



HAL
open science

L'évaluation en langues : quelles perspectives ?

Claire Tardieu

► **To cite this version:**

Claire Tardieu. L'évaluation en langues : quelles perspectives ?. Recherches en Didactique des Langues et Cultures - Les Cahiers de l'Acedle, 2006, 2, 10.4000/rdlc.5550 . hal-03281575

HAL Id: hal-03281575

<https://hal.science/hal-03281575v1>

Submitted on 8 Jul 2021

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

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



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

L'évaluation en langues : quelles perspectives ?

Claire Tardieu



Édition électronique

URL : <https://journals.openedition.org/rdlc/5550>

DOI : [10.4000/rdlc.5550](https://doi.org/10.4000/rdlc.5550)

ISSN : 1958-5772

Éditeur

ACEDLE

Ce document vous est offert par Bibliothèque Sainte-Barbe - Université Sorbonne Nouvelle Paris 3



Référence électronique

Claire Tardieu, « L'évaluation en langues : quelles perspectives ? », *Recherches en didactique des langues et des cultures* [En ligne], 2 | 2006, mis en ligne le 06 décembre 2006, consulté le 08 juillet 2021. URL : <http://journals.openedition.org/rdlc/5550> ; DOI : <https://doi.org/10.4000/rdlc.5550>

Ce document a été généré automatiquement le 8 juillet 2021.



Recherches en didactique des langues et des cultures is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

L'évaluation en langues : quelles perspectives ?

Claire Tardieu

1. Introduction

1

L'évaluation en langues a fait l'objet de nombreux travaux portant aussi bien sur les types d'évaluation et leurs fonctions que sur les modalités proprement dite. Pourtant, il semble que de nouvelles données soient à prendre en compte aujourd'hui. Cet article propose de mettre en relation des complexités repérées, dans la définition même du terme et de son objet, et les ouvertures offertes par de nouveaux outils comme le Cadre européen de référence en langues. Évaluer, c'est à la fois mesurer avec précision et estimer approximativement, et l'objet de l'évaluation – la compétence en langue – peut se décliner de différentes manières selon les références. Nous verrons en particulier quelles sont les retombées – directes ou indirectes – du cadre européen, à travers la banque d'outils d'évaluation, les tests fondés sur les scénarii, les textes officiels, le *Portfolio*, et les modes d'apprentissage coopératifs. Nous présenterons ensuite un outil de précision à destination des analystes ou des concepteurs d'items, la *Dutch CEFR Grid*. Nous esquisserons finalement des perspectives pour l'avenir, nous basant sur des projets en cours et la nécessaire évolution des pratiques et des conceptions qui les sous-tendent.

2. Complexité terminologique

2

Le *Petit Robert* donne le XIV^{ème} siècle pour l'apparition du mot *évaluer*, qui tirerait son origine d'*avaluer*, 1283, de l'ancien français *value* – valeur, prix. Il s'agit en fait d'*esvaluer* (1366) "déterminer la valeur, le prix de quelque chose". On peut aussi considérer que

évaluer vient du latin *valere* : être fort, valoir. Le *Petit Robert* nous donne la définition suivante du verbe *évaluer* :

1. Porter un jugement sur la valeur, le prix de. **V. Estimer, priser.** *Faire évaluer un meuble, un tableau, par un expert. V. Expertiser.* "Il essaya d'évaluer la fortune paternelle et quelle en serait sa part ; mais il n'avait là-dessus aucune donnée précise". (Mart. Du G.). **V. Calculer, chiffrer.** *Évaluer un bien au-dessous de sa valeur. V. Surévaluer ; sous-évaluer.* Sa maison est évaluée un million, à un million (**V. Valoir**). *L'arbitre a évalué le dommage.* Déterminer (une quantité) par le calcul sans recourir à la mesure directe. *Évaluer un volume, le débit d'une rivière. V. Jauger.*
2. Par ext. Fixer approximativement. **V. Apprécier, estimer, juger.** *Évaluer une distance à vue d'œil. Foule, assistance évaluée à deux mille personnes environ. – (Abstrait) "Fernando Lucas savait accepter tous les risques dont il avait, depuis longtemps, évalué la quantité et la qualité" [...].*¹

3

Plusieurs notions sont présentes ici, de manière parfois contradictoire. Ainsi, remarque-t-on d'un côté les notions de calcul, de chiffre, de prix – qui renvoient à un jugement objectif fondé sur des critères externes ; d'autre part, le sens d'estimation approximative, laissant envisager une marge d'erreur possible, inévitable même, et des critères de jugement plus flous, voire subjectifs. On retrouve la même ambiguïté du terme dans la définition du *Trésor de la Langue Française 2004* : "Déterminer, délimiter, fixer avec précision" et "Conjecturer, faire l'estimation d'une quantité, d'une durée qui n'est pas encore vérifiable".² Selon la même source, on peut trouver un élément de définition supplémentaire : *reconnaître la valeur de, être sensible aux qualités de, aimer (quelqu'un)*.

4

Encore une fois, nous voyons ici l'affirmation d'une dimension subjective, voire affective, ainsi que l'idée que l'on ne peut évaluer que ce qui est mesurable (quantitativement, qualitativement), ce qui implique les notions de normes pré-établies et de critères.

5

Si l'on relie à présent ces définitions au domaine de la didactique, elles s'avèrent très éclairantes :

Tableau 1 – inventaire des notions dans les définitions citées

Notions	Implications
calcul, chiffre, mesure	jugement précis, avéré, objectif, relatif à des critères, une norme
estimation approximative	jugement global, prédictif, subjectif
estimation valorisante	jugement subjectif, affectif

6

7

Ces listes assez contradictoires nous amènent à poser plusieurs questions. Peut-on évaluer justement ? Quel est l'objet de l'évaluation ? Cet objet est-il quantifiable ? En quels termes ? En chiffre ? Par des appréciations ? Le jugement fait-il appel à des critères externes ? Est-il biaisé, dans un sens positif ou négatif ?

