

La validité de l'évaluation sommative du FLE en Algérie :
Les questionnaires de la compréhension de l'écrit au cycle moyen
The validity of the summative evaluation of French as a foreign language in Algeria: the questions of reading comprehension in the middle cycle

LAIDOUDI Assia^{1,*}, KHADRAOUI Errime²

assia.laidoudi@univ-msila.dz

¹Research laboratory: Literature Teaching Strategies: A Notion in Motion, University Of Batna 2 (Algérie)

Receipt date: 06 / 08 / 2019; Acceptance date: 29/ 09/ 2019; Publishing date: 31 / 10 / 2020

Abstract. This article, which focuses on the summative evaluation of French Foreign Language, examines the validity of content of questionnaires designed to assess the comprehension of texts in the middle cycle. The results obtained from the analysis of the questionnaires collected revealed that the items developed by the teachers-designers do not contribute to proving the competence of the learners to construct the meaning of the texts read. These questionnaires are therefore not sufficiently valid to evaluate this competence.

Keywords. validity; summative evaluation; reading; reading comprehension; French as a Foreign Language.

Résumé. Le présent article qui porte sur l'évaluation sommative du FLE vise à étudier la validité du contenu des questionnaires destinés à évaluer la compréhension de l'écrit au cycle moyen. Les résultats obtenus, suite à l'analyse des questionnaires recueillis ont révélé que les items élaborés par les enseignants-concepteurs ne contribuent pas à apporter des preuves sur la compétence des apprenants au niveau de la construction du sens des textes lus. Ces questionnaires ne sont pas donc suffisamment valides pour évaluer ladite compétence.

Mots-clés .validité ; évaluation sommative ; lecture ; compréhension de l'écrit ; FLE.

*corresponding author

I. Introduction.

Il n'est pas aisé d'évaluer les compétences des apprenants en langue étrangère ; c'est une réalité que les théoriciens, les spécialistes -à l'instar de Hadji (2015) et Rey (2015)- ainsi que les praticiens ne peuvent nier. Cette complexité de l'évaluation est particulièrement due à la difficulté de concevoir des outils suffisamment valides pour mesurer ce que l'on veut mesurer. La question qui se pose à ce niveau concerne alors la validité de la mesure de ces diverses compétences notamment en l'absence d'outils appropriés. Par validité, nous entendons à l'instar de plusieurs chercheurs (De Landsheere, 1976, Noël-Jothy et Sampsonis, 2006, Tagliante, 2005, Mègre et Patrick, 2015...) la capacité du test (toute évaluation dans un sens large) à mesurer ce que nous voulons mesurer. Cette définition de la validité telle que nous l'avancions se rapporte à un type particulier de validité : la validité du contenu.

Il importe de préciser que cette question de validité se pose notamment avec certaines compétences. Ainsi, les compétences expressives, tant à l'oral qu'à l'écrit, semblent-elles plus privilégiées par la recherche que les compétences de compréhension pour plusieurs raisons. Outre le fait que l'évaluation de la compétence expressive intègre celle de la compréhension de la consigne, la raison indéniable renvoie à la complexité du concept de compréhension et à la diversité des stratégies permettant de développer cette compétence chez les apprenants. S'agissant de la compréhension de l'écrit, la littérature scientifique en la matière aborde également la notion de lecture. Ce concept est d'une complexité terminologique connue de tous. En effet, il recouvre une infinité d'acceptions allant du simple décodage des mots à l'interprétation des textes lus (Cuq, 2003, Robert, 2002). Dans notre étude, nous l'évoquons dans le sens de compréhension et afin d'éviter toute confusion, nous nous servirons du concept lecture-compréhension, qui sous-entend une distinction des autres lectures dont un texte écrit est susceptible.

En plus de ces éléments, les processus et les stratégies mis en œuvre par les apprenants pour accéder au sens ne sont pas clairement définis, ce qui rend plus difficile la conception d'un instrument de mesure permettant de refléter la compétence de compréhension chez l'apprenant. En effet, il est difficile de concevoir des activités qui permettent de collecter des preuves pertinentes sur la compétence de lecture des évalués.

Tous ces éléments ont amené plusieurs auteurs, tels que Beaume (1988) et Magherbi, Rocher, Gyselinck, Troseille, Tardieu (2009), à consacrer leurs écrits à la validité de l'évaluation de cette compétence. Notre étude se rapproche du même sujet : elle porte sur cette compétence pour étudier la validité de son évaluation dans le système éducatif algérien, plus particulièrement au cycle moyen. Elle sera centrée sur l'analyse des items proposés en évaluation sommative du FLE pour juger leur capacité à refléter la compétence de lecture-compréhension. Nous répondrons alors à la question suivante : **le questionnaire proposé en évaluation sommative permet-il de mesurer la compétence de compréhension de l'écrit chez les apprenants du FLE au cycle moyen ?**

Cette question devient encore plus intéressante si nous considérons le contexte actuel où une nouvelle approche a été intégrée créant un changement remarquable au niveau de l'évaluation. Avec l'intégration de l'Approche Par les Compétences, abrégée en APC dans le reste de notre étude, qui a rénové les pratiques pédagogiques et évaluatives, l'évaluation des diverses compétences se complexifie davantage et il devient d'une énorme difficulté de concevoir des outils de mesure suffisamment valides pour révéler les compétences des évalués, notamment que jusqu'ici l'évaluation scolaire porte particulièrement sur des items de restitution et

d'application qui deviennent peu significatifs avec la nouvelle approche. Avant de nous intéresser à ce qui doit être évalué dans la lecture-compréhension, nous aborderons des précisions sur la notion de validité, qui a suscité plusieurs réflexions évoquées dans ce qui suit.

