



**ASp**  
la revue du GERAS

**27-30 | 2000**  
**Réflexions sur le cognitif, Approche du discours scientifique, Ouvertures sociologiques, Études de genre, études de discours**

---

## L'évaluation de la compréhension et la problématique de la restitution

Shaeda Isani

---



### Édition électronique

URL : <http://journals.openedition.org/asp/2148>

DOI : 10.4000/asp.2148

ISBN : 978-2-8218-0382-4

ISSN : 2108-6354

### Éditeur

Groupe d'étude et de recherche en anglais de spécialité

### Édition imprimée

Date de publication : 31 décembre 2000

Pagination : 261-272

ISSN : 1246-8185

### Référence électronique

Shaeda Isani, « L'évaluation de la compréhension et la problématique de la restitution », *ASp* [En ligne], 27-30 | 2000, mis en ligne le 28 janvier 2011, consulté le 19 avril 2019. URL : <http://journals.openedition.org/asp/2148> ; DOI : 10.4000/asp.2148

---

Ce document a été généré automatiquement le 19 avril 2019.

Tous droits réservés

---

# L'évaluation de la compréhension et la problématique de la restitution

Shaeda Isani

---

## Le profil du public concerné : LANSAD universitaire pré-professionnel

- 1 Le public concerné par notre réflexion est celui des universitaires pré-professionnels LANSAD, terme qui a remplacé celui de « non-spécialiste », considéré à juste titre comme péjoratif et peu clair. Le public LANSAD universitaire pré-professionnel est un public vaste et hétérogène qui inclut les étudiants en 2<sup>e</sup> et 3<sup>e</sup> cycles des universités de langues et lettres (Langues étrangères appliquées), des Sciences pures et appliquées, des Sciences sociales, sans oublier la Formation continue.
- 2 La spécificité de ce public en ce qui nous concerne, enseignants et chercheurs en langues étrangères, réside dans la finalité de leur apprentissage en langues étrangères : un outil de communication qui permet de fonctionner correctement dans un contexte socioprofessionnel. Nous avons déjà eu l'occasion ailleurs (Isani 1996) de définir ce que nous entendons par domaine socioprofessionnel et nous nous bornerons par conséquent à le résumer brièvement ici en disant premièrement, que s'il ne s'agit pas de l'anglais de Shakespeare, il ne s'agit pas non plus de l'anglais de Ken Loach et, deuxièmement que, s'il ne s'agit pas de l'anglais réduit au jargon technique, il ne s'agit pas non plus de l'anglais du tourisme.
- 3 Eu égard à la finalité professionnelle de son apprentissage en langues étrangères, la question d'un système d'évaluation et de certification qui fournira, à la fois au futur employeur et au candidat, un « cliché » du profil langagier du candidat et de sa capacité à fonctionner dans ce milieu socioprofessionnel, revêt une importance capitale. Or, compte tenu du fait que la plupart des employeurs sont relativement peu au courant des spécificités et des nuances des différents types de certifications en langues étrangères existantes aujourd'hui<sup>1</sup>, il devient clair que tout système de certification à objectif

professionnel se doit d'être d'une grande lisibilité pour le futur employeur, sans besoin de décodage de « scores » ou de transposition de « compétences ». Cet objectif nous oriente vers le choix d'un système d'évaluation directe, où les tâches proposées reflètent les tâches réelles que le candidat sera appelé à faire.

- 4 Une deuxième raison qui nous pousse vers le choix d'un système d'évaluation directe est le phénomène de *washback* qui est l'effet que produit un système d'évaluation sur les pratiques pédagogiques et les objectifs d'apprentissage. Cet effet peut être négatif (lorsque le candidat passe son temps à maîtriser l'art du QCM) ou positif lorsqu'il se consacre à l'acquisition d'une compétence orale. Compte tenu de son pouvoir d'infléchir les pratiques pédagogiques, aucun enseignant ne peut y être insensible.

