

**Évaluer le français au lycée, évaluer les littéracies universitaires  
Analyse diachronique de l'examen régional du baccalauréat marocain  
(2018-2025) et hypothèse de la rupture évaluative secondaire-supérieur**

**Evaluating French at high school, evaluating academic literacies :  
a diachronic analysis of the Moroccan regional baccalaureate examination  
(2018-2025) and the hypothesis of an evaluative rupture between secondary  
and tertiary education**

**Manal Antoine**

Docteure en didactique du français

*Université Hassan II de Casablanca, Faculté des Lettres et des Sciences Humaines de Mohammedia,*

**Date de soumission :** 15/05/2026

**Date d'acceptation :** 27/07/2026

**Pour citer cet article :**

Antoine. M. (2026) «Évaluer le français au lycée, évaluer les littéracies universitaires : Analyse diachronique de l'examen régional du baccalauréat marocain (2018-2025) et hypothèse de la rupture évaluative secondaire-supérieur», Revue Internationale du chercheur « Volume 7 : Numéro 3 » pp : 48-71

## Résumé

La rupture entre la maîtrise du français acquise au baccalauréat marocain et les compétences littéraciques attendues à l'université est généralement traitée dans la littérature comme un déficit linguistique des bacheliers. Cet article propose d'en déplacer le diagnostic en l'interrogeant comme un effet de la dissociation entre deux régimes évaluatifs distincts. L'étude s'appuie sur l'analyse diachronique d'un corpus de treize épreuves de l'examen régional de fin de première année du baccalauréat marocain, recueillies de 2018 à 2025 dans six académies régionales (Rabat-Salé-Kénitra, Casablanca-Settat, Tanger-Tétouan-Al-Hoceima, Fès-Meknès, Laâyoune-Sakia Elhamra et Béni-Mellal-Khénifra), complétée par une enquête par entretiens semi-directifs menée auprès de cinquante enseignants du cycle secondaire qualifiant. Les épreuves sont caractérisées selon une grille à quatre critères articulant niveau taxonomique (Bloom, 1956), part de la production singulière (Scallan, 2004), degré d'authenticité de la situation évaluative (De Ketele, 2010) et degré de prévisibilité du sujet de production écrite ; la fiabilité de cette grille a été éprouvée par une validation inter-juges (coefficient de Fleiss). Les résultats mettent au jour une stabilité structurelle massive du modèle évaluatif sur sept années : la production singulière représente en moyenne 19 % des questions de l'étude de texte, les treize épreuves relèvent structurellement de l'évaluation artefactuelle au sens de De Ketele, et aucun sujet de production écrite n'atteint le niveau 3 de prévisibilité (sujet véritablement non préparé en classe). Les entretiens confirment ce dernier résultat : quarante-trois enseignants sur cinquante indiquent que les thèmes des sujets de production écrite se répètent et qu'ils les préparent en classe avec leurs élèves. Une enquête exploratoire menée auprès de quarante enseignants universitaires conforte ce décalage : la quasi-totalité d'entre eux observe que les bacheliers savent restituer des connaissances mais peinent à les analyser et à produire un raisonnement personnel. Articulés au cadre théorique des littéracies universitaires (Pollet, 2014 ; Reuter, 2012 ; Delcambre & Lahanier-Reuter, 2012), ces résultats soutiennent que la rupture secondaire-supérieur n'est pas réductible à un déficit linguistique : elle tient à un défaut d'articulation entre deux dispositifs évaluatifs qui ne mesurent pas les mêmes compétences.

## Mots-clés :

Évaluation certificative, littéracies universitaires, baccalauréat marocain, transition secondaire-supérieur, didactique du français, production écrite, taxonomie de Bloom.

## Abstract

The gap between the French language proficiency acquired at the Moroccan baccalaureate and the literacy skills expected at university is generally interpreted in the literature as a linguistic deficit of secondary school graduates. This article proposes to shift this diagnosis by examining it as an effect of the dissociation between two distinct evaluative regimes. The study is based on the diachronic analysis of a corpus of thirteen examinations from the regional baccalaureate of the first year of Moroccan high school, collected between 2018 and 2025 across six regional academies (Rabat-Salé-Kénitra, Casablanca-Settat, Tanger-Tétouan-Al-Hoceima, Fès-Meknès, Laâyoune-Sakia Elhamra and Béni-Mellal-Khénifra), supplemented by a semi-structured interview survey of fifty upper-secondary school teachers. The examinations are characterized according to a four-criteria grid articulating taxonomic level (Bloom, 1956), the proportion of singular production (Scallon, 2004), the degree of authenticity of the evaluative situation (De Ketele, 2010), and the degree of predictability of the writing topic; the reliability of this grid was tested through an inter-rater validation (Fleiss' kappa). The results reveal a massive structural stability of the evaluative model over seven years: singular production represents on average 19% of the text analysis questions, all thirteen examinations belong structurally to artifactual evaluation in De Ketele's sense, and no writing topic reaches level 3 of predictability (a truly non-prepared topic). The interviews confirm this last result: forty-three teachers out of fifty report that the themes of the writing topics recur and that they prepare them in class with their students. An exploratory survey of forty university teachers supports this gap: almost all of them observe that secondary school graduates can recall knowledge but struggle to analyse it and to produce personal reasoning. Linked to the theoretical framework of academic literacies (Pollet, 2014 ; Reuter, 2012 ; Delcambre & Lahanier-Reuter, 2012), these results support the claim that the secondary-tertiary gap is not reducible to a linguistic deficit: it stems from a lack of articulation between two evaluative devices that do not measure the same competencies.

**Keywords :** Certificative evaluation, academic literacies, Moroccan baccalaureate, secondary-to-tertiary transition, French didactics, written production, Bloom's taxonomy.

## Introduction

La transition entre l'enseignement secondaire et l'enseignement supérieur constitue, dans les systèmes éducatifs francophones, un point de tension récurrent dont la littérature documente les manifestations multiples (Delcambre & Lahanier-Reuter, 2010 ; Pollet, 2014 ; Boch & Rinck, 2010). Au Maroc, cette transition est particulièrement sensible en langue française, qui fonctionne simultanément comme objet d'enseignement au lycée et comme langue de scolarisation à l'université pour les filières scientifiques, économiques et techniques. Les difficultés rencontrées par les bacheliers à leur entrée à l'université sont attestées de longue date. Chami (2004, p. 65) notait déjà que « la majorité des élèves marocains qui accèdent aujourd'hui à l'Université ont une connaissance très approximative de la langue française aussi bien à l'oral qu'à l'écrit ».

Le diagnostic dominant pour rendre compte de ces difficultés mobilise le registre du déficit linguistique (Mehdaoui, 2022). Ce diagnostic rejoint les travaux qui analysent la « fracture linguistique » entre le secondaire et l'université au Maroc, imputée au changement de langue d'enseignement des disciplines scientifiques (Messaoudi, 2016). Selon cette lecture, les étudiants entrent à l'université avec une maîtrise insuffisante du français, ce qui les empêche d'accéder aux discours académiques et de produire les écrits attendus dans leurs disciplines. Les dispositifs institutionnels mis en place pour y répondre relèvent majoritairement de cette lecture déficitaire : modules de français transversal, remédiation linguistique, certifications de niveau adossées au CECRL. Or, ces dispositifs produisent des résultats limités. La littérature documente une baisse persistante du niveau des étudiants malgré les cours de remédiation mis en place depuis 2003 puis renforcés en 2009, et souligne une fracture entre l'enseignement du français au secondaire et les cours de langue à l'université (Mehdaoui, 2022).

