compréhension écrite • Interaction orale • Production écrite |
| Stades | <ul style="list-style-type: none"> • Élémentaire (0+ - 2)) • Intermédiaire (3-4) • Supérieur (5-6) | <ul style="list-style-type: none"> • Élémentaire (A1-A2) • Indépendant (B1-B2) • Expérimenté (C1-C2) | <ul style="list-style-type: none"> • Élémentaire (1-4) • Intermédiaire (5-8) • Avancé (8-12) | <ul style="list-style-type: none"> • Débutant (1-4) • Intermédiaire (5-8) • Avancé (9-12) |
| Niveaux | 7 niveaux | 6 niveaux | 3 niveaux | 12 niveaux |
| Langue d'origine | Français | Langue de chaque pays européen | Français | Français |
| Traduction | - | oui | Traduction des CLB | - |
| Consultables sur : | http://www.fda.cciip.fr | http://www.coe.int | http://www.language.ca | http://www.immigration-quebec.gouv.qc.ca |

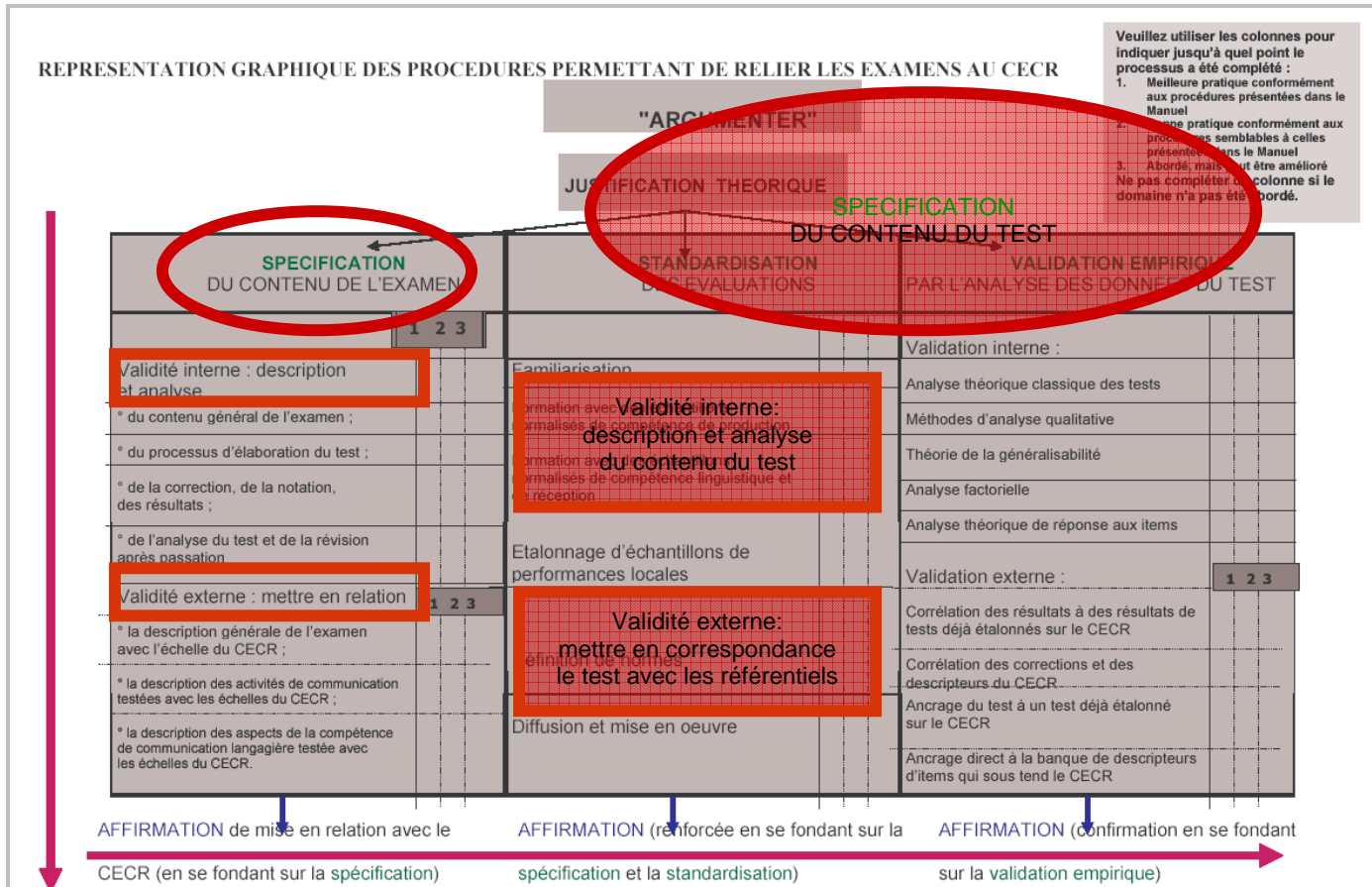
Ce tableau résume bien les ressemblances et les divergences entre référentiels. Par exemple, l'approche de la langue : elle est abordée de façon différente selon qu'il s'agit du TEF, du CECR, des NCFLS ou des SLC : le découpage des compétences diffère au niveau des habiletés productives ou réceptives et la notion de communication est plus ou moins importante.

