

Évaluation et contrôle

Michael FLEMING

ETUDE PRELIMINAIRE
Langues de scolarisation

Division des Politiques linguistiques, Strasbourg
www.coe.int/lang/fr

Évaluation et contrôle

Michael Fleming, *Université de Durham, Royaume-Uni*

Conférence intergouvernementale
Langues de scolarisation: vers un Cadre pour l'Europe

Strasbourg, 16-18 octobre 2006

Division des Politiques linguistiques, Strasbourg

Les vues exprimées dans la présente étude sont celles de l'auteur ; elles ne reflètent pas nécessairement celles du Conseil de l'Europe.

Toute correspondance relative à cette publication ainsi que toute demande de reproduction ou de traduction totale ou partielle du document doivent être adressées au Directeur de l'éducation scolaire, extrascolaire et de l'enseignement supérieur du Conseil de l'Europe (F-67075 Strasbourg Cedex) .

La reproduction d'extraits est autorisée, sauf à des fins commerciales, à condition que la source soit mentionnée.

Division des Politiques linguistiques

DG IV – Direction de l'éducation scolaire et extra-scolaire et de l'enseignement supérieur

Conseil de l'Europe, F-67075 Strasbourg Cedex

Sommaire

Introduction 6
1. Buts de l'évaluation 6
2. Méthodes d'évaluation 8
3. La langue en tant que matière scolaire 10
4. Langue(s) d'enseignement (des autres matières) 11
5. La ou les langues de scolarisation dans le contexte du plurilinguisme 11

Synthèse

Le contrôle occasionne souvent des tensions et des désaccords en raison des fonctions diverses et exigeantes que l'intéressé est censé remplir. Ces divergences peuvent se compliquer dans le cadre de la langue en tant que matière scolaire, car l'éventail des buts et produits y est particulièrement étendu. Il importe d'examiner et d'expliciter les valeurs qui sous-tendent et informent un système de contrôle, faute de quoi les apprenants risquent d'en subir sans le savoir des conséquences fâcheuses. L'un des rôles d'un Cadre pour la ou les langues de scolarisation serait d'expliciter les différentes fonctions de contrôle et d'illustrer comment les intégrer dans la pratique afin d'élever les normes.

Introduction

Peu d'éducateurs nieraient l'importance essentielle de l'évaluation. Depuis l'enseignant dans sa classe jusqu'aux personnes chargées de formuler une politique, tout le monde a intérêt à savoir si l'enseignement a été couronné de succès ou non et dans quelle mesure il l'a été. L'attention que les chercheurs accordent à l'évaluation depuis quelques années a fait apparaître toute une terminologie visant à faciliter la communication et à clarifier la mise au point de la politique comme de la pratique. Les expressions techniques abondent et sont devenues très familières : par exemple, évaluation normative ; évaluation critériée ; formative et sommative ; valeur ajoutée ; basée sur la compétence (évaluation de la capacité)[1]. Malgré le partage d'une même terminologie et le dialogue, l'évaluation est souvent un foyer de désaccords dans le monde de l'éducation, ce qui entraîne parfois des retranchements et des désaffections. Il importe donc qu'un Cadre pour la ou les langues de scolarisation traite des questions relatives à l'évaluation, ainsi qu'à l'amélioration des normes qualitatives éducationnelles. Ce bref document a pour objet non pas d'examiner en détail les points techniques relatifs à l'évaluation, mais d'exposer en termes généraux les principales questions et sources de désaccord.

