

Claire Tardieu

IUFM de Paris – Université Paris 4 Sorbonne

Corriger ou évaluer ?

Mots-clés : Corriger, évaluer, noter, *CECRL*, psychologie cognitive

Résumé : Corriger, évaluer, noter, faute, erreur, etc. Autant de termes utilisés couramment par les uns et les autres sans que le sens en soit toujours précisé. Or, les mots pour le dire ne sont pas anodins. La question des significations –celles du dictionnaire et, au-delà, des sens attribués (Vygotski *in* Vergnaud 2000 : 64) – est d’autant plus importante dans le domaine de l’enseignement des langues que les conceptions sont invitées à évoluer avec l’adoption du *Cadre Européen Commun de Référence pour les Langues (CECRL)* et l’influence des études en psychologie cognitive. Le *Cadre Européen* nous encourage à renouveler notre vision de l’évaluation ; la psychologie cognitive souligne l’importance du climat de confiance et de l’estime de soi pour l’apprentissage (de La Garanderie 1987 ; Bandura 2007). Comment définir ces différents termes ? Quels rapports entretiennent-ils les uns avec les autres ? Certaines traditions de la correction semblent aujourd’hui remises en cause par le discours institutionnel lui-même : comment justifier ces incitations aux changements ? Quelles sont les pistes de réponses offertes par certaines sciences contributives de la didactique des langues ?

Correction or assessment?

Key words: Correction, assessment, marking, *CEFRL*, cognitive psychology

Abstract: Correction, assessment, marking, mistake, error, and so on. All words everyone commonly uses in teaching without necessarily taking the time to define them precisely. Yet, the words we choose to express these issues are not neutral. The question of their meanings and of the connotations they convey (Vygotski *in* Vergnaud 2000: 64) is all the more important in the field of language teaching with the adoption of a *Common European Framework of Reference for Languages (CEFRL)* and the increasing influence of cognitive psychology. Indeed, both invite us to modify our ideas. The *European Framework* encourages us to reconsider our basic conceptions of assessment. current research in cognitive psychology emphasizes the importance of establishing a climate of confidence and of enhancing self-esteem to improve learning (de La Garanderie 1987; Bandura 2007). How does one define these various terms? What relationship do they have with each other? Some traditional ways of assessing have been questioned within the institutional discourse itself: how can we justify this invitation to change? What answers do contributory sciences have to offer foreign language didactics?

Corriger ou évaluer ?

Lors d'un séjour au Danemark, l'une de mes collègues eut l'occasion de visiter des établissements scolaires. Elle assista à un cours de musique où les élèves apprenaient une chanson. C'était, selon elle, une vraie cacophonie, mais le professeur danois continuait à faire chanter les élèves comme si de rien n'était. À la fin de la séance, ma collègue lui demanda pourquoi il laissait les enfants chanter faux sans les corriger. Il répondit, pourquoi les corriger ? Si je les corrige, ils perdront le plaisir de chanter. De toute façon, dans quinze jours, ils chanteront juste. Et effectivement, quinze jours plus tard, les élèves chantaient juste.

Ce récit a de quoi laisser perplexe. Est-ce à dire que l'apprentissage se fait naturellement sans qu'il soit nécessaire de corriger ? N'est-ce pas là faire l'amalgame entre apprentissage naturel et apprentissage en milieu institutionnel et, d'une certaine manière, donner corps à la théorie de Krashen *et al.*¹, rejetée par de nombreux chercheurs en didactique des langues ? Avant même de pouvoir s'interroger sur la pertinence de telle ou telle pratique par rapport à telle autre, encore faut-il s'entendre sur le choix des termes.

Corriger, évaluer, noter, faute, erreur, etc. Autant de termes utilisés couramment par les uns et les autres sans que le sens en soit toujours précisé. Or, les mots pour le dire ne sont pas anodins. Que signifient véritablement ces expressions dans la bouche du professeur : « J'ai corrigé vos devoirs », « je vais vous rendre vos évaluations », « il me faut une note d'oral » ou, dans la bouche de l'élève : « ce sera noté, Madame ? », ou encore dans le discours institutionnel : « Tous les exercices utilisés pour une évaluation ne doivent pas forcément faire l'objet d'une notation. » (*Anglais – Classe de seconde générale et technologique, Accompagnements des programmes*, 2003 : 24).

* Claire Tardieu est Professeur des universités à l'IUFM (Institut de formation des maîtres) de Paris Sorbonne où elle enseigne la didactique des langues. Elle dirige un séminaire à Paris 3 Sorbonne Nouvelle (EA 4398 PRISMES – LILT). Elle a écrit plusieurs articles et ouvrages sur la didactique de l'anglais. Nommée en tant qu'experte française dans le domaine de l'évaluation par le Ministère de l'Éducation, elle a pris part à deux projets internationaux : la Dutch CEFR Grid et EBAFLS et mène actuellement le projet CEF-ESTIM au Centre Européen des Langues Vivantes. <cltardieu@orange.fr>.

¹ Krashen (1983) privilégie « l'acquisition » par rapport à « l'apprentissage » et minimise l'importance du *monitoring*.

La question du choix des termes, de leur signification et, au-delà, de leur sens attribué² est d'autant plus importante dans le domaine qui nous concerne que les conceptions sont invitées à évoluer avec l'adoption du *Cadre Européen Commun de Référence pour les Langues* (désormais *CECRL*) et l'influence des études en psychologie cognitive. Le *CECRL* développe une vision positive de l'évaluation ; la psychologie cognitive souligne l'importance du climat de confiance et de l'estime de soi pour l'apprentissage. Sans doute l'explicitation des termes utilisés en matière d'évaluation constitue-t-elle un préalable à toute réflexion véritable sur l'orientation nouvelle à donner ou non aux pratiques enseignantes. Comment définir ces différents termes ? Quels rapports entretiennent-ils les uns avec les autres ? Certaines traditions de la correction semblent aujourd'hui remises en cause par le discours institutionnel lui-même : comment justifier ces incitations aux changements ? Quels sont les apports de certaines sciences contributives de la didactique des langues ? Autant d'interrogations auxquelles nous nous efforcerons d'apporter des éléments de réponses.