8

Cossu et Favel soulèvent en particulier le problème de la mesure :

Le paradoxe se retrouve lorsqu'il s'agit d'évaluer dans le système scolaire, car en effet il existe une distance entre la nature du système généralement utilisé pour évaluer, c'est-à-dire la notation (le plus souvent chiffrée), et la nature de ce qui est évalué, c'est-à-dire des productions scolaires parfois difficilement chiffrables. Certaines productions sont mesurables quantitativement (c'est le cas des tests contraignants demandant une seule réponse par question) et peuvent être évaluées par un chiffre, d'autres relèvent davantage du domaine de l'estimation, du jugement qualitatif, de l'appréciation non chiffrée (c'est le cas de l'originalité, la variété de la langue, l'élégance, la prise de risque, etc., dans les productions libres). (Cossu & Favel 1998 : 133-34).

9

Le problème se pose avec plus d'acuité dans un enseignement qui vise à développer des compétences de communication. Sibylle Bolton pose le problème de la fiabilité et de la validité des tests :

[...]i l'on va dans le sens de l'objectif de "compétence communicative interactive", les tests doivent laisser la plus grande place possible à des réponses individuelles créatives. Mais on accroît alors la marge d'interprétation dans la notation des résultats oraux et écrits, autrement dit, on diminue la fiabilité de la mesure. (Bolton, 1991 : 4-5)

10

Tout l'art de l'évaluateur consistera précisément à rendre de plus en plus mesurable cet obscur objet de l'évaluation.

11

En fait, le terme *évaluer* est relativement récent en didactique. Le concept lui-même est issu des sciences sociales behavioristes aux États-Unis, l'apprentissage signifiant un changement de comportement mesurable. Niklas décrit ainsi la démarche :

On soumet donc l'écopier à certains stimuli (offres d'apprentissage) afin qu'il arrive à un objectif fixé d'avance (objectif d'apprentissage). [...] Le dernier degré de la théorie behavioriste de l'apprentissage est finalement celui de l'évaluation : on vérifie si l'objectif d'apprentissage a été atteint donc si l'élève a appris ce qu'il devait apprendre. (Niklas, 2003 : 173).

12

En France on commence à utiliser l'évaluation dans les années soixante-dix, pour accompagner les méthodes novatrices (pédagogie par objectifs en particulier), ce qui semble logique. En effet, dès lors que l'objet à mesurer (connaissances et compétences définies par des référentiels d'apprentissage) devient palpable, des tests d'évaluation peuvent voir le jour.

3. Complexité de l'objet à évaluer

13 Le terme "évaluer" se définit aussi par son objet : performance individuelle, collective, institutionnelle. Nous nous pencherons ici sur son objet dans le domaine de l'enseignement / apprentissage des langues. On peut dire qu'aujourd'hui on évalue à la fois des connaissances et des compétences à travers des performances réalisées par les élèves dans la langue-cible. Un aperçu historique des méthodologies montre que ces notions ne recouvrent pas la même réalité. En fait, en fonction du statut de la langue au cœur de l'apprentissage, l'objet d'évaluation varie lui aussi.

Tableau 2 - évolution de l'objet de l'évaluation

Types de méthode	Conception de la langue et de l'enseignement / apprentissage	Objet de l'évaluation
Méthode active	Langue, objet d'apprentissage, fin en soi. L'écrit est privilégié.	Connaissances linguistiques (lexicales, grammaticales), et culturelles (littérature, histoire, arts). Compétences de compréhension et de production écrites (commentaires de textes), de transposition (thème et version).
Méthode audiovisuelle	Langue : ensemble de structures permettant un apprentissage psittaciste.	Connaissances linguistiques (structures, lexique) acquises mécaniquement. Compétences "pré-communicatives" de compréhension et de production orales guidées (l'écrit étant un "oral scripturé", selon Bailly [1997 : 53]).
Approche communicative et cognitive	Langue : moyen de communication authentique, outil à maîtriser	Connaissances linguistiques (lexicales, grammaticales, fonctionnelles, phonologiques, culturelles, méthodologiques). Compétences de communication orale et écrite dans une perspective d'autonomie. Méta-compétences de réflexion sur la langue, l'apprentissage, la culture.

14

15

16

17

Ainsi, il paraît plus difficile d'évaluer une compétence communicative (par le biais d'épreuves requérant l'utilisation de la L2 dans des situations réelles et interactives), qu'une compétence "pré-communicative", selon l'expression de Bolton (1991 :11) – à

savoir la maîtrise d'éléments lexicaux ou syntaxiques isolés.³ Que signifie véritablement évaluer une compétence de communication ?

18

Tout d'abord, comment peut-on définir "l'objet" compétence ? En linguistique chomskyenne, le terme réfère à une capacité sous-jacente, opposée à la performance, qui est la manifestation langagière de la compétence. Dans l'approche communicative, la compétence est définie comme la capacité à utiliser un langage approprié dans des situations et des dispositifs de communication variés (Bachman, 1990). Philippe Perrenoud montre bien combien il est vain d'opposer connaissances et compétences :

Pour comprendre le monde et agir sur lui, ne faut-il pas, à la fois, s'approprier des connaissances étendues et construire des compétences susceptibles de les mobiliser à bon escient ? La figure de l'expert l'atteste : c'est à la fois un savant, un érudit, quelqu'un qui a " lu tous les livres " et accumulé des trésors de connaissances au gré de son expérience, et quelqu'un qui a du flair, de l'intuition, du sens clinique, du know-how et l'ensemble des capacités qui permettent d'anticiper, de prendre des risques, de décider, bref d'agir en situation d'incertitude. (Perrenoud, 1997 : 12)

