
Réflexion autour d'un exemple d'évaluation de la production orale en Langue 2 par les tâches

Muriel Grosbois



Édition électronique

URL : <http://journals.openedition.org/rdlc/2141>

DOI : 10.4000/rdlc.2141

ISSN : 1958-5772

Éditeur

ACEDLE

Référence électronique

Muriel Grosbois, « Réflexion autour d'un exemple d'évaluation de la production orale en Langue 2 par les tâches », *Recherches en didactique des langues et des cultures* [En ligne], 7-2 | 2010, mis en ligne le 01 octobre 2010, consulté le 19 avril 2019. URL : <http://journals.openedition.org/rdlc/2141> ; DOI : 10.4000/rdlc.2141

Ce document a été généré automatiquement le 19 avril 2019.



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

Réflexion autour d'un exemple d'évaluation de la production orale en Langue 2 par les tâches

Muriel Grosbois

1. Introduction

- 1 Les employeurs ont désormais tendance à exiger une preuve de compétences en langues. Cela s'applique également aux futurs enseignants qui, depuis le décret ministériel du 28 mai 2010, doivent justifier d'un certificat de compétences en langues de l'enseignement supérieur, pré-requis pour être titularisé¹. C'est dans cette optique que nous nous concentrerons ici sur l'évaluation d'une compétence clé pour les professeurs des écoles, celle de la production orale en Langue 2 (L2).
- 2 Nous présenterons plus particulièrement l'évaluation que nous avons mise en place pour mesurer l'évolution de cette compétence à l'issue d'une formation dont nous avons fait l'hypothèse qu'elle servirait le développement de la production orale en L2 des apprenants.
- 3 Cet article sera aussi l'occasion de nous interroger sur les conditions de la généralisation de l'évaluation interne proposée. En d'autres termes : quelles pistes peut-on dégager d'une évaluation locale, avec les limites qu'elle comporte ? Notre réflexion nous conduira ensuite à considérer comment les résultats de l'analyse à laquelle nous nous sommes livrée peuvent s'inscrire dans le nouveau contexte de formation.

2. Formation et évaluation : une approche par tâches

- 4 Nous commencerons par articuler formation et évaluation, en replaçant la présente recherche dans son contexte, à la fois théorique et institutionnel.

2.1. Contexte de la recherche

- 5 Nous avons proposé à de futurs professeurs des écoles de l'université Paris Sorbonne - IUFM, dans le cadre de leur cours d'anglais L2, de participer à la création d'une ressource numérique en collaborant à distance avec des partenaires dont ils étudient la langue (Grosbois, 2009). Des définitions de la tâche que donnent Breen (1987 : 23), Ellis (2003 : 16) et Nunan (2004 : 4), nous retiendrons d'abord que la tâche est dotée d'un objectif et repose sur un ensemble d'actions réalistes qui font appel à l'usage de la langue sans pour autant que le résultat soit nécessairement d'ordre linguistique.
- 6 Ellis (2003 : 16) distingue également entre les tâches qui attirent l'attention sur le sens ("*unfocused tasks*") et les tâches qui portent davantage sur la forme ("*focused tasks*"). Guichon résume la distinction en ces termes : les micro-tâches "*attirent intentionnellement l'attention des apprenants sur une forme spécifique*" alors que les macro-tâches "*sont inclusives et confrontent l'apprenant incidemment à des aspects particuliers de la langue en même temps qu'il est en train de résoudre un problème plus général*" (Guichon, 2004 : 138).
- 7 Si l'on prend appui sur ces définitions, le formateur se doit donc de proposer des macro-tâches qui s'apparentent le plus possible à des pratiques sociales, pouvant servir d'objectifs aux apprenants, et susceptibles de déclencher des interactions entre les participants. Dans un second temps, il importe de relever les problèmes langagiers qu'elles engendrent. Il est alors possible de mettre en place une structuration de l'apprentissage en proposant des micro-tâches d'accompagnement qui prolongent l'apport des tâches réalistes en permettant un travail plus spécifique sur la langue, en fonction des besoins observés. Macro-tâches et micro-tâches contribuent ainsi conjointement à l'apprentissage de la L2.
- 8 La formation qui consiste, pour les futurs enseignants, à élaborer collectivement une ressource numérique destinée à des élèves de l'école primaire, relève de l'apprentissage par la tâche. En effet, le scénario de formation mis en place repose sur le projet de création multimédia avec des partenaires anglophones, macro-tâche dont découle l'ensemble des autres tâches. Ces dernières consistent essentiellement à :
 - aller individuellement à la recherche d'informations auprès des partenaires du *King's College* de Londres au moyen du courriel ;
 - comprendre les messages écrits envoyés par les partenaires anglophones ;
 - échanger oralement avec les membres du groupe-classe les informations recueillies individuellement ;
 - opérer collectivement des choix en vue de l'élaboration du contenu de la ressource multimédia.
- 9 Ces tâches ont été complétées par des micro-tâches ciblées sur les besoins langagiers spécifiques.
- 10 Pour valider ou invalider l'hypothèse de recherche selon laquelle un tel scénario de formation était susceptible de contribuer au développement de la production orale en L2 des participants, il était nécessaire d'évaluer leur production orale en amont et en aval de la formation.
- 11 Sachant que l'évaluation se doit d'être en adéquation avec les principes mêmes de l'apprentissage, proposer une formation en L2 qui repose sur des tâches situées socialement et professionnellement (Collins, Brown & Newman, 1989) exécutées en