1. Buts de l'évaluation

Dans sa formulation la plus simple, l'évaluation informe sur le point de savoir si l'enseignement ou l'apprentissage a été couronné de succès ou non. Toutefois, elle fournit des informations à plusieurs auditoires potentiels n'ayant pas tous les mêmes exigences précises. Les enseignants ont besoin de s'informer périodiquement sur la manière dont se développent les connaissances, les aptitudes et la compréhension des élèves, afin de savoir comment adapter leur enseignement et quel genre de feedback s'impose pour améliorer l'apprentissage des élèves. Les directeurs d'établissement et les décideurs nationaux ont besoin d'informations supplémentaires et plus étendues sur la qualité de l'éducation, dans une école pour les uns, dans un pays pour les autres. Les données comparatives nécessaires à cette fin doivent présenter un degré élevé de fiabilité et d'uniformité. S'agissant de la langue en tant que matière scolaire, satisfaire à cet impératif est extrêmement malaisé, car il est difficile de créer des épreuves qui soient gérables tout en étant fidèles aux objectifs de la matière. Les employeurs et la société dans son ensemble ont besoin d'informations fiables pouvant contribuer à certifier un niveau de réussite et à servir de base pour la sélection. Les parents ont besoin d'informations qui les aident à apprécier le degré de réussite et les limites de leurs enfants. Enfin, les apprenants eux-mêmes ont besoin de savoir comment ils progressent et comment améliorer leurs résultats, mais aussi – parfois – d'être protégés des effets potentiellement démotivants d'une évaluation négative.

Employée en rapport avec l'évaluation, la notion de « responsabilité » renvoie à la mise en place de systèmes d'évaluation extérieurs au processus d'apprentissage, en tant que moyen de « mettre de l'ordre » dans les normes afin d'assurer le fonctionnement efficace du système d'enseignement. Mais le terme peut s'employer dans un sens plus large et plus positif lorsqu'il s'applique aux différentes obligations

[1] Pour une liste commentée de types d'évaluations, voir *Cadre européen commun de référence pour les langues*, Conseil de l'Europe, 2001, Editions Didier, p. 139 à 145.

incombant à tous. Les enseignants ont une responsabilité vis-à-vis des apprenants, mais aussi de la société et de ses besoins. Les décideurs ont manifestement un devoir à remplir vis-à-vis des citoyens et doivent veiller à ce que le système d'enseignement produise des résultats, mais ils ont aussi des responsabilités vis-à-vis de chaque apprenant, et c'est dans cette optique qu'il leur faut envisager les conséquences des mesures prises par eux. Interprétée ainsi, la notion de responsabilité amène chacun à regarder au-delà de ses intérêts purement personnels pour envisager la question dans sa globalité. On doit relier la responsabilité à un processus consistant à partager des perceptions et à favoriser la compréhension. Il importe donc de ne pas exagérer les disparités entre « parties prenantes » potentielles ; ce que celles-ci ont en commun, c'est le souci fondamental de voir l'évaluation contribuer à élever le niveau de réussite et à améliorer l'apprentissage. Pour que se relâchent les tensions liées à tout ce qui concerne l'évaluation, il faut commencer par développer la compréhension d'autres points de vue. Un Cadre pour la ou les langues de scolarisation contribuerait dans une large mesure à préciser les diverses fonctions de l'évaluation et la manière dont celles-ci peuvent s'intégrer dans la pratique.

Il peut être utile aussi d'établir une distinction selon que l'évaluation est effectuée à des fins sommatives ou appréciatives[2]. L'évaluation sommative consiste à juger sommairement le niveau d'un apprenant à un moment donné, ce qui aboutit d'ordinaire à une notation. Le contrôle évaluatif, en revanche, a pour but de fournir des informations plus générales afin de mettre en lumière les niveaux moyens de capacité atteints par des groupes dans une école, une autorité de l'enseignement ou un pays. Cette distinction est utile, car il est possible de remplir la fonction évaluative du contrôle en pratiquant un échantillonnage au lieu d'employer les données sommatives afférentes à une population entière. En ce qui concerne les grands élèves, il est probable que l'on prendra comme données évaluatives les données sommatives obtenues par le biais des systèmes nationaux d'évaluation souvent utilisés à la fin de la scolarité obligatoire. En ce qui concerne les élèves plus jeunes, en revanche, il peut suffire de tirer des données évaluatives d'un échantillonnage pour que les fonctions sommatives et les fonctions formatives de l'évaluation puissent fusionner entre elles. Lorsque les données d'évaluation servent à comparer les progrès de plusieurs cohortes d'élèves, on voit se généraliser le recours à la méthode de la « valeur ajoutée », qui tient compte des niveaux-planchers à partir desquels les intéressés ont commencé de progresser.

