



HAL
open science

Un sujet CLES conçu pour le niveau B2 peut-il permettre de valider un niveau B1 ?

Claire Tardieu

► **To cite this version:**

Claire Tardieu. Un sujet CLES conçu pour le niveau B2 peut-il permettre de valider un niveau B1 ?. Monica Masperi, Cristiana Cervini, Yves Bardière (eds. Évaluation des acquisitions langagières : du formatif au certificatif, 32, , pp.A297-A323, 2021, mediAzioni. hal-03832182

HAL Id: hal-03832182

<https://univ-sorbonne-nouvelle.hal.science/hal-03832182>

Submitted on 6 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Un sujet CLES conçu pour le niveau B2 peut-il permettre de valider un niveau B1 ?

Version auteur

CLAIRE TARDIEU – UNIVERSITÉ SORBONNE-NOUVELLE
PRISMES – SESYLIA (EA4398)
claire.tardieu@sorbonne-nouvelle.fr

Résumé

Cet article concerne le Certificat de Compétences en Langues de l'Enseignement Supérieur (CLES) qui s'inscrit dans la démarche actionnelle et emprunte la forme holistique du scénario. Le CLES B2 correspond au niveau B2 du CECRL, niveau minimum normalement requis en France à l'issue de la licence. Or, il s'avère que le CLES B2 qui évalue la compréhension orale, la compréhension écrite, l'expression écrite et l'interaction orale a la particularité de ne pas pratiquer la compensation entre les compétences, créant ainsi une difficulté supplémentaire dans l'attractivité du certificat comparé à d'autres qui évaluent un nombre plus restreint de compétences ou proposent un score final. Une des solutions consisterait à délivrer un niveau B1 ou B1+ à des candidats qui auraient échoué au niveau B2. Est-il possible de produire un certificat à B1, voire B1+ à partir de sujets qui ont été créés pour le niveau B2 uniquement ?

Après un rappel des enjeux actuels du testing, nous présenterons la méthodologie de notre étude qui fait appel à la grille CEF-ESTIM (TARDIEU, HILDEN, REICHERT, LEHMANN 2009) pour déterminer ce qui relève de B1 ou de B2 dans un sujet CLES 2. Enfin, certaines recommandations pourront être faites pour une utilisation future.

Mots clés : évaluation en langues ; CLES ; CEF-ESTIM ; testing ; enseignement supérieur

Abstract

This article concerns the language competence certificate for higher Education (CLES) which is part of the action-oriented approach and is designed in the

holistic form of a scenario. CLES B2 corresponds to level B2 of the CEFR, the minimum level normally required in France at the end of the Licence (3-year undergraduate degree). However, it turns out that the CLES B2 which assesses oral comprehension, written comprehension, written expression and oral interaction has the particularity of not allowing the compensation between the skills, thus creating an additional difficulty in the attractiveness of the certificate compared to others which assess a smaller number of skills or offer a final score. One of the solutions would be to issue a B1 or B1 + level to candidates who have failed at B2 level. Thus, our research aims to find out whether it is possible to produce a B1 or B1+ certificate from CLES tests that have been created for B2 level only.

After providing a brief overview of the contemporary challenges of language testing, we will present the methodology applied to our study which uses the CEF-ESTIM grid (TARDIEU, Hilden, Reichert, Lehmann 2009) to determine what falls under B1 or B2 in a test at B2.

Finally, some recommendations will be made for future uses.

Key words: language assessment; CLES; CEF-ESTIM; testing; higher education.

Introduction

A l'heure où sont prises des mesures susceptibles d'entériner la suprématie des tests privés dispensés par les multinationales de l'évaluation (Mac Gaw, 2017) dans l'enseignement supérieur public français, le présent article vise à s'interroger sur un certificat public dont les caractéristiques éthiques et scientifiques présentent un intérêt certain. Le Certificat de Compétence en

Langues de l'Enseignement Supérieur créé en 2000, modifié en 2007 et en 2016, doit s'efforcer de gagner le soutien politique qui lui fait encore défaut malgré une large implantation sur le territoire national et la création du réseau européen NULTE (Network of University Language Testers in Europe)¹. Le succès croissant du CLES² auprès des étudiants s'explique par sa qualité mais aussi par son coût moindre (CLES B1 : 80 € max, CLES B2 : 100€ max, CLES C1 : 150 € max) et son accès souvent gratuit (offert par les universités). À titre indicatif (2021), le TOEIC coûte 132€ à quoi s'ajoutent 16€ pour l'obtention du certificat officiel ; le TOEFL, 220€ et l'IELTS, 200€). C'est également un certificat dont l'approche holistique garantit un niveau dans un ensemble d'activités langagières et qui n'a rien à envier aux autres certificats, en termes de qualité scientifique et d'intégrité. Pour autant, c'est peut-être sa trop grande rigueur qui le dessert, face à des pratiques certificatives moins exigeantes (absence d'évaluation des compétences orales, ou certifications de type *achievement test* (Bachman, 1990) au terme d'une formation ciblée, ou encore de type QCM ne permettant pas d'évaluer la capacité à communiquer réellement). Le comité scientifique et pédagogique du CLES a proposé plusieurs solutions pour mener à bien la récente réforme et rendre ce certificat plus attractif. L'une d'entre elle est de se donner la possibilité, en cas d'échec global d'un candidat au test (absence de compensation entre les compétences) d'attester sa réussite au niveau immédiatement inférieur (par exemple, B1 au lieu de B2). Les nouveaux sujets seront construits pour permettre une telle différenciation interne. Mais qu'en est-il des sujets existants dans la banque-ressource des centres CLES ? Seront-ils rapidement jugés obsolètes ? Peut-on leur donner une seconde vie ? Notre démarche s'inscrit donc en aval de la conception des sujets. La question posée est la suivante : Est-il possible de produire un résultat différencié (B2 ou B1) à partir de sujets qui ont été créés pour le niveau B2 uniquement ?

¹ Officialisé le 08/09/2018, NULTE réunit six pays européens (Allemagne, Espagne, France, Pologne, République tchèque, Royaume-Uni, Slovaquie), soit 250 universités en Europe et 300 000 étudiants

² À ce jour, le CLES dispose de 58 centres de passation répartis en 11 pôles sur toute la France et La Réunion. Abou Dhabi, Nouméa. Il concerne environ 25000 étudiants par an.

Au plan méthodologique, nous proposons d'utiliser la grille CEF-ESTIM³ (TARDIEU et al. 2009), disponible en [PDF](#) sur le site du CELV (Centre Européen des Langues Vivantes), pour déterminer ce qui relève de B1 ou de B2 dans un même sujet. La grille CEF-ESTIM s'appuie en particulier sur la Dutch CEF Grid (Alderson et al., 2004), permettant d'estimer la difficulté d'un texte en fonction de 11 dimensions repérées dans les descripteurs du CECRL (2001) (voir annexe 6). La grille CEF-ESTIM ajoute la possibilité d'estimer a priori le niveau requis par les activités proposées sur un texte ainsi que les stratégies requises, toujours à partir des descripteurs du CECRL. Elle a été conçue pour les enseignants qui auraient besoin d'aide dans la planification de leurs cours et dans la mise en œuvre du CECRL dans l'enseignement. C'est un outil simple et utile permettant de partager son expérience et ses idées avec d'autres enseignants. Elle offre en outre la possibilité d'archiver des ressources pédagogiques, facilitant ainsi, pour une équipe d'enseignants, le recueil d'une banque d'activités liées aux niveaux du CECRL, ce grâce à une procédure de validation qualitative.