Cet article propose de déplacer le diagnostic. Plutôt que de penser la rupture secondaire-supérieur comme un effet d'un déficit linguistique des bacheliers, nous l'examinons comme un effet de la dissociation entre deux régimes évaluatifs qui ne mesurent pas les mêmes compétences. Cette hypothèse alternative s'inscrit dans le programme de recherche des littéracies universitaires, qui pose précisément que les difficultés des étudiants à l'entrée à l'université ne sont pas réductibles à un déficit linguistique mais relèvent d'une acculturation inachevée aux discours académiques (Delcambre & Lahanier-Reuter, 2012 ; Reuter, 2012).

Pour mettre cette hypothèse à l'épreuve, nous conduisons une analyse diachronique des épreuves de l'examen régional de fin de première année du baccalauréat marocain, qui constitue

l'évaluation certificative la plus proche de l'entrée à l'université. Nous mobilisons un corpus de treize épreuves recueillies entre 2018 et 2025 dans six académies régionales différentes, que nous caractérisons à l'aide d'une grille d'analyse à quatre critères. Cette grille articule des références canoniques de la didactique de l'évaluation (Bloom, 1956 ; Scallon, 2004 ; De Ketele, 2010) et un critère portant sur la prévisibilité des sujets de production écrite. Cette analyse documentaire est complétée par une validation inter-juges du codage et par deux enquêtes de terrain (l'une auprès d'enseignants du secondaire, l'autre auprès d'enseignants universitaires) destinées à éprouver, respectivement, la prévisibilité des sujets et la réalité du décalage perçu à l'entrée du supérieur.

Cet article se distingue par trois éléments convergents. Premièrement, le corpus : échelonné sur sept années et plusieurs académies régionales, il constitue, à notre connaissance, le corpus diachronique le plus étendu mobilisé pour étudier l'évaluation certificative du français au baccalauréat marocain. Deuxièmement, la grille d'analyse : la combinaison simultanée des quatre critères retenus n'a pas, semble-t-il, été appliquée à un corpus d'examens régionaux marocains. Troisièmement, la mise en dialogue avec le champ des littéracies universitaires : là où les analyses existantes des examens marocains restent essentiellement centrées sur la didactique du secondaire, nous relions le diagnostic du secondaire au questionnement universitaire sur les littéracies attendues.

La question de recherche est la suivante : dans quelle mesure les épreuves de l'examen régional du baccalauréat marocain en langue française, telles qu'elles se déploient de 2018 à 2025, évaluent-elles les compétences attendues à l'entrée à l'université dans le cadre des littéracies universitaires ? L'article s'organise comme suit. Le cadre théorique articule la didactique de l'évaluation et les littéracies universitaires (section 2). La méthodologie expose le corpus, l'enquête par entretiens et la grille d'analyse (section 3). Les résultats sont présentés à travers l'analyse de trois épreuves emblématiques et un tableau récapitulatif (section 4). La discussion interprète ces résultats à la lumière de l'hypothèse de la rupture évaluative (section 5). La conclusion formule les implications didactiques et les perspectives de recherche.

## **2. Cadre théorique**

Le cadre théorique de cet article croise deux champs complémentaires : la didactique de l'évaluation et les littéracies universitaires. Nous présentons successivement chacun de ces deux ancrages avant d'explicitier leur mise en relation au service de notre hypothèse.

### ***2.1. La didactique de l'évaluation : fonctions et caractérisation***

La didactique de l'évaluation s'est constituée comme champ de recherche autonome à partir des années 1980, dans le sillage des travaux de Cardinet (1986), Allal (1979, 2007) et Hadji (1989). Cette littérature distingue classiquement plusieurs fonctions de l'évaluation : diagnostique, formative, sommative, certificative. La fonction certificative, qui nous intéresse spécifiquement, vise à attester de la maîtrise des compétences au terme d'un cycle de formation et conditionne le passage au cycle suivant. Au Maroc, l'examen régional de fin de première année du baccalauréat relève typiquement de cette fonction : il atteste de la maîtrise du français en vue de la deuxième année du baccalauréat et, de manière médiate, de l'entrée à l'université.

Pour analyser systématiquement une évaluation certificative, nous mobilisons trois cadres conceptuels complémentaires. Le premier est la taxonomie des objectifs éducatifs proposée par Bloom (1956), qui hiérarchise six niveaux du plus simple au plus complexe : connaissance, compréhension, application, analyse, synthèse, évaluation. Cette taxonomie, bien que critiquée pour son caractère parfois schématique (Anderson & Krathwohl, 2001), demeure un outil descriptif solide pour situer le niveau cognitif sollicité par une question d'épreuve.

La deuxième référence est celle de Scallon (2004) sur la planification de l'évaluation certificative dans une approche par compétences. Scallon identifie trois types d'activités cognitives : la restitution (mobilisation de connaissances mémorisées ou recopiage d'éléments du texte support), l'application (mise en œuvre d'une procédure connue dans un contexte donné), et la production singulière (élaboration d'un discours ou d'un point de vue non prédéfini par le dispositif d'enseignement antérieur). Cette distinction est particulièrement opératoire pour notre propos parce qu'elle permet de qualifier le type d'activité demandée à l'élève, au-delà du seul niveau cognitif en jeu.

La troisième référence est celle de De Ketele (2010) sur l'authenticité de la situation évaluative. De Ketele oppose les évaluations authentiques, où la situation reproduit une pratique disciplinaire ou professionnelle réelle (complexité non préparée, ressources multiples, singularité de la réponse comme objet d'évaluation), aux évaluations artefactuelles, où la situation est construite pour les besoins propres de l'évaluation (tâche préparée, ressources prédéterminées, conformité au format valorisée). Il importe de souligner que cette distinction n'introduit pas une hiérarchie normative : toute évaluation certificative à fort enjeu social comporte une part artefactuelle nécessaire pour garantir l'équité entre les candidats, la comparabilité des résultats et la fiabilité de l'évaluation.

## ***2.2. Le champ des littéracies universitaires***

Le champ des littéracies universitaires s'est constitué dans le monde francophone à partir des années 2000, sous l'influence des travaux anglophones initiés par Lea et Street (1998) puis Lillis et Scott (2007). Dans le contexte francophone, il s'est développé notamment autour des contributions de Delcambre, Reuter et Lahanier-Reuter à Lille, de Pollet à Namur, et de Boch et Rinck à Grenoble (Boch & Rinck, 2010 ; Delcambre & Lahanier-Reuter, 2012 ; Pollet, 2014 ; Reuter, 2012). Ce champ se caractérise par un déplacement épistémologique majeur : il ne s'agit plus de repérer des lacunes méthodologiques, des déficits langagiers ou des écarts cognitifs des étudiants par rapport à des normes hypothétiques, mais de comprendre les logiques qui déterminent leur rapport aux pratiques langagières disciplinaires. Au Maroc, ce questionnement commence à être investi par la recherche didactique nationale : El Gousairi (2019), analysant les écrits scientifiques de futurs enseignants, montre que l'entrée dans les littéracies universitaires suppose une socialisation progressive aux pratiques discursives de l'écrit académique.