communiquant et en négociant du sens dans l'interaction, implique qu'il est pertinent d'évaluer, d'une part, sur la base de réalisation de tâches, et, d'autre part, sur un axe linguistique et pragmatique. Ce positionnement prend appui sur Bachman (1990) pour qui l'utilisation de la langue, révélée dans les interactions, ne se limite pas au code linguistique mais intègre la dimension pragmatique, composant essentiel de la compétence. Springer (2004) résume en ces termes cette approche holistique :

Bachman et Palmer (1990, 1996) proposent un modèle dynamique en trois parties. La partie langagière comporte un élément organisationnel, qui renvoie à la composante linguistique et à la composante discursive, et un élément pragmatique, qui renvoie globalement à la composante sociolinguistique. La composante stratégique, qui est en quelque sorte séparée du noyau langagier et sert d'interface avec les réalités de mise en œuvre en situation réelle, représente le troisième élément. [...] Il ne s'agit donc plus, dans ce modèle, de chercher à évaluer séparément chaque composante, mais de concevoir la compétence à communiquer langagièrement de manière globale (Springer, 2004 : 160).

- 12 L'évaluation de la compétence s'inscrit ainsi dans une optique interactionniste, reprise par le Cadre Européen du Conseil de l'Europe (2001), et à propos de laquelle Springer (2004) explique :

Le Cadre Européen du Conseil de l'Europe propose une approche légèrement différente tout en s'inscrivant dans cette option actionnelle de la communication. [...] Le modèle du Conseil de l'Europe accorde à la composante sociolinguistique un rôle clé, dans la mesure où on souhaite privilégier la dimension culturelle et situationnelle. C'est donc elle qui articule la composante linguistique et la composante pragmatique. On peut simplifier ces deux modèles en disant que le premier prend une entrée plus franchement interactionniste, la langue est une activité cognitive essentiellement située (on s'intéresse aux stratégies de communication et aux conditions de l'interaction) alors que le deuxième modèle valorise une entrée culturelle (la langue comme produit d'une culture) tout en s'inscrivant dans la même approche interactionniste (Springer, 2004 : 161).

- 13 C'est le premier modèle auquel Springer se réfère que nous avons retenu, afin qu'il y ait adéquation entre la formation proposée et les principes d'évaluation. Dans le scénario de formation tel qu'il a été conçu, le développement de l'interlangue se construit en effet dans l'interaction dialogique puisque la macro-tâche multimédia engendre des activités qui supposent une intention commune et un effort mutuel et coordonné de résolution de problème, qui déclenchent des interactions verbales et une négociation de sens, susceptibles de contribuer à l'acquisition de la L2 des participants.
- 14 Quel outil utiliser en adéquation avec ce cadre théorique ? Le DLC (Diplôme de Compétence en Langues), qui repose sur un scénario simulant un jeu de rôle professionnel, s'inscrit dans cette perspective interactionnelle et acquisitionnelle de l'évaluation. Il en va de même pour le Cles (Certificat de Compétences en Langues de l'Enseignement Supérieur), comme l'exprime Springer (2001, 2003).
- 15 Évaluer selon ces principes suppose donc, soit de retenir un scénario parmi ceux déjà proposés par le DCL ou le Cles, soit d'en concevoir un, et de prévoir les procédures d'implémentation ainsi que la mesure de la performance.
- 16 Cela nécessite aussi de considérer les contraintes et exigences auxquelles on ne peut se soustraire dans un contexte institutionnel donné.

2.2 Contraintes et exigences

- 17 Parmi les contraintes institutionnelles, nous retiendrons d'abord la difficulté, sur une formation courte (32 heures), de soumettre les apprenants à un scénario complet du DCL ou du Cles. De plus, les thèmes abordés dans les *scenarii* existants se sont avérés trop éloignés de l'optique professionnelle des apprenants pour être retenus.
- 18 Nous avons donc œuvré dans l'esprit du DCL et du Cles, mais élaboré des tâches évaluatives en nous référant à la littérature du domaine, dont Chalhoub-Delville (2001) qui, pour la conception de tâches visant à évaluer la production orale, explique que la tâche doit encourager l'expression individuelle, être contextualisée, offrir des situations qui ont du sens pour l'apprenant, ce qui est en accord avec les principes de la formation telle que nous l'avons conçue et le cadre théorique de référence auquel nous nous sommes adossée.
- 19 Ce choix corrobore la proposition de Bachman et Palmer (1996) de s'appuyer sur la notion de "*tâche communicative*", c'est-à-dire une activité dans laquelle des locuteurs interagissent avec un but particulier dans un contexte spécifique.
- 20 Ajoutons que, à la difficulté de concevoir une tâche évaluative selon les principes que nous venons d'énoncer, s'ajoute, dans notre cas, la nécessité de concevoir deux tâches évaluatives équivalentes, afin de déceler une éventuelle évolution de la production orale des apprenants entre le début et la fin de la formation. Rappelons que procéder ainsi à un pré-test et un post-test visait à nous permettre d'obtenir un retour sur l'efficacité du scénario de formation proposé.