I.1. La validité : une notion complexe

La notion de validité, un concept-clé de notre étude, a intéressé plusieurs auteurs qui ont formulé des définitions variées. Nous les classons dans deux grandes catégories : les définitions qui ont conçu la validité en tant que qualité du test et celles qui l'ont présentée comme une caractéristique de l'interprétation de la mesure.

Pour la première catégorie, des auteurs tels que Buckingham et al. (1921), Kelley (1927) ont présenté la validité comme une propriété du test, une qualité de celui-ci qui détermine sa capacité à mesurer (évaluer) ce que nous voulons mesurer. Il s'agit de préciser si les instruments de mesure retenus donnent une description précise des compétences des apprenants évalués. Contrairement à cette première tendance réfutée, les auteurs André, Loye et Laurencelle, reprennent Cronbach (1971) qui affirme que « *l'on ne valide pas le test, mais l'interprétation des mesures qu'il fournit* » (2015, 132). Autrement dit, la validité ne se rapporte plus au test qui mesure les compétences, mais aux résultats auxquels ce test aboutit. Cette deuxième tendance, qui a été fructueuse car elle a donné lieu à plusieurs modèles de validation connus et appliqués comme celui de Kane (2006), est également contestée par Borsboom, Mellenbergh et Van Heerden (2004) qui estiment que le concept de validité « *était correctement formulé, par exemple, par Kelley (1927, p. 14) lorsqu'il a déclaré qu'un test est valide s'il mesure ce qu'il prétend mesurer* » (2004, 1061).

Une lecture de deux tendances présentées ci-dessus nous mène à affirmer que la notion de validité est globale. En effet, elle comprend les deux tendances supra citées dans la mesure où pour aboutir à des résultats valides, il est nécessaire que les instruments de mesure (outils évaluatifs) sélectionnés soient valides. Ceci dit, avant de valider les résultats, il faut valider d'abord le test qui permet de les collecter.

Dans ce travail, nous nous intéressons à la validité du questionnaire, un instrument de mesure retenu pour évaluer la compétence de lecture-compréhension chez les apprenants algériens.

I. 2. Pour évaluer la lecture-compréhension...

La compétence de lecture-compréhension mobilise des processus cognitifs complexes, des stratégies diverses et des connaissances linguistiques et sociales variées (Cuq, 2003). Contrairement aux compétences expressives où ces stratégies mises en œuvre par les apprenants sont prises en considération en évaluation, la compétence de compréhension requiert une évaluation du produit de ce processus qui constitue une preuve de compétence. Dans ce contexte, Tagliante avance que « *les stratégies de lecture qui sont mises en œuvre ainsi ne peuvent être mesurées, c'est uniquement leur résultat qu'on évalue* » (2005, 74). Mais, en quoi consiste le produit de la lecture-compréhension des textes ?

Pour déterminer ce produit, nous nous référons à la définition que Le Ny donne de la compréhension pour en expliciter le produit. L'auteur avance :

La compréhension d'un énoncé, c'est-à-dire son traitement cognitif dans l'esprit/cerveau d'un compreneur, y produit causalement, d'abord une suite de sous-processus et d'états mentaux transitoires, qui aboutissent finalement à la construction d'une représentation sémantique terminale, mentale, qui est composée et structurée : c'est cette représentation qui constitue le sens de l'énoncé(2005, 104).

La compréhension consiste en un ensemble de processus de traitement cognitif des structures linguistiques qui constituent le texte pour aboutir à la construction de représentations sémantiques. Ce sont alors ces représentations sémantiques, produits de cette interaction texte-lecteur, qui constituent des éléments à mesurer et apportent des preuves de compréhension.

Rejoignant cette perspective qui évalue la compréhension à travers les représentations sémantiques, Kintsch (1988) - cité par Magherbi, Rocher, Gyselinck, Trosseille et Tardieu (2009)- a proposé un modèle comprenant trois niveaux de représentations qui ont permis d'établir trois catégories de questions qui correspondent aux niveaux de représentations suivants :

- **Le niveau de surface** : ce niveau qui est le plus superficiel correspond à l'analyse des structures du texte. Les questions de compréhension correspondantes sont orientées aux informations directement repérées : thème, idée générale...
- **La base de texte prépositionnelle** : ce deuxième niveau qui est centré sur la décomposition des propositions du texte pour accéder au sens, ce qui requiert l'investissement des ressources déjà maîtrisées.
- **Le modèle de situation** : dans ce dernier niveau, le compreneur ne se limite pas au traitement des informations incluses dans le corps du test dans la mesure où des informations extratextuelles sont invoquées pour la construction de représentations plus profondes.