Quatre caractéristiques des littéracies universitaires, telles que les décrivent ces travaux, retiennent particulièrement notre attention. Premièrement, elles supposent la confrontation à des textes nouveaux, jamais étudiés en classe préalablement. L'étudiant doit produire une compréhension réflexive et critique de ces textes, mobiliser des grilles de lecture disciplinaires, et construire son interprétation. Deuxièmement, elles appellent la production d'écrits singuliers qui ne se réduisent pas à la restitution d'arguments préconstitués. L'étudiant doit problématiser, argumenter, prendre position, en engageant son propre point de vue énonciatif. Troisièmement, elles supposent l'autonomie dans la mobilisation des ressources : l'étudiant choisit ses lectures, construit sa bibliographie, organise son travail, sans encadrement étroit. Quatrièmement, elles exigent une maîtrise des codes discursifs propres à chaque discipline : chaque champ universitaire a ses propres genres textuels, ses propres formes argumentatives, ses propres conventions énonciatives (Lea & Street, 1998).

## ***2.3. Articulation des deux ancrages : l'hypothèse de la rupture évaluative***

L'articulation des deux ancrages permet de formuler l'hypothèse alternative que nous défendons dans cet article. Le diagnostic dominant pose que les étudiants entrent à l'université avec un déficit linguistique qui les empêche d'accéder aux littéracies universitaires. Nous proposons une lecture alternative que nous formulerons comme hypothèse de travail : les étudiants entreraient à l'université avec une maîtrise du français qui correspondrait précisément

à ce que le dispositif évaluatif antérieur aurait évalué, et donc constitué comme objet d'apprentissage.

Si cette hypothèse est juste, trois conséquences empiriques devraient pouvoir être observées dans les épreuves de l'examen régional du baccalauréat : ces épreuves devraient privilégier les niveaux taxonomiques inférieurs de Bloom au détriment des niveaux supérieurs, accorder une place limitée à la production singulière au sens de Scallon, et relever majoritairement de l'évaluation artefactuelle au sens de De Ketele. C'est ce faisceau d'hypothèses que notre étude empirique se propose de mettre à l'épreuve dans les sections 4 et 5.

### **3. Méthodologie**

#### **3.1. Le corpus documentaire**

Le corpus mobilisé dans cette étude rassemble treize épreuves de l'examen régional de fin de première année du baccalauréat marocain en langue française, recueillies sur la période 2018-2025. Les épreuves proviennent de six académies régionales : Rabat-Salé-Kénitra (2018, 2019, 2021, 2025), Casablanca-Settat (2018, 2025), Tanger-Tétouan-Al-Hoceima (2018, 2023, 2024), Fès-Meknès (2018, 2024), Laâyoune-Sakia Elhamra (2018) et Béni-Mellal-Khénifra (2018). Le tableau 1 (section 4.2) récapitule la composition exhaustive du corpus.

Le choix de la période 2018-2025 répond à trois critères : la disponibilité publique des épreuves, une étendue temporelle suffisante pour examiner la stabilité diachronique du modèle évaluatif, et l'inclusion de plusieurs académies régionales permettant d'interroger la stabilité interrégionale du dispositif. Pour la majorité des épreuves du corpus (neuf sur treize), nous disposons également de la grille de correction officielle élaborée par les Académies Régionales d'Éducation et de Formation. Cette grille permet une analyse comparée entre ce qui est demandé à l'élève et ce qui est explicitement valorisé par le correcteur. Le choix des six académies retenues relève en outre d'un échantillonnage de convenance, déterminé par l'accès au terrain : ce sont les académies dans lesquelles nous disposons de contacts professionnels parmi les enseignants de français, ce qui a rendu possible, dans un second temps, la conduite de l'enquête (section 3.2). Cette sélection ne constitue pas un échantillon représentatif de l'ensemble des académies du Royaume ; les résultats ne prétendent pas à une généralisation statistique, mais à la mise au jour d'une régularité structurelle sur un corpus diachronique étendu. La sélection des treize épreuves procède de la combinaison de plusieurs critères : la disponibilité matérielle des sujets que nous avons pu collecter sur la période ; le souci de couvrir plusieurs académies et plusieurs sessions afin d'éprouver la stabilité interrégionale et



diachronique du dispositif ; ainsi que la diversité des œuvres support mobilisées (*La Boîte à merveilles*, *Antigone*, *Le Dernier Jour d'un condamné*), afin que les régularités observées ne soient pas tributaires d'un texte particulier.

Le corpus comporte également les Orientations Pédagogiques officielles pour l'enseignement du français au cycle secondaire qualifiant (Ministère de l'Éducation Nationale, 2007), qui constituent le cadre prescriptif de référence. Au cours de la première année du baccalauréat, le programme officiel impose trois œuvres : *La Boîte à merveilles* » d'Ahmed Sefrioui, *Antigone* » de Jean Anouilh et *Le Dernier Jour d'un condamné* de Victor Hugo. Chaque Académie Régionale élabore son épreuve à partir de ces œuvres, en proposant un extrait d'une œuvre ou, comme dans l'épreuve de Béni-Mellal-Khénifra 2018, deux extraits de deux œuvres différentes dans une approche comparative. La réorientation de la discipline vers l'œuvre littéraire intégrale, introduite par la réforme de 2007, a constitué un infléchissement majeur vers l'appropriation de la langue-culture (Sadiki, 2016).

### **3.2. L'enquête par entretiens**

L'analyse documentaire du corpus a été complétée par une enquête par entretiens semi-directifs, conduite spécifiquement pour cette étude au cours de l'année scolaire 2024-2025. Cinquante enseignants de français du cycle secondaire qualifiant ont été interrogés. Ils exerçaient dans les six académies dont sont issues les épreuves du corpus, et leur ancienneté s'échelonnait de cinq à vingt-cinq ans, assurant une diversité de profils. Leur recrutement, par réseau professionnel (anciens camarades de formation, collègues d'établissement, contacts noués lors de formations continues et au sein de groupes professionnels d'enseignants de français), relève d'un échantillonnage de convenance qui ne prétend pas à la représentativité statistique. Cette enquête visait à vérifier, auprès des acteurs du terrain, le degré de prévisibilité des sujets de production écrite, qui constitue le quatrième critère de notre grille d'analyse.

Le dispositif d'entretien reposait sur une démarche de validation par les praticiens : les sujets de production écrite du corpus ont été présentés aux enseignants, qui étaient invités à indiquer si ces sujets, ou des sujets de thématique équivalente, avaient déjà été travaillés avec leurs élèves au cours de l'année, et s'ils les estimaient prévisibles. Les réponses recueillies sont mobilisées dans la section 4 pour étayer le critère de prévisibilité.

### **3.3. La grille d'analyse à quatre critères**

La grille d'analyse articule quatre critères systématiques, chacun adossé à une référence canonique. Leur mise en relation constitue le principal outil méthodologique de l'étude.