4. La mise en correspondance du TEF avec les SLC

4.1. Schématisation d'une procédure pour relier les examens au CECR

La méthodologie adoptée pour cette mise en correspondance des spécifications du TEF avec les Standards linguistiques canadiens s'appuie principalement sur une démarche proposée par le Conseil de l'Europe dans le manuel « Relier les examens de langue au cadre européen de référence pour les langues » et dont la procédure est présentée dans le schéma n°3.

Schéma 3 – Représentation graphique des procédures permettant de relier les examens au CECR



Source : Conseil de l'Europe (2003, p 4). Relier les examens de langue au Cadre européen de référence pour les langues: apprendre, enseigner, évaluer.

4.2. Démarche adoptée pour la mise en correspondance du TEF avec les SLC

Démarche

La démarche de recherche adoptée repose

- d'une part sur une étude qualitative relative à la mise en correspondance théorique des objectifs du TEF avec les descripteurs proposés par les Standards linguistiques canadiens,
- et d'autre part sur une analyse quantitative réalisée à partir du classement des prototypes d'items et des unités réelles du TEF.

Corpus étudié

Cette étude a été établie à partir de quatre corpus :

- les spécifications du test ;
- les unités spécifiques du TEF (unité prototypées en termes de compétences et d'objectifs visés) ;
- les unités réelles (items conçus pour les jeux d'épreuves et réalisés à partir des spécifications définis dans l'unité spécifique) ;
- les descriptions de compétences et d'échelles contenues dans les Standards linguistiques canadiens.

Déroulement

Le tableau n°5 retrace les différentes phases suivies pour la mise en correspondance de l'échelle TEF avec les niveaux de compétence des Standards linguistiques canadiens.

Tableau 5 – Présentation des phases et étapes de la mise en correspondance du TEF avec les SLC

| | PHASES | ETAPES |
|---|--|--|
| 1 | Synthèse des SLC par compétence et par niveau | <ul style="list-style-type: none"> • Information et formation de l'équipe TEF aux outils SLC ; • Réalisation de tableaux synoptiques types de la synthèse des SLC ; • Lecture approfondie des SLC et sélection de produits et normes de compétence les plus pertinents en regard du TEF et édition des tableaux synthétiques. |
| 2 | Regroupement des SLC aux points de césure | <ul style="list-style-type: none"> • des SLC aux points de césure CIC ; (<i>agglomérat et écarts*</i>) • Liste de contrôle du TEF. |
| 3 | Mise en correspondance théorique des deux cadres théoriques TEF et SLC | <ul style="list-style-type: none"> • Réalisation d'un tableau synthétique ; • Mise en correspondance théorique des deux cadres. |
| 4 | Mise en correspondance théorique du cadre TEF, des questions TEF et des SLC | <ul style="list-style-type: none"> • Mise en correspondance des deux cadres avec les questions du TEF (objectifs et niveaux associés) ; • Description et comparaisons ; • Analyse et jugement. |
| 5 | Mise en correspondance pratique du cadre TEF, des questions TEF et des SLC | <ul style="list-style-type: none"> • Protocole d'action ; • Compilation de données et du matériel pour l'expertise ; • Classement des unités types et réelles par des experts (analyse des données) ; • Interprétation des résultats. |