1. Définitions des termes³

Évaluer

« Évaluer » vient de *value* (ancien français) qui signifie : « valeur, prix » ; ou encore du latin *valere* : « être fort, valoir ». Ces deux étymologies semblent renvoyer à deux notions complémentaires, l'une quantitative, l'autre qualitative, l'une matérielle, l'autre morale ou spirituelle. Dans le dictionnaire *Trésor de la langue française* on propose les trois significations suivantes :

- « déterminer, délimiter, fixer avec précision » ;
- « conjecturer, faire l'estimation d'une quantité, d'une durée qui n'est pas encore vérifiable » ;
- « reconnaître la valeur de, estimer ».

On remarque trois notions distinctes : détermination, conjecture, reconnaissance.

La première renvoie à un calcul précis, à une mesure ; la deuxième à une estimation approximative ; la troisième à une appréciation. Ces trois notions semblent se contredire partiellement. Comment concilier en effet « détermination précise » et « estimation approximative » ? La première définition suggère un jugement précis, avéré, objectif, relatif à une norme. La seconde entend plutôt un jugement global, prédictif et subjectif ; quant à la troisième, elle autorise la dimension subjective, voire affective.

² Vygotski distingue la signification du mot (celle du dictionnaire) de son sens qui, lui, représente tous les faits psychologiques que ce mot fait apparaître dans notre conscience (Vergnaud 2000 : 64).

³ On se référera ici à trois dictionnaires : *Littré*, *Petit Larousse* et *Trésor de la langue française*, aux définitions souvent complémentaires.

Historiquement parlant, le terme est relativement récent dans le domaine de l'enseignement-apprentissage (Tardieu 2005 : 9). Il apparaît avec le courant behavioriste initié par Watson dès 1913, appliqué à l'apprentissage des langues par Skinner (1957) qui l'utilise pour référer à la mesure (signification 1) d'un changement de comportement initié par l'apprentissage, comme nous le rappelle Niklas :

On soumet donc l'écolier 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. (2003 : 173)

Les machines à apprendre de Skinner (1969) permettent à l'apprenant d'obtenir un feedback sur ses réponses qui joue le rôle d'agent de renforcement et l'autorise à poursuivre plus avant son apprentissage. En parallèle, à partir des années 1950, les travaux de Tyler (1950) et de Bloom (1969) sur la pédagogie de maîtrise développent l'idée d'un apprentissage par objectifs. Évaluer est l'opération par laquelle on détermine la congruence entre la performance et les objectifs (Hadji 1990 : 31). En 1971, Bloom reprend la notion d'évaluation formative proposée par Scriven quelques années plus tôt (1967) pour désigner la fonction de régulation de l'apprentissage (cité par Hadji 1997 : 8).

En France, on commence à « évaluer » dans les années 1970 en parallèle avec la création de référentiels d'apprentissage et l'introduction de « la pédagogie par objectifs », ceci essentiellement dans les lycées professionnels (Quivy & Tardieu 2002 : 82-83). Puis le terme se généralise, apparaît dans les textes officiels de 1985 pour le collège, 1987 pour le lycée, décliné en évaluation formative et sommative, puis diagnostique dans les années 1990 (avec l'instauration des livrets d'évaluation de début de seconde en 1992) et la création au ministère d'un Département de l'Évaluation et de la Prospective (DEP) en 1987⁴, aujourd'hui DEPP : Direction de l'Évaluation, de la Prospective et de la Performance.

On oppose parfois le terme de « contrôle » qui mesurerait des écarts par rapport à une norme ou demeurerait extérieure à l'élève, à celui d'« évaluation », d'ordre plus qualitatif et associant l'élève. Ainsi, pour Ardoino et Berger, « contrôler » n'est pas « évaluer » (1986 : 121). Pour Hadji, ces deux univers sont complémentaires.

En fait, c'est ce terme de « contrôle » qui est utilisé aujourd'hui pour désigner la stratégie cognitive de vérification par l'apprenant lui-même, ce qui rejoint les notions d'évaluation formatrice et d'auto-régulation, « pièce maîtresse de tout le dis-

⁴ La DEP fut créée en 1987 sous ce nom mais existait précédemment sous diverses appellations en tant que service statistique du ministère. C'est en 1975/1976 seulement qu'est conçue et réalisée la première évaluation d'élèves en France. Par la suite, la DEP est devenue DPD (Direction de la programmation et du développement) en 1997, a repris le nom de DEP en 2003. C'est en 2006 que le deuxième P fut ajouté avec la restructuration du ministère en avril, passant sous la coupe du nouveau Secrétariat général.

positif pédagogique », selon Nunziati (1990 : 57 cité par Hadji 1997 : 63). Le *CECRL* emprunte largement à ces conceptions (*CECRL* 2001 : 75).

Dans tous les cas, évaluer consiste pour Hadji à « gérer le probable », à « peser le présent pour peser sur l'avenir » (1990 : 16-17), que les décisions relèvent de l'apprenant lui-même ou d'instances autres. En termes plus contemporains, on pourrait dire qu'évaluer avant, pendant ou après l'apprentissage relève de l'art de la gouvernance.

En ce sens, évaluer peut être mesurer, contrôler, vérifier, réguler, apprécier, et ne devrait pas être confondu avec deux autres termes plus anciens dans le répertoire institutionnel que sont « corriger » et « noter ».

Corriger

« Corriger » vient du latin *corrigere*, de *cum rigere* pour *regere* : « étendre, dresser ». Le *Littré* donne « ramener au bien ce qui est mal, à la règle ce qui s'en écarte, redresser. Redresser ce qui est fautif ou défectueux. Tempérer, adoucir par quelque mélange de certaines substances. Punir. » On relève les connotations morales du terme et l'idée de châtement corporel : « corriger un enfant », « donner une bonne correction ». La définition qu'en donne le *Trésor de la langue française* atténue quelque peu ces traits. « Corriger » signifie : « Faire disparaître ou relever un écart par rapport à une norme en vue de la rétablir ou de la faire respecter, amender, améliorer » et, en particulier dans le domaine de l'enseignement : « Relever les écarts en vue de noter et faire respecter la norme ».