19

La notion de situation d'incertitude paraît tout à fait appropriée pour la communication en langue étrangère. C'est bien parce que l'on doit faire face à des situations toujours nouvelles que l'apprentissage par cœur de structures ne fonctionne pas. De même les savoirs ou les savoir-faire développés dans la méthode active, si pointus fussent-ils, ne permettraient pas d'échanger librement avec un interlocuteur. "[L]es compétences [manifestées dans des actions] ne sont pas elles-mêmes des connaissances, elles utilisent, elles intègrent, elles mobilisent des connaissances" (Perrenoud, 1997 : 8). Autrement dit, les connaissances n'ont de valeur que dans leur mise en relation permanente. Perrenoud explique que certains préfèrent conserver le terme de connaissances et le décliner, au lieu de recourir au terme de compétences. Ainsi, les sciences cognitives distinguent-elles trois grands types de connaissances : les connaissances déclaratives stockées sous forme de concepts, de propositions (faits, lois, régularités, etc.), les connaissances procédurales – actions potentielles, dynamiques, implicites et automatiques (procédures à suivre, opérations) et les connaissances conditionnelles qui "*précisent les conditions de validité des connaissances procédurales*" (Perrenoud, 1997 : 9) – c'est-à-dire qui permettent leur utilisation à bon escient. C'est la mise en synergie de ces trois types de connaissances qui confère une réelle compétence langagière. Par exemple, la compréhension d'un segment de texte requerra à la fois : des connaissances lexicales, grammaticales, culturelles ; des connaissances de procédures comme l'inférence (permettant de deviner le sens d'un mot inconnu grâce au contexte) et la déduction (permettant de déduire le sens d'un mot inconnu à partir de règles grammaticales connues et sollicitées) ; et enfin des connaissances de conditions (permettant d'utiliser l'inférence là où le contexte fait sens, la déduction là où l'on reconnaît un mot dérivé, etc.).

20

Si pour les cognitivistes le terme *connaissances* est devenu générique, pour d'autres, c'est le terme de *compétence* qui est retenu. Ainsi dans le *Cadre européen commun de référence en langues (Cecrl)*, on note que la compétence est définie comme "*un ensemble de savoirs, de savoir-faire, de savoir-être et de savoir-apprendre*". Les savoirs étant définis comme des "*connaissances déclaratives*", les savoir-faire comme des "*connaissances procédurales*", les savoir-être comme des "*dispositions individuelles, des traits de personnalité, des dispositifs d'attitude*" ; ces savoir-être ne sont pas posés "*comme des attributs permanents d'une personne*

" mais comme *"sujets à des variations"*. Enfin les savoir-apprendre *"mobilisent tout à la fois des savoir-être, des savoir et des savoir-faire et s'appuient sur des compétences de différents types"*. En la circonstance, "savoir-apprendre" peut aussi être paraphrasé comme *"savoir / être disposé à découvrir l'autre"* (Cecrl, 2001 : 16-17). Vus sous un certain angle, les savoir-être et les savoir-apprendre relèveraient des "connaissances conditionnelles" dans le langage cognitiviste.

21

Le Cecrl fait aussi la distinction entre l'évaluation du savoir (ou du niveau) – qui est *"l'évaluation de l'atteinte d'objectifs spécifiques"*, *"centrée sur le cours"*, correspondant à *"une vue de l'intérieur"* – et l'évaluation de la capacité (mise en œuvre de la compétence ou performance), c'est-à-dire *"de ce que l'on peut faire ou de ce que l'on sait en rapport avec son application au monde réel"* (Cecrl, 2001 : 139). L'évaluation de la capacité correspond *"à une vue de l'extérieur"*. Idéalement, dans le contexte d'enseignement et d'apprentissage de l'approche communicative, l'écart entre l'évaluation du savoir et l'évaluation de la capacité devrait être infime.

Tableau 3 – l'objet d'évaluation dans l'approche actionnelle du Cecrl

Approche actionnelle européenne ⁴	Objet d'évaluation
Langue : moyen d'interaction sociale	Compétences linguistiques (lexicales, grammaticales, sémantiques, phonologiques, orthographiques, ortho épiques). Compétences socio-linguistiques (marqueurs linguistiques de relations sociales, conventions de politesse, expressions de sagesse populaire, différences de registres, dialectes et accents). Compétences pragmatiques (compétence discursive, compétence fonctionnelle). (Cecrl, 2001 : 17-18)

22

Dans le tableau 3, on remarque que l'approche actionnelle européenne privilégie une définition sociale de la communication interculturelle, en définissant quatre grands domaines de communication (personnel, public, professionnel, éducationnel), et en négligeant peut-être ce que Michael Byram appelle le domaine documentaire – dans lequel la culture constitue *"l'ensemble des productions intellectuelles et créatives, et dans lequel se trouvent enregistrées dans le détail la pensée et l'expérience humaine"* (Byram, 1992 : 111-12).

23

Il faut émettre certaines réserves concernant l'évaluation de la compétence de communication et la notion même de test. En effet, comme le souligne Mothe :

[...] est logique de penser, dans une perspective d'évaluation communicative, que les épreuves spécifiques de grammaire ou de vocabulaire, si elles peuvent avoir un intérêt pédagogique, ne sauraient prétendre vérifier dans quelle mesure les objectifs terminaux, à savoir l'acquisition d'un instrument de communication, ont

été atteints. Elles sont donc peu valides dans une évaluation terminale (à l'examen par exemple) et insuffisantes pour une évaluation formative. (Mothe, 1981 : 40).

24

Et pour Bolton (1991), seuls les tests de niveaux de compétence et les tests de contrôle permettent véritablement d'évaluer la compétence de communication, la communication étant conçue comme l'usage de la parole pleinement intégrée dans la situation discursive. On peut ainsi proposer un tableau des types de test en relation avec l'évaluation :

Tableau 4 – types d'évaluation et types de test

Moment d'intervention (dans l'apprentissage)	Type	Fonction	Tests
Avant	Évaluation organisationnelle ou diagnostique	Prédictive, prospective Centrée sur le projet	Les tests d'aptitude Les tests de placement Les tests diagnostiques
Pendant	Évaluation formative	Régulatrice Centrée sur le processus	Les tests de progrès
Après	Évaluation sommative ou certificative	Récapitulative, rétrospective Centrée sur le résultat	Les tests de contrôle Les tests de niveaux de compétence (Bolton, 1991: 12)

25

Il n'est pas sûr que seuls les deux derniers types de test évaluent une compétence de communication. On va voir en effet que l'approche actionnelle du *Cadre* a des répercussions qui ne sont pas toujours celles attendues.