## Les critères docimologiques : un besoin de modestie

- 5 Il n'existe pas d'évaluation parfaite. De plus, se donner comme objectif de pratiquer une évaluation directe de la compétence de communication ne fait qu'accroître la complexité et la fragilité de l'épreuve. Tenter d'y incorporer les différentes sous-composantes du critère de validité (*face validity, content validity, construct validity, empirical validity*), en même temps que le critère de fiabilité – tout en sachant que la corrélation entre le critère de validité et le critère de fiabilité est nécessairement faible – nous conduit inéluctablement à construire une usine à gaz, ce qui met à mal le critère de facilité de mise en œuvre. Il est donc, préférable de rester modeste et de se contenter de ce qui est raisonnable et réalisable dans le contexte d'une évaluation de masse. À l'intérieur de ces contraintes, il est parfaitement envisageable d'élaborer un système d'évaluation qui garde toute sa crédibilité en proposant comme épreuves des tâches aussi proches de la vie réelle que possible, qui reste fiable en assurant le même résultat d'un correcteur à un autre, d'une séance à une autre, grâce à un référentiel et un barème et, finalement, qui permet une facilité de mise en œuvre.

### La compréhension écrite et la compréhension orale : similarités

- 6 Bien qu'il existe des différences importantes entre la compréhension écrite et la compréhension orale – caractère éphémère du discours oral et la nécessité de traiter l'information en quasi simultané et en temps réel, présentiel de l'interlocuteur, et redondance et relâchement du discours oral pour ne citer que les principales – et compte tenu du fait que de nombreuses considérations restent communes à l'évaluation des deux compétences, nous avons choisi, dans le cadre de cette étude, de traiter les deux compétences ensemble.

## Le statut de la compréhension dans un contexte de compétence communicative : le paradoxe

- 7 Ayant défini le public concerné, il s'agit de s'interroger sur la place de la compréhension écrite et orale par rapport à ses besoins, son apprentissage et, bien entendu, l'évaluation de ses compétences. Une première réflexion nous conduit à penser que dans le contexte d'un système d'évaluation holistique des compétences de communication, la compréhension, écrite ou orale, n'a pas sa place en tant que tâche langagière isolée. Dans la vie réelle, il est rare que la compréhension soit une activité en soi et pour soi. Elle conduit dans la plupart des cas, à des activités autres auxquelles elle sert de support.

L'évaluer comme compétence en soi risquerait par conséquent, de porter atteinte à l'authenticité de l'épreuve car, isolée, elle ne constitue pas une tâche langagière communicative dans l'acception générale du terme « communicatif ».

- 8 Si telle est notre hypothèse, il faudra admettre qu'il n'y a pas de place pour une épreuve de compréhension en soi dans le contexte d'une évaluation des compétences communicatives, cette épreuve devant être forcément rattachée à une activité de production dont elle sert de point de départ. C'est ici que nous nous heurtons à la complexité du profil du public LANSAD et à la grande diversité de ses besoins car une des caractéristiques de ce public *est précisément une demande forte pour une évaluation de la compréhension, notamment de l'écrit, en isolé.*
- 9 Qui sont les candidats qui cherchent à valider seulement la compréhension ? Dans le domaine du public pré-professionnel universitaire LANSAD, la demande de ce type de validation partielle est, en général, restreinte aux étudiants (et ils ne sont pas aussi rares que l'on pourrait croire) qui ont étudié l'anglais seulement en troisième langue ou encore qui ne se sont pas remis à l'anglais depuis le baccalauréat (facultés où les langues vivantes restent une option) et qui, se trouvant à la veille d'une embauche, souhaitent faire valider leurs compétences en anglais, aussi partielles soient-elles.
- 10 La question qui se pose est de savoir si un système d'évaluation destiné à valider les compétences communicatives du public pré-professionnel LANSAD, tel le CCLES<sup>2</sup>, doit oui ou non tenir compte de cette demande spécifique. Peut-on raisonnablement valider une compétence « non communicative » dans le contexte d'une évaluation des compétences de communication en langue étrangère ?
- 11 Dans le cadre de notre réflexion actuelle sur l'élaboration d'un système d'évaluation validant les compétences de communication des étudiants pré-professionnels LANSAD, nous avons décidé que, compte tenu du fait qu'il s'agit d'une demande forte et légitime, nous ne pouvons pas ne pas en tenir compte et qu'il nous incombe, par conséquent, d'élaborer un système d'évaluation qui permette d'évaluer à la fois les compétences partielles de la compréhension écrite et orale et la capacité à *communiquer* en langue étrangère. Heaton (1989 : 163) nous conforte dans notre choix :

Communicative language testing has recently introduced another dimension (...) namely, profile reporting. In order to obtain a full profile of a student's ability in the target language, it is necessary to assess his or her performance separately for each of the different areas of communication. [...] Furthermore, performance is assessed according to the purpose for which language is used. [...]. The object of sub-tests through which performance is assessed is to indicate the extent of the learner's mastery of the various language skills which he or she will acquire for a particular purpose.