Source : Documents CCIP-DRI/E-TEF

La formation des évaluateurs

Une attention toute particulière a été accordée au recrutement et à la formation des experts évaluateurs. Les critères de sélection ont porté principalement sur la compétence des évaluateurs qui devait avoir un minimum de dix années d'expertise de l'enseignement des langues et être spécialisé en linguistique, didactique et évaluation des langues. Ces critères de choix ont été déterminés en fonction des champs disciplinaires que recouvrait cette expérience (Cf. schéma n°1).

Sept experts, les mêmes tout au long de l'expérimentation, ont participé, soit :

- 3 experts extérieurs (évaluation externe) ;
- 1 expert extérieur, mais assurant tout au long de l'année la qualité du test (évaluation mixte) ;
- 3 experts internes (évaluation interne) et les concepteurs du TEF.

Quant à la formation, elle a été dispensée en amont de l'expérimentation à partir de fiches de synthèse des SLC. Les experts avaient également reçus une copie du référentiel SLC de manière à se l'approprier, deux mois avant l'expérimentation.

4.3. Mise en correspondance théorique : l'exemple de la compréhension écrite

Méthodologie

Cette partie correspond aux phases 3 et 4 décrites dans le tableau n°5. La démarche adoptée a consisté en :

- la collecte de données (sélection de descripteurs SLC appropriés au TEF) ;
- l'organisation de ces données (mise en correspondance) ;
- l'interprétation des résultats.

Résultats

La mise en correspondance du cadre TEF et des SLC est présentée par compétence. L'exemple proposé dans le tableau n°6 concerne une unité type de la compréhension écrite codée CEAU10⁺. Ces éléments de description sont extraits de la table de spécification et mis en regard des descripteurs de performance des SLC les plus pertinents pour le TEF.

Après mise en correspondance lors de la phase 3, on constate que l'ensemble des objectifs des questions TEF recouvre une partie des descripteurs de performance relatifs à chaque niveau de compétence des SLC. Les domaines d'application sélectionnés dans la table de spécification du TEF recouvrent bien ceux des SLC et aux mêmes niveaux. L'introduction des SLC insiste particulièrement sur cet aspect : « *Les SLC mettent l'accent sur les tâches reliées à la collectivité, au travail et aux études* » (Standards linguistiques canadiens, 2002, page iv).

Tableau 6 – Mise en correspondance théorique du cadre TEF, des questions TEF et des SLC : la compréhension écrite / lire**Compréhension écrite / lire**

| CADRE THEORIQUE ET QUESTIONS DU TEF | | | CADRE DES SLC | |
|-------------------------------------|-------------|--|------------------------|---|
| Code item TEF | Niveaux TEF | Objectifs des questions du TEF pour le candidat | Niveaux SLC | Descripteurs de performance globale pour l'apprenant |
| CEAU1 | TEF 0+ | Reconnaître les situations élémentaires de la vie courante : - identifier la nature et la fonction d'un document très court de l'environnement quotidien (pancartes, étiquettes ...); - repérer des mots isolés. - 20 mots maximum. | Niveau I Standard 1 | « Peut apparier des illustrations simples et de courtes phrases écrites contenant des mots familiers » (p 28). De 1 à 5 phrases. |

Source : Documents CCIP-DRI/E-TEF

PHASE 3

Après mise en correspondance, lors de la phase 4, des items avec les descripteurs présélectionnés en phase 3, on constate également une bonne concordance, comme l'illustre l'encart n°1. Il y a parfaite adéquation entre le descripteur SLC du tableau n°6 et l'item proposé dans l'encart 1, tant au niveau de l'habileté linguistique que de l'aspect psycholinguistique (longueur des phrases et traitement cognitif).