Nous faisons l'hypothèse que la grille CEF-ESTIM, qui n'est pas conçue pour le *testing*, peut constituer un bon outil d'estimation empirique permettant de déterminer parmi les sujets CLES déjà créés à B2 ceux qui sont susceptibles de fournir simultanément des informations sur le niveau B1 d'un candidat. Elle permettrait donc d'opérer un tri préliminaire à une validation ultérieure avec des candidats réels.

En premier lieu nous rappellerons quelle est la situation du CLES au sein du *testing*, à la fois au plan économique et politique et au plan socio-pédagogique. Puis nous expliciterons la méthodologie d'analyse utilisée sur un sujet CLES B2 disponible en ligne (2019). Enfin, il s'agira d'étudier de près les résultats et de faire des préconisations pour d'autres sujets potentiels.

³ Acronyme franco-anglais composé de *CEF* pour Common European Framework et *Estim* pour estimation.

1. La situation du CLES au sein du Testing

1.1. Aspects politico-économiques

Le *testing* en langue, et particulièrement en anglais, représente un marché phénoménal sans que la question de la valeur pour l'être humain soit réellement posée, comme le souligne Shohamy : « Test results are endowed with economic value and thus can provide a means for controlling knowledge » (Shohamy, 2007, p. 374, cité par Mac Gaw, 2017, p. 16). Et, elle poursuit en dénonçant l'utilisation détournée qui est faite de l'outil certificatif et la main-mise qui en découle sur le système éducatif : *“Tests, originally developed for democratizing purposes, have been utilized in undemocratic acts as they have become tools for centralized authority for power and control and for the manipulation of educational systems”* (*ibid.* : 387).

En effet, les tests génèrent des intérêts financiers colossaux et adoptent des stratégies commerciales très invasives sur les marchés nationaux sans forcément satisfaire aux exigences des usagers. Par exemple, le TOEIC ne permet pas d'évaluer les compétences d'expression orale, la plupart des candidats ne passant que la version de base (*listening and reading*)⁴. Certaines écoles⁵ lui préfèrent le TOEFL ou le IELTS ou encore le CLES.

Aux États-Unis, la société Pearson, qualifiée de « Godzilla de l'évaluation » (métaphore tirée de l'article du 21.01.2015, publié par Fortune et intitulé : *“Everybody hates Pearson”*) réalise un chiffre d'affaires de neuf milliards de dollars par an (Mac Gaw, 2017, p. 187).

Les deux principales sociétés qui se partagent le marché mondial des tests pour l'anglais actuellement sont ETS (*Educational Testing Service*) pour les États-Unis et *Cambridge Assessment* pour le Royaume-Uni. Cambridge Assessment est une organisation qui se définit comme « à but non lucratif ». Son organe

⁴ L'épreuve d'expression orale se déroule face à ordinateur et ne permet pas l'interaction avec un interlocuteur.

⁵ À titre d'exemple, l'école des Mines ParisTech a abandonné le TOEIC.

Cambridge English language assessment (CE) est présent dans 130 pays au monde et réalise un chiffre d'affaires considérable (MacGaw, 162).

On comprend aisément que la puissance de ces certificats ne se limite pas à la phase de *testing* mais influe véritablement sur l'enseignement, la production de manuels en amont, bref sur la politique éducative des langues en général, et ce, au niveau mondial.

Qu'en est-il de la qualité intrinsèque de ces certificats ?

1.2. Aspects socio-pédagogiques

Selon la classification de Spolsky (1981), on note trois générations de test : pré-scientifique, structuraliste psychométrique et sociolinguistique-intégrative. Attachons-nous à la troisième génération.

1.2.1. Approche sociolinguistique-intégrative

On peut citer le TOEFL comme exemple de test de troisième génération qui parvient à associer l'approche structuro-psychométrique précédente utilisant les items discrets et l'approche intégrée introduisant les *skills*⁶ (dont l'expression orale).

Pour Horner, « ce type d'évaluation paraît à la fois valide, fiable, et faisable au sens où « les examinateurs ne disposent que d'un temps limité, ne voient qu'un échantillon limité de performances et qu'il y a des limites au nombre et à la nature des catégories qu'ils peuvent manipuler comme critères » (Horner, 2010 : 17).

Mais cette approche se révèle à son tour limitée, à mesure qu'évolue la conception de la communication et de l'apprentissage des langues qui en découle.

⁶ Cette approche intégrée demeure toutefois modeste avec 4 tâches sur 6 en production orale, 1 sur 2 en production écrite, mais aucune intégration entre les quatre activités langagières (CE, CO, EO, EE) et aucune thématique commune.

1.2.2. Les limites de l'approche sociolinguistique intégrative

Bachman (1990 : 82) considère qu'un test alliant activités langagières et composants linguistiques (items discrets) ne permet pas de déterminer la nature du lien entre les deux types. D'autre part, tous les aspects de la compétence à communiquer langagièrement ne sont pas mesurés. On peut citer à cet égard la synthèse réalisée par Springer (2000) dans sa « marguerite de la communication » (reprenant l'évolution de Chomsky (1965) à Canale et Swain (1980) en passant par Hymes (1972) et bien sûr les travaux de Oller (1976) et Bachman (1990). Selon Oller, la compétence linguistique n'est pas la somme d'éléments séparés. Il forge le concept d'une compétence holistique : « The Unitary Competence Hypothesis ». Ainsi, la production écrite ou orale n'est pas séparée de la réception, car le locuteur opère simultanément le contrôle de sa production pour s'assurer d'être compris. De même, en phase de réception, l'auditeur ou le lecteur formulent en permanence des hypothèses sur la suite du message.

Enfin, il apparaît que les tests formalisés ne peuvent mesurer la dimension socio-pragmatique de la communication, ni les dimensions cognitive et intellectuelle et que leur échappe ainsi la capacité de communication réelle chère à Bachman. Bachman distingue en effet la compétence à communiquer langagièrement (*communicative language ability*) de la communication véritable (*actual communication*). La première incluant une dimension liée aux savoirs - *language knowledge dimension* - comprenant elle-même des compétences langagières organisationnelles et pragmatiques, des compétences stratégiques et des mécanismes psycho-physiologiques ; la seconde se caractérisant par sa dimension « d'utilisation réelle » (Bachman, 2007, p. 54).