Ces trois niveaux, très utiles à la conception des tests, peuvent être rapprochés aux trois catégories proposées par Tagliante (2005) :

- **La compréhension littérale** : les questions qui sont regroupées sous cette première catégorie sont posées sur les informations explicites du texte.
- **L'interprétation** : sous cette deuxième catégorie, nous rangeons les questions qui permettent de recueillir les interprétations conçues de ce qui est lu.
- **L'analyse et la critique** : cette troisième catégorie suscite les réactions et les attitudes du lecteur-compreneur suite au traitement des informations des deux univers (textuel et extratextuel).

Une lecture studieuse des catégories proposées par les deux chercheurs révèle un lien étroit entre elles. En effet, les deux premiers niveaux proposés par Kintsch (le niveau de surface et la base de texte prépositionnelle) correspondent au premier niveau conçu par Tagliante (la compréhension linéaire). Pour ce qui est du dernier niveau de Kintsch (le modèle de situation), il est identique au deuxième niveau de Tagliante (l'interprétation). Enfin, le dernier niveau de Tagliante (analyse) est le plus profond dans la lecture-compréhension des textes lus.

I. 3. Plusieurs dispositifs pour évaluer la lecture-compréhension

Dans les pratiques courantes, la maîtrise de la lecture-compréhension est prouvée par la réussite des questions posées sur un texte lu par les apprenants, qui présente une situation donnée. Ainsi, le questionnaire continue-t-il à être le dispositif le plus répandu en dépit de certains biais qui altèrent la validité du jugement porté à l'égard des compétences mesurées (Beaume, 1988). En effet, dans cette méthode de recueil des preuves de compétence, la rédaction des items nécessite une prudence et une réflexion pour éviter l'ambiguïté qui fausse les données obtenues sur la capacité mesurée. De même, nous aboutissons à une évaluation de

la compréhension des consignes dont la compréhension conditionne la capacité de sélectionner les réponses correspondantes aux questions proposées.

Mais, il est important de préciser que d'autres dispositifs de mesure assurent l'évaluation de cette compétence de lecture-compréhension. Dans ce contexte, Ecalle (2010) donne un aperçu très utile sur ces dispositifs. Bien que ceux-ci soient essentiellement conçus pour évaluer la compréhension des phrases, ils sont jugés généralisables pour accéder au sens des textes étant donné que la compréhension des phrases assure celle du texte qu'elles constituent. Entre autres, citons :

- **Le modèle de Maeder (2006)**

Dans son Test de Compréhension Syntaxique (TCS), des activités d'appariement et de représentation sont intégrées. Les élèves sont appelés à relier les énoncés ayant un sens identique ou à concrétiser des extraits lus pour prouver leur compréhension.

- **Le modèle de Khomsi (1999)**

Centré sur les images, son test Lecture de Mots et Compréhension- Version corrigée (LCR-R) consiste à repérer parmi une série d'images proposées celle qui correspond à ce qui est lu.

Ce modèle s'approche de celui conçu par Lecoq (1996) où le repérage de l'image s'appuie sur la mémorisation des structures.

En dépit de leur diversité, les dispositifs cités par Ecalle ne peuvent être appliqués dans le cadre d'une évaluation sommative où le nombre des apprenants est considérable, ce qui rend le questionnaire le modèle le plus pratique.

2. Méthodologie

Dans cette étude, nous avons opté essentiellement pour la démarche analytique qui consiste à étudier les examens de FLE proposés au cycle moyen. Cette analyse sera renforcée par un volet descriptif pour donner un aperçu général sur leur structure.

Dans le système éducatif algérien, l'évaluation du FLE demeure écrite sous forme d'examens (devoirs et compositions) répartis tout au long de l'année scolaire. Tout examen est constitué de deux parties : la première est consacrée à la compréhension du texte écrit proposé alors que la seconde expose une situation pour la rédaction de textes (souvent du même type) par les apprenants.

Pour la première partie sur laquelle notre étude porte, elle consiste en un questionnaire, le dispositif qui persiste depuis plusieurs années. En effet, plusieurs questions sont proposées : certaines sont relatives au texte lu alors que d'autres concernent les points grammaticaux abordés durant la séquence d'enseignement/apprentissage (une unité pédagogique).

Pour étudier la validité du questionnaire, il est important d'étudier si les questions qui le constituent évaluent la compétence ciblée. Rappelons que nous nous intéressons dans le cadre de notre étude à la validité du test, la capacité de ses items à évaluer la lecture-compréhension. Pour ce faire, les items des examens recueillis sont d'abord recensés puis répartis sur les catégories que nous avons proposées en nous appuyant sur les modèles de Kintsch et de Tagliante (cf. 2.3). Intégrées dans le volet compréhension, elles s'opposent aux items-ressources qui assurent la vérification des savoirs et des savoir-faire grammaticaux (vocabulaire, syntaxe, conjugaison et orthographe).

Ainsi, les catégories suivantes sont retenues : Ressources, Niveau de surface, Base de texte prépositionnelle, Interprétation, Analyse et Critique.