3. Tâches évaluatives, outils et critères d'évaluation

- 21 Comment avons-nous conçu les tâches évaluatives et quels outils et critères d'évaluation avons-nous choisis ?

3.1. Élaboration des tâches évaluatives : étapes méthodologiques

- 22 Nous nous sommes appuyée sur Ellis (2003 : 303-304) qui définit ainsi les étapes méthodologiques de l'élaboration d'une tâche évaluative.
- Définir l'objectif de l'évaluation.
 - Recenser des ressources potentielles et prendre en compte les contraintes éventuelles.
 - Recenser les besoins en vue de la conception des tâches langagières.
 - Établir un cahier des charges incluant l'élaboration du contenu des tâches évaluatives (en fonction des conclusions des étapes précédentes), les procédures d'implémentation ainsi que la méthode et les critères d'évaluation.
 - Initier les formateurs externes.
 - Tester les tâches.
 - Apporter les modifications suite aux problèmes éventuellement rencontrés lors du test.
 - Proposer la nouvelle version de la tâche évaluative aux formateurs externes.

Ayant observé ces principes, nous avons élaboré deux formats de tâches.

3.2. Formats de tâches

- 23 Conformément au cadre théorique retenu, nous devons élaborer des tâches qui permettent de vérifier la capacité des apprenants à interagir oralement en L2. Elles devaient aussi permettre une évaluation sur un axe à la fois linguistique et pragmatique, en effectuant des tâches situées invitant à communiquer et négocier du sens dans l'interaction.
- 24 D'où l'idée de procéder à une interview autour de thèmes que nous avons préalablement définis, liés au domaine professionnel des apprenants. Et plutôt que de mener nous-même l'entretien, nous avons proposé à ces derniers d'opérer par dyade, réduisant ainsi le caractère artificiel et anxiogène de la situation d'évaluation, comme l'explique Ellis (2003 : 294-295) en se référant aux travaux de Fulcher (1996) et Berkoff (1985).
- 25 Ainsi le premier type de tâche revêt-il la forme d'une interview menée par les apprenants eux-mêmes. Le support pour ce type de tâche, proposé au début de la formation, figure en annexe 1. Nous le résumons ainsi.
- 26 Les apprenants ne se connaissent pas en début d'année, or ils vont être amenés à travailler ensemble dans le cadre de la création multimédia collective sur laquelle repose la formation. La mise en situation consiste à les inviter à se poser mutuellement des questions pour savoir s'ils peuvent faire équipe dans le cadre du projet commun.
- 27 Ce type de tâche a l'avantage de présenter un caractère d'authenticité, en revanche, il est empreint de variabilité car la production orale est fonction des questions formulées par le pair (qui varient d'un individu à l'autre). En d'autres termes, si les interactions que ce premier type de tâche suscite sont révélatrices de la compétence de l'apprenant puisque le contexte favorise l'authenticité de la parole, elles dépendent aussi des répliques de l'interlocuteur, ce qui introduit de la variabilité dont il importe de tenir compte.
- 28 Il nous a donc paru nécessaire d'ajouter un second type de tâche qui consiste, lui, à rendre compte oralement d'un enregistrement oral. Moins authentique que le premier car ne reposant pas sur des interactions suscitées par une situation dialogique, il présente l'avantage de reposer sur un contenu de départ identique pour tous. Le support pour ce type de tâche, proposé au début de la formation, figure en annexe 2. Il s'agit d'un sujet d'épreuve orale proposé en 2003 dans les académies de Paris et Versailles pour le concours de recrutement des professeurs des écoles. Il traite d'un thème en rapport avec leur domaine de professionnalisation : l'impact de l'ordinateur sur les enfants, notamment en classe (ce qui constitue aussi une entrée en matière pour le projet multimédia).
- 29 La combinaison de ces deux formats de tâches devait nous permettre d'avoir une vision fine des progrès individuels en termes de production orale.
- 30 Précisons aussi que, comme nous souhaitons évaluer les apprenants en amont et en aval de la formation pour essayer de déceler l'évolution de leur production orale, nous nous sommes efforcée, pour les deux types de tâches que nous venons de présenter, de proposer des supports de nature et de difficulté comparables au début et à la fin de la formation. En effet, bien que ne nous situant pas dans un protocole de quasi-expérimentation, nous devons veiller à prendre ces précautions pour pouvoir mettre en regard les résultats obtenus. Aussi avons-nous conçu les supports en parallèle dès le départ. Ceux utilisés pour l'évaluation en fin de formation figurent en annexe 3 et 4.