Ainsi, le terme « corriger » semble aller de pair avec la signification 1 de « évaluer » : « déterminer, délimiter, fixer avec précision » ou encore avec la définition de « contrôle ». Il peut impliquer la référence aussi bien à des critères de performance qu'à une norme curriculaire pré-établie et appelle la rigueur de la note. Bien sûr, « corriger » ne signifie pas nécessairement « noter ».

Noter

« Noter » vient du latin *notare* : « faire une marque sur ». Si l'on se réfère au *Petit Larousse* – « estimer, apprécier le travail, la conduite de quelqu'un » – il semble que « noter » se rapproche davantage de la troisième signification de « évaluer » : « reconnaître la valeur de, estimer ». Mais « noter », c'est aussi, pour *Littré* « Marquer d'une manière défavorable, noter d'infamie, couvrir de honte ». L'ambivalence de la signification se traduit en termes de valeur dans l'attribution de « bonnes » ou de « mauvaises » « marques ».

Pourtant, les travaux de Noizet & Caverni (1978)⁵ ont mis à mal le lien entre valeur et note. Et bien avant eux, dès 1920, d'une manière plus large, Thorndike nommait « effet de halo » le biais cognitif qui affecte notre perception des personnes ou des

⁵ Cités par de Peretti *et al.* 1998 : 45.

firmes commerciales. Soit, mais l'on pourrait arguer que, dans le cas d'une évaluation anonyme comme le baccalauréat, la subjectivité de la correction ne peut jouer. Or, comme le rappellent de Peretti *et alii*, le point de départ de la recherche sur l'évaluation scolaire est précisément la critique des notes mises au Baccalauréat en France, avec Laugier & Weinberg en 1938, puis Piéron en 1963, qui fut l'inventeur, à la fin des années 20, du terme « docimologie » pour désigner « la science et la technique des examens » (Laugier *et alii* 1934 : 24). Piéron avait calculé qu'il faudrait... « 127 correcteurs pour obtenir la valeur vraie d'une dissertation philosophique et 13 pour un devoir de mathématiques, pourtant réputées sciences exactes ! » (1963 : 23, cité par Pasquier 2005 : 17).

Malgré tout, la note est souvent perçue comme une garantie de précision, émanant d'un barème, et quand les enseignants s'avisent de lui substituer une simple appréciation, une explication, un encouragement, ce sont souvent les élèves et les familles qui réclament la marque du chiffre sur la copie. « Noter », comme « corriger » renverrait donc à la signification 1 de « évaluer » dans les représentations habituelles mais à son sens 2 selon la critique docimologique.

Si les termes « corriger » et « noter » peuvent se combiner avec plusieurs significations de « évaluer », leur sens attribué, le plus souvent non questionné, risque de borner leur utilisation à une habitude sociale. Or, c'est à un renouvellement des pratiques évaluatives que l'Institution elle-même nous convie dans *l'Accompagnement des programmes pour la classe de Seconde* :

Rappelons que l'évaluation diagnostique ne donne pas lieu à une note globale. Il en est de même de l'évaluation formative, qui doit être positive et incitative en cela qu'elle est conçue comme un élément de la formation des élèves. Au-delà du simple constat, les appréciations orales ou écrites du professeur peuvent également renseigner l'élève de manière nuancée sur ses réussites et ses points faibles, et lui donner des conseils pour qu'il puisse progresser » (2003 : 24).

Comment renouveler des pratiques fondées sur des concepts qui paraissent faussés au départ, sinon en questionnant le sens caché qui les sous-tend ?

2. Remises en cause diverses

Premier sens caché : l'idée de la norme

« Evaluer » (signification 1), « corriger », « noter », tous ces termes impliquent l'idée d'une norme préexistante. Il s'agit des règles de la langue 2 (désormais L2) telles que la définissent les ouvrages de référence. Or, les travaux de Bourdieu ont montré les liens entre légitimité linguistique et légitimité sociale (2001 : 78). Ainsi, ce dernier critique-t-il l'abstraction opérée par Saussure et lui oppose-t-il une sociologie structurale de la langue qui doit se donner pour objet *la relation qui unit des systèmes structurés de différences linguistiques sociologiquement pertinentes et des systèmes également structurés de différences sociales* (2001 : 83). Autrement dit, le « bon usage » est toujours celui de quelqu'un. Les réformes dans l'enseignement des langues depuis les années 80, de l'approche communicative

première génération à orientation fonctionnelle et sémantique à l'approche communicative deuxième génération à orientation cognitive (grammaire énonciative et stratégies d'apprentissage) à l'approche dite actionnelle d'aujourd'hui (orientation pragmatique et sociolinguistique), ont progressivement entériné cette conception de la L2 en s'efforçant, bon an mal an, de prendre en compte non l'objet langue en soi mais son emploi, dans des contextes socioculturels repérés.

Aujourd'hui, d'autres remises en cause de l'idée de norme sont menées, auxquelles on n'avait peut-être pas songé avant qu'une langue – l'anglais, ou plutôt l'anglais américain – ne prenne un statut de langue de communication à l'échelle du monde. D'ailleurs, le nom même de cette langue est en péril puisqu'on l'appelle aussi *Glo-bish*. La question de l'identité culturelle d'une langue globale se pose en effet et c'est cette question que soulève Thomas Paikeday dans son ouvrage *The Native Speaker is Dead* (1985). Les cercles de locuteurs anglophones qu'il circonscrit débordent largement le cadre de la *sceptred isle* usant de la langue de Shakespeare... Les locuteurs natifs possèdent-ils encore leur langue ou leur échappe-t-elle comme le latin échappa aux Romains et finit par se vulgariser jusqu'à perdre son statut de langue de diffusion ?