Sous la première catégorie (Ressources), nous retenons les questions de maîtrise grammaticale (fonctions et natures des mots, identification des rapports logiques, reconnaissance des modes et des temps verbaux...). La deuxième catégorie intègre les questions dont les réponses sont directement extraites du texte (les personnages, les dates, le temps et le lieu des événements, les circonstances, les qualités, les défauts...). S'agissant de la troisième, elle inclut les questions où les ressources sont exploitées pour trouver les réponses appropriées (repérage de causes à travers les articulateurs exprimant ce rapport, identification des habitudes grâce aux temps verbaux, ordre des événements à travers les connecteurs chronologiques...). Les items de la catégorie Interprétation exigent une interprétation des informations que le texte comporte pour des réponses prouvant sa compréhension (déduction de causes ou de conséquences qui ne sont pas abordées, proposition de solutions en cohérence avec le texte...). Enfin, la dernière catégorie (Analyse) englobe les questions qui suscitent la réaction du compreneur (position de l'auteur envers un sujet traité, jugement à l'égard des idées du texte...).

Il est nécessaire d'établir une distinction entre deux catégories parmi celles que nous avons listées : Ressources et Base de texte prépositionnelle. En effet, les deux intègrent les savoirs et les savoir-faire, mais à des fins différentes. Ainsi, la première catégorie (Ressources) évalue la maîtrise sans apporter des preuves sur la compréhension du texte lu. Cependant, la seconde vérifie la capacité de l'apprenant à se servir de ces outils linguistiques pour accéder au sens.

Notre corpus recueilli est constitué d'une quarantaine d'examens officiels et trimestriels de divers établissements : dix sujets sont retenus pour chaque niveau. A ce niveau, nous tenons à préciser que les examens (trimestriels et officiels) du cycle moyen en Algérie obéissent à des critères précis et communs à tous les établissements du territoire national. Ces critères se rapportent au nombre et à la nature des questions, aux points de langue intégrés, aux textes retenus...

Il en est pour la structure des tests qui est identique à tous les niveaux du cycle : un questionnaire (sur un texte proposé) suivi d'une situation de rédaction.

Le nombre des items qui constituent le questionnaire varie de 6 à 10, voire plus, selon les objectifs des concepteurs, la longueur du texte et la note attribuée. Outre leur nombre, leur nature est également diverse, car ils comprennent des questions ouvertes ou fermées, des questions à choix multiples, des activités d'appariement (association), des textes lacunaires...

La variation de la nature des items est nécessaire pour répondre à l'hétérogénéité des apprenants qui résolvent dans leur examen des activités adéquates à leurs stratégies et styles d'apprentissage.

3. Résultats et discussion

Suite à l'analyse des items que les questionnaires visant la mesure de la compétence de lecture-compréhension, nous avons obtenu les données que nous exposerons dans ce qui suit. Dans un premier lieu, nous jugeons utile de présenter les proportions de deux éléments : Ressources et le volet Compréhension (Niveau de surface, Base de texte prépositionnelle,

Interprétation, Analyse). En effet, la confrontation de ces deux éléments apporte quelques informations permettant de répondre à la question principale de notre étude sur la validité des questionnaires proposés

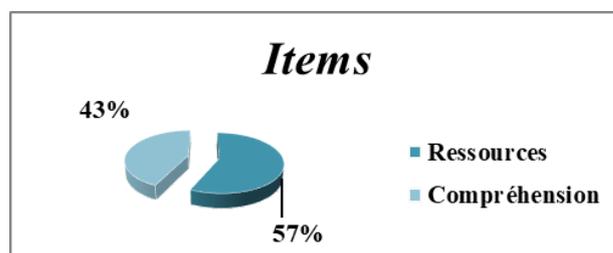


Figure (1) : Proportion Ressources/Compréhension

Cette première figure nous permet de constater que les questionnaires de compréhension qui constituent les tests du cycle moyen consacrent plus d'items aux ressources (savoirs et savoir-faire grammaticaux) .

Dans une deuxième figure, nous exposons les diverses catégories en détaillant le volet Compréhension. Ceci nous permet la confrontation des catégories du volet Compréhension pour étudier l'importance accordé à chacun des niveaux de compréhension.

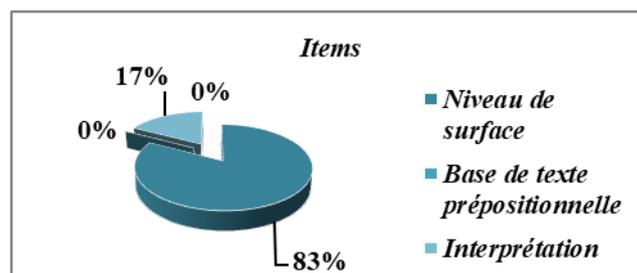


Figure (2) : la répartition des catégories du volet Compréhension

Tout comme la première figure, la deuxième apporte des informations sur l'importance de chacun des niveaux de compréhension. En effet, le premier niveau de Kintsch (Niveau de surface) se trouve en première position avec 115 items. La deuxième position est occupée par le niveau (Interprétation) dans lequel nous avons rangé 24 items. Le niveau de Kintsch (Base de texte prépositionnelle) et le niveau (Analyse) inspiré des travaux de Tagliante (2005) ne sont pas concrétisés dans les items des questionnaires analysés.