L'enregistrement correspondant à l'annexe 4 est un des supports utilisés pour l'épreuve orale de 2002 pour le concours de recrutement des professeurs des écoles dans l'académie de Paris. Les deux enregistrements que nous avons sélectionnés (annexes 2 et 4) présentent les caractéristiques communes suivantes.

- 31 Même durée d'enregistrement : 1min 45s.
- 32 Même nature : une interview. Deux personnes interagissent sous forme de dialogue, avec un accent britannique standard. Les informations sont d'ordre factuel et ne présentent pas de difficultés particulières. La vitesse d'élocution est similaire dans les deux cas.
- 33 Les thèmes diffèrent mais tous deux sont en rapport avec le domaine de l'enseignement.
- 34 Examinons maintenant la grille d'évaluation de la production orale que nous avons utilisée.

3.3. Critères d'évaluation

- 35 Nous avons retenu la grille du DCL (spécifique à l'anglais) car elle offre un cadrage plus détaillé que celle du Cles, ce qui convient à notre volonté de procéder à une évaluation la plus fine possible, à une sorte de "radiographie" de la production des apprenants, tout en restant de l'ordre de la macro-analyse.
- 36 Nous avons adapté la grille du DCL à nos besoins en regroupant les critères de la façon suivante.

Pour la compétence linguistique :

- le groupe verbal (des énoncés déclaratifs et interrogatifs) ;
- la syntaxe des énoncés déclaratifs ;
- le questionnement ;
- le groupe nominal ;
- le lexique ;
- la phonologie.

- 37 À noter que nous avons limité le critère "syntaxe" aux énoncés déclaratifs, car la syntaxe des énoncés interrogatifs s'avère être intégrée au critère "questionnement".

Pour la compétence pragmatique :

- la quantité ;
- l'efficacité ;
- l'interaction ;
- le repérage d'informations.

- 38 À chacun de ces critères correspondent des zones de niveau de compétence (qui vont de 1 à 5 pour les critères linguistiques et de 0 à 5 pour les critères pragmatiques) et qui permettent de situer l'interlangue à un moment donné sur la base du degré d'opérationnalité des apprenants. (Notre grille d'évaluation figure en annexe 5).
- 39 L'interaction communicative déclenchée par les tâches ainsi que les critères linguistiques et pragmatiques auxquels correspondent des niveaux de compétences devaient nous permettre d'évaluer l'état de l'interlangue des apprenants de manière positive, par rapport à des degrés d'opérationnalité, conformément aux principes d'acquisition et en nous plaçant dans une approche globale des compétences d'utilisation de la langue.
- 40 Nous avons ensuite passé les productions orales des apprenants au crible de la grille que nous avons établie.

- 41 Le recueil et l'exploitation des données a rendu possible la mesure de l'évolution de la production orale de manière individuelle (annexe 6) et collective (annexe 7).

4. Résultats et analyse

- 42 Nous commencerons par analyser l'adéquation entre l'évaluation et les besoins, avant de considérer les questions de validité, fiabilité, faisabilité, inhérentes à toute réflexion sur l'évaluation, comme le rappelle Roots-Buck (2005 : 120, 130).

4.1. Adéquation entre l'évaluation et les besoins

- 43 Nous avons d'abord constaté que l'évaluation mise en place a permis un retour sur l'hypothèse de recherche.
- 44 En effet, pour chaque apprenant, nous avons pu établir les valeurs de progression, de régression ou de stabilité en comptabilisant, pour les 10 critères (linguistiques et pragmatiques) de la grille d'évaluation, le numéro de la zone que nous avons sélectionnée, en amont et en aval de la formation.
- 45 Nous avons ensuite récapitulé l'ensemble des données de tous les apprenants pour aller au-delà d'une vue détaillée et individuelle et accéder à une vue globale.
- 46 Cette évaluation nous a permis de valider notre hypothèse de recherche. En effet, nous avons pu conclure que le scénario de formation proposé avait opéré comme levier d'apprentissage de l'anglais et contribué au développement de la production orale en L2 (annexe 7), même si la stabilité l'a emporté sur les progrès. Des modules de 32 heures sont sans doute trop courts pour qu'une progression plus significative de l'interlangue puisse se produire, comme le soulignent les chercheurs du domaine, dont Chapelle (2003 : 119-120).
- 47 Nous avons aussi été en mesure de mettre les résultats obtenus en regard des tâches proposées pour ensuite procéder à une analyse fine de la conception du scénario de formation, afin de pouvoir l'ajuster. En effet, après avoir classé les critères de la grille d'évaluation en fonction de l'évolution constatée pour le groupe d'apprenants, nous avons obtenu les résultats présentés en annexe 8. Ces résultats s'avèrent être en adéquation avec les tâches proposées, qui consistaient essentiellement à aller à la recherche d'informations en questionnant les partenaires anglophones, puis à comprendre les informations obtenues pour en rendre compte et en débattre en groupe-classe, en vue de l'objectif commun de création de la ressource multimédia. Les apprenants ont ainsi probablement eu tendance à accorder une priorité au sens plutôt qu'à la forme, leurs difficultés ne leur permettant guère de gérer les deux simultanément. D'où les faibles progrès pour les critères "syntaxe des énoncés déclaratifs", "groupe verbal des énoncés déclaratifs et interrogatifs" et "groupe nominal". On note aussi que la différence de progression entre les critères pragmatiques et les critères linguistiques est somme toute peu marquée, ce qui semble rejoindre les propos de Springer pour qui "*l'aspect pragmatique ne peut augmenter qu'avec la confiance linguistique*" (Springer, 2003 : 10).
- 48 Côté apprenants, on remarquera qu'ils se sont déclarés motivés par la situation d'évaluation dans laquelle ils opéraient en dyade, c'est-à-dire lorsque nous leur proposons une "*évaluation solidaire*" plutôt que "*solitaire*", pour reprendre l'expression de Springer² (2009). Ils ont également montré un grand intérêt pour les grilles d'évaluation