Un autre phénomène de « relativisation » de la norme est une conséquence indirecte et non intentionnelle du *CECRL* qui abandonne la seule référence au « niveau-seuil » élaborée dans les années 1970 au profit d'une référence déclinée en trois types d'utilisateurs : élémentaire (A1 et A2), indépendant (B1 et B2), expérimenté (C1 et C2). Le corrélat de cette évolution est l'idée généreuse de « compétence partielle », ouvrant la voie à la différenciation et encourageant l'évaluation positive des performances du locuteur plurilingue (on songe ici en particulier à la formulation choisie dans les descripteurs : « Peut... » ou « Je peux... »).

À titre d'exemple il est intéressant de noter qu'un locuteur capable de « varier l'intonation et placer l'accent phrastique correctement afin d'exprimer de fines nuances de sens » se situe au niveau C1 et C2, tandis que le niveau A2 (attendu en fin de 5^e) est décrit comme suit : « la prononciation est en général suffisamment claire pour être comprise malgré un net accent étranger mais l'interlocuteur devra parfois faire répéter » (*CECRL* 2001 : 92).

Si la norme de référence pour l'apprentissage n'est plus celle du natif, mais celle du descripteur de chaque niveau, pourquoi évaluer les élèves de l'école primaire censés atteindre A1 en fin de CM2 ou ceux du collège, visant A2 en fin de 5^e, B1 en fin de 3^e, ou même ceux du lycée attendus à B1 ou B2, à l'aune des descripteurs de C1 ou C2 ? S'interroger ainsi ne signifie pas que l'on renonce à toute velléité d'enseigner la L2 du locuteur « natif », mais seulement que les exigences en matière d'évaluation devraient être différenciées en fonction des objectifs de l'enseignement.

Beaucoup d'efforts sont fournis actuellement pour lier les tests ou les examens à chaque niveau du *CECRL*. On peut citer à titre d'exemple les échantillons produits

par *ESOL Examinations* de l'Université de Cambridge sous le label du Conseil de l'Europe qui montrent des locuteurs de différents pays du monde évalués en anglais en expression orale en continu et en interaction orale, à différents niveaux du *CECRL* (2004). Des projets européens comme celui de la *Dutch CEFR Grid* (2004) (Tardieu 2006) ou de *EBAFLS*⁶ (2007) s'interrogent sur la possibilité d'établir des corrélations fiables entre niveaux de tests et niveaux du *CECRL*. Et toutes les certifications en langues d'aujourd'hui ont pour ambition de certifier à un niveau précis du *CECRL*, ce qui montre bien qu'en matière d'évaluation la norme du « natif » n'est plus la seule retenue.

Mais alors, quelle note apposer au travail d'un élève au niveau A1, A2 ou B1 ? Note ou « marque d'infamie » pour qui n'a pas encore atteint la maîtrise parfaite ou « marque de gloire » en rapport avec la réussite au niveau escompté ? Réponse de l'inspection générale :

Si l'on peut se réjouir que les professeurs de langue hésitent à donner une note inférieure à 5/20, on observe aussi l'existence d'un idéal inatteignable qui correspondrait à 20/20, mais qui reste mal défini. [...]

La note en langue devrait correspondre à l'estimation donnée à une performance appréciée en rapport avec un niveau cible. Elle devrait renseigner sur les progrès accomplis et sur la façon dont les objectifs sont atteints. [...] Lorsque l'objectif est clairement fixé, la note est une indication du chemin parcouru car elle mesure le degré d'atteinte de cet objectif et uniquement celui-ci. On pourrait alors envisager, pour chaque activité langagière, d'accompagner la note du niveau visé (A1, A2, B1, B2, etc.) et même d'attribuer la note maximale lorsque les objectifs caractérisant un niveau ont été atteints. (2007 : 11)

Première implication : si l'on doit conserver la tradition française qui consiste à mettre des notes tout en adoptant l'optique du *CECRL*, l'attribution de bonnes, voire d'excellentes notes devrait devenir monnaie courante. Ce qui va dans le sens de la démarche d'Antibi qui vise à éradiquer la « constante macabre », cette « proportion de mauvaises notes » que l'enseignant se croit obligé d'utiliser pour demeurer crédible (Antibi 2003 : 15). Un tel changement de mentalité mettrait un terme au « découragement » légitime des élèves puisque, comme le souligne le rapport de l'Inspection générale, les efforts fournis par ces derniers ne seront plus « rendus dérisoires par la confrontation de leur propre niveau avec celui du locuteur natif, idéal et inatteignable » (2007 : 15).

Deuxième sens caché : l'erreur est constitutive de l'apprentissage mais doit être éradiquée

On s'efforce depuis l'adoption de l'approche communicative et cognitive (au milieu des années 1980) de préférer au terme de « faute » celui d'« erreur » en arguant que le premier véhicule des connotations morales et induit des sentiments de culpabilité, alors que le second, purement technique, fournit un renseignement à la

⁶ *Building a European Bank of Anchor Items for Foreign Language Skills.*

fois didactique et psycholinguistique. Ainsi, l'erreur est-elle non seulement inévitable mais constitutive du processus d'apprentissage selon la théorie constructiviste. Pour Gauthier, « l'élève doit se colleter avec l'erreur » (1987 : 22) et, selon Berthoud, c'est « en se trompant qu'on apprend » (1987 : 12). Le *CECRL* reprend à son compte cette différenciation, conservant le terme de « faute » pour désigner les lapsus ou erreurs non liées à un défaut de compétence : « les fautes pour leur part, ont lieu quand l'utilisateur/apprenant est incapable de mettre ses compétences en œuvre, comme ce pourrait être le cas pour un locuteur natif » (1987 : 118), et réservant le terme d'erreur à ce qui est causé « par une déviation ou une représentation déformée de la compétence cible. Il s'agit alors d'une adéquation de la compétence et de la performance de l'apprenant qui a développé des règles différentes de celles de la L2 » (1987 : 118). Ici est fait implicitement référence aux travaux de Corder (1966) et de Selinker (1992) sur l'interlangue – langue transitoire de l'apprenant intermédiaire entre L1 et L2. Dans ces conditions pourquoi ne pas aller jusqu'à dire, comme Narcy-Combes, que le terme d'erreur lui-même est inapproprié puisqu'il s'agit d'un processus constitutif d'un système en construction (2005 : 102-103) ?