Ces données suscitent un nombre de questionnements auxquels nous tenterons de répondre dans notre analyse. Ceux-ci concernent notamment le privilège accordé aux ressources dans les questionnaires de compréhension et à certains niveaux de compréhension. En y répondant, nous préciserons si nous évaluons les connaissances des apprenants ou leurs compétences. De même, nous expliquerons les raisons de la préférence des questions du (Niveau de surface).

L'analyse des questionnaires collectés révèle que le taux des items qui portent sur les ressources est clairement plus important que celui du volet Compréhension. Ce premier constat, qui altère considérablement la validité du dispositif mis en œuvre, mérite des éclaircissements étant donné que les raisons de la prédominance des savoirs et savoir-faire grammaticaux doivent être expliquées.

3.1. Ressources et Compréhension

La priorité accordée aux items centrés sur les ressources révèle un enseignement décontextualisé de la grammaire en classe de FLE. Il est connu que l'APC prône un enseignement contextualisé de la grammaire de façon à conduire les apprenants à intégrer leurs savoirs et savoir-faire dans des productions concrètes (représentations sémantiques au niveau de la compréhension et énoncés rédigés au niveau de la production). Bien qu'elle soit adoptée dans le système éducatif algérien, il semble qu'il y ait un décalage entre pratiques pédagogiques et pratiques évaluatives. Autrement dit, elle n'est pas appliquée à bon escient vu que ce qui est évalué est le reflet de ce qui est enseigné.

En plus de la difficulté d'application de l'APC décelée à travers l'évaluation, le taux important accordé aux ressources assure que les concepteurs de tests n'ont pas la formation requise pour concevoir des examens révélateurs des compétences effectives de leurs apprenants. Ce problème de formation se confirme si nous nous penchons sur les examens élaborés avant l'intégration de la nouvelle approche. En effet, ceux-ci manifestent la réticence des enseignants du cycle qui demeurent prisonniers de la structure traditionnelle d'examens. D'ailleurs, les changements restreints introduits n'en touchent que la forme. Ainsi, le modèle classique, appliqué depuis les pédagogies précédentes, subdivise les items en trois parties : un questionnaire de compréhension, un questionnaire de grammaire et une consigne de rédaction. En préconisant un enseignement contextualisé de la grammaire, le questionnaire de grammaire, intitulé *le fonctionnement de la langue*, a été omis pour plus que des items supplémentaires soient consacrés à la compréhension. D'ailleurs, les connaissances ne sont pas évaluées, car elles constituent une composante de la compétence. À travers une évaluation de celle-ci, c'est la capacité des apprenants à transférer leurs connaissances en tenant compte des paramètres d'une situation de communication donnée qui est jugée.

Incapables de se détacher du modèle classique des examens, les enseignants continuent à évaluer les connaissances grammaticales des apprenants en les insérant dans la rubrique *Compréhension* et ce, à travers des phrases tirées du support textuel proposé.

La prise en compte de ces données nous encourage à contester la validité des questionnaires conçus pour évaluer la lecture-compréhension. En effet, le modèle appliqué à la conception des examens est plus valide avec la pédagogie par objectifs (PPO). Bien que des cas pareils soient souvent imputés à la formation des enseignants, le manque de dispositifs et d'instruments de mesure suffisamment valides pour évaluer exactement les diverses compétences ne peut être nié. En effet, les recherches nombreuses en évaluation scolaire n'ont pas encore répondu à tous les questionnements soulevés par les praticiens en évaluation scolaire.

Notre premier jugement sur la validité du questionnaire dans les épreuves de lecture-compréhension peut être contesté dans la mesure où une évaluation des ressources est justifiée dans le cadre de l'APC. Mais, cette évaluation des ressources donne-t-elle des preuves sur la compétence des évalués ? Les items sélectionnés évaluent-ils la compétence que nous voulons évaluer ?

Pour répondre à ces questions, nous nous servons d'exemples tirés des examens qui forment notre corpus. Il est important de mentionner que nous ne focalisons pas notre analyse sur la correction linguistique des consignes étant donné que c'est leur aspect argumentatif qui est

utile à notre étude. Les ressources évaluées englobent des savoirs et des savoir-faire en vocabulaire, grammaire, conjugaison et orthographe.

En premier lieu, nous proposons les premières questions extraites d'un examen destiné aux élèves de 4^{ème} année moyenne.

b)-« Allez, achève-la, ce n'est pas la peine de la faire souffrir, dit mon père. » (1pt)

Le verbe introducteur de parole exprime :

-une demande - une déclaration - une interrogation

7)- Indique la nature grammaticale de chacune des expressions soulignées dans le texte. (1.5)

8)-Transforme au discours indirect :

Mon père dit : « La mésange à tête noire est un oiseau ami de l'homme. Elle ne touche pas à ses fruits. »

9)-Forme des adverbes à partir des adjectifs suivants :

Méchant, nuisible, prudent.