4. Incidences du Cadre européen sur les pratiques

4.1. La banque d'outils

26

En premier lieu on pourrait évoquer la banque d'outils d'évaluation élaborée par la Division de l'Évaluation et de la Prospective, et qui s'efforce de mettre en ligne des outils diagnostiques permettant d'évaluer trois compétences (compréhension de l'oral, compréhension de l'écrit, production écrite). On peut citer en exemple un outil, créé en allemand, évaluant la capacité à reconnaître les nombres dans le contexte d'une

conversation téléphonique entre un directeur d'école passant commande de fournitures scolaires à un libraire.⁵

27

Ces outils s'inscrivent dans certaines échelles du *Cecrl*, au niveau A2. Par exemple, "*Peut comprendre des expressions et des mots porteurs de sens relatifs à des domaines de priorité immédiate (par exemple, information personnelle et familiale de base, achats, géographie locale, emploi)*" (55). Ou encore au niveau A1 : "*Peut se débrouiller avec les nombres, les quantités, l'argent, l'heure*" (66). Si le cadre n'offre pas toujours de descripteurs au niveau A1, le *Portfolio* peut venir pallier certains manques : par exemple, ici, au niveau A2 : "*Je comprends quand dans une phrase, quelqu'un utilise des chiffres et des nombres pour dire son âge, son numéro de téléphone, un prix, ... et quand il parle clairement et lentement.*" Les outils qui vont être mis en ligne prochainement s'efforcent d'intégrer davantage les niveaux du cadre.

4.2. CLES, DCL et scénarii

28

On peut bien sûr citer aussi les travaux autour du CLES (*Certificat en langues de l'enseignement supérieur*) défini par le BO n° 25 du 29 juin 2000 dont l'expérimentation a été pilotée par Claude Springer à l'université de Strasbourg. Le CLES se décline en trois niveaux de qualification (1er, 2ème, 3ème niveaux), définis respectivement par référence aux niveaux B1, B2 et C1 du Conseil de l'Europe. Les compétences évaluées – par le biais du scénario – sont la compréhension orale, écrite, la production écrite en situation, l'expression orale et l'interaction orale. On trouve en ligne sur le site (<http://webcles.u-strasbg.fr/home.php>) des exemples de tests pour s'entraîner.

29

Un autre diplôme, le DCL (*Diplôme de compétence en langue*), davantage destiné au monde de l'entreprise, est également basé sur le principe des scénarii.

Le diplôme de compétence en langue valide un savoir-faire fondé sur un savoir. Il ne se fonde pas sur le décompte de manques dans l'ordre du seul savoir par rapport à une norme linguistique idéale, mais évalue de façon positive la compétence en langue de candidats par référence au degré d'opérationnalité dans l'accomplissement d'une tâche. C'est l'efficacité et la qualité de la performance qui permettent l'attribution de l'un des cinq degrés du diplôme. Les épreuves de l'examen reposent sur l'exploitation d'un scénario unique construit à partir de supports authentiques permettant la simulation de situations de communication réelles. (*BOEN n° 22 du 30 mai 2002*)

30

L'épreuve de l'examen, d'une durée de trois heures se présente sous la forme d'un scénario, selon une logique analogue à celle des études de cas : une situation à découvrir, des tâches à effectuer, des choix à faire, une solution à proposer dans le cadre d'une mission confiée au candidat. Les tâches s'appuient sur des documents écrits, sonores et / ou audiovisuels. L'épreuve se décompose en cinq phases dont chacune permet de tester plus particulièrement une compétence à l'écrit ou à l'oral, à travers un certain nombre d'activités : recueil et tri d'informations, formulation d'un problème et choix de solutions, argumentation. Dans les phases 1, 2 et 3, le candidat recueille des informations d'après des documents écrits et sonores et répond par écrit à un questionnaire écrit concernant ces documents. Il complète ensuite ses informations à partir d'un entretien

téléphonique avec un interlocuteur (durée totale 1h40, dont 10 minutes maximum d'entretien téléphonique). Dans la phase 4, le candidat présente oralement à un interlocuteur la solution qu'il a retenue au problème posé dans le cadre de sa mission. Un temps d'échange lui permet de défendre son point de vue (préparation : 20 minutes ; durée totale présentation + échange : 20 minutes). À partir des éléments recueillis au cours des phases antérieures, le candidat rédige dans la phase 5 un document proposant la solution retenue au problème posé dans le cadre de sa mission (durée 40 minutes). Un exemple de scénario est téléchargeable à <http://www.d-c-l.net/telechargements.htm>.

31

Les pratiques évaluatives illustrées par le CLES et le DCL auront certainement une incidence directe sur les pratiques d'enseignement/apprentissage. Il faut espérer que l'institution revoie également la forme de l'examen terminal en langues dans le secondaire, au moins pour un certain nombre de séries au "Bac".

5. Textes officiels, Portfolio européen des langues

32

On peut citer aussi les tentatives actuelles pour intégrer les niveaux du cadre dans les textes officiels français (niveau A1 en fin d'école primaire, B1 en fin de collège et B2 en fin de lycée), ou encore les recommandations pour intégrer le *Portfolio* dans les pratiques évaluatives.