Le premier critère est le niveau taxonomique sollicité, selon la taxonomie de Bloom (1956). Pour chaque question de chaque épreuve, nous identifions le niveau cognitif dominant : connaissance, compréhension, application, analyse, synthèse ou évaluation. Lorsqu'une question fait appel à plusieurs niveaux, nous retenons le plus élevé effectivement exigé pour répondre correctement.

Le deuxième critère est la part respective de la restitution, de l'application et de la production singulière, selon Scallon (2004). Pour chaque question, nous identifions le type d'activité cognitive dominant. La restitution regroupe les questions qui invitent l'élève à mobiliser des connaissances mémorisées ou à recopier des éléments présents dans le texte support. L'application concerne les questions qui font appel à une procédure connue (transformation grammaticale, identification d'une figure de style, repérage d'un champ lexical). La production singulière, enfin, désigne les questions qui demandent à l'élève d'élaborer un discours ou un point de vue non prédéfini, typiquement les questions de réaction face au texte et la production écrite finale.

Le troisième critère est le degré d'authenticité de la situation évaluative, selon De Ketele (2010). Il porte non sur chaque question individuelle mais sur l'économie générale de l'épreuve. Nous qualifions chaque épreuve d'« authentique » ou d'« artefactuelle » selon que la situation évaluative reproduit (ou non) une situation discursive réelle. Trois indicateurs sont examinés : le statut du texte support, le format des questions et le statut du sujet de production écrite.

Le quatrième critère porte spécifiquement sur le degré de prévisibilité du sujet de production écrite. Nous proposons une échelle à trois niveaux. Le niveau 1 correspond au sujet entièrement prévisible : le thème a été traité en classe lors de l'étude de l'œuvre, les arguments « pour » et « contre » ont été collectivement élaborés, et l'élève peut produire une copie acceptable en restituant ce qui a été préparé. Le niveau 2 correspond au sujet partiellement prévisible : le thème général a été travaillé en classe à propos de l'œuvre au programme, mais l'angle précis demandé par le sujet appelle une reconfiguration de l'argumentation. Le niveau 3 correspond au sujet véritablement ouvert : le thème n'a pas été préparé en classe, ni à propos de l'œuvre ni dans le cadre d'exercices d'argumentation, et l'élève doit construire son argumentation à partir de son expérience ou de ses connaissances générales sans appui préparatoire. L'attribution d'un

niveau de prévisibilité à chaque sujet a été confortée par les données issues de l'enquête par entretiens décrite en section 3.2.

### **3.4. La démarche d'analyse**

La démarche d'analyse a été conduite en deux temps. Dans un premier temps, chacune des treize épreuves du corpus a été soumise à une analyse question par question, en appliquant systématiquement les deux premiers critères (Bloom et Scallon). Lorsque l'épreuve était accompagnée d'une grille de correction officielle, cette grille a été utilisée pour confirmer ou nuancer l'attribution. Les pourcentages de questions mobilisant les niveaux taxonomiques supérieurs de Bloom et la production singulière au sens de Scallon ont été calculés sur la base du nombre total de questions de l'étude de texte de chaque épreuve.

Dans un second temps, chaque épreuve a fait l'objet d'une caractérisation globale selon les troisième et quatrième critères (authenticité, prévisibilité). Cette caractérisation a été conduite par analyse holistique de l'épreuve dans son ensemble. Le critère de prévisibilité a été appliqué en tenant compte des thématiques effectivement travaillées en classe à propos des trois œuvres au programme, et les données issues de l'enquête par entretiens ont permis de conforter les attributions retenues.

Pour la présentation des résultats, trois épreuves emblématiques ont été sélectionnées pour faire l'objet d'une analyse détaillée. Ces trois épreuves ont été choisies pour leur représentativité du corpus diachronique : Rabat-Salé-Kénitra 2018 (point d'ancrage initial), Fès-Meknès 2024 (introduction d'une question intégrant la dimension plurilingue), et Casablanca-Settat 2025 (épreuve la plus récente présentant plusieurs nouveautés). Les dix autres épreuves sont synthétisées dans un tableau récapitulatif (tableau 1, section 4.2).

### **3.5. La validation du codage**

La grille reposant sur des jugements analytiques, sa fiabilité a été éprouvée par une validation inter-juges. Quatre codeurs ont caractérisé indépendamment les cent vingt-six questions d'étude de texte du corpus, selon les deux premiers critères (niveau taxonomique de Bloom et production singulière au sens de Scallon), sans concertation préalable ni accès à notre analyse. En cas de double cotation, la valeur la plus basse a systématiquement été retenue.

L'accord interjuge a été estimé au moyen du coefficient de Fleiss, adapté à la comparaison de plus de deux codeurs. Sur le niveau taxonomique de Bloom, il s'établit à  $\kappa = 0,69$ , valeur substantielle au sens de Landis et Koch (1977) ; il atteint  $\kappa = 0,80$ , soit un niveau quasi parfait,

lorsque les six paliers sont regroupés selon la dichotomie qui fonde l'analyse, opposant les niveaux inférieurs (connaissance, compréhension, application) aux niveaux supérieurs (analyse, synthèse, évaluation). Le critère de production singulière obtient un accord quasi parfait ( $\kappa = 0,89$ ).

L'écart entre l'accord fin et l'accord regroupé tient à la difficulté, bien documentée, de départager des paliers taxonomiques contigus ; les divergences se concentrent à l'intérieur d'un même bloc et n'affectent pas la dichotomie inférieurs/supérieurs sur laquelle repose l'interprétation des résultats. Cette convergence des quatre codeurs conforte la robustesse de la caractérisation proposée.

### **3.6. L'enquête auprès des enseignants universitaires**

Afin de documenter le versant universitaire de la transition, nous avons conduit une enquête exploratoire par questionnaire en ligne auprès d'enseignants du supérieur accueillant des élèves en première année. Quarante enseignants y ont répondu, entre le 12 et le 17 mars 2026, dans sept villes du Royaume (Rabat, Casablanca, Fès, Marrakech, Tanger, Agadir et Oujda). L'échantillon couvre une diversité de filières (facultés des sciences économiques, juridiques et sociales, facultés des sciences et techniques, départements de lettres et de français, modules de langue et écoles d'ingénieurs) ainsi que tous les degrés d'ancienneté. Le recrutement s'est appuyé sur plusieurs canaux (courrier électronique institutionnel, groupes professionnels en ligne, intermédiaire des chefs de département, diffusion directe).

Cet échantillonnage, de convenance et non probabiliste, ne prétend pas à la représentativité statistique ; les résultats qui en sont issus dégagent des tendances et n'ont pas vocation à être généralisés. Le questionnaire transpose, du côté des attentes universitaires, les quatre critères de notre grille : niveaux cognitifs sollicités, valorisation d'une production personnelle, authenticité des tâches et prévisibilité des sujets. Il visait à mettre en regard ce que le baccalauréat évalue et ce que l'université attend, sans prétendre établir un lien de causalité entre les deux.

## **4. Résultats**

### **4.1. Trois épreuves emblématiques**

#### **4.1.1. Rabat-Salé-Kénitra 2018 : le modèle évaluatif au début de la période**

L'épreuve de Rabat-Salé-Kénitra de 2018 prend pour support un extrait de *La Boîte à merveilles* d'Ahmed Sefrioui. L'extrait choisi est étudié systématiquement en classe lors du parcours

annuel sur cette œuvre. L'épreuve est structurée en deux parties : étude de texte (10 points) et production écrite (10 points). L'étude de texte se subdivise en trois sous-parties explicitement annoncées : contextualisation du texte (2 points), analyse du texte (6 points) et réaction face au texte (2 points).