1.2.3. L'approche sociolinguistique intégrative de type actionnel ou « scénarisé »

Afin de compléter la classification de Spolsky, il paraît possible d'ajouter désormais une quatrième approche, ne faisant pas appel à la psychométrie, que l'on pourrait appeler : « l'approche sociolinguistique intégrative de type actionnel ou approche par scénario ». Cette approche s'appuie sur une

conception socio-pragmatique de la communication appliquée à l'enseignement dont on trouve les fondements chez Widdowson, Nunan, Ellis, entre autres. Elle est connue sous le nom de Scenario-Based Language Assessment (Purpura, 2016). Le Certificat en Langues de l'Enseignement Supérieur (CLES) ou le Diplôme de Compétences en Langue (DCL), appartiennent à cette catégorie. Les épreuves se présentent sous forme de scénarios et intègrent en particulier l'interaction orale.

1.2.4. Le CLES

Le Certificat de Compétence en Langues de l'Enseignement Supérieur créé en 2000, modifié en 2007 et en 2016, constitue « une évaluation de la compétence de communication de l'évalué en tant que sujet communiquant » (Bourguignon & Narcy-Combes, 2003 : 2).

Inauguré en 2000⁷, révisé en 2007⁸ et inscrit, momentanément, avec le C2i2e, comme un passage obligé des concours d'enseignement, en quoi consiste exactement le Certificat de Compétences en Langues de l'Enseignement Supérieur ?

« Le CLES se décline en trois niveaux de qualification (1er, 2ème, 3ème degrés), définis respectivement par référence aux niveaux B1, B2 et C1 du Conseil de l'Europe. Le 1er degré se situe au-delà du niveau d'exigence moyen de la langue vivante 1 au baccalauréat⁹. »

⁷ B.O.E.N. n° 25 du 29 juin 2000

<http://www.education.gouv.fr/bo/2000/25/sup.htm>

⁸ Voir B.O.E.N. n° 20 du 17 mai 2007 : <http://www.education.gouv.fr/bo/2007/20/MENS0700723A.htm>

⁹ B.O.E.N. n° 20 du 17 mai 2007

Scenario type B2

Durée totale : 3h - Simulation d'une situation réaliste autour d'une thématique générale et de documents pour dégager une problématique qui conduira à une négociation

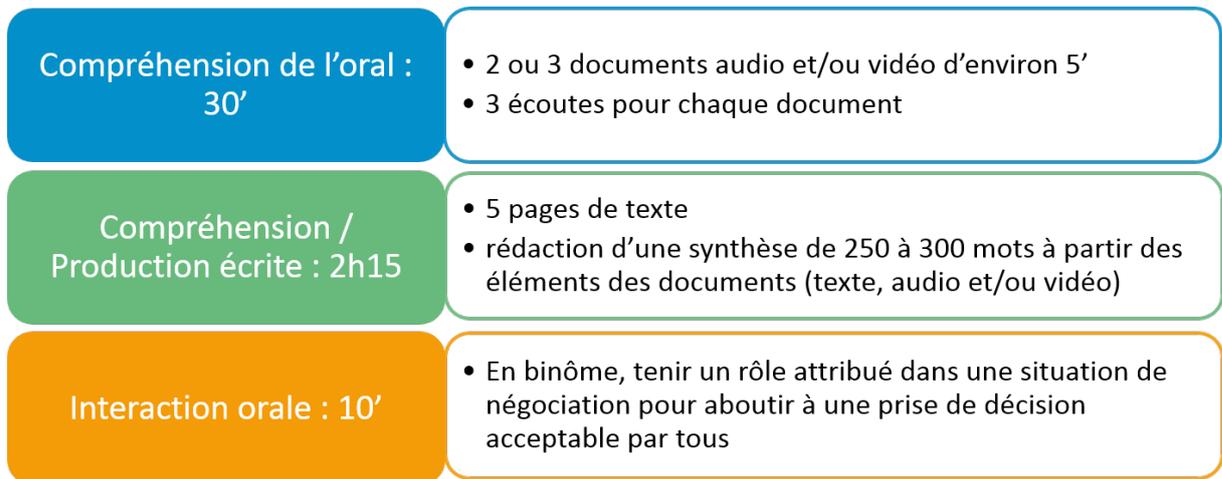


Figure 1 : Métastructure d'un scénario CLES B2¹⁰.

Scenario type B1

Durée totale : 2h - Simulation d'une situation réaliste à l'étranger

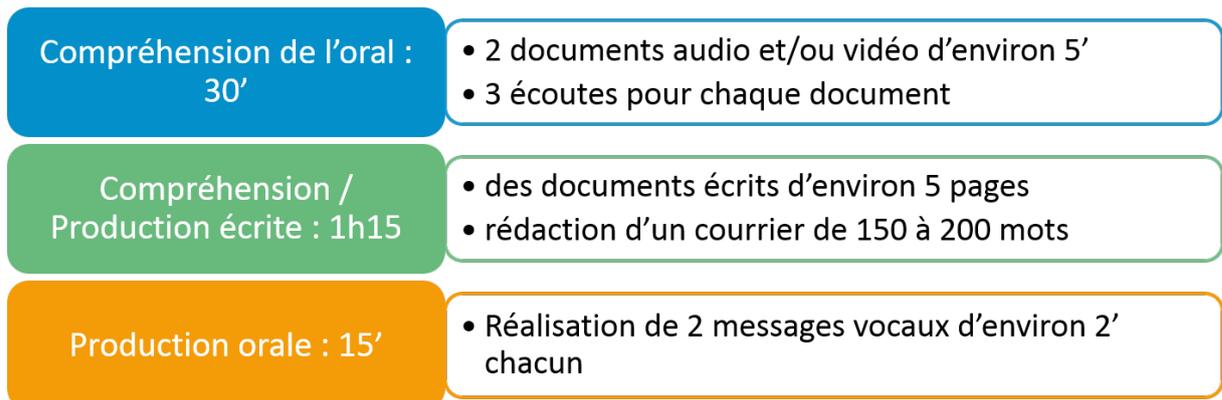


Figure 2 : Métastructure d'un scénario CLES B1¹¹.

¹⁰ [Épreuves - Certificat de compétences en langues de l'enseignement supérieur \(certification-cles.fr\)](http://epreuves-certificat-de-compétences-en-langues-de-l'enseignement-supérieur/certification-cles.fr)

¹¹ *ibid.*

Si l'on compare les deux métastructures, on remarque que la différence se situe essentiellement dans la production écrite : un courrier de 150 à 200 mots pour le CLES B1 contre une synthèse de 250 à 300 mots pour le niveau B2 ; et dans le type de production orale demandée : deux productions orales en continu de 2' chacune pour le CLES B1 contre une interaction orale entre pairs de 10' visant l'atteinte d'un consensus. Le CLES B2 emprunte la forme d'un scénario qui confronte les candidats à un problème à la résolution duquel concourent les différentes activités de lecture, d'écoute, de prise de parole en interaction et d'écriture. La structure du CLES B1, de par sa forme même, semble un peu moins exigeante en termes de capacité de synthèse.