#### **Quel type de texte et quel type d'image ? Pratiques et représentations**

- 12 Les notions de texte et d'image ont beaucoup évolué depuis l'arrivée des nouvelles technologies et l'utilisation de l'hypertexte et de l'image virtuelle – pour ne nommer que ceux-là – commence à se répandre. Il pourrait donc paraître quelque peu paradoxal – pour ne pas dire démodé – de limiter notre réflexion aux notions plus traditionnelles du texte et de l'image, c'est-à-dire, le texte comme **article extrait de la presse écrite non spécialisée** et l'image comme **séquence vidéo**. Et pourtant, malgré les avancées faites dans le domaine des nouvelles technologies et dans l'enseignement des langues, il reste indéniable que la grande majorité des enseignants du domaine LANSAD exerçant en

milieu universitaire aujourd'hui continuent de privilégier l'article de presse comme support écrit et la séquence vidéo comme support audiovisuel.

- 13 Dans le contexte du choix du texte et de l'image, il est intéressant d'ouvrir une parenthèse sur les représentations qui existent dans un milieu universitaire de langues et lettres concernant le type de texte à privilégier. Ici surgit un problème de **culture « ethnique » et professionnelle**. Lors des réunions pédagogiques toutes langues confondues, nous avons souvent eu l'occasion de procéder à un recensement du type de texte utilisé comme support pédagogique et/ou évaluatif. Ce recensement a permis de révéler la brèche qui sépare, d'une part, les anglicistes de leurs collègues hispanistes, germanistes et italianistes et, d'autre part, ceux d'une formation littéraire de ceux d'une formation en Sciences du langage, Didactique des langues étrangères et autres. Il est apparu clairement que certains types de texte, couramment exploités par des anglicistes et par des enseignants non littéraires, étaient perçus comme n'étant pas assez *dignes* de servir de support pédagogique. Ainsi, tout support « fonctionnel » (publicités, horaires, enregistrements du standard automatisé, etc.), couramment utilisé en Formation continue ou en deuxième langue ou encore dans les « universités LANSAD », est perçu comme « inappropriés » pour le public d'une université de langues et lettres où le texte écrit doit pouvoir faire valoir ses titres de noblesse et provenir de préférence d'une œuvre littéraire (« plus récente » éventuellement pour les LANSAD) ou, à la rigueur, d'un journal de qualité. La même divergence existe par rapport à la séquence vidéo : le film documentaire est perçu comme un outil pédagogique plus valable qu'un film « commercial ».
- 14 À ce sujet, il est à noter que si la séquence vidéo est assez largement exploitée dans le contexte pédagogique de la salle de classe, elle disparaît quasi totalement dans la situation d'évaluation pour être remplacée au mieux, par la version audio en laboratoire de langue, ou, ce qui est plus souvent le cas, par la lecture oralisée en amphithéâtre. Les raisons en sont essentiellement matérielles, car projeter une séquence vidéo dans un amphithéâtre de 250 étudiants en situation d'examen reste problématique dans la plupart des cas aujourd'hui.

## Le problème de la restitution : « *muddied measurement* »

- 15 Nombre de spécialistes de l'évaluation en langue soulignent le problème fondamental qui sous-tend l'évaluation de la compréhension, celui de la restitution ou de l'obtention d'un échantillonnage mesurable. En production orale et écrite, l'évaluateur dispose d'un échantillonnage visible et tangible de ce que le candidat est capable de faire. Il en est tout autrement en ce qui concerne la compréhension qui est une activité difficilement observable. Si l'évaluateur peut, à un niveau débutant, recourir à la restitution non verbale par le biais des techniques de transfert de l'information, il se voit contraint de recourir à la production langagière dès qu'il s'agit d'évaluer des candidats possédant un niveau plus avancé. Ceci entraîne un croisement de compétences où l'évaluateur doit passer par le biais des compétences productives pour porter une appréciation sur la capacité de compréhension du candidat. Sachant que les capacités de compréhension dépassent souvent de loin celles de la production, nous prenons vite la mesure de la

contradiction et de la confusion qui en résultent, et que Weir (1993 : 72) qualifie de « *muddied measurement* » ou l'évaluation brouillée.