Encart 1 – Unité réelle du TEF

PHASE 4

Entrez sans sonner

Question 1

On peut lire cette inscription

- A. sur une porte.
- B. sur un téléphone.
- C. sur un ordinateur.
- D. sur un bus.

Source : Documents CCIP-DRI/E-TEF

Les difficultés rencontrées lors de cette mise en correspondance théorique reposent essentiellement sur des formulations floues, ne permettant pas aux évaluateurs de distinguer le degré de distinction entre deux descripteurs de niveaux différents. Par exemple, l'utilisation des verbes « reconnaître » et « comprendre » sont utilisés à des niveaux supérieurs, alors que les spécifications du test ne les introduisent qu'aux niveaux 0+ et 1. D'autre part, les descripteurs relatifs à un contexte culturel particulier ont été systématiquement exclus : à titre d'exemple, dans la compétence parler, le candidat « discute l'actualité canadienne ».

4.4. Mise en correspondance empirique

Protocole d'expérimentation

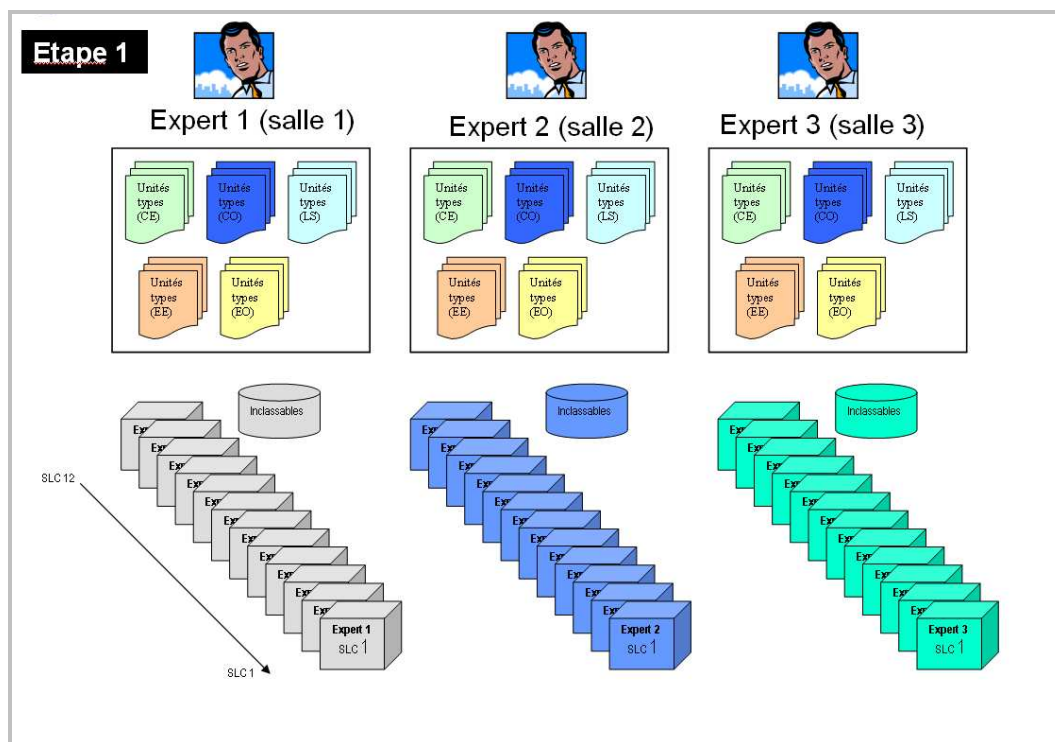
Suite à cette mise en correspondance théorique, une expérimentation a été menée afin d'envisager, de façon plus pragmatique, cette indexation d'échelles de compétence en français langue étrangère ou seconde. Cette partie correspond à la phase 4 décrite dans le tableau n°5.