Le premier item qui évalue la compétence lexicale est centré sur le sens du verbe introducteur souligné. Pour le concepteur, il est étroitement lié au texte étant donné que la phrase analysée en est extraite. Cependant, il est loin de donner des preuves sur la compréhension du texte lu. Il en est de même pour le deuxième item dont l'objectif est la vérification de la maîtrise du métalangage, les mots soulignés dans le texte étant de trois natures diverses (adjectif qualificatif, complément du nom, proposition relative). Il est connu que les diverses pédagogies souvent restrictives en matière de l'intégration du métalangage grammatical en classe de FLE. Pour le troisième item, il s'agit d'une question d'application permettant aux apprenants de prouver leur compétence. Enfin, le dernier item porte sur la formation des adverbes à partir des adjectifs qualificatifs proposés.

Ces quatre items permettent de vérifier les savoirs et les savoir-faire des apprenants sans donner un aperçu sur leur compétence en compréhension-lecture. Cette incapacité des items proposés en compréhension à mesurer ladite compétence est un indice du manque de validité du questionnaire élaboré.

Après avoir étudié les items de la 4^{ème} année, nous considérons dans ce qui suit les consignes que comporte l'examen de 3^{ème} année moyenne.

3. A partir du verbe « localiser », forme le nom d'agent et le nom d'action de la même famille

Nom d'agent : (1pt)

Nom d'action :(1pt)

4. Mets la phrase suivante au discours indirect : (1pt)

Les parents de la petite déclarent : « notre fille a disparu dimanche matin. »

5. Transforme la phrase suivante à la forme passive : (1pt)

Les éléments de la brigade ont arrêté l'adolescente.

.....

6. Accorde correctement les participes passés entre () : (3pts)

Les enquêteurs sont (arriver)à la maison de l'adolescente. Ils ont (retrouver).....le corps de la victime dans un sac et ils l'ont (ramener).....au médecin législateur.

Tout comme dans le premier exemple, dans celui-ci quatre items sont conçus pour évaluer la lecture-compréhension. Le premier item évalue la capacité des apprenants à dériver des noms à partir de verbes proposés. Cette activité relève d'un savoir ou d'un savoir-faire, mais elle ne se rapporte aucunement à une compétence. Si les verbes sont déjà traités dans le cadre du cours, cet item sera une évaluation de savoirs et sa résolution nécessite la restitution de ce qui a été appris. Dans le cas échéant, il sera question d'une évaluation de savoir-faire qui requiert l'application de ce qui a été appris. Le troisième item relève également de l'application : la démarche de transformation est appliquée à une nouvelle phrase. Enfin, le dernier item est complètement dissocié du texte dans la mesure où les phrases proposées sont décontextualisées.

Pour ce deuxième questionnaire, nous aboutissons au même résultat que celui du premier. Ainsi, ses items ne permettent pas de mesurer la capacité des apprenants à comprendre le support proposé. Ils ne donnent pas de preuves convaincantes sur la compétence évaluée. Nous poursuivons notre analyse avec des exemples extraits d'un examen de la 2ème année moyenne.

4. *Relève deux mots du champ lexical de « l'eau ». 2pts*

5. *Ecris au présent de l'indicatif : 4pts*

- *Les loups (dévorer)l'agneau.*

- *Les loups (être)dangereux.*

- *Les agneaux (avoir)soif.*

- *Les loups (finir)leur repas.*

6. *Relève la morale de cette fable.*

7. *« Ne discute pas avec moi, ordonna le loup en montrant ses dents. »*

Remplace le verbe de parole souligné par :

Demanda/questionna/supplia. (Entoure la bonne réponse). 1pt

Comme les deux examens étudiés ci-dessus, celui-ci porte sur des savoirs et des savoir-faire. En effet, la première consigne aboutit à une évaluation des connaissances lexicales des apprenants : elle vise à évaluer les savoirs. La deuxième évalue un savoir-faire, celui de la conjugaison des verbes en suivant les règles apprises pendant le cours. La quatrième question relative au vocabulaire vérifie les connaissances des apprenants (synonymie). Nous n'aborderons pas à ce niveau la troisième question qui est rangée parmi les questions de compréhension.

Nous exposerons en dernier lieu les dernières consignes proposées dans une évaluation destinée aux élèves de 1ère année moyenne

8- *Ponctue les phrases suivantes*

Quels sont les bienfaits de l'activité physique ()

Chaque semaine () je fais du footing ()

9- *Relève du texte un mot de liaison (un connecteur d'énumération).*

10- *Souligne la proposition subordonnée relative et entoure l'antécédent.*

La marche est une activité sportive qui permet d'éviter la prise de poids.

Ce dernier questionnaire soutient le jugement que nous avons émis sur le manque de validité des items proposés pour évaluer la lecture-compréhension. Tout comme les précédents, ils vérifient les savoirs (connaissances déclaratives) et les savoir-faire (connaissances procédurales) des apprenants qui ne doivent pas être confondus avec la compétence qui « *ne se réduit jamais à des connaissances procédurales codifiées et apprises comme des règles, même si elle s'en sert lorsque c'est pertinent* » (Brossard, 1999. In Médioni, 2016, 31).