- 16 Pour ceux qui prônent les activités langagières intégratives, le problème ne se pose pas car la compréhension est habituellement évaluée par le biais d'autres tâches. Tel est aussi le raisonnement qui sous-tend l'évaluation des compétences communicatives. Mais dans le cas de l'évaluation de la compréhension en tant que compétence isolée, il est impératif d'endiguer ce phénomène de contamination qui nuit à la validité de la tâche demandée et conduit à une mesure brouillée.

## Typologie des modes et des formats de restitution

- 17 La question qui se pose alors est de savoir si on peut obtenir une évaluation de la compréhension sans risque de contamination, tout en veillant au respect des trois critères docimologiques évoqués. Pour guider notre réflexion, une ébauche typologique des différents modes et formats de restitution de la compréhension peut s'avérer utile (voir tableau 1).

Tableau 1. Modes et formats de restitution de la compréhension

MODE	FORMAT
RESTITUTION NON VERBALE (transfert d'information)	<p>☒ exécution manuelle : notice, recette de cuisine</p> <p>☒ repérage : suivre un itinéraire, étiquetage</p>
RESTITUTION PASSIVE SANS PRODUCTION LANGAGIERE	<p>☒ QCM</p> <p>☒ exercice lacunaire (cloze) ouvert ou fermé...</p> <p>☒ vrai/faux</p>
RESTITUTION AVEC PRODUCTION LANGAGIÈRE	<p>☒ en LM</p> <p>☒ QROC (SAQ)</p> <p>☒ résumé</p>

### 1. Restitution non verbale : transfert d'information

- 18 Nous ne nous attarderons pas sur la restitution non verbale. Bien qu'elle constitue le seul mode de restitution que l'on pourrait considérer « pure » et non contaminée, elle est peu utilisée en milieu universitaire, en partie sans doute, pour les raisons évoquées précédemment concernant les représentations qui existent au sujet des supports pédagogiques.

### 2. Restitution passive sans production langagière

#### Le QCM

- 19 En ce qui concerne la restitution passive sans production langagière, le QCM est certainement le plus connu grâce aux examens du TOEFL et du TOIEC. Bien que le QCM soit le format d'évaluation le plus objectif (et donc, le plus fiable) et le plus facile à mettre en œuvre, nombreuses sont les voix qui s'élèvent contre les effets pervers que ce format d'évaluation engendre. Nous ne retiendrons ici que les trois principaux. Le premier, bien entendu concerne l'effet pervers que le QCM exerce sur le phénomène de **washback** déjà évoqué dont l'importance, faste et néfaste, n'est plus à démontrer. Nous nous accordons tous à dire qu'il s'opère un dérapage inquiétant lorsque la focalisation de l'apprentissage porte sur l'acquisition des trucs et des astuces du QCM au lieu de porter sur l'acquisition d'authentiques compétences langagières. Un autre inconvénient majeur réside dans le temps et l'effort qu'exige l'élaboration d'un bon QCM, même lorsqu'il s'agit d'un concepteur expérimenté. L'investissement requis n'est pas en proportion avec l'effet de *washback* escompté, et nombreux sont les enseignants qui estiment qu'il serait plus utile de consacrer ce temps et cet effort à la pédagogie plutôt qu'à un acte évaluatif sans grand retour. Et finalement, une analyse de plusieurs séries de QCM, même celles élaborées par des professionnels, révèle que trop souvent les questions se situent à un niveau de difficulté linguistique supérieur à celui du texte sur lequel l'épreuve de compréhension est censée porter. Nous aboutissons par conséquent, à une situation où un texte de niveau « indépendant » ou même « instructional » est suivi de questions qui se situent au niveau « frustrational ».<sup>3</sup> Prenons un exemple :

Figure 1. Exemple

***Gina Grant is the ideal Harvard prospect – and in fact, in December the university mailed her one of its precious few early admission letters.***

According to this sentence, what can be said about Harvard's admission letters?

- a) The precious ones are mailed early
- b) They are mailed early in December
- c) They are sent to the ideal prospects only
- d) Not a great many go out early

- 20 Les énoncés de la réponse non seulement noient le candidat, mais masquent également l'information essentielle de la phrase – celle que le candidat doit cibler – que pour l'Université de Harvard Gina Grant est une étudiante potentielle exceptionnelle.