L'application du premier critère (Bloom) fait apparaître la distribution suivante des dix questions de l'étude de texte. Trois questions se situent au niveau de la connaissance (identification de l'auteur et de l'œuvre, restitution d'informations encyclopédiques). Quatre portent sur la compréhension (situation du passage, repérage textuel). Une mobilise l'application (identification d'une figure de style sans justification). Deux, enfin, relèvent du niveau de l'évaluation (questions de réaction). Les niveaux supérieurs de Bloom représentent ainsi 30 % des questions. Le niveau de la synthèse est totalement absent de l'étude de texte.

L'application du deuxième critère (Scallion) montre que, sur les dix questions, cinq relèvent de la restitution pure, trois relèvent de l'application, et deux seulement relèvent de la production singulière. Cette dernière représente 20 % de l'épreuve. L'espace effectivement réservé sur la copie pour répondre à ces deux questions est de l'ordre d'une ligne et demie chacune, ce qui limite matériellement le développement argumentatif de l'élève.

L'application du troisième critère (De Ketele) montre que l'épreuve relève structurellement de l'évaluation artificielle. Trois indicateurs convergent : le texte support est extrait d'une œuvre étudiée en classe pendant plusieurs mois ; les questions reproduisent un format que les élèves ont rencontré dans tous leurs contrôles antérieurs ; et le sujet de production écrite porte sur un thème traité en classe lors de l'étude de l'œuvre.

L'application du quatrième critère (prévisibilité) confirme cette caractérisation. Le sujet proposé porte sur la place du père dans la famille et la capacité d'une mère à gérer son foyer en l'absence du père. Ce thème fait partie des sujets archétypaux traités en classe à propos de *La Boîte à merveilles*, et les arguments « pour » et « contre » sont collectivement élaborés au fil de l'année. Le sujet relève du niveau 1 (entièrement prévisible). La grille de correction officielle prévoit cinq points pour les critères de discours et cinq points pour les critères de langue. Aucun critère ne valorise explicitement la singularité du point de vue.

#### ***4.1.2. Fès-Meknès 2024 : des nouveautés ponctuelles dans un cadre stable***

L'épreuve de Fès-Meknès de juin 2024 prend également pour support *La Boîte à merveilles*, mais propose un extrait différent de l'épreuve de 2018. L'extrait choisi est le passage où Sidi

Mohammed revient à la maison fatigué après une journée au Msid, et où son père le félicite. La structure de l'épreuve est désormais simplifiée par rapport à 2018 : pas de tripartition explicite, simplement dix questions enchaînées dans l'étude de texte.

L'application du premier critère (Bloom) fait apparaître une distribution sensiblement différente. Une question se situe au niveau de la connaissance, deux portent sur la compréhension et trois mobilisent l'application. Deux relèvent du niveau de l'analyse, notamment la question 7 sur l'équivalent en français de « la kissaria » et l'interprétation du destinataire visé. Les deux dernières portent sur l'évaluation. Les niveaux supérieurs représentent 40 % de l'épreuve, soit dix points de pourcentage de plus que l'épreuve de 2018.

La question 7 mérite une attention particulière. Elle demande à l'élève de trouver dans le texte l'équivalent en français du mot marocain « la kissaria », puis d'identifier à quel lecteur le narrateur s'adresse en utilisant cet équivalent. Cette question intègre explicitement une dimension métalinguistique sur le contact des langues dans l'œuvre. Le corrigé officiel indique « Le marché des tissus » et « un lecteur étranger (ou français) ».

L'application du deuxième critère (Scallon) montre que la part de la production singulière reste de 20 % (deux questions sur dix). La progression observée au niveau de Bloom ne se traduit donc pas par un déplacement de la part de la production singulière, qui demeure structurellement marginale. L'application du troisième critère (De Ketele) montre que l'épreuve est artefactuelle : le texte support reste extrait du programme, les questions reproduisent les formats standards, le sujet de production écrite porte sur un thème travaillé en classe.

L'application du quatrième critère (prévisibilité) demande un examen attentif. Le sujet proposé est : « Aujourd'hui, certains lycéens et lycéennes se plaignent de plus en plus de l'École et s'y ennuiant. Et vous, partagez-vous l'avis de ces élèves ? » Ce sujet n'introduit pas de référence explicite à *La Boîte à merveilles*, mais le thème de la critique de l'école, du rapport au savoir et des attentes des lycéens est régulièrement travaillé en classe dans le cadre de la préparation à l'argumentation, notamment en lien avec les passages du roman consacrés au Msid et à l'apprentissage. Nous classons ce sujet au niveau 2 (partiellement prévisible) : le thème a été préparé en classe, mais l'angle précis demandé appelle une reconfiguration de l'argumentation.

#### **4.1.3. Casablanca-Settat 2025 : des nouveautés circonscrites**

L'épreuve de Casablanca-Settat de juin 2025 prend pour support *Le Dernier Jour d'un condamné* de Victor Hugo. L'extrait choisi est le passage du départ des galériens. L'épreuve

est accompagnée de définitions de termes spécifiques (gardes-chiourmes, carcan de la chiourme, galères).

L'application du premier critère (Bloom) fait apparaître la distribution suivante. Une question se situe au niveau de la connaissance et trois portent sur la compréhension. Quatre relèvent du niveau de l'analyse, dont la question 8 qui demande explicitement à l'élève d'identifier la figure de style et de préciser l'effet recherché. Cette demande de justification de l'effet est une nouveauté par rapport aux épreuves antérieures, où l'identification de la figure de style était demandée sans justification. Les deux dernières questions portent sur l'évaluation. Les niveaux supérieurs représentent 60 % de l'épreuve, soit le pourcentage le plus élevé du corpus.

L'application du deuxième critère (Scallon) montre toutefois que la part de la production singulière reste de 20 % (deux questions sur dix). Cette stabilité, malgré l'élévation du niveau taxonomique, est significative : les questions classées au niveau de l'analyse selon Bloom restent des questions d'application de procédures connues, pas des questions de production singulière. L'application du troisième critère (De Ketele) montre que l'épreuve demeure structurellement artefactuelle, le texte support restant extrait du programme.

L'application du quatrième critère (prévisibilité) appelle une analyse nuancée. Le sujet proposé est : « Certaines personnes pensent qu'il faut durcir les jugements et les peines pour réduire la délinquance et la criminalité. Qu'en pensez-vous ? » À première vue, ce sujet ne fait aucune référence explicite à l'œuvre étudiée et pourrait être interprété comme un sujet ouvert. Toutefois, le thème de la sévérité des peines, du rapport à la délinquance, de la justice répressive ou réhabilitatrice, fait partie du parcours pédagogique standard sur *Le Dernier Jour d'un condamné*. Le roman de Hugo est précisément un plaidoyer contre la peine de mort, et son étude conduit les enseignants à travailler en classe ces thématiques avec leurs élèves. Les arguments « pour » et « contre » la sévérité des peines sont donc largement préparés en classe à l'occasion de l'étude de cette œuvre. Nous classons ce sujet au niveau 2 (partiellement prévisible).