Pour rappel, la procédure de mise en correspondance des questions du TEF et des SLC a été assurée à la fois par un groupe interne (équipe pédagogique TEF et concepteurs TEF) et une équipe d'experts externes : le Service de Pédagogie Expérimentale de l'Université de Liège (SPE-ULG), un chargé de mission, le Groupe d'Études et de Recherche pour le Français Langue Internationale (GERFLINT).

En dehors du travail d'analyse du référentiel SLC, une mise à l'épreuve du classement des questions TEF a été réalisée sur deux jours consécutifs, au siège de la DRI/E-CCIP. La procédure suivante a été adoptée :

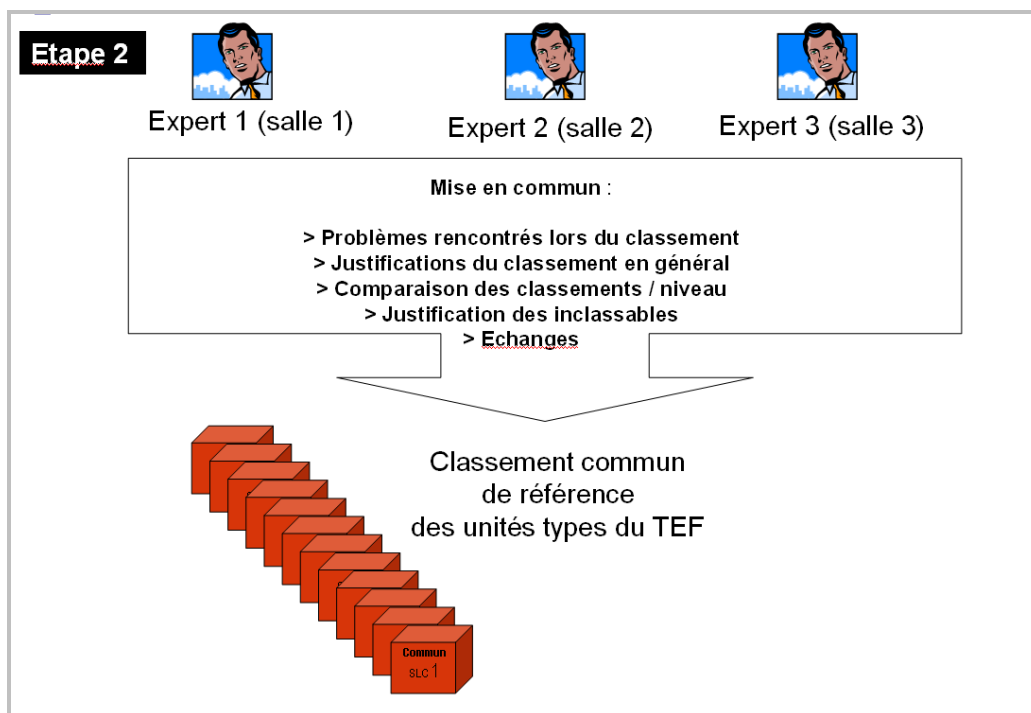
Le premier jour, la procédure de validation a été présentée aux trois experts externes, préalablement formés aux SLC. Une première expérimentation a alors été mise en œuvre durant laquelle chaque expert, isolé dans un local différent et face à 12 bacs correspondant aux 12 niveaux de référence des SLC, devait ranger les unités types du TEF couvrant les trois domaines de compétence suivants : compréhension écrite, compréhension orale et lexique/structure. Le schéma n°4 illustre cette partie de l'expérimentation.