#### **Exercice lacunaire**

- 21 Il existe de nombreuses variantes de l'exercice lacunaire dans les épreuves de compréhension : fermées, ouvertes, avec la première lettre indiquée, élimination ciblée de mots, élimination mécanique de mots, etc. Très utilisé dans le passé, l'exercice lacunaire semble être quelque peu délaissé aujourd'hui et l'explication se trouve très certainement dans l'évolution des pratiques pédagogiques vers l'approche communicative où l'exercice lacunaire est perçu comme étant trop éloigné de la tâche communicative authentique et de la compétence globale et comme une focalisation excessive sur les aspects microlinguistiques parcellisés de la langue (connaissances grammaticales et lexicales).
- 22 Malgré ces réticences, l'exercice lacunaire présente un certain nombre d'avantages, notamment une facilité d'élaboration et de mise en œuvre, facteurs non négligeables dans

un système d'évaluation de masse. Une question souvent soulevée par rapport à l'exercice lacunaire concerne sa fiabilité qui, bien que relativement élevée dans l'ensemble, peut présenter quelques faiblesses car il arrive fréquemment que le candidat fournisse une réponse qui pourrait convenir, mais qui n'est pas celle prévue par le(s) concepteur(s) qui conviendrait « mieux », en quelque sorte. Bien qu'un protocole d'évaluation et un barème de correction contribuent à circonscrire ce problème, la question surgit régulièrement notamment lorsqu'il s'agit d'une épreuve évaluée par plusieurs correcteurs et soulève le problème épineux de la fiabilité inter-correcteurs (inter-rater reliability).

### Vrai/faux

- 23 L'exercice vrai/faux est celui qui reçoit l'approbation d'un grand nombre de spécialistes dans le domaine de l'évaluation aujourd'hui car il présente à la fois un faible risque de contamination, un degré de fiabilité élevé et une grande facilité de mise en œuvre. Même l'objectif d'authenticité peut être considéré comme étant respecté dans la mesure où nous sommes souvent appelés à nous interroger sur notre compréhension du discours de l'autre. Dans le contexte d'une évaluation de la compréhension isolée, ce format présente un potentiel intéressant.
- 24 La principale faiblesse du format vrai/faux réside dans le fait qu'il permet au candidat de deviner ou de répondre au hasard, si l'on peut dire, avec 50 % de chances d'avoir une bonne réponse. Trois stratégies pour contrer cette faiblesse sont couramment mises en œuvre. La première consiste à noter sur un barème de 3 points pour chaque item : +3 pour une bonne réponse, -1 pour une mauvaise réponse et 0 pour aucune réponse. La deuxième consiste à demander au candidat de justifier son choix en indiquant le mot ou la phrase dans le texte qui valide sa réponse. Et finalement, une troisième stratégie consiste à encourager le candidat à nuancer sa réponse en élaborant une grille de réponse avec trois options (voir tableau 2).

Tableau 2. Grille de réponse

	vrai	faux	non explicité dans le texte
Question 1			
Question 2			
Question 3			
Question 4			
Question 5			

### 3. Restitution avec production langagière : le croisement des compétences et le phénomène de contamination

- 25 Dès que nous basculons du mode passif au mode actif dans la vérification de la compréhension écrite ou orale, nous nous trouvons confrontés au problème du croisement de compétences : la réception, généralement une compétence plus accessible, est évaluée par le biais de la production, une compétence considérée plus complexe, ou



encore, la compréhension orale est évaluée par le biais de la production écrite. Nous assistons alors à un décentrage de la focalisation qui, au lieu de porter exclusivement sur le fond, se déplace pour porter également sur la forme et nous obtenons un bel échantillonnage de « muddied measurement ». Weir (1993 : 74) nous met en garde :

If we want to test the construct of reading *per se*, then we must take steps to avoid the contamination of the measurement of this ability by the method adopted to assess it. We must try to test reading and only reading.