En réalité, le CLES B1 s'appuie sur la même finalité de résolution de problème mais dans un contexte beaucoup plus concret et moins sociétal. Ainsi, les deux niveaux répondent-ils à la volonté des concepteurs du CLES de lier les parties de l'évaluation grâce à une thématique filée tout au long de l'épreuve par les différents supports de compréhension et de production. Est suggérée ici une forme d'homologie de l'évaluation et de l'apprentissage. Si l'évaluation a tendance à formater l'apprentissage en amont, phénomène connu sous le nom de *washback effect* (Heaton, 1990), alors l'intérêt du CLES est évident. Il favorise l'implantation de méthodes d'enseignement fondées sur l'approche actionnelle.

On peut citer ici les différentes typologies de tâches non différenciables entre apprentissage et évaluation. L'une de ces typologies recensées par Ellis (2003) est la classification cognitive de Prabhu. Prabhu (1987) distingue trois types de tâches fondées sur la capacité cognitive impliquée : des activités proposant un déficit d'information, un déficit de raisonnement ou un déficit d'opinion. On voit bien comment le CLES peut s'emparer de ce genre de typologie en particulier au niveau 2 lorsque deux étudiants sont évalués en interaction. (Prabhu, 1987, p. 46, cité par Ellis, 2003, p. 213).

Ellis fait référence à un autre type de classification qui peut présider à la conception d'un projet pédagogique ou d'un scénario évaluatif, celle, psycholinguistique de Pica, Kanagy, and Falodun (1993), fondée sur les catégories interactionnelles : la relation entre les interactants (qui détient

l'information, qui la recherche, etc.) ; les exigences de l'interaction (l'information doit-elle être recherchée et fournie obligatoirement par les interactants ?) ; le but de l'interaction : l'accord ou le désaccord ; enfin, les résultats envisageables : une seule issue ou plusieurs issues possibles. (Ellis, 2003, p. 215).

Ainsi, les épreuves de CLES B2 requièrent souvent que les partenaires de l'interaction parviennent à un consensus à partir d'un jeu de rôle imposé sur un sujet de controverse complexe.

Ellis lui-même propose un cadre général en quatre points pour l'élaboration d'un cours s'appuyant sur l'approche par les tâches :

L'input, soit la nature des données fournies pour l'accomplissement de la tâche. Dans le cas du CLES, une thématique générale est proposée à laquelle tous les supports écrits et oraux renvoient. Il s'agit d'un input unifié ;

Les conditions, soit la manière dont l'information est présentée aux apprenants et la façon dont elle est utilisée. Les épreuves de CLES proposent des supports multimodaux ;

Les processus sollicités, soit la nature des opérations cognitives en jeu et le type de discours requis par la tâche (informatif, explicatif, narratif, argumentatif). Le fait que le CLES soit une épreuve composite favorise la multiplicité des sollicitations tant au plan des opérations cognitives que des types de discours ;

Les résultats, soit la nature du produit fini résultant de l'accomplissement de la tâche (Ellis, 2003 : 217). Le CLES requiert que l'on parvienne concrètement à une solution sous forme écrite et orale.

On discerne un fort potentiel d'homologie entre le scénario d'apprentissage et le CLES et, d'une certaine manière, entre évaluation formative et évaluation sommative (Springer 2001), le scénario évaluatif jouant sur les deux tableaux.

Pour résumer, le CLES appartient à une quatrième génération de tests qui favorise en amont un apprentissage proche de la vie réelle sollicitant des compétences langagières intégrées à des processus cognitifs et discursifs repérables et susceptibles d'être développés. Même si les compétences

demeurent au final évaluées séparément comme dans le cas de la précédente approche des *four skills*, le gain est notable en termes de cohérence et même de cohésion car les activités d'expression écrite (EE) et d'expression orale en interaction (EOI) s'appuient sur l'ensemble du dossier.

Sur un plan éthique et déontologique, en se démarquant de l'approche psychométrique décriée par certains (Huver 2010, Springer & Huver 2011), le CLES semble meilleur, du moins en termes de validité (sinon de fiabilité et d'équité), que les autres tests, à la fois pour les institutions et pour les candidats, car il rend véritablement compte des capacités réelles en langue dans des situations plausibles qui font pleinement sens. Scientifiquement parlant, la direction du CLES s'est doté d'un conseil scientifique international, et possède des outils performants de formation pour la conception des sujets et la correction des tests, notamment des grilles élaborées sous la supervision du Conseil scientifique. Pour autant, commercialement parlant, le CLES ne peut pas rivaliser avec les géants anglais ou américains¹². La réforme de 2016 (arrêté du 4 novembre 2016) qui élargit le CLES à la langue française et préconise la possibilité d'attribuer un niveau B1 aux candidats recalés au niveau B2 pourrait véritablement rendre le CLES intrinsèquement attractif.

L'idée est ainsi de valider plus de candidats et de modifier la réputation du CLES jugé parfois trop difficile.

Rappel de la question :

Un sujet CLES conçu pour le niveau B2 peut-il permettre de valider un niveau B1 ?

=> Est-il possible de produire une évaluation différenciée (B2 ou B1) à partir de sujets qui ont été créés pour le niveau B2 uniquement ?

¹² Le CLES est à but non lucratif et dispose de 25 k€ de fonctionnement, alors qu'une société comme ETS Global [éditeur des tests Toefl et Toeic] totalise près de 1 Md \$ de chiffre d'affaires (source : Yves Bardière, 2020)

2. Cadre méthodologique

La méthodologie utilisée sera la suivante : Sur un sujet B2 public (téléchargé sur le site du CLES), on utilisera la section B de la grille CEF-ESTIM concernant les supports vidéo et textuels pour s'assurer qu'ils correspondent aux spécifications de B2 et on recherchera d'éventuelles caractéristiques de B1 (*The Manual*, 2011). Ensuite on utilisera les sections C et D pour estimer le niveau des activités proposées aux candidats et des stratégies de communication à mettre en œuvre pour vérifier si certaines tâches pourraient être réalisées avec succès par des candidats au niveau B1. Enfin, au total, le dossier aura été visé par trois estimatrices différentes.

S'il se vérifie qu'avec un outil simple comme la grille CEF-ESTIM on peut opérer cette différenciation au sein d'un sujet, alors, une procédure scientifique de validation d'autres sujets à B2 pourra être envisagée. Plusieurs possibilités se présentent :

- le dossier CLES analysé est hétérogène : niveau B2 global mais niveau B2//B1 pour les supports et/ou les activités, ou au contraire niveau B2/C1 ;
- le niveau du dossier (supports + activités) est homogène à B2 ;
- la grille CEF-ESTIM n'est pas adaptée à ce type de différenciation au sein d'un même dossier.