#### **4.2. Synthèse récapitulative de l'ensemble du corpus**

Le tableau 1 ci-dessous récapitule les résultats obtenus en appliquant la grille à quatre critères aux treize épreuves du corpus.

Tableau 1. Caractérisation des treize épreuves du corpus selon la grille à quatre critères

| Académie / Année        | Œuvre support                 | Bloom : niveaux supérieurs | Scallon : production singulière | De Ketele     | Prévisibilité |
|-------------------------|-------------------------------|----------------------------|---------------------------------|---------------|---------------|
| Rabat-Salé-Kénitra 2018 | La Boîte à Merveilles         | 30%                        | 20%                             | Artefactuelle | Niveau 1      |
| Laâyoune-Sakia 2018     | La Boîte à Merveilles         | 40%                        | 20%                             | Artefactuelle | Niveau 1      |
| Casablanca-Settat 2018  | La Boîte à Merveilles         | 50%                        | 20%                             | Artefactuelle | Niveau 2      |
| Tanger-Tétouan 2018     | La Boîte à Merveilles         | 30%                        | 20%                             | Artefactuelle | Niveau 2      |
| Fès-Meknès 2018         | La Boîte à Merveilles         | 44%                        | 22%                             | Artefactuelle | Niveau 2      |
| Béni-Mellal 2018        | DJDC + La Boîte à Merv.       | 44%                        | 11%                             | Artefactuelle | Niveau 2      |
| Rabat-Salé-Kénitra 2019 | La Boîte à Merveilles         | 20%                        | 20%                             | Artefactuelle | Niveau 1      |
| Rabat-Salé-Kénitra 2021 | La Boîte à Merveilles         | 37,5%                      | 12,5%                           | Artefactuelle | Niveau 1      |
| Tanger-Tétouan 2023     | Antigone                      | 40%                        | 20%                             | Artefactuelle | Niveau 2      |
| Tanger-Tétouan 2024     | La Boîte à Merveilles         | 30%                        | 20%                             | Artefactuelle | Niveau 2      |
| Fès-Meknès 2024         | La Boîte à Merveilles         | 60%                        | 20%                             | Artefactuelle | Niveau 2      |
| Rabat-Salé-Kénitra 2025 | Le Dernier Jour d'un Condamné | 40%                        | 20%                             | Artefactuelle | Niveau 1-2    |
| Casablanca-Settat 2025  | Le Dernier Jour d'un Condamné | 60%                        | 20%                             | Artefactuelle | Niveau 2      |

La lecture du tableau 1 fait apparaître cinq résultats principaux. Premièrement, la part des niveaux taxonomiques supérieurs de Bloom varie entre 20 % et 60 %, sans jamais dépasser cette borne supérieure. La moyenne sur l'ensemble du corpus est d'environ 40 %, ce qui indique que la majorité des questions relèvent des niveaux inférieurs (connaissance, compréhension, application). Le niveau de la synthèse au sens de Bloom est quasiment absent de l'étude de texte sur l'ensemble du corpus : il n'est mobilisé que dans la production écrite finale, où il est encadré par un format normé.

Deuxièmement, la part de la production singulière au sens de Scallon se révèle remarquablement stable sur les sept années couvertes par le corpus : elle se situe entre 11 % et 22 %. La moyenne sur l'ensemble du corpus est de 19 %. Cela signifie que 81 % des questions relèvent de la restitution ou de l'application, et seulement 19 % de la production singulière.



Cette proportion demeure stable indépendamment des évolutions observées au niveau de Bloom et du choix de l'œuvre support.

Troisièmement, l'ensemble des treize épreuves relève structurellement de l'évaluation artefactuelle au sens de De Ketele. Tous les textes supports sont extraits du programme étudié en classe, conformément au cadre prescriptif officiel. Aucune épreuve du corpus ne propose une situation évaluative authentique au sens fort de De Ketele.

Quatrièmement, selon le critère de prévisibilité du sujet de production écrite, les treize sujets se distribuent comme suit : quatre sujets au niveau 1 (entièrement prévisibles), neuf sujets au niveau 2 ou 1-2 (partiellement prévisibles), aucun sujet au niveau 3. Le constat est net et constant : aucune des treize épreuves du corpus ne propose un sujet véritablement ouvert et non préparé en classe. Cette absence est massive et constante sur les sept années couvertes par le corpus.

Ce résultat est conforté par l'enquête par entretiens. Sur les cinquante enseignants interrogés, quarante-trois confirment que les thèmes des sujets de production écrite se répètent d'une année à l'autre et qu'ils les ont préparés en classe avec leurs élèves. Cette convergence entre l'analyse documentaire et le témoignage des praticiens renforce la solidité du constat : la prévisibilité des sujets n'est pas seulement une caractéristique repérable dans les énoncés, elle correspond à une réalité des pratiques de préparation en classe.

Cinquièmement, aucune des treize épreuves n'évalue les compétences que les Orientations Pédagogiques de 2007 mentionnent pourtant comme constitutives du profil de sortie attendu : la réception et la production de l'oral, le travail en autonomie, la compréhension réflexive de textes nouveaux.

## **5. Discussion**

Les résultats présentés dans la section précédente permettent de dégager trois constats convergents : la stabilité diachronique du modèle évaluatif, son caractère structurellement artefactuel, et l'hypothèse de la rupture évaluative comme alternative au diagnostic du déficit linguistique.

### ***5.1. Une stabilité diachronique remarquable***

Le premier résultat empirique est la stabilité diachronique du modèle évaluatif sur la période 2018-2025. Cette stabilité est d'autant plus remarquable que la période couverte par notre corpus correspond à des transformations institutionnelles importantes du système éducatif

marocain : réformes curriculaires partielles, mise en œuvre du plan stratégique 2015-2030, débats sur la place du français à l'école. Or, en dépit de ces transformations institutionnelles, le format de l'épreuve, la structure des questions, la part de la production singulière et le degré d'authenticité de la situation évaluative présentent une remarquable continuité.

Cette stabilité s'explique, à notre sens, par la combinaison de plusieurs facteurs institutionnels. Le premier facteur est l'encadrement curriculaire par les Orientations Pédagogiques de 2007, qui imposent un canevas précis pour l'examen normalisé. Le deuxième facteur est la nature même de l'évaluation à fort enjeu social : un examen certificatif comme l'examen régional du baccalauréat est nécessairement standardisé pour garantir l'équité entre les candidats, et cette standardisation s'accommode mal de transformations rapides. Le troisième facteur est l'interdépendance entre l'évaluation et les pratiques d'enseignement : les enseignants préparent leurs élèves au format de l'examen, ce qui produit une cristallisation des pratiques autour des compétences effectivement évaluées.