#### En LM : situation endolingue

- 26 Le recours à la langue maternelle du candidat est un mode d'évaluation de la compréhension qui existe encore (cf. l'examen de sortie de l'ENA pour la deuxième langue, épreuve dite de « connaissance passive »), mais qui semble avoir perdu du terrain dans les pratiques pédagogiques et évaluatives en milieu universitaire – la popularisation de la méthode directe ne doit pas être étrangère à ce phénomène. Le recours à la langue maternelle reste cependant un mode d'évaluation qui présente l'avantage de ne pas être contaminé par le recours aux compétences productives dans la langue étrangère. Il est à ce titre, utile dans l'évaluation de la compréhension des candidats dont la maîtrise de la langue étrangère est quasi inexistante au niveau de la production, mais partielle et vérifiable, donc validable, au niveau de la compréhension, notamment de l'écrit. Un autre avantage intéressant du recours à la langue maternelle est qu'il correspond à une réalité de la vie quotidienne dans la mesure où le passage d'une langue à l'autre est une activité constante et presque automatique.
- 27 L'inconvénient du recours à la langue maternelle comme mode de restitution en situation d'évaluation réside dans le fait qu'il ne peut s'appliquer qu'en situation endolingue et encore ! Si nous examinons le profil de notre public LANSAD universitaire, nous découvrons que bien qu'il s'agisse d'un enseignement-apprentissage endolingue, la notion du français comme langue maternelle est à nuancer. Sans aborder le cas complexe des candidats issus de l'immigration, nous évoquerons simplement le nombre toujours croissant d'étudiants étrangers LANSAD (européens, mais également asiatiques et africains) qui suivent des cours d'anglais dans un cadre universitaire. La question est de déterminer si ces étudiants sont pénalisés ou non par une évaluation de leur compréhension d'une langue étrangère (L3) à travers une restitution en une langue (L2) qui est, elle aussi, une langue étrangère pour eux.

#### Le QROC

- 28 Plusieurs variantes du QROC (Questionnaire à Réponse Ouverte Courte) ou SAQ (*Short Answer Question*) sont utilisées dans le contexte de l'évaluation de la compréhension écrite et orale, allant de la simple WH- question (*who, when, where, why, what and how, etc.*) aux réponses exigées sous forme de grille de réponse (tabular restitution). Pour simplifier, nous pouvons regrouper sous le terme générique de QROC **tout exercice qui exige du candidat de rédiger une réponse courte dans un espace limité clairement indiqué**. Le principe de base consiste à obtenir des réponses brèves et à réduire au minimum l'activité rédactionnelle.
- 29 Le QROC bénéficie d'une place privilégiée dans le contexte de l'évaluation des compétences communicatives à travers des tâches langagières authentiques car tout en respectant les trois principaux critères docimologiques (validité : il s'agit d'une activité authentique, fiabilité : il y a un faible potentiel d'écart entre correcteurs, et facilité de mise en œuvre), il permet aussi au concepteur de l'épreuve de calibrer la difficulté des questions et ainsi de cibler différents niveaux à partir d'un seul et même texte.

- 30 La question soulevée fréquemment dans l'exploitation de ce type de format de restitution concerne la politique à adopter concernant la qualité purement linguistique (grammaire, orthographe, ponctuation) de la réponse donnée. Bien que dans des situations d'évaluation autres que celles de la compréhension ces aspects formels de la réponse fournie par le candidat soient pris en considération dans la base de notation, il est clair que dans le contexte de l'évaluation de la compréhension *per se*, cette pratique constituerait une interférence et nuirait par conséquent, à la validité de l'épreuve.

#### **Le résumé écrit : le croisement des compétences par excellence**

- 31 Le résumé écrit et le résumé oral sont exploités tous les deux comme format de restitution pour des épreuves de compréhension. Notre enquête nous révèle que s'il est très utilisé en milieu universitaire, il existe néanmoins un paradoxe dans son mode d'exploitation car le résumé oral semble être plus souvent exploité dans les épreuves visant à évaluer la compréhension écrite et le résumé oral semble être davantage privilégié pour l'évaluation de la compréhension orale. Ayant déjà par ailleurs traité le domaine de l'évaluation de la compréhension écrite à travers le résumé oral (Isani : 1996) nous nous bornerons essentiellement au sujet du résumé écrit dans le cadre de cette étude.
- 32 Il s'agit d'une lecture faite à haute voix d'un article de la presse écrite, suivie d'une deuxième lecture, parfois avec un lecteur différent. Le candidat a comme consigne de prendre des notes pendant les deux lectures et de rédiger ensuite un résumé dans un nombre de mots limités.
- 33 Cette pratique appelle un certain nombre de remarques. La première concerne le format de restitution habituellement choisi : le résumé écrit. Comme nous avons déjà eu l'occasion de le signaler (Isani 1996), le résumé, écrit encore plus qu'oral, est un mode d'évaluation qui présente des problèmes majeurs de fiabilité car il ne s'agit plus d'évaluer la compréhension orale *per se* mais de l'évaluer premièrement, à travers la production écrite et deuxièmement, sous forme de résumé, un genre complexe et varié qui obéit à des règles discursives *sui generis*. Même avec une grille d'évaluation très fine, il s'agit d'une tâche difficile, complexe et consommatrice de temps pour l'évaluateur car il doit évaluer en même temps la restitution des items de compréhension ciblés, la conformité aux règles discursives du genre et l'expression écrite et ses aspects formels. Dans la complexité du processus nous perdons de vue que l'objectif de départ était d'évaluer la compréhension de l'oral du candidat ce qu'illustre avec clarté ces deux exemples d'annotations portées par le correcteur en bas de la copie : « Vous ne savez pas faire un résumé » ou « Ceci n'est pas un bon résumé ». Tous ces éléments nous conduisent à conclure que si le résumé écrit présente un intérêt certain dans le domaine de l'enseignement aux spécialistes, il s'agit d'un format trop lourd à gérer dans le contexte de l'évaluation de masse du public LANSAD universitaire.
- 34 La deuxième question soulevée par le résumé écrit concerne la lecture oralisée sous forme de monologue qui jusqu'à récemment fut souvent critiquée comme étant une tâche langagière dont l'authenticité était restreinte au contexte du cours magistral universitaire. Et effectivement, *a priori*, nous pourrions nous demander si écouter un monologue, sans interaction de la part de celui qui écoute, relève des activités communicatives de tous les jours. Cependant, une réflexion plus poussée démontre qu'écouter un monologue est effectivement devenu une situation de communication assez courante de nos jours : nous écoutons tous la radio, le journal télévisé, le standard