On peut aussi s'interroger sur le type d'« erreurs » considérées comme telles par les enseignants et sur les focalisations privilégiées. La propension à corriger les erreurs linguistiques (grammaticales et lexicales en particulier) au détriment des erreurs sociolinguistiques ou socioculturelles, ou encore, pragmatiques, semble avérée au grand dam de certains, dont Chen-Géré (2004), pour qui les malentendus viennent le plus souvent comme leur nom l'indique d'erreurs du deuxième ou du troisième type...

Les auteurs du *CECRL* dressent la nomenclature des pratiques observées en Europe (2001 : 118). Les « mesures à prendre » étant ainsi déclinées des plus strictes aux plus laxistes en passant par les plus soucieuses d'équilibre. Ainsi, pour les uns « toutes les fautes et les erreurs doivent être corrigées immédiatement par l'enseignant », pour les autres « les erreurs devraient être acceptées comme "langue transitoire" et ignorées », tandis que pour d'autres encore, « les fautes qui ne sont guère que des lapsus doivent être ignorées mais les erreurs systématiques doivent disparaître », ou bien « toutes les erreurs devraient être relevées et corrigées lorsque cela n'interfère pas avec la communication (par exemple, en séparant l'objectif de correction de celui d'aisance) ».

Un enseignant peut assez facilement repérer dans cette liste (dont ne figurent ici que des extraits) ses pratiques les plus courantes. On devine en filigrane, outre la référence implicite à Corder et à Selinker, les deux axes de l'apprentissage des langues : *accuracy* et *fluency*. Si l'on considère que la mission de l'enseignant est de permettre à l'élève d'améliorer ses productions en L2, pourquoi ne pas simplement ménager des phases différentes alternant entraînement à la correction de la langue et entraînement à l'aisance ? Ainsi, l'élève saurait ce qui est tour à tour privilégié : la production qualitative de phrases grammaticales qui véhiculent

l'intention de communication du locuteur et possèdent une précision lexicale, une cohérence et une cohésion discursives adéquates, faisant appel à l'intelligence verbale en particulier, ou la production quantitative de L2 avec mélodie, rythme, débit, gestuelle appropriés, sollicitant d'autres formes d'intelligence (musicale, kinésique, interpersonnelle, etc.).

Une exigence qualitative exclusive risque de réduire considérablement la quantité d'expression chez le plus grand nombre. A contrario, une exigence quantitative accrue ne favoriserait-elle pas à terme la qualité de l'expression ? C'est le pari tenté et tenu, depuis 1993, par Kaid Slimane à l'université d'Oran avec des étudiantes en anglais de L1, L2, L3, à travers l'expérience du *go-round copy-book* (in Tardieu 2004). Les étudiantes sont invitées à écrire à tour de rôle dans un journal confidentiel sur des sujets de leur choix, l'objectif étant d'obtenir des productions écrites de plus en plus longues. L'enseignante elle-même s'interdit toute correction, mais pas toute implication puisqu'elle répond régulièrement. L'expérience montre invariablement, sur les trois années, un progrès significatif de l'expression écrite, d'abord au plan quantitatif, puis qualitatif.

Deuxième implication : si l'on doit conserver la tradition française de la correction systématique de l'erreur, du moins devrait-on cantonner cette pratique à des phases précises de l'apprentissage et favoriser davantage la « socialisation en langue étrangère » (Kramsch 1984 : 72).

Troisième sens caché : en classe, on est là pour apprendre. À l'extérieur de la classe, c'est différent.

Aujourd'hui, est remise en cause la notion même de classe. Au sens organisationnel du terme, d'abord, avec la création et la mise en œuvre non encore généralisée des groupes de compétences (2006), en application du *Plan de rénovation de l'enseignement des langues vivantes* (2005). Les trois unités de composition, de temps, et d'action disparaissent. Selon les périodes, les groupes sont remaniés, les enseignants changent, les activités aussi en fonction des compétences visées.

Au sens physique ensuite : la salle de classe est censée ne plus imposer ses limites spatio-temporelles. La démarche actionnelle au cœur du *CECRL* devrait normalement se trouver relayée par une pédagogie de projet, de tâche signifiante où l'apprentissage se fait de manière fortuite et non intentionnelle (Ellis 2003 : 3) à travers des tâches non focalisées (Ellis 2003 : 16), où le sens de l'apprentissage n'est plus finalisé par le texte mais par une macro-tâche, c'est-à-dire « un ensemble d'actions réalistes conduisant à une production langagière non limitée à l'univers scolaire » (Demaizière & Narcy-Combes 2005 : § 6.4). Comment dès lors évaluer de manière scolaire des réalisations non purement scolaires ? Comment noter les productions des élèves dans le cadre d'un projet collaboratif avec des partenaires hors les murs, d'une construction à visée sociale ? Imaginons par exemple des élèves organisant avec leur professeur un voyage dans un pays étranger dont ils étudient la langue. À leur retour, ils sont sollicités par la mairie de leur ville pour venir

présenter leur voyage et donner des conseils pratiques à un groupe de personnes âgées désirant se rendre dans le même pays. Les élèves préparent une présentation, un glossaire, des conseils pratiques, des expressions utiles, etc. Comment évaluer ce travail ? Si le professeur ne le note pas, quelle valeur lui sera attribuée ? S'il le note, quelle échelle utilisera-t-il, en fonction de quels critères ?

Howard Gardner (1996) dirait sans doute que l'évaluation doit se faire ici par « la preuve sur le terrain ». Autrement dit, si les personnes âgées sont contentes de leur après-midi et si leur futur voyage s'en trouve facilité, alors la tâche est « réussie ». Si au contraire elles s'ennuient très vite et ne tirent aucun profit de la séance, alors la tâche est ratée. Pour Howard Gardner, il existe deux conceptions de l'évaluation : le test formalisé et la preuve sur le terrain. Il ajoute :

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. (1996 : 115)

Pourtant, le système scolaire réclame la notation et la remarque de Gardner ne peut s'appliquer tout à fait.