Les nouveautés observées dans le corpus restent ponctuelles et n'affectent pas le cœur du modèle évaluatif. L'introduction d'une question intégrant la dimension plurilingue à Fès-Meknès 2024 (question sur « la kissaria »), la demande de justification de l'effet stylistique à Casablanca-Settat 2025, la codification croissante du format des questions de réaction : ces nouveautés méritent d'être notées, mais elles ne remettent pas en cause la part structurelle de la production singulière (qui demeure à 19 % en moyenne) ni le caractère artefactuel de la situation évaluative (qui vaut pour les treize épreuves). Le choix de l'œuvre support varie d'une académie à l'autre et d'une année à l'autre, mais cette variation s'opère à l'intérieur du même cadre prescriptif standard du programme officiel.

### ***5.2. Un modèle structurellement artefactuel : une caractérisation et non une critique***

Le deuxième résultat empirique est le caractère structurellement artefactuel du modèle évaluatif au sens de De Ketele. Il importe de préciser d'emblée la portée de ce résultat. Le terme « artefactuel », chez De Ketele, n'est pas péjoratif : il caractérise une configuration évaluative dans laquelle la situation est construite pour les besoins propres de l'évaluation, ce qui est précisément la condition de toute évaluation certificative à fort enjeu social. L'équité entre les candidats, la comparabilité des résultats et la fiabilité de l'évaluation exigent une dimension artefactuelle importante. Notre analyse ne suggère donc nullement que le dispositif évaluatif au lycée marocain devrait être transformé pour devenir authentique au sens fort de De Ketele.

Une telle transformation, qui supposerait par exemple de proposer aux élèves des textes inédits non étudiés en classe lors de l'examen certificatif, présenterait au contraire des inconvénients majeurs. Elle accentuerait probablement les inégalités scolaires, puisque seuls les élèves disposant d'un capital culturel familial fort pourraient affronter sereinement un texte inédit dans une situation à fort enjeu. Elle introduirait une double sanction pour les élèves les plus vulnérables, qui seraient confrontés à un objet d'évaluation qu'ils n'ont pas eu le temps de travailler en classe.

Notre analyse caractérise donc le dispositif évaluatif tel qu'il est, sans préjuger de ce qu'il devrait être. Ce qui importe pour notre analyse, ce n'est pas que le dispositif soit artefactuel en soi, mais que les compétences qu'il évalue diffèrent structurellement des compétences attendues à l'université.

### ***5.3. L'hypothèse de la rupture évaluative***

Le troisième résultat empirique permet de formuler explicitement l'hypothèse alternative que nous défendons. La rupture observée à l'entrée à l'université ne se réduirait pas, fondamentalement, à un déficit linguistique des bacheliers : elle tiendrait aussi à une dissociation entre deux régimes évaluatifs qui ne mesurent pas les mêmes compétences.

Le régime évaluatif du baccalauréat marocain, tel que notre corpus diachronique le documente, valorise structurellement la restitution de connaissances mémorisées sur des textes étudiés en classe, l'application de procédures connues, et la conformité à des formats normés sur des sujets dont les thématiques sont travaillées en classe à propos des œuvres au programme. Ce régime évaluatif produit, par ses propres mécanismes, des bacheliers qui maîtrisent ces compétences : ils savent mobiliser leurs connaissances sur les œuvres au programme, identifier des figures de style et transformer des phrases au discours indirect, rédiger un texte argumentatif conforme aux attendus formels du barème.

Le régime évaluatif universitaire, tel que les littéracies universitaires le caractérisent, attend des compétences fondamentalement différentes : la compréhension réflexive de textes nouveaux, la production singulière d'un point de vue argumenté, la problématisation située, l'engagement énonciatif, la maîtrise des codes discursifs disciplinaires, l'autonomie dans la mobilisation des ressources. Ces compétences ne sont pas évaluées par l'examen régional du baccalauréat. Elles ne sont donc pas non plus, par effet retour sur les pratiques d'enseignement, constituées comme objets d'apprentissage explicites au lycée.

Cette opposition doit toutefois être nuancée. La frontière entre les deux régimes n'est pas étanche : certaines évaluations universitaires demeurent elles aussi largement reproductives, tandis que des pratiques d'écriture singulière existent ponctuellement au lycée. Notre propos ne pose donc pas une dichotomie absolue, mais une tendance dominante de part et d'autre, suffisamment marquée pour être structurante.

Les bacheliers entreraient ainsi à l'université avec la maîtrise du français que le dispositif évaluatif antérieur a précisément valorisée : maîtrise en partie inadéquate aux attentes universitaires, non parce qu'elle serait déficitaire en elle-même, mais parce qu'elle ne mobilise pas les compétences que l'université attend.

Cette lecture est confortée par l'enquête exploratoire menée auprès de quarante enseignants universitaires (section 3.6). La totalité des répondants observe que les nouveaux bacheliers savent restituer des connaissances mais peinent à les analyser et à produire un raisonnement personnel, et rapporte ces difficultés, au moins en partie, à la manière dont les élèves ont été évalués au lycée. À l'inverse de la prévisibilité qu'ils jugent caractéristique du baccalauréat, les enseignants décrivent leurs propres évaluations comme largement imprévisibles et tournées vers les niveaux taxonomiques supérieurs : ce contraste, perçu des deux côtés, conforte l'hypothèse d'une discontinuité entre régimes évaluatifs.

Interrogés sur l'origine des difficultés, les enseignants ne les imputent qu'exceptionnellement au seul déficit de langue, la très grande majorité y voyant soit une rupture dans les façons d'évaluer et de penser le savoir, soit la conjonction des deux facteurs. Les rares répondants privilégiant le facteur linguistique enseignent presque tous la langue ou la communication, ce qui suggère que la perception de la difficulté dépend en partie de la discipline d'appartenance.

Cette analyse permet de comprendre pourquoi les dispositifs de remédiation linguistique déployés à l'université produisent des résultats limités (Mehdaoui, 2022). Ces dispositifs ne s'attaquent peut-être pas au problème pertinent. Le niveau B1, calibré sur les compétences communicatives générales du CECRL, ne mesure pas les compétences littéraires attendues à l'université. La remédiation linguistique conventionnelle ne peut donc pas combler une rupture qui n'est pas, fondamentalement, linguistique.

#### **5.4. Implications didactiques**

L'hypothèse de la rupture évaluative ouvre des perspectives didactiques précises pour repenser l'articulation secondaire-supérieur au Maroc. Trois voies complémentaires peuvent être envisagées.

La première voie consiste à enrichir les pratiques d'enseignement pendant l'année scolaire pour y intégrer des activités d'écriture singulière et de confrontation à des textes nouveaux, sans que ces activités soient nécessairement évaluées à l'examen certificatif. Le dispositif certificatif demeure tel qu'il est, garantissant l'équité de l'évaluation à fort enjeu, mais les pratiques pédagogiques quotidiennes s'élargissent pour préparer les élèves à des compétences qui ne sont pas directement évaluées.

La deuxième voie, qui nous paraît la plus prometteuse, consiste à introduire à l'université des dispositifs explicites d'acculturation aux littéracies universitaires. Ces dispositifs ne se conçoivent pas comme une remédiation linguistique mais comme une socialisation aux pratiques discursives disciplinaires. Cette voie s'inscrit dans les propositions élaborées dans la littérature francophone (Pollet, 2014 ; Boch & Rinck, 2010) et suppose une reconnaissance institutionnelle du caractère non strictement linguistique des difficultés à l'entrée à l'université.