Une autre distinction utile est faite entre « évaluation » et « suivi ». (On notera que certains auteurs appellent « évaluation » ce qui est désigné ici sous le vocable « suivi »[3].) Si l'« évaluation » sert souvent à désigner un processus de mesure plus formel ou l'attribution d'une note, d'un niveau ou d'une mention, le « suivi » est un terme plus large qui englobe de manière moins précise la notion de jugement. Cette distinction est particulièrement appropriée dans le contexte de l'enseignement d'une langue en tant que matière scolaire, où les objectifs sont forcément variés et complexes (voir le document de Laila Aase). Il est difficile, par exemple, d'attribuer

[2] à ne pas confondre avec l'opposition entre types d'évaluation sommatifs et formatifs (voir note 1 en bas de page).

[3] On notera aussi, toutefois, l'emploi suivant du terme « évaluation » : « Évaluation est un terme plus large que contrôle. Tout **contrôle** est une forme d'évaluation, mais dans un programme de langue, la compétence de l'apprenant n'est pas la seule chose évaluée – par exemple, la rentabilité de certains matériels pédagogiques ou méthodes, le type et la qualité du discours effectivement produit, la satisfaction de l'enseignant et celle de l'apprenant, l'efficacité de l'enseignement, etc. » (*Cadre européen commun de référence pour les langues*, Conseil de l'Europe, 2001, Editions Didier, p. 135).

formellement un niveau ou une mention à des usages créatifs de la langue. De même, si le maître peut être en mesure de juger d'une façon informelle l'enthousiasme d'un élève pour la lecture, il lui est malaisé d'attribuer un niveau précis ou une mention précise à une aptitude de ce genre en fonction des critères qui la dénotent. Le suivi effectif consiste à guetter tous les signes pouvant fournir une indication quelconque sur la réussite ou le manque de succès d'un élève. S'y attache, en outre, la connotation d'un processus réparti dans le temps. Pour être efficace, l'enseignant doit être conscient de tous les indicateurs relatifs à la compréhension de l'élève, que des évaluations formelles aient lieu ou non. Dans la classe de langue, ces indicateurs peuvent être les suivants : participation aux discussions de groupe ; conversations individuelles, informelles ; lecture à haute voix ; théâtre ; dessins et diagrammes. Néanmoins, ces activités ne se prêtent pas toutes facilement à une évaluation formelle. Un processus de suivi débouchera inévitablement sur une évaluation formative.

Un Cadre pour la ou les langues de scolarisation devrait traiter de la manière dont le feedback oral et verbal donné aux élèves peut servir le mieux à améliorer l'apprentissage dans le contexte du processus d'évaluation formative, en puisant dans les résultats des recherches les plus récentes. Il existe un danger, par exemple, de faire aux élèves des remarques excessivement négatives ne pouvant que les décevoir et les démotiver. Mais il y a aussi le danger de leur faire des commentaires irréalistes à force d'être trop positifs, ce qui risque de susciter en eux un sentiment d'autosatisfaction.

2. Méthodes d'évaluation

L'évaluation de la langue en tant que matière scolaire a toujours revêtu une forme très simple : les élèves avaient à faire un petit devoir écrit leur valant une mention ou une note. Cela leur permettait (ainsi qu'au maître) de juger de leur rang par rapport au reste du groupe (évaluation normative). Étant donné, toutefois, l'absence de critère précis, cette information indiquait rarement de quelle manière ils pouvaient progresser dans leur apprentissage. De même, l'épreuve proprement dite traduisait une conception très étroite de la compétence langagière désirée ; elle était fréquemment axée sur la connaissance de la forme et des structures de la langue (syntaxe et grammaire), de même que sur un éventail étroit d'emplois de la langue (souvent, une simple rédaction). Or, il importe que tout système d'évaluation prenne en compte le caractère très vaste de la langue en tant que matière scolaire et la complexité des buts de cette matière.