Schéma 4 – Protocole d'action pour le classement des unités types



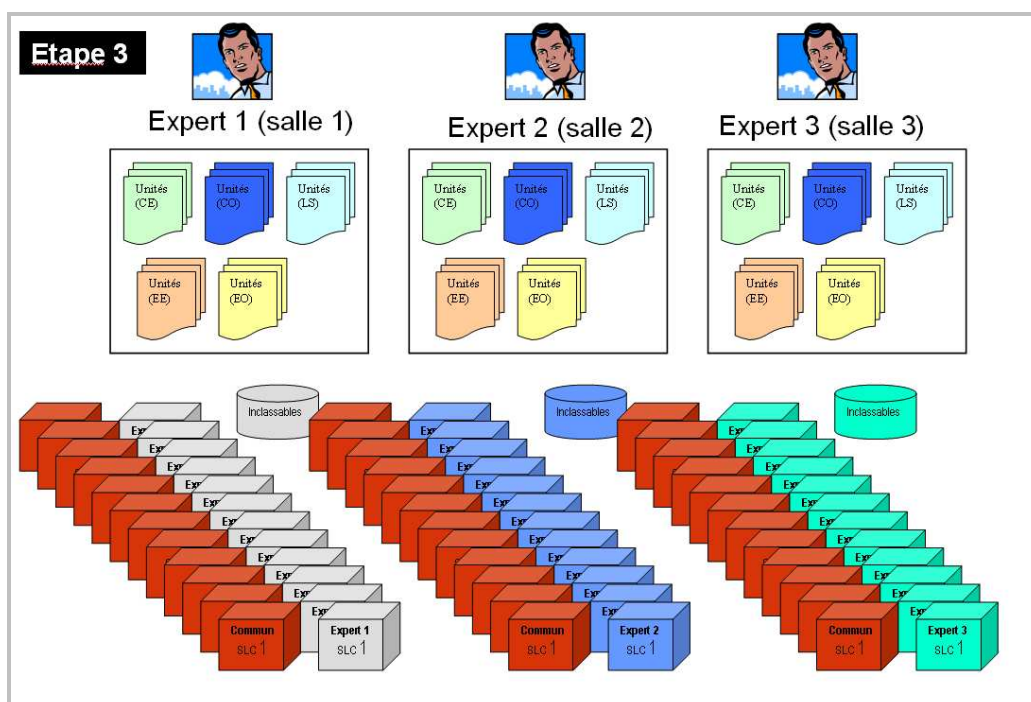
Après la première expérimentation, les résultats des 3 experts ont été confrontés au classement réalisé par l'équipe interne de manière à obtenir une concordance optimale et commune des 4 classements, comme le montre le schéma n°5.

Schéma 5 – Protocole d'action pour le classement commun de référence des unités types



Le deuxième jour a été consacré au classement des unités réelles. Il a été réalisé à partir du classement commun effectué la veille et obtenu sur la base d'un commun accord entre les 3 experts et l'équipe interne. 294 unités, extraites de 4 jeux réels, déjà utilisés, ont alors été classées par les trois experts dans un même temps, mais dans une salle différente pour éviter l'effet de contamination.

Schéma 6 – Protocole d'action pour le classement des unités réelles.



| |
|-----------|
| Résultats |
|-----------|

Les taux de concordance des classements des 294 unités sont présentés dans les trois tableaux n°7, n°8 et n°9. La valeur 0 signifie que les trois experts et le groupe interne ne parviennent pas, pour une unité donnée, à fournir une réponse identique en termes de classement ; la valeur 2 signifie qu'au moins deux experts ou un expert et le groupe interne fournissent un classement identique de l'unité¹, la valeur trois que trois experts ou deux experts et le groupe interne fournissent une réponse identique ; la valeur 4, que tous les experts et le groupe interne fournissent une réponse identique. Les résultats présentés sont agrégés au niveau du système de points accordés par CIC (cf. tableau 3).