et l'évolution qu'elles permettaient de déceler. Nous pouvons donc penser qu'elles pourraient constituer des supports à l'apprentissage, mais à condition qu'elles soient facilement interprétables par les apprenants et qu'elles soient fiables. Ce qui nous invite à nous intéresser maintenant aux limites de l'évaluation locale proposée.

4.2. Limites de l'évaluation locale

- 49 Passons maintenant notre évaluation au crible des critères de validité, fiabilité et faisabilité.
- 50 Commençons par le critère de validité, que nous examinerons sous l'angle de la "validité théorique", la "validité de contenu" et la "validité faciale".
- 51 Concernant la validité théorique, que plusieurs auteurs dont Bachman (1990 : 21) considèrent comme essentielle, nous soulignerons que les tâches évaluatives que nous avons élaborées sont en adéquation avec la théorie de l'apprentissage qui sous-tend le scénario de formation proposé et nous adopterons la conclusion de Roots-Buck :
- On peut constater la validité théorique de cette tâche car :
- elle fait appel à des stratégies compensatoires qui révèlent l'état de développement de l'interlangue ;
 - elle place le sens au centre de l'activité et non pas la langue en tant que système ;
 - son exécution indique la performance dans le monde réel (Roots-Buck, 2005 : 181).
- 52 À propos de la validité de contenu, nous nous associerons également à Roots-Buck en avançant que le contenu des tâches évaluatives proposées est "*en grande partie, en lien direct avec le domaine des apprenants concernés*" (Roots-Buck : 2005, 123).
- 53 Pour ce qui relève de la validité faciale c'est-à-dire la pertinence de l'évaluation pour les différents acteurs impliqués tels qu'apprenants, institutions, employeurs... comme l'explique Roots-Buck (op.cit.), nous soulignerons la corrélation entre l'évaluation et les besoins des professeurs des écoles en regard de leur emploi futur.
- 54 Que peut-on maintenant dire de la fiabilité de notre évaluation ?
- 55 Certes, lors d'une interview, le rôle joué par l'interlocuteur peut nuire à la performance de l'autre. Mais, comme nous l'avons déjà signalé, nous avons veillé à compenser cet aspect par une épreuve supplémentaire d'un format différent.
- 56 Nous n'avons pas non plus fait passer deux épreuves de même niveau à un même candidat pour voir s'il obtenait le même résultat, ni comparé les performances de plusieurs candidats de même niveau sur une même épreuve. Mais cela était inutile dans la mesure où nous proposons une évaluation où plusieurs résultats étaient possibles pour le même candidat (puisque'il y avait interaction orale et donc une grande divergence de réponses).
- 57 Enfin, il faut reconnaître qu'un biais a pu être introduit par le fait que le formateur était aussi l'évaluateur. Mais là encore, des précautions ont été prises puisque, d'une part, l'évaluateur ne jouait aucun rôle dans l'interaction orale, et d'autre part, les productions orales ont été enregistrées et transcrites afin de pouvoir sélectionner avec précision les zones relatives aux différents critères des grilles d'évaluation.
- 58 Cette dernière précision nous conduit à considérer maintenant le critère de faisabilité, pour les apprenants et pour l'institution.
- 59 Côté apprenants, il s'avère que les consignes et le contenu étaient à leur portée, la tâche ayant au préalable fait l'objet d'un test sur un échantillon d'apprenants autres que ceux concernés par le scénario de formation, à l'instar de ce que préconise Ellis (2003 :

303-304). Ce test nous avait d'ailleurs conduit à procéder à quelques ajustements sur la formulation des consignes (qu'une simplification du lexique employé a permis de clarifier).