La multiplicité des objectifs, qui reflète celle des auditoires potentiels, illustre la difficulté qu'il y a de concevoir des instruments d'évaluation. Deux grandes tendances existent à cet égard. L'une met l'accent sur l'évaluation *de* l'apprentissage, aux fins de laquelle des mesures fiables et objectives s'imposent en priorité. On insiste, à cet égard, sur les jugements sommatifs, ce qui peut nécessiter en pratique des épreuves et examens plus formels, ainsi qu'une notation destinée à assurer la sûreté du processus. L'autre méthode consiste à abandonner l'évaluation *de* l'apprentissage au profit d'une évaluation *pour* l'apprentissage, en vue de laquelle on insiste beaucoup plus sur le feedback afin d'améliorer les réalisations. Cette dernière méthode peut être appliquée au moyen d'un travail en classe et d'une évaluation à l'aide d'un portfolio, aux fins desquels il est possible de rassembler diverses informations reflétant le caractère vraiment très vaste de la matière en question. Autre avantage d'une telle démarche : il est possible d'y englober différentes formes d'auto-évaluation pouvant constituer des moyens utiles de motiver les apprenants et de les amener à réfléchir

sur leurs progrès. L'auto-évaluation encourage les élèves à assumer la responsabilité du processus d'apprentissage, bien qu'à cette fin, ils aient intérêt à se former aux techniques pertinentes pour que la méthode soit efficace. Plus la méthode d'évaluation est générale (et comprend le jugement d'une série de réalisations différentes dans divers contextes), plus on peut dire qu'elle est pertinente dans la matière considérée. Des tensions apparaissent cependant, car il est parfois difficile de comparer avec un minimum de précision les résultats d'une méthode générale d'évaluation. Si elle est guidée par d'étroites préoccupations comptables, la recherche de méthodes d'évaluation « objectives » et fiables présente plusieurs dangers. L'enseignement axé sur l'examen ne pose pas forcément problème si les épreuves sont complexes et très variées, mais il peut s'avérer difficile, en pratique, de gérer celles qui sont excessivement compliquées. Des épreuves trop étroites et simplistes peuvent nuire à l'enseignement.

Le système d'évaluation idéal refléterait toute la complexité du langage en tant que matière scolaire et motiverait les apprenants en suscitant un feedback utile, tout en fournissant aux autres parties prenantes (par exemple, les décideurs et les employeurs) les informations dont elles ont besoin. Une approche intégrée de l'évaluation garantirait le maintien d'un équilibre entre les différents objectifs et méthodes, de telle sorte qu'aucun objectif n'ait d'influence indue et négative sur l'ensemble du système.

La notion de « transparence » comporte l'idée-clé selon laquelle les personnes soumises à l'évaluation sont au courant des critères utilisés pour les juger et de la manière dont on les juge. La connaissance de ces critères peut améliorer les réalisations et la motivation, mais là encore, dans le contexte de la langue en tant que matière, tout est plus complexe qu'il n'y paraît de prime abord. On estime communément que les élèves apprennent le mieux lorsqu'ils savent ce qu'ils s'efforcent d'obtenir et comment l'obtenir. Bien que ce point de vue soit largement fondé, il admet des exceptions. Étant donné que le développement du langage peut se décrire à certains égards comme un processus « naturel », les apprenants n'ont pas toujours besoin, pour faire des progrès, de se concentrer entièrement sur tel ou tel aspect de leurs réalisations, et une concentration excessive sur ces dernières risque même de se révéler gênante pour eux : le langage d'un locuteur peut sembler affecté, et un scripteur à qui l'on a dit de rechercher l'effet en employant davantage d'adjectifs peut écrire avec un style emprunté et artificiel. Sans remettre en cause la nécessité de la transparence comme principe, ces remarques soulignent le fait que dans la pratique pédagogique, tout principe doit s'interpréter et s'appliquer avec prudence.

L'utilisation des compétences pour décrire les produits évaluables s'est considérablement développée ces dernières années. Il est bon de noter que les auteurs emploient le terme « compétence » dans des sens différents, ce qui peut être une source de confusion. Parfois, d'une manière très générale, il est synonyme de « capacité », comme dans l'expression « compétence langagière ». D'autres auteurs emploient ce mot pour décrire de vastes modes ou domaines langagiers tels que la lecture, l'écriture, l'expression écrite et la compréhension de l'oral. Plus communément, toutefois, les « compétences » renvoient à des actions précises que l'apprenant doit accomplir et que l'on peut évaluer pour mettre en lumière les résultats qu'il a obtenus dans une matière donnée. Les partisans de l'usage des relevés de compétence aux fins d'évaluation et d'élaboration des plans d'étude soulignent que leur avantage tient surtout à la clarté et à la transparence qu'ils apportent à la spécification des produits de l'apprentissage. Ceux qui critiquent l'approche fondée sur la compétence estiment qu'étant trop étroits et trop spécifiques, les relevés de

réalisation ne reflètent ni tout l'éventail, ni toute la subtilité de ce qui est en jeu dans le développement du langage. Il y a donc lieu de mettre en balance les avantages et les inconvénients de cette formule.