Puisqu'il doit porter une note sur le bulletin, notre professeur pourra sans difficulté distribuer des notes individuelles lors de micro tâches, intentionnelles et focalisées, préparatoires à l'exposé (prises de parole en continu sur une diapositive, exercices de phonologie, rédactions de résumés, etc.). Dans ce cas, les critères scolaires habituels s'appliqueront parfaitement. Mais il aura sans doute à cœur de souligner l'investissement positif de ses élèves et, en se référant au sens 3 de « évaluer », de faire reconnaître au sein même de l'institution scolaire la valeur sociale, individuelle et collective, de leur travail. Pour ce faire, il ne manquera pas de lui attribuer une bonne note, même si la réussite véritable de la tâche ne peut être évaluée au final que par le public non professionnel bénéficiaire de la présentation.

Troisième implication : dans une perspective actionnelle, l'évaluation par la preuve sur le terrain devrait compter autant que l'évaluation « par la preuve scolaire ».

Mais comment cette bonne note, ces bonnes notes seront-elles perçues au sein de la communauté scolaire par des pairs qui n'ont pas l'habitude de voir tous les élèves d'une classe notés au-dessus de 10 ? Comment les parents réagiront-ils à leur tour, passé le premier instant de contentement de voir leur enfant enfin « valorisé » par le système ? Ne faut-il pas revenir au vieil adage *Qui bene amat bene castigat* ? A trop choyer les élèves par des encouragements, des appréciations positives, des notes mirobolantes, ne va-t-on pas les bercer d'illusions et ne rendre la chute que plus cruelle dans le rude monde qui les attend ?

On le voit donc, les orientations actuelles pour l'enseignement-apprentissage des langues remettent quelque peu en cause les conceptions sous-jacentes aux notions « évaluer, corriger, noter » et viennent perturber des pratiques bien ancrées et ad-

mises par la majorité des professeurs de langues interrogés⁷. Pourtant, la psychologie cognitive et la recherche sur l'imagination laissent entendre que ces pratiques non suffisamment questionnées ont des effets secondaires qui sont loin d'être inoffensifs sur l'apprentissage.

3. Apport de la psychologie cognitive

Les croyances d'efficacité personnelle

Dans l'anecdote citée au début, on remarque la volonté de l'enseignant de laisser les élèves prendre plaisir à chanter, mais aussi sa confiance dans le résultat final. Si l'on transpose ce type de comportement professoral à l'enseignement des langues, ne risque-t-on pas de propager l'erreur, de la voir proliférer puis se fossiliser ?⁸ La fossilisation correspondrait pour Gaonac'h à « une procéduralisation de connaissances encore insuffisamment étayées » (2006 : 65). Autrement dit, la création d'automatismes en *free style* sans contrôle ni l'appui de connaissances explicites. Le processus de fossilisation par lequel l'interlangue de l'apprenant semble définitivement figée a été observé en particulier chez les adultes migrants. Ce qui tendrait à conforter l'idée qu'en l'absence de correction, de reprise, de réflexivité induite, les progrès en L2 s'interrompent. Certes, mais encore faut-il qu'il y ait matière à fossilisation, autrement dit qu'un système d'interlangue soit construit...

À l'issue de la publication des résultats de l'enquête PISA⁹ qui montrait que contrairement à leurs homologues anglais enclins à s'exprimer sans réserve à l'écrit, les jeunes Français préféreraient rendre copie blanche plutôt que se tromper, Jean-Richard Cytermann, directeur de la programmation et du développement se demandait alors si la tradition française n'était pas de trop « stigmatiser l'erreur » (2002 : 17). Il y a fort à parier que des élèves invités à prendre la parole ou la plume en toute confiance sans risquer d'être interrompus ou systématiquement corrigés trouveraient plaisir à manipuler la L2, à la jouer, à se l'approprier et n'éprouveraient ni le sentiment de culpabilité lié à la faute ni la peur d'être puni liée aux sens cachés de correction et de notation. Au contraire, ils développeraient un sentiment d'estime de soi et de confiance dans leur potentiel d'apprentissage.

⁷ Voir le *Rapport de l'inspection générale sur l'évaluation en langues*, 2007.

⁸ Fossilisation : terme utilisé [par Selinker 1972] pour désigner un phénomène particulier observé chez certains apprenants, qui, à partir d'un certain niveau « acceptable », cessent de progresser, « malgré l'enseignement qu'ils reçoivent, ou malgré leur utilisation active de la langue dans des situations de communication. » [...] (Gaonac'h 2006 : 65).

⁹ PISA est une enquête menée tous les trois ans auprès de jeunes de 15 ans dans les 30 pays membres de l'OCDE et dans de nombreux pays partenaires. Elle évalue l'acquisition de savoirs et savoir-faire essentiels à la vie quotidienne au terme de la scolarité obligatoire. Les tests portent sur la lecture, la culture mathématique et la culture scientifique et se présentent sous la forme d'un questionnaire de fond. Lors de chaque évaluation, un sujet est privilégié par rapport aux autres. Les premières collectes de données ont eu lieu en 2000, les suivantes en 2003 et en 2006. La prochaine collecte est prévue pour 2009.

Le psychologue Albert Bandura, formé dans les années 1960 au comportementalisme qui soumet le sujet à l'environnement, s'en est éloigné pour développer l'approche sociocognitive dans laquelle l'individu est simultanément agent et objet, par la réflexion sur soi et l'influence sur soi. Les deux concepts clés de la théorie de Bandura sont « l'efficacité personnelle perçue » qui concerne les aptitudes du sujet, et « l'estime de soi » qui concerne la valeur personnelle. Ces deux éléments sont plus ou moins dissociés selon les sujets. Ainsi, quelqu'un peut avoir un sentiment d'efficacité personnelle faible dans certains domaines sans perdre pour autant l'estime de soi et inversement. Les travaux de Bandura ont contribué à montrer que les croyances d'efficacité personnelle sont un facteur prépondérant de l'échec ou de la réussite.