- 60 Côté institutionnel, il est important de considérer le fait que l'évaluation a eu lieu moitié sur le temps libre des apprenants, moitié sur les séances de formation (nous avons fixé la durée à 30 minutes maximum par apprenant, pour les deux types de tâches). Il va sans dire que cela se traduit par un coût en temps pour les formateurs, et donc un coût financier pour l'institution. En effet, pour évaluer les 16 apprenants, trois formateurs (préalablement initiés aux principes de l'évaluation afin d'assurer une cohérence d'ensemble) ont été mobilisés afin de respecter le créneau horaire imparti.
- 61 Cette réflexion nous permet de constater que si nous avons gagné en validité par rapport à ce qui avait été jusqu'alors proposé en interne, nous avons perdu en faisabilité.
- 62 À partir de notre analyse et des réserves qui en découlent, nous devons donc nous interroger sur la manière de procéder à une évaluation tout aussi rigoureuse mais faisable à plus grande échelle, en nous plaçant dans l'esprit de la certification qui est désormais exigée.

4.3. Propositions

- 63 Une des pistes serait peut-être à trouver dans la combinaison des propositions suivantes.
- Garder l'idée d'une évaluation signifiante qui intègre la dimension de réalisation de tâches en adéquation avec la formation elle-même.
 - Penser la formation sur le mode d'un dispositif hybride dans lequel le présentiel est en grande partie réservé aux tâches d'expression orale avec retour sur les difficultés rencontrées, la partie "à distance" pouvant être consacrée aux échanges écrits, aux réalisations concrètes, ainsi qu'aux micro-tâches spécifiques en fonction des besoins de chacun.
 - Trouver pour les grilles d'évaluation un compromis entre le détail proposé par les grilles du DCL et l'aspect binaire (validé/non validé, et donc réducteur) des supports d'évaluation de l'oral du Cles, c'est-à-dire adapter le degré de granularité en fonction des besoins (et donc en fonction d'un référentiel pour le métier de professeur des écoles) et ajuster les descripteurs.
 - Procéder à au moins une partie de l'évaluation de l'interaction orale pendant les séances en présentiel et impliquer les étudiants dans ce processus (en leur donnant les grilles d'évaluation) afin qu'ils développent un recul réflexif sur leurs besoins et les moyens de les atteindre.
 - Effectuer une autre partie de l'évaluation de la production orale à partir de fichiers sons enregistrés et mis à disposition des formateurs sur une plate-forme.
- 64 L'avantage serait de gagner en faisabilité, sans perdre en validité, mais aussi d'entraîner les étudiants, pendant leur formation, au format des épreuves de certification qui les attendent et aux exigences en matière d'emploi.

5. Conclusion

- 65 Ce travail d'évaluation interne, même avec ses limites, a permis, d'une part de contribuer à combler la pénurie d'exemples d'évaluation par les tâches soulignée dès 2001 par Chalhoub-Delville (2001 : 211) et qui vaut encore à l'heure actuelle, et, d'autre part, de

faire évoluer notre réflexion sur l'évaluation de la compétence de production orale par les tâches contextualisées : d'une évaluation locale se dégagent en effet des pistes pour une évaluation réaliste, rigoureuse et faisable à plus grande échelle.

- 66 Par ailleurs, sachant qu'il existe une influence des pratiques d'enseignement sur le mode d'évaluation et inversement, surtout dans la mesure où une certification est désormais un pré-requis pour devenir professeur des écoles titulaire, il importe de penser la formation et l'évaluation de cette dernière dans cet esprit.
- 67 Certes la formation et l'évaluation de la compétence de production orale en L2 sont coûteuses et par conséquent souvent écartées, alors qu'il s'agit là d'une compétence clé. Il nous semble néanmoins nécessaire d'accorder à cette compétence la place qu'elle mérite à la fois dans la formation et dans l'évaluation.