33

Le *Portfolio* est la propriété de l'apprenant, et se présente sous la forme d'un dossier dans lequel toute personne ayant appris une ou plusieurs langues (que ce soit dans le cadre de l'école ou en dehors), consigne ses connaissances et ses compétences ainsi que ses expériences culturelles. Il contribue ainsi à clarifier l'apprentissage. Commun à tous les pays européens dans sa structure même et dans ses références directes aux descripteurs du *Cecrl*, le *Portfolio* est donc compréhensible partout en Europe par les différents interlocuteurs : universités, employeurs, organismes divers. Il contribue ainsi à favoriser la reconnaissance de niveaux en langue et la mobilité des européens. Destiné à être complété, enrichi ou modifié au cours de la vie, le *Portfolio* est rempli par le détenteur lui-même, dans un esprit d'évaluation positive. Il permet d'exprimer ce que l'on sait faire dans une ou plusieurs langues. Cette dimension plurilingue valorise toutes les compétences si minimes soient-elles. Il ne s'agit plus seulement de valider un niveau de maîtrise d'une langue 2 uniquement lorsqu'on satisfait à certains critères, mais de valider le niveau de maîtrise d'une langue 2 tel qu'il est au moment où on l'évalue. Ainsi, on peut être capable de décliner son identité en anglais, espagnol et allemand, de comprendre une indication de chemin en anglais et en allemand, de lire un article de presse utilisant une langue simple en allemand, plus complexe en anglais, etc. Il est recommandé pour l'utilisation par des adultes (16 ans et plus) mais existe dans des versions adaptées pour le primaire ou le collège.

34

Le *Portfolio* se compose de trois éléments : le passeport de langues ; une biographie langagière ; un dossier présentant des documents relatifs aux expériences effectuées et aux compétences acquises. Le passeport de langues invite à auto évaluer pour chaque langue ses niveaux de compétence en compréhension et en expression en relation avec

ceux du *Cadre*. Les mots-clés choisis sont volontairement simples et compréhensibles : écouter, lire, prendre part à une conversation, s'exprimer oralement en continu, écrire. Dans l'exemple qui suit, l'apprenant peut inscrire des langues différentes dans les pointillés (AN pour anglais, AL pour allemand, etc.) – au total six langues. Puis il coche les cases correspondant à ses capacités dans chaque langue.

AL	AN	
						Je comprends quand on me salue, quand on me demande comment je vais ou quand on me dit au revoir
						Je reconnais des consignes claires et simples
						Je comprends des expressions familières et simples de la vie quotidienne (acceptation, refus, remerciements, etc.)

35

Cette grille permet à l'apprenant de définir ses compétences linguistiques selon des critères reconnus dans tous les pays européens et de compléter ainsi les certificats institutionnels.

36

La première partie de la biographie langagière permet au détenteur de relater ses expériences, ses stages, ses séjours à l'étranger, de faire état de lectures d'œuvres complètes, de films vus en version originale, de lectures de presse étrangère, de participation à des exposés, des associations de langue, des créations théâtrales, etc. Une deuxième partie permet des précisions sur les langues apprises à l'université ou dans un cadre structuré ou de manière plus libre (dans le cadre de la profession ou de rencontres avec des locuteurs natifs). Au collège, cette biographie permet à l'apprenant de noter plus précisément ses capacités et de planifier ses apprentissages. Elle joue donc aussi le rôle d'un référentiel de compétences. Au niveau A2 on peut lire ceci : "*Écouter et comprendre : je peux comprendre les numéros de téléphone ou encore, je peux comprendre des consignes et des instructions simples, par exemple pour aller d'un point à un autre à pied ou en transport en commun*". Contrairement au passeport qui n'est mis à jour qu'une ou deux fois par an, la biographie langagière accompagne l'apprentissage jour après jour. Le détenteur du *Portfolio* mettra dans son "dossier" tous les documents attestant du niveau de compétence exprimé (travaux réalisés, diplômes, attestations de séjours à l'étranger, etc.)

37

Un certain nombre de questions peuvent être posées, et en particulier celle de la fiabilité de ce type d'instrument. Sa fiabilité formelle, d'abord. Quelle est la relation entre le *Portfolio* et le *Cadre* ? Le *Portfolio* n'a-t-il pas tendance à simplifier, avec le risque d'une déperdition qualitative ? Si au contraire il propose des descripteurs au niveau A1 là où le cadre n'en proposait pas, sur quelles données s'appuie-t-il ? La question de la fiabilité des contenus se pose ensuite. Comment intégrer la valeur de l'autoévaluation dans un système institutionnel ? On peut se demander aussi dans quelle mesure il contribue à orienter et soutenir l'enseignement/apprentissage.

6. L'évaluation dans un mode d'apprentissage coopératif : Tandem

38

Le *Cecrl* peut avoir une incidence privilégiée aussi sur les modes d'apprentissage coopératifs, qui remettent quelque peu en cause l'évaluation traditionnelle, puisqu'ils font appel à des compétences plus diverses que l'apprentissage en classe, et présentent certaines caractéristiques de l'apprentissage "naturel" (Le Bihan, 2004 : 74).

L'apprentissage en tandem en est un exemple. Il peut se pratiquer en présentiel ou à distance (par le biais des Technologies de l'information et de la communication) ; il sollicite les compétences de l'oral ou de l'écrit. L'apprentissage en tandem se définit comme un mode d'apprentissage ouvert et coopératif, dans lequel deux personnes de langue maternelle différente travaillent en binôme pour apprendre plus sur la personne et la culture du partenaire ; pour s'entraider dans l'amélioration de leurs connaissances et compétences linguistiques ; pour échanger d'autres connaissances, par exemple dans le domaine de leur activité professionnelle (<http://tandem.ac-rouen.fr/index.html>). Les deux principes qui régissent l'apprentissage en tandem sont l'autonomie et la réciprocité. L'autonomie parce que ce sont les partenaires qui définissent leurs objectifs d'apprentissage, leurs thèmes de discussion, etc. (éventuellement avec l'aide d'un conseiller) ; la réciprocité parce que chacun s'engage à aider l'autre à apprendre. Concrètement, la règle de tandem consiste à utiliser pendant 50 % du temps sa langue maternelle (langue-cible de l'autre) et pendant 50 % du temps la langue maternelle de l'autre (langue-cible pour soi).