automatisé des services publics, le répondeur téléphonique, les annonces faites dans les lieux publics, les vidéos de démonstration, etc. Il est donc incontestable qu'aujourd'hui le fait d'avoir à écouter un monologue constitue une tâche communicative authentique.

- 35 Si écouter un discours monologue est donc réhabilité comme tâche langagière authentique, quelques problèmes persistent. Le premier concerne le support de l'épreuve : la lecture d'un article de la presse écrite à haute voix. Cette pratique constitue une distorsion fondamentale de la notion de document authentique. Un article de la presse écrite n'est pas conçu pour être lu à haute voix devant un auditoire. Une lecture de ce type – tronquée de tous les éléments facilitateurs inhérents au discours oral naturel (reformulation, répétitions, hésitations, faux départs, autocorrections, pauses, phatiques, mimo-gestualité, etc.) – est trop soutenue et trop dense pour être entendue et ne peut pas constituer, par conséquent, un support d'épreuve authentique.
- 36 La troisième question posée par ce type d'épreuve est le problème du nombre d'écoutes. Les pratiques actuelles semblent se généraliser autour de **deux écoutes**. Ceci soulève le problème de l'authenticité, et donc, de la validité, de l'épreuve. Avons-nous souvent l'occasion d'écouter et ensuite de réécouter un des exemples de monologues cités ci-dessus ? Parfois oui, si nous recomposons le numéro de téléphone pour nous offrir la possibilité de réécouter le message ou les consignes enregistrés. Mais dans la plupart des cas, cette possibilité d'une deuxième écoute n'existe pas. La question qui se pose alors est de savoir si la pratique de proposer deux écoutes est cohérente avec nos objectifs d'authenticité. La plupart des spécialistes dans le domaine de l'élaboration des épreuves de langue s'y opposent formellement en arguant que si la majorité des candidats se trouve dans l'incapacité d'extraire les informations ciblées après une seule écoute, il faut en tirer les conclusions nécessaires concernant le calibrage de l'épreuve et le niveau ciblé.

### La restitution... et après ?

- 37 Une fois l'information ciblée extraite et restituée, se pose la question primordiale de la suite à donner. Dans le cas des candidats cherchant à faire valider simplement leurs compétences réceptives, nous nous arrêterons là. Mais dans le cas des candidats cherchant à faire valider une compétence de communication globale, il s'agit d'aller plus loin et l'option qui nous semble la plus adaptée à nos objectifs est de réintégrer les compétences réceptives au sein de **la tâche communicative complexe** dont elles constituent une composante essentielle. Autrement dit, le candidat sera appelé à recueillir des informations de sources écrites et orales dans un but précis, éléments qu'il devra par la suite **sélectionner, organiser et mobiliser** pour l'accomplissement d'une tâche communicative complexe, sous forme d'une rédaction de rapport et d'une présentation et argumentation orales de son point de vue.

### Conclusion : quel(s) choix ?

- 38 Confronté à l'extrême complexité qui accompagne la conception, l'élaboration et la mise en œuvre d'un système d'évaluation qui se veut à la fois bon, bon marché et simple, quel concepteur-évaluateur ne s'est pas demandé s'il ne cherchait pas la quadrature du cercle et s'il ne serait pas plus simple d'adopter un système d'évaluation indirecte garantissant les critères de fiabilité et de facilité de mise en œuvre : le QCM !