BIBLIOGRAPHIE

- Bachman, L.F. (1990). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Bachman, L.F. & Palmer, A.S. (1996). *Language Testing in Practice*. Oxford: Oxford University Press.
- Berkoff, N. (1985). "Testing oral proficiency: A new approach". In Lee, Y. (dir.) *New Directions in Language Testing*. Oxford: Pergamon. pp. 93-100.
- Breen, M. (1987). "Learner contributions to task design". In Candlin, C. & Murphy, D. (dir.) *Language Learning Tasks*. Englewood Cliffs, NJ: Prentice-Hall. pp. 23-46.
- Chalhoub-Delville, M. (2001). "Task-based assessments: Characteristics and validity evidence". In Bygate, M., Skehan, P. & Swain, M. (dir.) *Researching Pedagogic Tasks, Second Language Learning, Teaching and Testing*. Harlow: Longman. pp. 210-228.
- Chapelle, C. (2003). *English Language Learning and Technology*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Collins, A., Brown, J. S. & Newman, S. E. (1989). "Cognitive Apprenticeship: Teaching the Crafts of Reading, Writing, and Mathematics". In Resnick, L. B. (dir.) *Knowing, learning, and instruction*. Hillsdale, NJ: Lawrence Erlbaum. pp. 453-494.
- Conseil de l'Europe. (2001). *Cadre européen commun de référence pour les langues*. Paris : Didier.
- Ellis, R. (2003). *Task-based Language Learning and Teaching*. Oxford: Oxford University Press.
- Fulcher, G. (1996). "Testing tasks: Issues in task design and the group oral." *Language Testing*, n° 13. pp. 23-51.
- Grosbois, M. (2009). "TIC, tâches et nativisation : impact sur la production orale en L2". *Alsic (Apprentissage des Langues et Systèmes d'Information et de Communication)*, vol. 12. pp. 19-39.
- Guichon, N. (2004). *Compréhension de l'anglais oral et TICE - Les conditions d'un apprentissage signifiant*. Thèse de doctorat sous la direction de Narcy-Combes, J.-P. Université de Nantes.
- Nunan, D. (2004). *Task-Based Language Teaching*. Cambridge: Cambridge University Press.

Roots-Buck, J. (2005). *Le scénario comme démarche d'enseignement et comme mode d'évaluation*. Thèse de doctorat sous la direction de Narcy-Combes, J.-P., université de Nantes.

Springer, C. (2001). "Diagnostic, bilan de compétences, certification : les nouveaux habits de l'évaluation". *Les langues modernes*, n° 2. pp. 48-60.

Springer, C. (2003). "Évaluation de la compétence et problématique de l'acquisition en L2 : Préliminaires pour une définition de profils prototypiques de compétence en L2". In "*Actes du XIe colloque international (1999) : Acquisition d'une langue étrangère : perspectives et recherches - Usages pragmatiques et acquisition des langues étrangères*". Saint-Chamas : M.L.M.S. éditeur.

Springer, C. (2004). *Plurilinguisme et Compétences : Décrire, entraîner et certifier*. Habilitation à diriger des recherches, université Marc Bloch, Strasbourg 2.

ANNEXES

Annexe 1

Tableau 1- Grille support de la tâche d'interview en amont de la formation

Getting to know each other
<p>During the year, you will have the opportunity to team up with your French classmates as well as with English partners. Together, you will create a multimedia resource to help French primary school children learn English.</p> <p>Today, you will be given twenty minutes to interview each other to find out how well you could pair on the project.</p> <p>Here are some topics you could discuss:</p> <ul style="list-style-type: none"> - type of studies/professional background; - favourite subjects: (music, drawing...); - level in English (strong points, weak points); - communication with native speakers of English; - work habits/availability; - work experience with children; - computer equipment and use; - views on teamwork versus individual work; - experience of working on a project. <p>You may ask as many questions as you like. Feel free to follow the order you want and to gather any other information that you think might be useful.</p> <p>One of you shall start the interview, obtain all the necessary information, and then be interviewed in turn. Or, if you prefer, you can alternate the questioning.</p> <p>You can rephrase your questions and ask your partner for extra explanations if necessary.</p>

Annexe 2

Tableau 2 – Enregistrement oral pour l'évaluation en amont de la formation

Children and computers
<p>Children click with computers earlier. We interviewed Tom Johnson, author of a report on children and new technologies.</p> <ul style="list-style-type: none"> - What's the main difference between adults and children regarding new technologies? - Well for example, when they acquire a new mobile phone, children start playing with it immediately, learning on the way. In contrast most adults will pour over the instructions even before turning the phone on. - How about computing? At what age is a child able to use a computer? - Amazingly enough, the age for computing is now two to four years. While adults struggle to relate a new technology to some experience they already know, children accept developments for what they are. - Are the children aware of the effects of computers on their lives? - Absolutely. They think they spend less time outdoors and they have friends who are addicted to computer games. Some even have bedrooms looking like mini media centres. - Daily life is constantly changing far faster than ever before. How far is school concerned? - At school too technology is having a huge impact. Wired classrooms could mean the end of the blackboard and chalk. In many schools introducing the Internet has increased the children's motivation. They interact more and discuss things more. - So it echoes government surveys that show a link between investment in computing and improved classroom performance.

Annexe 3

Tableau 3 - Grille support de la tâche d'interview en aval de la formation

Final interview

Our English class has come to an end. You will now be given twenty minutes to interview each other.

Here are some topics you could discuss:

- project presentation;
- introducing your English partner(s);
- e-mails received from your e-pals;
- placement experience;
- teaching English to primary school children in France;
- teaching in a primary school in England;
- use of computers for children to learn English;
- use of computers for adults to learn English;
- teamwork/partnership.

You may ask as many questions as you like. Feel free to follow the order you want.

One of you shall start the interview, obtain all the necessary information, and then be interviewed in turn. Or, if you prefer, you can alternate the questioning.

You can rephrase your questions and ask your partner for extra explanations if necessary.