39

Dans ces conditions, on comprend que l'objet de l'évaluation ne soit pas limité à la langue. On peut acquérir des connaissances culturelles, professionnelles, mais aussi des compétences dans le domaine social du comportement, de la relation à l'autre, ou encore dans celui de l'apprentissage : recherche d'information sur sa culture maternelle, demande de corrections ciblées adressées au partenaire, etc. Ce type d'apprentissage entre pairs nécessitera donc la mise en œuvre et l'évaluation de compétences au sens où l'entend le *Cecrl* : savoirs, savoir-faire, savoir-être et savoir-apprendre.

40

Le Bihan (2004) propose une grille d'autoévaluation en fin de programme qui se découpe en deux types d'acquisition : acquisitions linguistiques et culturelles, et acquisitions conceptuelles concernant la relation dans le tandem. Sous la rubrique "Acquisition linguistiques et culturelles", on trouve, par exemple : "J'ai acquis du vocabulaire nouveau" ou bien "J'ai acquis des structures nouvelles" – propositions assorties d'un choix possible entre pas du tout, un peu, beaucoup. La grille d'autoévaluation permet aussi de dresser un bilan sur "le domaine où j'ai le plus appris" (langue et vie quotidienne, systèmes administratifs et éducatifs, phonologie, culture et littérature, autres) ou sur "ce en quoi mon partenaire a le plus progressé" ou "ce en quoi j'ai le plus progressé".

41

Cette grille utilise le pronom personnel *je*, comme le *Portfolio européen*, se rangeant dans cette catégorie d'évaluation où l'apprenant est son propre évaluateur mais aussi l'évaluateur du partenaire. Le principe de réciprocité dans ce type d'apprentissage est sans doute à l'origine de ce processus d'interévaluation – d'évaluation réciproque ou

solidaire. Il s'agit surtout ici de prendre en compte les progrès remarquables du partenaire et non ses difficultés ou ses manques.

42

Pour ce type d'apprentissage en binômes – ou d'une manière plus générale d'apprentissage coopératif – on pourrait imaginer une grille pouvant se décliner en deux volets : mes progrès, les progrès de mon partenaire. Le tableau 5 propose une liste de catégories pour l'autoévaluation ; ce type de grille demande bien sûr à être expérimentée et réajustée selon les situations d'apprentissage.

Tableau 5 – évaluation des objets d'apprentissage en langue et en culture, échange en tandem

	langue (aspects linguistiques, sociolinguistiques et pragmatiques)	Culture
Savoirs	lexique faits de langue fonctions langagières phonologie rapports graphie/phonie	domaines • éducation et modes de vie • histoire et politique • arts et créativité • monde professionnel
savoir-faire	comprendre parler lire écrire interaction oral-écrit ou écrit-oral mettre en relation avec la L1 (2, 3...) (transposition, traduction)	analyser synthétiser élargir approfondir comparer : mettre en relation avec la culture 1
savoir-être : niveaux personnel et interpersonnel	prise en compte de la situation de communication adaptation du discours à l'interlocuteur confiance prise de risque respect solidarité	Curiosité empathie changement de conception accroissement enrichissement respect solidarité
Savoir-apprendre	demander/apporter de l'aide, des corrections ciblées, etc. rechercher en autonomie	demander/apporter de l'aide, des connaissances ciblées, etc. rechercher en autonomie

43

Les initiatives européennes en matière d'enseignement et d'évaluation des langues, depuis l'établissement d'un niveau-seuil dans les années soixante-dix, ne sont donc pas sans incidence sur les pratiques contemporaines ; elles invitent à ouvrir les modes de fonctionnement nationaux. La question qui demeure cependant, au vu de ces différents exemples, est comment ne pas rester dans l'approximation, dans le placage et dans la référence facile ?

7. La Dutch CEF Grid

44

Outil élaborée par un groupe d'experts autour de Charles Alderson de l'université de Lancaster, grâce à un financement du gouvernement hollandais, La *Dutch CEF Grid* offre un cadre d'analyse pour la création d'items d'évaluation référencés sur le *Ceclrl*, prenant en compte de multiples paramètres. Or, le *Ceclrl* ne semble pas offrir un instrument permettant de construire des items de compréhension orale et écrite et des tests (alors que pour la production, il offre des spécifications précises). Le projet consistait donc, après examen minutieux du *Ceclrl* à plusieurs étapes. Il fallait : développer une grille d'analyse des items et des tests en compréhension de l'oral et de l'écrit ; examiner un échantillon d'items et de tests censés se situer à tel ou tel niveau du *Ceclrl* ; examiner les points communs et les différences dans les spécifications de ces tests ; examiner comment les tests rendent opérationnels dans des tâches le développement des compétences de compréhension écrite et orale.

45

Au sein de ce projet, nous avons pu identifier quatre types de problèmes dans le *Ceclrl* : des inconsistances ; des problèmes de terminologie ; l'absence de définitions des termes utilisés ; d'autres failles. Comme inconsistances, par exemple, la mention du vocabulaire est présente à niveau B1, mais absente aux niveaux inférieurs ou supérieurs. L'opération "reconnaitre" est mentionnée aux niveaux A1, B1 et C1, mais absente en A2, B2 et C2. L'opération d'inférence n'apparaît qu'au niveau C1, et une fois en A2, mais ni en B1 ni en B2 ; la vitesse d'élocution n'est pas mentionnée au niveau B1. La capacité à saisir l'information dans les documents publicitaires n'est mentionnée qu'en A2, et ainsi de suite.

46

Du point de vue terminologique, examinons le cas des différents verbes utilisés pour désigner la compréhension : certains sont transitifs, d'autres intransitifs ; la question de la synonymie est donc posée.