Il est important de préciser que ce n'est pas l'évaluation des ressources dans le contexte scolaire que nous contestons. C'est notamment leur évaluation dans le cadre d'une évaluation sommative qui ne nous paraît pas fondée. En effet, les ressources doivent être évaluées dans le cadre de l'évaluation formative pour une remédiation aux lacunes constatées. Ce moment est d'ailleurs le plus propice pour la vérification du degré de maîtrise de ce qui est appris.

Que l'évaluation sommative intègre la vérification de la maîtrise des ressources reflète une négligence du rôle de l'évaluation formative, qui est une pratique préconisée vu ses avantages indéniables souvent évoqués.

Il est nécessaire de ne pas oublier que l'APC privilégie l'intégration des connaissances. Dans ce contexte, Delory considère que « *les connaissances grammaticales, au lieu d'être enseignées pour elles-mêmes – en référence à une discipline scientifique, la linguistique – sont intégrés à une compétence sociale de communication* » (2002, 22).

En somme, cette première partie de notre analyse contribue à répondre à la problématique de notre étude. En effet, nous pouvons affirmer suite à une analyse de notre corpus que les examens du cycle moyen ne sont pas suffisamment valides pour mesurer la compétence de lecture-compréhension chez les apprenants du cycle moyen vu la centration sur les ressources au détriment de la compréhension des supports.

3.2. Niveaux de la compréhension

Pour mieux argumenter notre affirmation, nous analysons les diverses catégories de questions de compréhension proposées. Ceci nous permet de vérifier si ces items portent sur les représentations sémantiques, qui constituent le produit du processus interactif de compréhension.

Comme nous l'avons mentionné en présentant les résultats auxquels notre étude a abouti, le niveau de surface est privilégié par les concepteurs. La quasi-totalité des items requiert des réponses directement extraites du texte. Voici quelques consignes qui relèvent de ce niveau :

1)- *Qui sont les personnages de cette fable ?*

.....

2)- *Où se déroulent les évènements de cette histoire ?*

.....

3)-*Réponds par vrai ou faux*

Le renard tomba face à face avec la panthère.

La panthère dévora le renard.

4) *Lis le texte puis complète les informations suivantes :*

Nom et prénom :

Date et lieu de naissance :

Activité:

Nationalité :

Date de décès:

Œuvres (livres) :

Cette catégorie qui correspond au premier niveau de la compréhension cible la compréhension du texte. Bien qu'elle soit superficielle, cette compréhension est ce que nous voulons mesurer à travers les questionnaires conçus, ce qui octroie aux examens conçus une certaine validité.

Cependant, la validité des items appartenant à ce premier niveau est contestée, car, outre les biais de forme et de fond dont ils ne sont pas exempts, ils ne peuvent pas apporter des preuves sur la compétence évaluée (Beaume, 1988). Dans certains cas, les réponses correctes données par les apprenants sont dans certaines situations dues à des facteurs autres que la compréhension. Ainsi, les apprenants donnent des réponses correctes aux questions de choix multiples ou aux questions vrai/faux sans pour autant comprendre la question : une réponse donnée au hasard pourrait être correcte.

Quant aux items du niveau (Interprétation) dont le nombre est restreint, ils permettent de juger de la capacité des élèves à se servir des informations du texte pour déduire des éléments sous-entendus.

Les deux niveaux omis des questionnaires destinés à évaluer la compréhension de l'écrit sont la base de texte prépositionnelle et l'analyse. Pour la catégorie la plus profonde (Analyse), son omission semble justifiée vu le niveau des apprenants en langue étrangère qui ne permet pas aux enseignants de sélectionner des questions de ce niveau.

Quant à celle de la base de texte prépositionnelle, son inexistence dans les items étudiés est expliquée dans la première partie de notre analyse : les ressources sont évaluées séparément de la compétence de compréhension. Alors que cette catégorie doit inclure des questions auxquelles les apprenants répondent en s'appuyant sur leurs connaissances (savoirs et savoir-faire), les enseignants continuent à élaborer des items dissociés l'un de l'autre comme dans le cadre de la PPO. Ainsi, négligent-ils qu'«*il s'agit maintenant de mettre l'accent sur les processus, la capacité de mobiliser des savoirs et des ressources pour réaliser une tâche complexe*» (Médioni, 2016, 31).

Brièvement, nous pouvons affirmer que les questionnaires conçus en vue de l'évaluation de la lecture-compréhension chez les apprenants du cycle moyen ne sont pas suffisamment valides. En effet, au lieu de mesurer cette compétence, plus que la moitié des items contrôlent la maîtrise des savoirs et savoir-faire. Ceux qui restent évaluent le niveau superficiel de compréhension, mais plusieurs biais peuvent altérer la mesure de la correction ou de la cohérence des représentations sémantiques des apprenants.