- 39 Mais il est évident que le concepteur-évaluateur-enseignant que nous sommes tous, se doit de résister à cette tentation de facilité et d'inscrire sa réflexion dans un cadre plus global qui dépasse le simple acte évaluatif. En même temps que le respect des critères de fiabilité et de facilité de mise en œuvre, il nous incombe de veiller au respect du critère de validité et ainsi au puissant phénomène de *washback* qu'il engendre.
- 40 À cet égard, nous pourrions nous féliciter du fait que nous avons une pléthore et non pas une pénurie à gérer en matière de choix de formats de restitution et envisager d'élaborer un système d'évaluation qui retiendra **plusieurs formats de restitution diversifiés**. Comme nous avons pu le constater, tous les formats d'épreuve présentent un certain nombre d'avantages et d'inconvénients. Pourquoi le concepteur ne pourrait-il pas exploiter plusieurs formats à *washback* positif à l'intérieur d'une seule épreuve ? Autrement dit, pourquoi les épreuves de compréhension écrite et orale ne pourraient-elles pas utiliser à la fois et en même temps, le vrai/faux, le QROC et, pourquoi pas à dose homéopathique ?, le QCM qui, ne l'oublions pas, est le format qui présente le degré de fiabilité le plus élevé. Disposer de plusieurs formats d'épreuves diversifiés faciliterait considérablement la tâche du concepteur-évaluateur sans pour autant nuire à notre objectif commun qui est d'ancrer les pratiques pédagogiques dans une perspective de communication authentique.

## BIBLIOGRAPHIE

Heaton, J. B. 1989. *Writing English Language Tests*. Londres : Longman.

Isani, Shaeda & Micheline Herino. 1994. « Quel(s) domaine(s) langagier(s) pour une évaluation en compétence en langue ? ». *ASp* 5-6, 249-256.

Isani, Shaeda. 1996. « Les épreuves de langue dans les concours de la fonction publique : domaines, critères et objectifs ». *ASp* 11-14, 351-361.

Weir, Cyril. 1993. *Understanding and Developing Language Tests*. New York : Prentice Hall International.

## NOTES

1. Nous rappelons à cet égard le colloque tenu à Aix-en-Provence en 1993 où la table ronde organisée entre employeurs et enseignants de langues universitaires a été révélatrice du peu de connaissances que les employeurs avaient par rapport aux différentes certifications existantes en langue étrangères.

2. Le CCLES : Certificat de Compétence en Langues de l'Enseignement Supérieur : il s'agit d'un projet actuellement à l'étude concernant la mise en place d'une certification des compétences langagières du public LANSAD universitaire.

3. J. B. Heaton (1989 : 131-132) distingue trois niveaux de difficulté dans la compréhension d'un texte écrit : « indépendant », « *instructional* » et « *frustrational* ». Un texte de niveau « indépendant » peut être abordé par l'étudiant en autonomie, un texte de niveau « *instructional* »

nécessite l'aide de l'enseignant, et un texte de niveau « *frustrational* » est celui qui dépasse complètement le candidat.

---

## RÉSUMÉS

Cet article se donne comme objectif d'explorer le problème de contamination inhérente à l'évaluation des compétences de compréhension de la langue étrangère. Nous examinons les pratiques actuellement en vigueur concernant la restitution active et passive et les différents formats utilisés, leur adéquation par rapport au contexte d'une évaluation des compétences communicatives et les trois critères docimologiques de validité, de fiabilité et de facilité de mise en œuvre.

This article sets out to explore the problem of contamination inherent to the evaluation of foreign language comprehension skills and information restitution. We examine current practices concerning the use of active and passive restitution and different formats used, assessing them in the context of the evaluation of communicative skills and the three docimological criteria of validity, reliability and practicality.

## INDEX

**Keywords :** cloze, evaluation (comprehension), evaluation (foreign language), multi-response question, question (short answer), representation of written and oral supports, restitution (of comprehension items), summary (written)

**Mots-clés :** évaluation (de la compréhension), évaluation (des compétences en langue étrangère), exercice lacunaire, QCM, QROC, questionnaire à réponse ouverte courte, représentation concernant le support d'épreuve, restitution (d'un item de compréhension), résumé

## AUTEUR

SHAEDA ISANI