Annexe 4

Tableau 4 – Enregistrement oral pour l'évaluation en aval de la formation

Volunteer work in Kenya

Our guest today is Tara Tweedy, eighteen, who works for the British voluntary youth service and is going to tell us about her experience as a volunteer in Africa last summer.

- Tara, what did your voluntary action consist of?

- I spent two weeks in Kenya as a voluntary teacher in a primary school with a group of ten other young people from my town.

- Had you ever taught before?

- Actually no, it was my first teaching experience. But I was quite used to dealing with kids and had already taken part in many place schemes at our local volunteer bureau.

- So you didn't find it too difficult to teach. English isn't those kids' native language.

- I must say it was a little difficult at first; but as we teachers felt terribly nervous... so we tended to speak much too fast; but the kids were so nice that we eventually relaxed and it went remarkably well.

- What did you actually do with them?

- We'd worked very hard to prepare lessons: comprehension, verbs, but we also focused on non-academic activities like portrait drawings and puppet making. We also organised mini Olympics and games.

- But surely you needed books, paper, pens and things like that?

- Yes, but we knew that the school couldn't afford what we needed. So the equipment was donated by schools in my town, Sutton.

- How did you pay for your trip?

- We spent months raising funds, and we managed to collect one thousand pounds each.

- What's your favourite memory of those two weeks?

- Oh I've got lots. I especially remember the tree planting ceremony that the school organised just before we left. I still can't believe that there are eleven trees in a schoolyard in Kenya planted in honour of us.

Annexe 5

Tableau 5 - Grille d'évaluation (sur la base du DCL)

[Données illisibles et inexploitables]

Annexe 6

Tableau 6 – Évolution de la production orale (niveau individuel). Exemple pour un apprenant.

Récapitulatif des données de début et de fin d'expérimentation		
Apprenant : A.		
	DÉBUT	FIN

	n° zone	n° zone
Groupe verbal (énoncés déclaratifs et interrogatifs)	2	2
Syntaxe (énoncés déclaratifs)	2	2
Questionnement	2	3
Groupe nominal	2	2
Lexique	2	3
Phonologie	3	3
Quantité	3	3
Efficacité	3	3
Interaction	3	3
Repérage d'informations	2	2
Évolution d'une zone à l'autre		
Progression	2	
Régression	0	
Stabilité	8	

Annexe 7

Tableau 7 - Évolution de la production orale (niveau global).

Évolution	Total	
Nombre de critères en progression	43	
Nombre de critères en régression	1	
Nombre de critères restés stables	116	
		Soit un total général de 160 critères (10 critères pour chaque apprenant).

Annexe 8

Tableau 8 – Classement des critères en fonction de l'évolution constatée.

Critères	Valeur de la progression
Lexique	+ 9
Repérage d'informations	+ 7
Questionnement	+ 6
Efficacité	+ 5
Interaction	+ 5
Syntaxe des énoncés déclaratifs	+ 3
Groupe verbal des énoncés déclaratifs et interrogatifs	+ 3
Phonologie	+ 3
Groupe nominal	+ 2
Quantité	+ 1

NOTES

1. Décret n° 2010-570 du 28 mai 2010 portant diverses dispositions statutaires applicables à certains personnels enseignants et d'éducation relevant du ministre de l'éducation nationale.
2. Expression employée lors d'une table ronde intitulée "Pourquoi et comment évaluer en langues ? ", au salon Expolangues, 2009.

RÉSUMÉS

Notre réflexion porte sur une évaluation par les tâches mise en place auprès d'un public de futurs enseignants, pour lesquels la compétence de production orale en Langue 2 s'avère être une compétence clé dans l'exercice de leur profession. Après avoir examiné l'élaboration de cette évaluation interne en la plaçant dans son contexte à la fois théorique et institutionnel, nous la soumettrons au crible des critères de validité, fiabilité et faisabilité. Nous nous interrogerons ensuite sur les conditions de sa généralisation en lien avec les nouvelles exigences en matière de certification en langues.

This study examines a task-based assessment set up for trainee teachers whose L2 speaking skills will be a key competency for their future profession. After explaining the process of elaboration of this assessment and placing it in its theoretical and institutional context, we analyze its

validity, reliability and feasibility. We then examine to what extent it could be generalized to comply with the recent requirements for language certification for teachers in France.

INDEX

Mots-clés : évaluation, tâches, production orale, formation des enseignants

Keywords : assessment, tasks, speaking, teacher training

AUTEUR

MURIEL GROSBOIS

Muriel Grosbois est maître de conférences à l'université Paris Sorbonne – IUFM.

Son laboratoire de rattachement est le DILTEC, groupe "Tâches et dispositifs", université Sorbonne nouvelle Paris 3.

Courriel : [muriel.grosbois\[at\]paris.iufm.fr](mailto:muriel.grosbois[at]paris.iufm.fr)

Toile : <http://muriel.grosbois.free.fr>

Adresse : 10, rue Molitor, 75016 Paris, France.