Les croyances d'efficacité constituent un facteur clé dans un système productif de compétence humaine. Des personnes différentes avec des aptitudes identiques, ou la même personne dans des circonstances différentes, peuvent donc obtenir des performances faibles, bonnes ou extraordinaires, selon les variations de leurs croyances d'efficacité personnelle. (63)

Quel impact à long terme peuvent avoir sur de jeunes ou moins jeunes élèves les corrections à répétition, les interruptions intempestives du discours, les annotations virulentes au stylo rouge ?

Un imaginaire positif

Les travaux de Bandura gagneraient à être mis en relation avec ceux de de La Garanderie, Malrieu et Osborn ayant trait au pouvoir de l'imagination sur l'apprentissage. Ainsi, de La Garanderie écrit-il :

Si l'on peut parler de paralysie de l'imagination, c'est uniquement de celle qui positivement construirait l'action, car celle qui va contribuer à l'empêcher est très active : je me vois ridicule, objet de moquerie, sujet de critiques par le moniteur [...] L'enseignant devra renseigner l'élève pour qu'il développe en lui « les bonnes images » et cesse d'être dupe « des mauvaises ». (1987 : 109)

L'idée est donc que pour parvenir à lever cette paralysie constatée chez un certain nombre d'élèves, il est nécessaire de développer un autre imaginaire, une autre vision de soi en rapport avec la L2 (Tardieu 2008 : 165). Pour Osborn :

Nos idées préconçues font obstacle à la solution des problèmes. Nous devons en quelque sorte nous dé-conditionner, perdre l'habitude de refouler notre production d'idées à l'intérieur des limites de notre expérience passée et laisser au contraire notre imagination courir, la bride sur le cou, à la recherche des moyens de résoudre de nouveaux problèmes. (Osborn, 1988 : 39)¹⁰

Malrieu définit l'acte d'imagination comme un acte de transfiguration du passé, résolument tourné vers l'avenir. Il écrit :

Là se trouve l'originalité de la temporalité des imaginaires : ils opèrent un dépaysement, ils nous font sortir des sentiers battus, ils nous transportent ailleurs. [...]

¹⁰ 1^e éd. 1959.

L'imaginaire révèle au rêveur qu'il est potentiellement autre qu'il ne croyait, à l'homme archaïque qu'il est en son corps terre et eau et feu, au lecteur d'Hölderlin qu'il peut être s'il le veut au contact des langages du ciel et de la réalité transcendante de la lumière ». (1967 : 71)

Alors, à quand cet ego langagier lumineux dans la classe de langue ?

Un gros travail reste à effectuer sur nous-même pour faire évoluer nos mentalités, traquer le sens caché des mots que nous utilisons couramment. Notre mission d'éducateur n'est-elle pas de donner confiance à la jeunesse et de la soutenir dans ses aspirations, plutôt que de la corriger systématiquement, de lui tronquer les ailes ?

Suite à une première étude sur la réussite des jeunes, la Fondation pour l'innovation politique, en partenariat avec l'Institut Kairos Future, a mené une enquête internationale sur « Les jeunes face à leur avenir » (2008) pour évaluer l'idée qu'elles se font du travail, de la famille, de la société et de leur avenir.

Cette enquête effectuée auprès de jeunes du monde entier a révélé la « dépression » de la jeunesse française et son manque d'espoir dans l'avenir. Là où les jeunes français ne sont que 39% à croire qu'ils peuvent changer le monde, les jeunes Danois sont 63%... En revanche, les jeunes français sont 54% (record mondial) à estimer que le regard des autres est déterminant. Est-ce vraiment là une surprise ? *Qui bene amat bene amat.*

Références bibliographiques

Ouvrages

Antibi, A. 2003. *La constante macabre*. Math'Adore, Antibes (3^e ed.).

Ardoino, J. & G. Berger. 1986. « L'évaluation comme interprétation ». *Pour n°107* : 120-127.

Bandura, A. 2007. *Auto-efficacité, le sentiment d'efficacité personnelle*, traduction de Jacques Lecomte, préface de Philippe Carré. Bruxelles : Éditions De Boeck Université.

Berthoud, A.-C. 1987. « Les erreurs des élèves : au panier ou sous le microscope ? ». *Les langues modernes n°5* : 11-17.

Bourdieu, P. 2001. *Langage et pouvoir symbolique*. Paris : Seuil.

Bloom, B. J. 1969. *Taxonomie des objectifs pédagogiques*, vol.1. Domaine cognitif. Montréal : Les Presses de l'université de Québec.

Bloom, B. J. 1971. *Handbook on formative and summative evaluation of student learning*. New York : McGraw-Hill Book Co.

Chen-Géré, M.-F. 2004. « L'évaluation de l'expression orale en anglais ». Académie de Paris. [en ligne] <<http://lve.scola.ac-paris.fr/anglais/evaloral.php>>.