3. La langue en tant que matière scolaire

La difficulté d'évaluer la langue en tant que matière scolaire tient, entre autres, à ce que le contenu en est extrêmement varié et complexe (voir l'étude de Florentina Sâmihaian), ce qui reflète d'ailleurs la complexité des objectifs. Chaque mode d'« écriture », de « lecture », d'« expression orale » et de « compréhension de l'oral » peut se décomposer en plusieurs sous-modes. Or, une matière tellement multidimensionnelle soulève la question de savoir si une tâche d'évaluation effectuée dans un sous-mode de lecture, par exemple, indique forcément une compétence pour lire. D'autre part, on ne voit pas bien s'il faut partir de l'idée que la capacité de lecture varie en fonction du type de texte (fiction, document, article de presse). L'évaluation de la lecture peut aisément rester au niveau superficiel si elle porte sur la mémoire ou la compréhension littérale plus que sur la compréhension profonde. Les questions à choix multiple portant sur un texte sont faciles à noter et peuvent présenter une fiabilité élevée (sur le plan technique), mais conviennent moins bien pour évaluer la réaction personnelle profonde de l'apprenant à un texte. Même avec de jeunes enfants, le processus de lecture représente davantage que le simple décodage d'un texte, et tout système d'évaluation doit en tenir compte.

Une question analogue se pose à propos de l'écriture. Les élèves doivent-ils être évalués sur un éventail de tâches diverses dans la mesure où l'écriture peut avoir plusieurs buts et s'adresser à différentes audiences ? L'évaluation de l'écriture semble aller davantage de soi que celle de la lecture et de l'expression orale, car du moins produit-elle un objet auquel on peut se référer après coup. Mais il est excessivement difficile de dire quel critère doit dominer lorsqu'on effectue un jugement (par exemple, précision de l'écriture ou valeur de ce qui est exprimé ?) et si le critère doit être choisi en fonction du type d'écrit dont il s'agit. Il serait faux de croire que la seule réaction à ce qu'écrit l'élève vient une fois le travail achevé. L'évaluation formative se traduisant par un dialogue sur le travail en cours constitue un important moyen d'améliorer les normes. Comme l'évaluation est toujours un choix, celui qui la pratique risque de valoriser par mégarde certains aspects de la matière par rapport à d'autres.

Il est extrêmement difficile d'évaluer l'expression orale et la compréhension de l'oral, car plus encore que pour d'autres aspects de la langue en tant que matière, la réalisation varie avec le thème, le contexte et le niveau de motivation. Une mauvaise réalisation dans un travail oral tient souvent à la nature de la tâche, à l'atmosphère de la classe et à la dynamique de groupe plus qu'à la compétence de l'élève ; pour évaluer valablement l'expression orale et la compréhension de l'oral, il faut chercher des indices dans diverses situations. On pourrait soutenir que l'évaluation de l'expression orale et de la compréhension de l'oral est si complexe, et le contenu de ces dernières si spécifique qu'il ne faut pas l'effectuer de façon formelle. Mais comme c'est l'évaluation qui détermine très souvent le programme d'enseignement et la manière dont il est appliqué, on a des raisons de penser que l'expression orale et la compréhension de l'oral doivent être évaluées en dépit des difficultés. La capacité d'exposer oralement un point de vue et de soutenir une argumentation est essentielle pour une véritable participation à la vie démocratique. Un Cadre pour la ou les langues de scolarisation contribuerait dans une large mesure à offrir des exemples pratiques de la manière dont tous les modes langagiers peuvent être évalués.