2.1. Le sujet CLES B2, support de l'analyse :

Le sujet choisi est disponible en ligne sur le site du CLES :

Il porte sur la problématique de la contrefaçon. Il est composé d'un certain nombre de documents (voir annexes 1, 2, 3, 4 et 5) :

- Document vidéo 1 : Counterfeit goods (2mn 54) (annexe 1)
- Document vidéo 2 : Seized counterfeit goods sent to Haïti in relief effort (2mn 13) (annexe 2)
- Dossier documentaire (annexe 3)

Text 1: Fake goods are proliferating to the alarm of companies and governments

Text 2: Brand piracy: faking it can be good

Text 3: Are donations of counterfeit goods to charities socially acceptable?

Text 4: Facing counterfeiting crackdown, Beijing vendors fight back

-Document de travail du candidat (annexe 4)

-Corrigé (annexe 5)

Tâche écrite :

Outre l'activité de compréhension orale (instrumentale également pour les tâches de production), on trouve ces deux énoncés : « Vous rédigerez un compte rendu des problèmes posés et des solutions envisagées en organisant les informations essentielles contenues dans les documents vidéo et écrits. »

« La chambre économique de New York vous a commandé un rapport sur la contrefaçon. »

Tâche orale : « Vous participerez à une discussion contradictoire avec un(e) autre candidat(e) sur le thème de la contrefaçon. Au cours de cette discussion, vous devrez présenter et défendre la position qui vous sera attribuée par tirage au sort et aboutir à un compromis. Vous réinvestirez les informations fournies par les documents qui ne seront plus à votre disposition pour cette activité. »

On peut noter que le test fait appel à des tâches fondées sur le déficit d'information (tâche de production écrite) et sur le déficit d'opinion (tâche d'interaction orale), deux des trois types de la classification cognitive de Prabhu (le troisième étant le déficit de raisonnement) (Prabhu, 1987, p. 46, cité par Ellis, 2003, p. 213). Il répond aussi à la classification psycholinguistique de Pica, Kanagy, and Falodun (1993) : l'information doit être recherchée et fournie obligatoirement par les interactants et le but de l'interaction est de parvenir à un accord (Ellis, 2003, p. 215).

Si l'on raisonne en termes de métastructure, un sujet similaire à B1 (outre des supports plus simples) proposerait la rédaction d'un courrier de 150 à 200 mots en tâche écrite et en tâche orale, l'enregistrement de deux messages vocaux, en rapport avec la thématique. Cependant, la question de recherche s'intéresse moins à la forme de l'évaluation elle-même qu'à la nature du dossier proposé. Dans quelle mesure la réussite ou l'échec à certaines questions ou activités permettent-ils de délivrer un niveau de compétence immédiatement inférieur ? Pour répondre à cette question, nous devons disposer d'un outil qui donne une estimation du niveau de compétence requis pour réussir la tâche (combinaison des supports, des activités et des stratégies). Notre hypothèse est que la grille CEF-ESTIM pourrait fournir un tel outil.

2.2. Méthodologie

2.2.1. La grille CEF ESTIM

La grille CEF-ESTIM, disponible sur le [site](#) du CELV (Centre Européen des Langues Vivantes) aide les enseignants à faire le lien, dans leurs activités de classe, avec les niveaux du Cadre européen commun de référence pour les langues (CECR). Cette grille vient en complément du projet DUTCH CECR GRID (2003-2004) dirigé par J. Charles Alderson (Université de Lancaster), qui proposait un outil en ligne aux évaluateurs et concepteurs d'examens. La grille CEF-ESTIM a une perspective plus large et vise une utilisation productive et interactive de la langue en classe.

La grille CEF-ESTIM a été conçue pour aider les enseignants dans la planification de leurs cours et dans la mise en œuvre du CECRL dans leur enseignement. C'est un outil simple et utile permettant de partager son expérience et ses idées avec d'autres enseignants. Elle offre également la possibilité d'archiver des matériels de classe, facilitant ainsi, pour une équipe d'enseignants, le recueil d'une banque d'activités calibrées aux niveaux du CECRL, ce grâce à une procédure de validation qualitative.

Le formulaire en ligne se compose de quatre sections. La première (A) recueille les informations contextuelles du dossier, la deuxième (B) permet l'estimation du niveau requis pour la compréhension d'un ou de plusieurs textes, la troisième (C), l'estimation du niveau requis pour accomplir les activités (en fonction des

objectifs visée et décrits dans le CECR) et la quatrième (D), l'estimation du niveau requis en termes de compétences et de stratégies (toujours en fonction des objectifs visée et décrits dans le CECR). (Voir annexe 6)

Une fois les données entrées, la grille calcule automatiquement ¹³le niveau estimé (ou la fourchette de niveau).

Laissons de côté la section A qui n'est pas concernée par l'estimation du niveau.

Voici les tables de correspondances de niveau entre les niveaux du CECR et les scores d'estimation CEF-ESTIM.

Table CEF-ESTIM des correspondances de niveaux :

Section B: estimated level of learner to cope with text(s)		
	A1	0-12
	A1/A2	13-16
	A2	17-24
	A2/B1	25-32
	B1	33-40
	B1/B2	41-48
	B2	49-56
	B2/C1	57-64
	C1	65-72
	C1/C2	73-80

Tableau n° 1 : La section B (texte(s)) de la grille CEF-ESTIM : correspondances avec les niveaux du CECR.

¹³ Correspondances établies par un ingénieur ECML (European Centre for Modern Languages).

Section C: estimated level of learner to cope with the communicative activities		
	A1	1-10
	A1/A2	11-20
	A2	21-30
	A2/B1	31-40
	B1	41-50
	B1/B2	51-60
	B2	61-70
	B2/C1	71-80
	C1	81-90
	C1/C2	91-100
	C2	101-110

Tableau n° 2 : La section C (activités) de la grille CEF-ESTIM : correspondances avec les niveaux du CECR.

Section D : level of communication strategies		
	A1	1-13
	A1/A2	14-26
	A2	27-39
	A2/B1	40-52
	B1	53-65
	B1/B2	66-78
	B2	79-91
	B2/C1	92-104
	C1	105-117
	C1/C2	118-130
	C2	131-143

Tableau n° 3 : La section D (stratégies) de la grille CEF-ESTIM : correspondances avec les niveaux du CECR.

2.2.2. La procédure

La grille CEF-ESTIM permet une estimation globale sur un dossier composé de supports audio/vidéo et textuels, et d'activités de CO, CE, EO, EOI, EE, et de médiation, portant sur lesdits supports. Cependant, il est possible d'obtenir un résultat chiffré sur une partie de la grille seulement. Nous avons donc d'abord appliqué la section B de la grille CEF-ESTIM sur les supports vidéo, puis sur les supports textuels pour vérifier leur homogénéité : types de textes, de discours, d'étendue de vocabulaire, etc. pour les 4 documents écrits et type de reportage pour les deux documents vidéo. Nous avons procédé de même en appliquant la section C de la grille sur les activités de CO, EE et EOI proposées dans le dossier, puis sur la section D. Enfin, nous avons appliqué la grille à l'ensemble du dossier pour obtenir l'estimation globale finale.