Tableau 6 – verbes utilisés pour décrire la compréhension, *Ceclrl*

niveau A2	niveau B1	niveau B2
understand	understand	understand
take	locate	scan
get	scan	monitor

follow	identify	obtain
identify	combine	select
infer	extrapolate	evaluate
	recognise	locate
		identify

47

L'absence de définitions des termes pose aussi un problème. Par exemple le terme *simple* est souvent utilisé dans les échelles, mais il n'est pas forcément simple de faire la distinction entre simple, moins ou très simple. Ce qui paraît simple dans la langue 1 ne l'est peut-être pas pour quelqu'un qui possède une autre langue 1. Le même type de question se pose avec des termes comme *familier, concret, factuel, complexe, court, long, etc.*

48

Enfin, une grande faille du *Cecrl* réside dans l'absence de toute référence à des opérations mentales dans les discussions de la compréhension : par exemple, une distinction entre fait et opinion, détails pertinents ou non pertinents, etc. Une autre notion manquante est celle de la tâche : que doit-on faire avec un texte ? Les tâches à accomplir ne sont pas systématiquement explicitées.

49

The Dutch CEF Grid se décompose en trois volets : texte, item et tâche – chacun avec des catégories à renseigner pour chaque support utilisé. Pour le volet *texte*, il faut identifier la source, le sujet, le domaine, le niveau du *Cecrl* visé. Dix dimensions (présentes dans le cadre sous des formulations plus ou moins précises) sont également à définir : authenticité, type de discours, nature du contenu (plus ou moins abstrait), longueur (nombre de mots pour la lecture, durée pour l'écoute), grammaire, vocabulaire, vitesse (seulement pour l'écoute), accent (marqué ou standard ; seulement pour l'écoute), clarté de l'articulation (écoute), nombre de participants (écoute).

50

Pour le volet *item*, il faut identifier : le comportement en jeu (reconnaître, inférer, évaluer) ; l'objet de la compréhension (une catégorie décrivant ce qui est lu ou écouté) ; la source de l'information (explicite dans le texte, ou seulement implicite ?) ; le format de l'item (QCM, lacunaire, etc.) ; l'estimation par rapport aux niveaux du *Cecrl* (en se fondant sur une compilation des échelles par activité, sur les "je peux" de *Dialang*⁶ et sur les niveaux de ALTE). Pour la *tâche*, la combinaison du texte et de l'item donneront une estimation de niveau global.

51

On trouve sur le site de Lancaster University des illustrations concrètes, ainsi qu'un module d'entraînement à destination des concepteurs d'items (<https://www.lancaster.ac.uk/linguistics/>). Un outil comme la *Dutch CEF Grid* favorise le croisement des paramètres de difficulté entre texte, item, et tâche. Rapporté à la classe, c'est un outil qui peut permettre de situer des contenus et des actions pédagogiques dans le cadre de l'apprentissage et non seulement celui de l'évaluation.

8. Nouvelles perspectives

52

Comment concrètement envisager l'avenir ? Plusieurs pistes se dessinent. Tout d'abord, la mutualisation des pratiques évaluatives ne peut que bénéficier aux usagers si l'on considère que les matériaux d'évaluation sont encore culturellement marqués. En ce sens, les tentatives contemporaines pour constituer des banques d'items à l'échelle européenne – le projet Ebafls (*European Anchorbank For Language Skills*), le travail de Sylvie Lepage au CIEP, de Neus Figueras à Barcelone – laissent augurer d'un enrichissement au niveau des pratiques de chaque pays non seulement au plan quantitatif mais aussi qualitatif.

53

Ensuite, il faudra définir la place que l'on accorde réellement aux outils autoévaluatifs comme le *Portfolio*. Seront-ils pris en compte comme de véritables attestations de pratiques et de niveaux (en complément des travaux et diplômes évalués de manière externe), ou demeureront-ils accessoires par rapport aux pratiques de classe et aux tests institutionnels et commerciaux ?

54

Enfin, les modes d'évaluation centrés sur la réalisation de projets – où les individus se révèlent toujours plus performants que dans des tests formalisés – vont-ils se développer et modifier en amont des pratiques de classes dépendantes, par la force des choses, de la forme de l'examen terminal ?

55

Howard Gardner fait la distinction entre deux conceptions de l'évaluation : le test formalisé et la preuve sur le terrain. Autrement dit, un mode qui relève d'une certaine sophistication, un autre davantage ancré dans l'apprentissage sur le tas. On pourrait objecter que les tests formalisés offrent une fiabilité et une objectivité que les évaluations sur projet ou action ne peuvent pas garantir. Cependant, tout instrument possède son biais. Gardner montre que les tests formalisés ne permettent d'évaluer que des formes très limitées d'intelligences, privilégiant ainsi certains individus au détriment des autres. Or, l'évaluation pourrait jouer un rôle social plus performant : "*Évaluer, c'est recueillir des données sur les compétences et les potentiels des individus dans le double objectif de leur en faire retour utilement et de procurer des informations indispensables à la communauté environnante*" (Gardner, 1986 : 114).

56

En définitive, l'évaluation doit avant tout s'avérer utile à l'individu et à la communauté. Elle devrait permettre la mesure positive du potentiel de chacun, pour le bénéfice de tous.

Quand l'évaluation se fonde progressivement dans le paysage, elle n'a plus besoin d'être distinguée des autres activités de la classe. À l'instar de ce qui se passe dans un bon apprentissage, les enseignants et les élèves sont toujours en train d'évaluer. Il n'y a plus alors nécessité d'enseigner pour l'évaluation parce que l'évaluation est omniprésente ; le besoin de tests formalisés pourrait bien disparaître. (Gardner, 1986 : 115)

57

Ceci n'est peut-être qu'une utopie. En attendant, on peut envisager qu'à l'avenir, les diplômes nationaux en langue perdent de leur valeur au profit d'outils produits par

chaque pays calibrés sur le Cadre et qui permettraient au sein de la complexité initiale de remplir un rôle de mesure précise aux côtés d'outils relevant de l'estimation approximative comme les documents d'autoévaluation.