4. Conclusion

La validité de l'évaluation d'une compétence quelconque est liée à l'approche appliquée à son enseignement : un enseignement dans le cadre de l'APC exige une adaptation des pratiques évaluatives aux pratiques pédagogiques. Cependant, un certain décalage au niveau de l'évaluation de la compréhension de l'écrit a été révélé par notre étude. En effet, les dispositifs appliqués à son évaluation (les questionnaires) ne permettent pas de donner un aperçu sur la compétence des apprenants.

Ce manque de validité que notre article a démontré nous conduit à exprimer la nécessité d'une refonte qui pourrait être radicale des pratiques évaluatives afin de les rendre conformes aux diverses exigences scolaires et sociales.

Références bibliographiques

- André, Nathalie. Loye, Nathalie. Laurencelle, Louis. (2015). La validité psychométrique : un regard global sur le concept centenaire, sa genèse, ses avatars. *Mesure et évaluation en éducation*, vol. 37, n° 3, pp 125-148.
- Beaume, Edmond. (1988). Comment évaluer la lecture en primaire ? *Communication et langages*, 77, pp 32-42.
- Borsboom, Denny. Mellenbergh, Gideon. J. Van Heerden, Jaap. The concept of validity. *Psychological Review*, vol. 111, n° 4, p. 1061–1071. Consulté le 12 décembre 2018, sur URL: <https://doi.org/10.1037/0033-295X.111.4.1061>
- Brossard, Luce. (1999). Construire des compétences. Tout un programme ! Entrevue avec Philippe Perrenoud. *Vie pédagogique*, n° 112, pp 16-20. Dossier « Faire acquérir des compétences à l'école ».
- Buckingham, Burdette Ross. McCall, William .A. Otis, A.S. Rugg, M.R. Curtis, S.A. (1921). Report of the standardization committee. *Journal of Educational Research*. 4.
- Cronbach, Lee Joseph. (1971). Test validation. In Robert Thorndike (Ed.), *Educational Measurement*. Washington, D.C. : American Council on Education.
- Cuq, Jean-Pierre. (2003). *Dictionnaire de didactique du français langue étrangère et seconde*. Paris: CLE international.
- De Landsheere, Gilbert. (1976). *Evaluation continue et examens - Précis de docimologie*. Paris: Nathan.
- Ecalte, Jean. (2010). L'évaluation de la lecture et des compétences associées. *Revue Française de Linguistique Appliquée*, vol. XV, pp. 105-120.
- Gerard, Jean-Marie. (2006). L'évaluation des acquis des élèves dans le cadre de la réforme éducative en Algérie. In Toualbi-Thaâlibi, Tawil (Dir.), *Réforme de l'éducation et innovation pédagogique en Algérie*. Alger : Casbah Éditions.
- Hadji, Charles. (2015). *L'évaluation à l'école pour la réussite de tous les élèves*. Paris : Nathan.
- Kelley, Truman Lee. (1927). *Interpretation of educational measurements*. Yonkers: World Book.
- Khomsi, Abdelhamid. (1999). *Epreuve d'évaluation de la compétence en lecture*. Paris : ECPA.
- Kintsch, Walter. (1988). The Role of Knowledge in Discourse Comprehension: A Construction-Integration Model. *Psychological Review*, vol. 95, n° 2, pp. 163-182.
- Le Ny, Jean-François. (2005). *Comment l'esprit produit le sens*. Paris : Odile Jacob.
- Lecoq, Pierre. (1996). *L'ECOSSE : une épreuve de compréhension syntaxico-sémantique*. Lille : PUS.
- Maeder, Christine. (2006). *TCS : Test de Compréhension Syntaxique*. Isbergues : Ortho-Editions.
- Magherbi, Hakima. Rocher, Thierry. Gyselinck, Valérie. Trosseille, Bruno. Tardieu, Hubert. (2009). Evaluation de la compréhension de l'écrit chez l'adulte. *Economie et Statistique*, n° 424-425, pp. 63-86.
- Médioni, Maria-Alice. (2016). *L'évaluation formative au cœur des apprentissages : Des outils pour la classe et pour la formation*. Lyon : Chronique sociale.

- Mègre, Bruno. Riba, Patrick. (2015). *Démarche qualité et évaluation en langues*. Paris : Hachette.
- Noel-Jothy, Françoise. Sampsonis, Béatrix. (2006). *Certifications et outils d'évaluation en FLE*. Paris : Hachette.
- Paquay, Léopold. Carlier, Ghislain. Collès, Luc. Huynen, Anne-Marie. (2002). *L'évaluation des compétences chez l'apprenant : Pratiques, Méthodes et Fondements*. Louvain-la-Neuve : UCL Presses Universitaires de Louvain.
- Rey, Bernard. (2015). Peut-on évaluer les compétences scolaires ? *Recherches*, n° 63, pp 9-21, L'évaluation.
- Robert, Jean-Pierre. (2008). *Dictionnaire pratique de didactique du FLE*. Paris : OPHRYS.
- Scallon, Gérard. (2000). *L'évaluation formative*. Québec : Editions Pédagogiques Inc.
- Tagliante, Christine. (2005). *L'évaluation et le Cadre européen commun de référence*. Paris : CLE international.