Un problème est survenu dans le mode d'emploi de la grille dû au fait que la section B ne permet pas de différencier le niveau des supports audio ou écrits. Or, dans le cas du présent dossier, comme le montrent les résultats ci-dessous, les documents vidéo et textuels ne sont pas de même niveau. Nous avons donc été obligés de faire la moyenne du score obtenu dans la section B pour obtenir un résultat juste.

Deux autres évaluatrices ont été sollicitées, Raili Hilden (Université d'Helsinki, Finlande), et Monique Reichert (IMAC, Luxembourg), toutes deux spécialistes et membres de l'équipe créatrice de la grille CEF-ESTIM. Elles ont procédé à une estimation séparée, sans connaître les résultats obtenus par les autres.

Nous présenterons d'abord les résultats détaillés de l'auteure de l'article (estimation 1), suivis des résultats comparés obtenus par les deux autres estimatrices (estimations 2 et 3).

3. Résultats estimation 1 :

3.1. Estimation du niveau du sujet CLES 2

Section B

- Estimation du niveau requis pour les ressources documentaires vidéo :

52 points => B2 (49-56)

- Estimation du niveau requis pour les ressources documentaires textuelles :

63 points => B2/C1 (57-64)

=> Estimation globale des supports : 57 points B2/C1 (57-64)

Section C

=> Estimation globale des activités : 70 points B2 (61-70)

Section D

=> Estimation globale des compétences et stratégies : 92 points B2-C1(92-104)

L'estimation du niveau global du dossier CLES est la suivante : **B2**

Il est à noter que le calcul automatique ne produit pas de score final. Si tel était le cas, on trouverait sans doute un score à la limite de B2/C1.

Qu'en est-il des deux autres estimations ?

3.2. Comparaison avec les deux autres estimations

3.2.1. Présentation des résultats et analyse

Estimatrices	Supports vidéo	Supports Textuels	Supports Ensemble	Activités Ensemble	Stratégies	Estimation finale
	Section B				Section D	
Estimatrice 1 (France)	B2 (52)	B2/C1 (63)	B2/C1 (57)	B2 (70)	B2/C1 (92)	B2
Estimatrice 2 (Finlande)	B2 (52)	B2/C1 (57)	B2 (53,5)	B2 (70)	B2/C1 (92)	B2
Estimatrice 3 (Luxembourg)	B1/B2 (47)	B2/C1 (63)	B2 (55)	B2 (63)	B2 (83)	B2

Tableau 4 : Résultats des trois estimations pour le dossier *Counterfeiting*

On note une assez grande adéquation entre les avis.

Les principaux points de convergence concernent :

La différence de niveau entre le niveau des textes (B2/C1) et celui des documents vidéo (B2, ou B1/B2), le niveau des textes écrits jugés à B2/C1, le niveau des activités situé à B2 par l'ensemble des estimatrices, et le niveau global du dossier.

Les principaux points de divergence concernent :

Le niveau des supports vidéo :

L'estimatrice 3 les considère accessibles au niveau B1/B2 alors que les deux autres estimatrices (1 et 2) les situent au niveau B2. Ceci pourrait s'expliquer par la différence culturelle. En effet, le projet EBAFLS (*Building a European bank of anchor items for foreign language skills*) (2004-2007) qui concernait 8 pays d'Europe et visait à créer une banque européenne d'items d'ancrage pour les tests nationaux a montré que le niveau B1 dans un pays ne correspond pas forcément au niveau B1 d'un autre pays. On peut penser que la collègue du Luxembourg, pays trilingue (français, allemand et luxembourgeois), attribue un niveau B1/B2 à des vidéos qui, en France, ne seraient accessibles qu'au niveau B2. Cependant, la collègue de Finlande qui est un pays bilingue voire trilingue (Finlandais, anglais, suédois) devrait faire une estimation similaire, ce qui n'est pas le cas. La différence principale concerne la caractéristique B 9 sur la vitesse de l'enregistrement où elle a coché « fast » là où la collègue Luxembourgeoise a coché « normal », et B1 où la première a coché comme l'estimatrice 1 « sometimes unclearly articulated », là où l'estimatrice 3 a coché « normally articulated ». On peut penser que la variable culturelle entre moins en ligne de compte ici que la méthode d'estimation elle-même. De fait, il semble difficile à plusieurs estimatrices d'avoir exactement le même jugement sur toutes les caractéristiques, lorsque ce jugement est porté individuellement. Un consensus pourrait plus facilement s'obtenir par une discussion entre les trois de type « standard setting ».

Le niveau d'ensemble des supports :

Concernant le niveau requis pour l'ensemble des supports. L'estimatrice 1 obtient une moyenne de 57 points soit l'exacte limite inférieure de B2/C1. L'estimatrice 2 obtient 53,5, soit B2, et l'estimatrice 3 55 points, également B2. Il est à noter que c'est son estimation à B1/B2 des supports vidéo qui maintient son jugement à B2, car elle situe la difficulté des supports écrits à B2/C1 avec le même score élevé de 63 points que l'estimatrice 1.

On voit ici les limites de la grille CEF-ESTIM qui suppose un niveau homogène des supports ou requiert une utilisation plus complexe des sections, avec l'obligation de reporter la moyenne dans la section B.

Le niveau des compétences et stratégies

Enfin, même si l'on obtient un niveau B2 global pour les activités de communication, les estimatrices 1 et 2 obtiennent un score de 92, soit la limite inférieure de B2/C1 (91-104) pour les compétences et stratégies requises. Ceci en raison du fait que les activités de CE, EE et EOI portent sur l'ensemble des documents et non sur des ressources séparées, ce qui nous semble complexifier la tâche et requérir non seulement du temps supplémentaire mais aussi une capacité de synthèse présente dans les descripteurs au niveau C1, bien que déjà attendue au niveau B2 (CECRL, 2001, p.65).

À titre d'exemple, en CE, la question 4c se présente ainsi :

Quote a sentence that means:

'Fighting counterfeiting takes time'

Text No. _____ Lines _____

La plupart des questions demandent ainsi à chaque fois une relecture des 4 textes. Ce n'est pas forcément une tâche difficile, mais c'est une tâche qui prend plus de temps et qui tient compte de capacités de repérage et de mémorisation plus complexes que si elle portait sur un seul texte. Cette estimation un peu plus haute que les trois autres pour la section D de la grille se situe néanmoins juste au-dessus de B2, à la limite inférieure de C1.

La concordance globale entre les avis semble néanmoins montrer que la grille CEF-ESTIM constitue une aide valable à l'estimation du niveau d'un sujet CLES.

En outre, le fait que l'on puisse entrer toutes les informations dans la grille et obtenir une estimation globale rapide rend son utilisation très pratique. Enfin, la possibilité d'utiliser la grille par morceaux comme nous l'avons fait également

permet de faire émerger une certaine granularité au sein d'un même dossier (ex : un niveau plus élevé des supports textuels par rapport aux supports vidéo), ce qui la rend utile à notre recherche.

3.2.2. Résultat global

Globalement, le résultat qui se dessine est que le dossier requiert un niveau B2, voire B2+, autrement dit, il semble intrinsèquement hors de portée d'un candidat à B1. Une solution serait peut-être à chercher dans la modulation du barème.