Tableau 7 - Accords entre les 3 experts et le positionnement des différentes unités de compréhension écrite (CE) par le groupe interne dans le système de points (0, 1, 2 ou 4) de CIC.

| Accords | Fréquence | % | Fréquence cumulée | % cumulé |
|---------|-----------|-------|-------------------|----------|
| 4 | 20 | 27,78 | 72 | 100,00 |
| 3 | 24 | 33,33 | 52 | 72,22 |
| 2 | 28 | 38,89 | 28 | 38,89 |
| 0 | 0 | 0 | 0 | 0 |

Tableau 8 - Accords entre les 3 experts et le positionnement des différentes unités de compréhension orale (CO) par le groupe interne dans le système de points (0, 1, 2 ou 4) de CIC.

| Accords | Fréquence | % | Fréquence cumulée | % cumulé |
|---------|-----------|-------|-------------------|----------|
| 4 | 32 | 27,35 | 117 | 100,00 |
| 3 | 48 | 41,03 | 85 | 72,65 |
| 2 | 37 | 31,62 | 37 | 31,62 |
| 0 | 0 | 0 | 0 | 0 |

Tableau 9 - Accords entre les 3 experts et le positionnement des différentes unités de lexique et structure (LS) par le groupe interne dans le système de points (0, 1, 2 ou 4) de CIC.

| Accords | Fréquence | % | Fréquence cumulée | % cumulé |
|---------|-----------|-------|-------------------|----------|
| 4 | 28 | 26,67 | 105 | 100,00 |
| 3 | 36 | 34,29 | 77 | 73,33 |
| 2 | 40 | 38,10 | 41 | 39,05 |
| 0 | 1 | 0,95 | 1 | 0,95 |

Les taux d'accord impliquant trois ou quatre experts par groupe interne dépassent 60% pour la compréhension écrite et le lexique/structure, pour atteindre près de 70% en compréhension orale ; une seule unité a conduit à un problème important de classification, en lexique/structure.

¹ La valeur 2 est aussi accordée lorsque 2 groupes distincts de deux experts et/ou groupe interne s'accordent chacun sur un classement identique.

Bibliographie

Psychométrie

- Demeuse, M., Henry, G. (2004). L'analyse classique d'items. In M. Demeuse (éd.). *Introduction aux théories et aux méthodes de la mesure en sciences psychologiques et en sciences de l'éducation*. Liège (Belgique) : Les Editions de l'Université de Liège.
- Laurencelle, L. (1998). *Théories et techniques de la mesure instrumentale*. Sainte-Foy : Presses de l'Université du Québec.
- Leclercq, D., Denis, B., Jans, V., Poumay, M., Gilles, J.L. (1998). Chapitre 7. L'amphithéâtre électronique. Une application : le LQRT-SAFE. In D. Leclercq (éd.). *Pour une pédagogie universitaire de qualité*. Sprimont (Belgique) : Pierre Mardaga éditeur.

Linguistique et didactique des langues

- Baddeley, A. : *La mémoire humaine, théorie et pratique* (chapitres sur l'effet de récence et la mémoire), Presses Universitaires de Grenoble, 1993.
- Citoyenneté et Immigration Canada. *Standards Linguistiques Canadiens 2002*.
- Conseil de l'Europe, Division des Politiques Linguistiques (2003). *Relier les examens de langue au Cadre européen de référence pour les langues : apprendre, enseigner, évaluer (CECR) – Manuel avant-projet*. Strasbourg : DGIV/EDU/LANG.
- Cuq J.-P., Grucq I. (2003). *Cours de didactique du français langue étrangère et seconde*. Collection FLE. Presses Universitaires de Grenoble. Chapitre II.
- Lussier, D. (1992). *Évaluer les apprentissages dans une approche communicative*. Collection « f » autoformation. Paris : Editions Hachette.
- Mothe, J.-C. (1975). *L'évaluation par les tests dans la classe de français*. Paris : Editions Hachette / Larousse.
- Scovel, T. *Psycholinguistics*. Series Editors H.G. Widdowson. Oxford University Press.

Économie des langues

- Delamotte, E. (1999). *Le commerce des langues*. Collection CREDIF Essais. Paris : Editions Didier.