La troisième voie consiste à construire des dispositifs explicites d'articulation entre les deux régimes évaluatifs, notamment par des dispositifs de transition en deuxième année du baccalauréat ou en première année universitaire. Ces dispositifs, qui restent à inventer dans le contexte marocain, pourraient s'inspirer des initiatives mises en place dans d'autres contextes francophones : programmes de formation à la rédaction universitaire, ateliers d'écriture, modules de méthodologie disciplinaire.

#### **Conclusion**

Cet article a présenté l'analyse diachronique d'un corpus de treize épreuves de l'examen régional de fin de première année du baccalauréat marocain en langue française, recueillies entre 2018 et 2025. L'analyse, conduite à l'aide d'une grille à quatre critères articulant Bloom, Scallon, De Ketele et la prévisibilité des sujets, a fait apparaître trois résultats convergents : la stabilité diachronique massive du modèle évaluatif sur sept années, son caractère structurellement artefactuel au sens de De Ketele, et l'écart systématique entre les compétences évaluées et celles que les Orientations Pédagogiques de 2007 mentionnent comme constitutives du profil de sortie attendu.

Articulés au cadre des littéracies universitaires, ces résultats conduisent à reformuler le diagnostic dominant : la difficulté des bacheliers tiendrait moins à un déficit de langue qu'à la dissociation entre deux régimes évaluatifs (un régime certificatif structurellement artefactuel au lycée, un régime universitaire qui suppose l'authenticité discursive et la production singulière). Posée comme hypothèse de travail, cette lecture éclaire les limites des dispositifs de remédiation purement linguistiques et ouvre des perspectives didactiques fondées sur l'articulation institutionnelle entre les deux cycles.

Plusieurs limites de cette étude doivent être reconnues. Le corpus, bien qu'étendu sur sept années et plusieurs académies régionales, reste partiel : son extension à d'autres académies (notamment les régions amazighophones) et à des sessions de rattrapage offrirait l'occasion d'éprouver la stabilité de nos résultats. L'enquête conduite auprès des enseignants universitaires, de nature exploratoire, repose sur un échantillon de convenance de taille limitée. Enfin, notre étude documente principalement le profil de sortie du lycée ; la caractérisation des pratiques évaluatives universitaires marocaines, à partir de productions et d'épreuves réelles, reste à mener dans le cadre d'une recherche ultérieure de plus grande ampleur.

Ces limites dessinent les prolongements de recherche que nous envisageons. À court terme, l'extension du corpus aux épreuves de 2026 et aux sessions de rattrapage consoliderait l'analyse diachronique. À moyen terme, une étude empirique du versant universitaire, portant sur les attentes et les pratiques évaluatives effectives à l'entrée dans les filières, viendrait compléter le diagnostic et mettre pleinement à l'épreuve l'hypothèse de la rupture évaluative. À long terme, l'élaboration de propositions didactiques concrètes pour construire l'articulation institutionnelle entre les deux cycles, en concertation avec les acteurs du terrain marocain, constituerait une retombée pratique de ce travail.

## Bibliographie

ALLAL L. (1979). « Stratégies d'évaluation formative : conceptions psychopédagogiques et modalités d'application ». In L. Allal, J. Cardinet & P. Perrenoud (éd.), *L'évaluation formative dans un enseignement différencié*. Berne : Peter Lang, p. 153-183.

ALLAL L. (2007). « Régulations des apprentissages : orientations conceptuelles pour la recherche et la pratique en éducation ». In L. Allal & L. Mottier Lopez (éd.), *Régulation des apprentissages en situation scolaire et en formation*. Bruxelles : De Boeck, p. 7-23.

ANDERSON L. W. & KRATHWOHL D. R. (éd.) (2001). *A Taxonomy for Learning, Teaching, and Assessing : A Revision of Bloom's Taxonomy of Educational Objectives*. New York : Longman.

BLOOM B. S. (1956). *Taxonomy of Educational Objectives : The Classification of Educational Goals. Handbook 1 : Cognitive Domain*. New York : David McKay.

BOCH F. & RINCK F. (éd.) (2010). « Énonciation et rhétorique dans l'écrit scientifique ». *Lidil*, n° 41. Grenoble : ELLUG.

CARDINET J. (1986). *Pour apprécier le travail des élèves*. Bruxelles : De Boeck.

CHAMI M. (2004). « Quel français aujourd'hui à l'université marocaine ? ». *Revue Langues et Littératures*, vol. XVIII, p. 63-69.

DE KETELE J.-M. (2010). « Ne pas se tromper d'évaluation ». *Revue française de linguistique appliquée*, vol. XV, n° 1, p. 25-37.

DELCAMBRE I. & LAHANIER-REUTER D. (2010). « Les littéracies universitaires : influence des disciplines et du niveau d'étude dans les pratiques de l'écrit ». *Forumlecture.ch*, n° 3, p. 1-16.

DELCAMBRE I. & LAHANIER-REUTER D. (2012). « Littéracies universitaires : présentation ». *Pratiques*, n° 153-154, p. 3-19. Metz : CRESEF.

EL GOUSAIRI A. (2019). « Recherche et formation en didactique du français à l'université marocaine : tensions, obstacles et problèmes de posture ». *Le français aujourd'hui*, n° 204, p. 81-96.

HADJI C. (1989). *L'évaluation, règles du jeu : des intentions aux outils*. Paris : ESF Éditeur.

LANDIS J. R. & KOCH G. G. (1977). « The measurement of observer agreement for categorical data ». *Biometrics*, vol. 33, n° 1, p. 159-174.

LEA M. R. & STREET B. V. (1998). « Student Writing in Higher Education : An Academic Literacies Approach ». *Studies in Higher Education*, vol. 23, n° 2, p. 157-172.

LILLIS T. & SCOTT M. (2007). « Defining Academic Literacies Research : Issues of Epistemology, Ideology and Strategy ». *Journal of Applied Linguistics*, vol. 4, n° 1, p. 5-32.

MEHDAOUI M. (2022). « Le français à l'université marocaine. Difficultés d'apprentissage : le cas des anaphores ». *Revue Texte et Discours en Français Langue Étrangère (TDFLE)*, n° 80, p. 1-14.

MESSAOUDI L. (2016). « La fracture linguistique dans l'enseignement scientifique au Maroc : quelles remédiations ? ». *Langues, Cultures et Sociétés*, vol. 2, n° 1, p. 69-85.

MINISTÈRE DE L'ÉDUCATION NATIONALE DU ROYAUME DU MAROC (2007). *Les Orientations Pédagogiques générales et les programmes spécifiques pour l'enseignement de la langue française dans le cycle secondaire qualifiant*. Rabat : Direction des Curricula.

POLLET M.-C. (2014). *L'écrit scientifique à l'aune des littéracies universitaires. Approches théoriques et pratiques*. Namur : Presses universitaires de Namur.

REUTER Y. (2012). « Les didactiques et la question des littéracies universitaires ». *Pratiques*, n° 153-154, p. 161-176. Metz : CRESEF.

SADIQUI M. (2016). « Didactique du français au lycée au Maroc. Entre didactique des langues et didactique des disciplines scientifiques ». *Le français à l'université*, 21-01.

SCALLON G. (2004). *L'évaluation des apprentissages dans une approche par compétences*. Montréal : Éditions du Renouveau pédagogique.