3.3. Le Barème du CLES

Qu'en est-il du barème du CLES 2 ?

En compréhension orale, le total des points s'élève à 25. Il en faut 15 pour obtenir la compétence.

En compréhension écrite, le total est de 20 points. Il en faut seulement 12 pour obtenir la compétence¹⁴.

Dans quelle mesure, en modulant ce barème, pourrait-on délivrer un CLES B1, voire B1+ ou B1/B2 à un.e candidat.e qui aurait échoué dans une ou plusieurs compétences ?

3.4. Discussion des résultats sur le dossier *Counterfeiting*

Pour ce faire, il faudrait proposer un infra-barème permettant d'attribuer un certificat au niveau B1. Ainsi pour valider le CLES 1, il suffirait d'obtenir 10/25 en CO et 8/20 en CE à un certificat CLES B2. Ajoutons à cela que les sections C et D de la grille CEF-ESTIM permettent de décrire les compétences d'expression et

¹⁴ Depuis l'écriture de cet article, les sujets CLES ont été harmonisés (pour renforcer la fiabilité) de la même façon sur toutes les langues et, désormais, la CO est obligatoirement sur 20, la CE sur 25.

d'interaction attendues aux niveaux B1 et B2, à la fois en termes linguistiques et pragmatiques. C'est sans doute envisageable. De fait, la coordination nationale du CLES a opté pour attribuer le niveau B1 aux candidats à B2 qui obtiendraient un score de 10/20 en CO et 12/25 (contre 15) en CE.

Cependant, le dossier documentaire n'étant pas conçu a priori pour un candidat au niveau B1, le type de supports étant bien ancré à B2, voire B2/C1, selon la typologie de textes présentés dans le CECR, un biais majeur pourrait rendre la tentative caduque. En effet, comment savoir si le candidat possède le niveau B1 alors que sa réussite ou son échec dépendent d'un dossier qui dépasse ses compétences ?

De plus, le questionnement combiné des 4 textes en CE, l'utilisation des 6 ressources pour l'EE et l'EOI complexifient la tâche largement au-delà d'un niveau de compétences à B1. Et la combinaison d'activités sur des supports requérant un niveau B2 délivrera difficilement des résultats à B1, mais plus logiquement au niveau B2, voire B2/C1.

Par conséquent, vouloir délivrer un niveau B1 par modulation du barème semble irréaliste, sur ce sujet en tout cas. Cela reste une option possible sur un sujet CLES B2 « plus facile ». À noter que le comité de validation national s'efforce d'unifier la difficulté des sujets au sein d'un même niveau.

Cependant on peut inverser la question et se demander s'il serait possible de délivrer un certificat au niveau C1 en modulant le nombre de réponses correctes. Il s'agirait cette fois de relever le niveau d'exigence en proposant par exemple : 20/25 en CO et 16/20 en CE. Dans ce cas, on se référerait aux descripteurs de C1 pour évaluer l'EOI et la PE, par le biais des sections C et D de la grille CEF-ESTIM. Cela semble a priori possible.

Il paraît plus facile de resserrer le barème pour délivrer un certificat de niveau supérieur que de le relâcher pour délivrer un certificat de niveau inférieur, et ceci est dû à la fois à la difficulté des supports et à leur traitement combiné. Néanmoins, le CLES C1 se doit d'évaluer la capacité à comprendre des textes plus spécialisés qu'à B2. Le problème subsiste donc.

3.5. Remarques

3.5.1. Remarques sur le sujet CLES B2 :

Il semblerait que ce soit principalement la dimension holistique du CLES qui joue sur l'élévation de niveau, c'est-à-dire, la combinaison des supports et des compétences ainsi sollicitées, même si les activités de réception et de production ou d'interaction sont évaluées séparément. Mais c'est aussi tout à l'honneur du CLES que de proposer de véritables scénarios sur des sujets de société qu'on ne peut aborder que sous l'angle de la complexité. En ce sens, l'obtention du CLES B2 est une garantie pour l'employeur que la personne est capable de gérer un dossier relativement complexe en L2 faisant appel à plusieurs compétences simultanément.

3.5.2. La granularité de l'estimation

Dans quelle mesure la grille CEF-ESTIM peut-elle servir d'outil de différenciation au sein d'un même dossier CLES B2 anglais ? Dans quelle mesure n'est-elle pas mieux adaptée à un format non holistique (puisque centrée sur le texte et les activités qui s'y rapportent et non directement sur un scénario global – voir 2.2.) ? En fait, elle s'applique très bien à chaque morceau du dossier, permettant de déceler des différences de niveau, comme ici un éventuel B1/B2 pour les documents vidéo, de produire des scores moyens et au final une estimation globale.

3.5.3. L'estimation collaborative

Enfin, dernière remarque et non des moindres, la fiabilité de la grille CEF-ESTIM conçue pour des enseignants en exercice augmente avec l'estimation en équipe. Ceci est une caractéristique majeure des procédures de standard setting en général (entre 3 et 5 personnes) (ALTE, 2011 : 29). Pour relier véritablement des

examens au CECRL), une douzaine d'experts est idéalement requise (Conseil de l'Europe, 2009 : 39). Ici, les sujets ont déjà été standardisés par l'équipe CLES en amont. Une estimation experte individuelle suivie d'une mise en commun semble la procédure la plus souhaitable.

3.6. Réponse à la question initiale

Il s'agit maintenant de tirer des conclusions de cette analyse qui demeure limitée à un dossier CLES B2 mais ouvre des pistes de réflexion.

Un sujet CLES conçu pour le niveau B2 peut-il permettre de valider un niveau B1 ? Autrement dit, est-il possible de produire une évaluation différenciée (B2 ou B1) à partir de sujets qui ont été créés pour le niveau B2 uniquement ?

La réponse est la suivante, forcément indirecte : en fait, tout dépend de la combinaison triple entre le niveau des textes de support (écrits, audio et vidéo), le niveau des activités proposées (tâches de réception, production et interaction) et des stratégies liées et le barème choisi. De cette conclusion, on peut dresser une liste de préconisations, liste qui ne prétend pas résoudre la complexité de la chose, mais offre néanmoins des possibilités.

3.7. Préconisations

Textes	Tâches	Niveau global probable de l'évaluation	Autre niveau possible d'évaluation avec modulation du barème
B1/B2	B2	B2	B1
B2	B1/B2	B2	B1
B2	B2	B2	/
B2/C1	B2	B2	C1
B2	B2/C1	B2	C1

Plus généralement, la modulation du barème semble une solution sous certaines conditions : un B1 ou B1+ ou B1/B2 pourrait être délivré si l'un ou l'autre des composants est inférieur à B2 (textes ou tâches) ; un C1, B2+ ou B2/C1, pourrait être délivré si l'un ou l'autre des composants est supérieur à B2. Mais ceci relève de la conjecture et reste à tester sur le terrain.