- Corder, P. 1966. *Error analysis and interlanguage*. Oxford : Oxford University Press.
- Cytermann, J.-R. 2002. « Évaluation des connaissances et des compétences des élèves de 15 ans : questions et hypothèses formulées à partir de l'étude de l'OCDE ». Paris : *Les rencontres de la Desco*.
- Demaizière, F. & J.-P. Narcy-Combes. 2005. « Méthodologie de la recherche didactique : nativisation, tâches et TIC ». *ALSIC vol.8, n°1* : 45-64.
- Ellis, R. 2003. *Task-based Language Learning and Teaching*, Oxford : Oxford Applied Linguistics.
- Gaonac'h, D. 2006. *L'apprentissage précoce d'une langue étrangère*. Paris : Hachette Education.
- Gardner, H. 1996. *Les intelligences multiples*. Paris : Retz.
- Gauthier, A. 1987. « De l'intérêt qu'il y a pour l'élève à se colleter avec les difficultés de la langue qu'il s'efforce de s'approprier ». *Les langues modernes n°5* : 18-22.
- Hadji, C. 1990. *Evaluation. Les règles du jeu*. Paris : ESF.
- Hadji, C. 1997. *L'évaluation démystifiée*. Paris : ESF, coll. Pratiques & enjeux pédagogiques.
- Kaid Slimane, N. 2004. « The go-round copy-book : une expérience à vivre ». Colloque international *Langues et Modernité*, 13-15 mars 2004. Université d'Oran ES-SENIA.
- Kramsch, C. 1984. *Interaction et discours dans la classe de langue*. Paris : Hatier-Crédif.
- Krashen, S. & T. D. Terrell. 1983. *The natural approach: Language acquisition in the classroom*. Hayward, CA : Alemany Press.
- La Garanderie (de), A. 1987. *Comprendre et imaginer, les gestes mentaux et leur mise en œuvre*. Paidoguides, Le Centurion.
- Laugier, H., H. Piéron, Mme H. Piéron, E. Toulouse & D. Weinberg. 1934. *Etudes docimologiques sur la perfection des examens et concours*. Paris : Conservatoire des arts et métiers.
- Laugier, H. & D. Weinberg. 1938. *Recherches sur la solidarité et l'indépendance des aptitudes intellectuelles d'après les notes des examens écrits au baccalauréat*. Paris.
- Malrieu, P. 1967. *La construction de l'imaginaire*. Bruxelles : Charles Dessart.
- Narcy-Combes, M.-F. 2005. *Précis de didactique. Devenir professeur de langue*. Paris : Ellipses.

- Niklas, H. 2003. In Demorgon, J., E.-A. Lipiansky & B. Müller. *Dynamiques interculturelles pour l'Europe*. Paris : Editions Economica, Collection Anthropos.
- Noizet, G. & J.-P. Caverni. 1978. *Psychologie de l'évaluation scolaire*. Paris : PUF.
- Nunziati, G. 1990. « Pour construire un dispositif d'évaluation formatrice ». *Cahiers pédagogiques n°280* : 47-64.
- Osborn, A. F. 1988. *Créativité. L'imagination constructive*. Paris : Bordas. (Trad. *Applied Imagination* par Charles Scribner's Sons, New York (1^e éd. 1959; 2^e éd. 1964).
- Paikeday, T. 1985. *The native speaker is dead!* Missisauga, Ontario : Paikeday.
- Pasquier, D. 2005. *Les compétences à apprendre. L'évaluation chez l'adulte*. Paris : L'Harmattan.
- Peretti (de), A., J. Boniface & J. A. Legrand. 1998. *Encyclopédie de l'évaluation en formation et en éducation, Guide pratique*. Paris : ESF Editeur.
- Piéron, H. 1963. *Examens et docimologie*. Paris : PUF.
- Quivy, M. & C. Tardieu. 2002. *Glossaire de didactique de l'anglais*. Paris : Ellipses.
- Scriven, M. 1967. « The methodology of evaluation ». In *Perspectives of curriculum evaluation*. Chicago : Rand Mc Nally, 39-83.
- Selinker, L. 1972. « Interlanguage ». *International Review of Applied Linguistics*, 10, (3) : 209-231.
- Selinker, L. 1992. *Rediscovering interlanguage*. Londres : Longman.
- Skinner, B. F. 1957. *Verbal Behaviour*. Englewood Cliffs (N.J.) : Prentice Hall.
- Skinner, B. F. 1969. *La révolution scientifique de l'enseignement*. Coll. Psychologie et sciences humaines. Bruxelles : Charles Dessart, éditeur.
- Tardieu, C. 2004. « Quelques réflexions sur le compte rendu de l'évaluation des compétences en anglais des élèves de 15 à 16 ans ». *Les langues modernes n°2* : 66-78.
- Tardieu, C. 2005. *L'épreuve de didactique aux concours. L'évaluation en langues*. Paris : Ellipses. Collection CAPES/Agrégation langues.
- Tardieu, C. 2006. « The Dutch Grid, a practical tool for evaluating CEFR levels », (3-4), *New Standpoints*, n°28, May, June.
- Tardieu, C. 2008. *La didactique des langues en quatre mots clés*. Paris : Ellipses.
- Thorndike, E. L. 1920. A constant error on psychological ratings. *Journal of Applied Psychology*, 4 : 25-29.

Tyler, R. 1950. *Basic principles of curriculum and instruction*. Chicago : University Press.

Vergnaud, G. 2000. *Lev Vygotski, Pédagogue et penseur de notre temps*. Paris : Hachette éducation.

Enquêtes, rapports, directives

Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer. 2001. Conseil de l'Europe, 1^e éd. 1996, 2^e éd. corr. 1998. Paris : Didier.

Dutch CEFR Grid. 2004. <www.ling.lancs.ac.uk/cefgrid>.

EBAFLS. 2007. <<http://www.cito.nl>>.

ESOL Examinations, University of Cambridge. 2004. *Samples of oral production, illustrating for English, the levels of the Common European Frame of Reference*. Strasbourg : Council of Europe.

Groupes de compétences, circulaire n°2006-093 du 31-5-2006. B.O.E.N. n°23 du 8 juin 2006.

Inspection générale. 2007. *Rapport sur l'évaluation Etats des lieux et perspectives d'évolution*. Janvier. N°007-009.

« Les jeunes face à leur avenir » 2008. Résultat d'une enquête internationale menée par la Fondation pour l'innovation politique, en partenariat avec l'Institut Kairos Future. *Magazine L'EXPRESS*. 3 janvier.

O.C.D.E. 2003. *Rapport PISA*.

<http://www.oecd.org/document/31/0,3343,fr_2649_201185_39720927_1_1_1_1,00.html>.

Plan de rénovation des langues vivantes. B.O.E.N. n°21 du 26 mai 2005. <http://eduscol.education.fr/D0067/accueil.htm>

Programmes de langues, palier 1, collège. B.O.E.N. Hors série n°6 d 25 août 2005. Ministère de l'éducation Nationale.

Programmes de langues pour la classe de Seconde, B.O.E.N. Hors série, n°7 du 3 octobre 2002. Ministère de l'éducation Nationale.

Accompagnement des programmes d'anglais pour la classe de Seconde. 2003. Paris : CNDP.