BIBLIOGRAPHIE

- Bachmann, L. F. (1990). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Bailly, D. (1997). *Didactique de l'anglais, 1 : Objectifs et contenus de l'enseignement*. Paris : Nathan.
- Banqoutils, Banque d'outils d'évaluation en ligne : www.banqoutils.education.gouv.fr. (lien valide le 06/12/2006).
- Bulletin officiel de l'Éducation nationale* (2002). N° 22 du 30 mai.
- Bolton, S. (1991). *Évaluation de la compétence de communication en langue étrangère*. Paris : Didier.
- Byram, M. (1992). *Culture et éducation en langue étrangère*. Paris : Didier.
- Cadre européen commun de référence pour les langues* (2001). Paris : Didier.
- Cossu, Y., Favel, J.J.(1998). *L'enseignement de l'anglais*. Paris : Nathan.
- Gardner, H. (1996). *Les intelligences multiples*. Paris : Retz.
- Le Bihan, J.C. (2004). "La méthodologie Tandem dans la formation des professeurs des écoles". *Les langues modernes*, n° 3, pp. 72-83.
- Mothe, J.-C. (1981). "Évaluer les compétences de communication en milieu scolaire". *Le Français dans le Monde* 165, pp. 63-72.
- Niklas, H. (2003). "L'évolution historique de la notion d'évaluation". In Demorgon, J., Lipiansky, E.-M., Muller, B., Niklas, H. (dir.), *Dynamiques interculturelles pour l'Europe*, pp. 169-76. Paris : Éditions Anthropos.
- Perrenoud, P. (1998). *Construire des compétences dès l'école*. Paris : ESF éditeur.
- Portfolio européen*, (2001). Paris : CNDP / Didier.
- Puren, C. (2002). "Perspectives actionnelles et perspectives culturelles en didactique des langues-cultures : vers une perspective co-actionnelle / co-culturelle". *Les langues modernes*, n° 3, pp. 55-71.
- Puren, C. (2004). "Innovation et cohérence en didactique en langue". *New Standpoints*, 12 avril, pp. 2-7.
- Rapport de l'inspection générale sur la place de l'oral dans les enseignements de l'école primaire au lycée*. [ftp://trf.education.gouv.fr/pub/edutel/syst/igen/placeoral.pdf](http://trf.education.gouv.fr/pub/edutel/syst/igen/placeoral.pdf) (lien valide le 06/12/2006).

NOTES

1. Edition 1990, p. 715.
2. <http://atilf.atilf.fr/dendien/scripts/tlfiv5/visusel.exe?26;s=4127825505;r=2;nat=;sol=0>
3. Christian Puren (2004) distingue la centration sur le document de la centration sur l'échange.
4. Ce que Christian Puren appellerait "*une perspective co-actionnelle - co-culturelle*" (Puren, 2002 : 65).
5. <http://www.banqoutils.education.gouv.fr/fic/E3DIPCLA4.pdf>
6. Dialang est un outil d'évaluation élaboré à partir du Cefcl. Disponible en ligne, Dialang permet à tout apprenant de connaître son niveau par langue dans trois compétences (compréhension de l'écrit, expression écrite, compréhension de l'oral), et en grammaire et vocabulaire. Cet outil est téléchargeable gratuitement dans 14 langues : danois, hollandais, anglais, finlandais, français, allemand, grec, islandais, irlandais, italien, norvégien, portugais, espagnol et suédois (<http://www.dialang.org>). À la différence du Portfolio où l'individu est amené à s'autoévaluer en utilisant principalement son propre jugement, ou à la différence du contexte scolaire où les compétences de l'élève sont mesurées par le professeur en fonction de critères qui peuvent rester internes au projet pédagogique ou aux traditions du pays, Dialang est un outil permettant l'évaluation autonome en référence à une norme européenne.

RÉSUMÉS

L'évaluation en langues a fait l'objet de nombreux travaux portant aussi bien sur les types d'évaluation et leurs fonctions que sur les modalités. Pourtant, il semble que de nouvelles données soient à prendre en compte aujourd'hui. Il s'agira ici de mettre en relation des questions repérées, dans la définition même du terme et de son objet, et les ouvertures offertes par de nouveaux outils comme le Cadre européen de référence en langues.

On verra en particulier quel est l'apport du cadre européen en la matière, apport direct ou indirect, à travers l'élaboration de tests comme Dialang ou le DCL, et deux applications concrètes : l'évaluation dans l'apprentissage en Tandem et la Dutch CEFR Grid, outil à destination des concepteurs d'items. Enfin, on esquissera des perspectives pour l'avenir.

The issue of language assessment has already been discussed both in terms of types (or function), and modes. Yet new elements need to be taken into account. Our aim here is to establish connections between obvious questions concerning the definition of the term "assessment," and its object, and the possibilities offered by new tools such as the European Language Framework. We will examine in particular how the Framework has improved the language assessment, in a direct or indirect way, through testing devices like Dialang or the DCL. We will also examine two concrete examples: tandem learning, and the Dutch Framework grid – a tool aimed at helping item conceptors. Finally we will outline future directions for assessment.

INDEX

Mots-clés : évaluation, compétence, cadre européen commun

Keywords : language assessment, skill, European Framework for Languages

AUTEUR

CLAIRE TARDIEU

Claire Tardieu est professeur des universités à l'IUFM de Paris et coauteur d'ouvrages de didactique des langues pour les concours. Elle enseigne à Paris III et Paris IV dans le cadre de la préparation à l'épreuve sur dossier et à Paris VII, où elle anime un séminaire de didactique au sein du Master 2 de linguistique.

Courriel : [claire.tardieu\[at\]numericable.fr](mailto:claire.tardieu@numericable.fr)