3.8. Pistes de réflexion

Plus simplement, on pourrait aussi séparer les compétences comme d'autres tests le font sans pour autant cesser de délivrer un certificat global.

Ainsi pourrait-on imaginer que des candidats qui valident 2 compétences sur 4 au CLES 2 se voient délivrer un certificat à B1/B2 (ou B1+) ; et ceux qui valident 3 compétences sur 4, se voient gratifier d'un certificat B2- (mieux que B2/B1).

De manière plus précise, on pourrait aussi appliquer la grille d'évaluation et le barème de B1 sur les compétences non réussies par un.e candidat.e au CLES B2 et lui délivrer un B1/B2 ou un B2/B1.

Dans tous les cas, la délivrance d'un certificat à un niveau réussi (B1, B1/B2, B2/B1 ou B2) est la meilleure garantie pour renforcer l'attractivité du CLES auprès de la communauté étudiante avec pour motto : « Le CLES un certificat auquel il est impossible d'échouer ».

Références

ALTE (2011) *Manual for Language Test Development and Examining. For use with the CEFR*, On behalf of the Language Policy Division, Council of Europe.
<https://rm.coe.int/manual-for-language-test-development-and-examining-for-use-with-the-ce/1680667a2b>

Alderson Charles, Figueras Neus, Kuijper Henk, Nolte Günther, Takala Sauli and Tardieu Claire (2006) *Analysing Tests of Reading and Listening in Relation to the*

Common European Framework of Reference: The experience of the Dutch CEFR Construct Project, *Language assessment Quarterly*, 3 (1), 3-30.

Bachman Lyle F. (1990) *Fundamental Considerations in Language Testing*, Oxford: Oxford University Press.

- (2007) "What is the construct? The dialectic of abilities and contexts in Defining constructs in language assessment", in J. Fox, M. Wesche, D. Bayliss, L. Cheng, C.E. Turner & C. Doe (eds.), *Language testing reconsidered*, Ottawa, CA: University of Ottawa Press, 41-72.

Bardière, Yves (2020) News Tank, Higher Ed & Research, 15 mai 2020 à 15 h 06 - Actualité n° 183127.

Bourguignon Claire, Narcy-Combes Jean-Paul (2003) « Un point sur l'approche qualitative de l'évaluation », *ASP*, 39-40, 171-181. Disponible en ligne <https://doi.org.ezproxy.univ-paris3.fr/10.4000/asp.1377>

Canale Michael, Swain Merrill (1980) "Theoretical bases of communicative approaches to second language teaching and testing", *Applied Linguistics*, 1 (1), 1-47, disponible en ligne <https://doi.org/10.1093/applin/l.1.1>

Chomsky Noam (1965) *Aspects of the Theory of Syntax*, Cambridge, Massachusetts: M.I.T. Press.

Conseil de l'Europe (2009) *Relier les examens de langues au Cadre européen commun de référence pour les langues : Apprendre, enseigner, évaluer (CECRL). Un manuel*, Division des politiques linguistiques : Strasbourg.

Ellis Rod (2003) *Task-based Language Learning and Teaching*, Oxford: Oxford University Press.

Heaton John B. (1990) *Classroom Testing*, Harlow, UK: Longmann.

Horner David (2010) *Le Cecrl et l'évaluation de l'oral en anglais*, Paris : Belin, Coll. Guides de l'enseignement.

Huver Emmanuelle (2010) « Quelles représentations de l'évaluation chez les enseignants de langues dans l'enseignement supérieur ? », in F. Dervin et E. Suomela-Salmi (Eds.) *New approaches to assessing language and (inter-) cultural competences in higher education / Nouvelles approches de l'évaluation des compétences langagières et (inter)culturelles dans l'enseignement supérieur*, Bern : Peter Lang, 81-96.

Huver Emmanuelle, Springer Claude (2011) *L'évaluation en langues*, Paris : Les Éditions Didier.

Hymes Dell. H. (1972) "On Communicative Competence", in J.B. Pride and J. Holmes (eds) *Sociolinguistics. Selected Readings*, Harmondsworth: Penguin, 269-293 (Part 2).

Mac Gaw Stéphanie (2017) *La certification en langues, prisme révélateur d'enjeux politiques, économiques et sociaux : une étude par le CLES (Certificat de Compétences en Langues de l'Enseignement Supérieur)*, sous la direction de Georges Moracchini et Annick Rivens-Mompean, Université Pasquale Paoli, Corte, Corse, 22 novembre 2017.

Nunan David (1989) *Designing Tasks for the Communicative Classroom*. Cambridge: Cambridge University Press.

Oller John W. (1976) Evidence for a general language proficiency factor: an expectancy grammar", *Die Neuren Sprachen*, 75, 165-174.

Purpura James E. (2016) Second and foreign language assessment. *Modern Language Journal*, 100 (Supplement 2016), 190-208. DOI: 10.1111/modl.123080026-7902/16/190-208

Pica Teresa, Kanagy Ruth & Falodun Joseph (1993) "Choosing and using communication tasks for second language research and instruction", in S. Gass and G. Crookes (ed.): *Task-based Learning in a Second Language*, Clevedon: Multilingual Matters.

Shohamy Elana (2007), "Language tests as language policy tools", *Assessment in Education*, 14, 1, 117-130.

Spolsky Bernard (1981) "Some ethical questions about language testing", in C. Klein-Braley & D. K. Stevenson (eds) *Practice and problems in language testing*, Frankfurt am Main: Verlag Peter D. Lang, 5-30.

Springer Claude (2001) « Diagnostic, bilan de compétences, certification : les nouveaux habits de l'évaluation », *Langues modernes*, 2, 48-60.

Tardieu Claire (2013) « Les tests de langues : une nécessité ? », in E. Huver & A., Ljalikova, *L'évaluation en didactique des langues/cultures. Continuités, tensions, ruptures, Le Français dans le monde, Recherches et Applications*, 53, 67-79.

- (2013) « Testing et certification ». *Apprendre les langues autrement, Recherches en didactique des langues et des cultures, Les Cahiers de l'ACEDLE*, 10 (2), 237-251.

- (2010) « Votre B1 est-il mon B1 ? L'interculturel dans les tests d'évaluation en Europe », in *Cahiers de l'ACEDLE*, 7, 2, 225-239, consulté le 28 août 2020 : <http://acedle.org/spip.php?article3103> (ACL).

- Tardieu Claire, Hilden Raili, Reichert Monique & Lehmann Magdolna (2009) *La Grille CEF-ESTIM : Grille d'estimation des niveaux du CECRL pour les enseignants/The CEF-ESTIM Grid : CEFR : level estimation grid for teachers*. Disponible en ligne/Available online: <http://cefestim.ecml.at/>

Téléchargement direct :
https://www.ecml.at/Portals/1/3MTP/CEFESTIM/CEF_ESTIM_03_2011.pdf?ver=2020-08-21-091737-530

Widdowson Henry G. 2002 [1990], *Aspects of Language Teaching*, Oxford: Oxford University Press.