4. Langue(s) d'enseignement (des autres matières)

Quelles sont, en matière d'évaluation, les incidences d'une politique promouvant la langue dans la totalité du programme ? Si certains aspects de la langue sont enseignés dans différentes matières, il se pose deux questions essentielles : D'abord, la méthode employée pour l'évaluation de la matière reflète-t-elle les réalisations de l'élève dans la langue ? Ensuite, les matières contribuent-elles à l'évaluation de la langue en tant que matière scolaire ? En un sens, on peut répondre à la première question par l'affirmative, car la compréhension de la matière et la capacité de l'exprimer ne sont pas faciles à séparer de la compétence dans l'emploi de la langue. C'est une autre affaire, toutefois, de savoir si ce qui est parfois appelé les éléments « superficiels » de la langue (orthographe, grammaire, ponctuation) doit être pris en compte lorsqu'on évalue formellement les résultats d'un élève dans une matière donnée. La question peut se formuler concrètement dans les termes suivants : y a-t-il une proportion de la note qui rende compte de la précision d'emploi de la langue ? Si tel est le cas, les apprenants risquent d'être pénalisés deux fois. En outre, cette simple question pratique masque des difficultés. Comment, par exemple, détermine-t-on la « précision » d'emploi de la langue ? Quel degré de tolérance mérite, par exemple, l'emploi d'un dialecte dans le discours écrit ? Depuis peu, il est devenu commun, dans certains contextes, de parler de *adéquation* plutôt que de précision d'emploi de la langue. Mais le problème pratique demeure : ne pas reconnaître l'importance d'un usage approprié de la langue en ignorant les « mésusages » peut avoir des conséquences néfastes pour les réalisations de l'élève.

La seconde question mentionne l'apport éventuel d'autres matières à l'évaluation de la langue. Des difficultés concrètes existent manifestement à cet égard, par exemple dans la mesure où il n'ira pas de soi pour un professeur d'histoire de contribuer à évaluer les réalisations de ses élèves dans la langue en tant que matière. Si l'on donne à la langue d'enseignement (des autres matières) l'importance qui lui revient, elle doit forcément avoir un certain impact sur la manière dont la langue est évaluée. Deux facteurs sous-jacents sont à mentionner à cet égard : les frontières entre matières scolaires et le point de savoir jusqu'où la compétence langagière est facilement séparable du contexte dans lequel la langue est employée ; ainsi, pourrait-on dire des élèves qui apprennent l'histoire qu'ils apprennent aussi le langage de l'histoire.

5. La ou les langues de scolarisation dans le contexte du plurilinguisme

Lorsqu'on envisage la langue dans le cadre plus vaste de la volonté d'appliquer une politique de plurilinguisme, on perçoit mieux certaines des questions à prendre en compte aux fins d'évaluation. Il convient de prêter attention, d'une part au large éventail de contextes dans lesquels la langue peut s'employer, formellement ou non (par exemple, le langage écrit revêt différentes formes : courriels, articles, rapports, études universitaires), d'autre part aux divers buts dans lesquels on peut l'employer (par exemple, pour persuader, informer, débattre, rendre compte). La notion de plurilinguisme qui, dans son sens premier, désigne le fait de parler plusieurs langues, peut s'étendre à la langue en tant que matière. Ceux qui apprennent une langue ont le droit d'être exposés à une vaste gamme d'*usages* langagiers pouvant apparaître comme l'un des aspects d'une politique de plurilinguisme ; ils sont en droit également de développer tous ces usages. L'évaluation n'est pas un processus neutre, car elle fait inévitablement appel à des valeurs particulières qu'il faut examiner sous un angle spécifique dans la mesure où la diversité des façons d'aborder l'évaluation peut avoir des conséquences non souhaitées. Les conceptions des connaissances et de

l’apprentissage varient en fonction de l’approche retenue, d’où le risque de favoriser un apprentissage essentiellement superficiel. L’évaluation peut aussi exacerber le désavantage que certains apprenants subissent lorsqu’ils passent des épreuves culturellement tendancieuses ou risquant de sous-estimer la compétence de certains élèves en raison de la trop faible portée qu’elles ont. Un cadre pour la ou les langues de scolarisation servirait à expliciter les diverses fonctions de l’évaluation et certaines des conséquences non souhaitées de cette dernière, ainsi qu’à faire la démonstration des moyens concrets de traiter les conséquences en question.