



ISSN 1901-3809

ISSN en ligne 2261- 2807

L'évaluation de l'expression écrite en FLE à l'Université – une épreuve de langue fiable et valide ?

Jan Lindschouw

Université de Copenhague, Danemark

janl@hum.ku.dk

Résumé

Le présent travail étudie une épreuve de langue intitulée *Traduction danois-français et expression écrite* au quatrième semestre des études de français à l'Université de Copenhague à partir de cinq notions clés venant du domaine scientifique sur l'évaluation, à savoir *applications pratiques, fiabilité, validité, authenticité et influence de l'épreuve sur l'enseignement*. L'article conclut que cette épreuve pose des défis par rapport à plusieurs de ces paramètres, ce qui est dû notamment à la construction de l'épreuve elle-même qui repose implicitement sur deux conceptions de la langue divergentes et aux objectifs imprécis du programme d'études. L'article analyse également deux exemples de pratique, sous forme de deux copies qui ont reçu la même note, mais par des voies différentes, afin d'étudier ce que cela implique pour la fiabilité et la validité de l'épreuve.

Mots-clés : évaluation d'une épreuve de langue, traduction, expression écrite, fiabilité, validité

Evaluation of Written Expression in French as a Second Language at the University - a Reliable and Valid Test?

Abstract

This article analyses a language test entitled *Translation Danish-French and Written Expression* during the fourth semester of French Studies at the University of Copenhagen on the basis of five key notions from the scientific literature on language assessment: practicality, reliability, validity, authenticity and washback. The article concludes that this language test presents various challenges with respect to several of these notions, which is primarily due to the construct of the test, relying implicitly on two divergent theories of language and to the imprecise objectives of the study program. This article also studies two cases in which the papers received the same mark in different ways, in order to discuss what this implies for the reliability and the validity of the test.

Keywords : language testing, translation, written expression, reliability, validity

1. Introduction

Les évaluations d'épreuves de langue au Danemark reposent très souvent sur l'intuition des examinateurs, leurs expériences et leurs éventuelles conceptions de la langue divergentes. Le problème concerne en particulier les épreuves où sont testées les compétences linguistiques au sens large telles que l'expression orale et écrite. Que cela implique-t-il de pouvoir parler ou écrire le français ? Faut-il être en mesure de communiquer sans fautes grammaticales ? Faut-il avoir un vocabulaire large et varié ? Faut-il être capable de communiquer de façon pragmatiquement adéquate ? Ou doit-on maîtriser différentes stratégies communicatives ? Toutes ces questions ne trouveront probablement pas les mêmes réponses selon l'examineur interrogé. Si les examinateurs ne peuvent pas se mettre d'accord sur l'importance à accorder à ces dimensions, le même type d'épreuve risquera de recevoir différents jugements d'une année à l'autre et entre les examinateurs. À ces problèmes s'ajoute le fait que les objectifs des programmes d'études sont souvent vagues et imprécis pour ce qui est de l'évaluation de l'expression orale et écrite.

Dans cet article, nous allons analyser une épreuve de langue intitulée *Traduction danois-français et expression écrite* au quatrième semestre des études de français à l'Université de Copenhague à partir de notions clés venant du domaine scientifique sur l'évaluation. En outre, nous allons étudier deux copies qui ont reçu la même note afin de discuter ce que cela implique pour le type d'épreuve en question que l'on puisse obtenir la même note par des voies différentes. Notre analyse est fondée sur l'examen d'été en 2013 pour lequel un extrait du texte intitulé *Bryllupsrejse* 'Voyage de noces' de Naja Marie Aidt (2006 : 35-36) est utilisé.

2. Descriptif de l'épreuve

L'épreuve de traduction et d'expression écrite se divise en deux parties : un texte à traduire du danois au français (1100 unités) et une rédaction en français (200 mots au minimum). Selon le profil de compétences de cette matière, l'étudiant doit notamment comprendre « les problèmes liés à la traduction entre la syntaxe danoise et française ainsi que l'argumentation académique¹ », et apprendre à « exploiter son savoir linguistique pour faire des traductions du danois au français² » et à « employer son savoir linguistique pour présenter un point de vue de façon convaincante dans un texte français³ » (Programme d'études, 2012 : 17). Ces compétences sont approfondies dans les objectifs de cette matière. Selon ceux-ci, l'étudiant doit apprendre à « traduire un texte non technique et au niveau demi avancé en un français adéquat et correct⁴ », à « créer une argumentation académique en français écrit⁵ » et à « s'exprimer de façon nuancée, grammaticalement correcte et communicativement adéquate en français⁶ » (Programme d'études, 2012 : 17-18).

Dans l'ensemble, ces objectifs restent imprécis et sont par conséquent difficilement mesurables pour un évaluateur. Que signifient les termes *demi avancé* et *adéquat* ? Pour qui doit-on présenter un point de vue de façon convaincante et comment en mesurer l'effet chez l'interlocuteur ? Et que couvre le terme *s'exprimer de façon communicativement adéquate* ? S'agit-il d'une ou de toutes les compétences dites communicatives (Martinez, 1996 : 77) ?

Généralement, on peut dire que les deux parties de cette épreuve reposent sur deux conceptions de la langue divergentes. La partie *traduction* est fondée sur une conception de la langue formelle selon laquelle la précision grammaticale et la connaissance de la morphologie et la syntaxe françaises sont au centre. En revanche, la partie *expression écrite* repose sur une conception de la langue fonctionnelle et communicative (Canale, 1983, Lund, 2015 : 106-108) selon laquelle la compétence grammaticale n'est pas suffisante pour maîtriser la langue étrangère. Il faut aussi acquérir les compétences discursive, pragmatique et socioculturelle.

En raison des renseignements vagues du programme d'études, il est fort probable que l'évaluation des copies des étudiants risque d'être aléatoire, ce qui implique que la fiabilité de cette épreuve est mise en question. Il en est de même pour sa validité, car si les composantes linguistiques examinées dans l'épreuve ne sont pas évidentes, l'objectif de l'épreuve est par conséquent peu clair.

Afin d'étudier systématiquement les défis de cette épreuve linguistique, nous allons par la suite l'analyser à partir des notions clés empruntées au domaine scientifique sur l'évaluation, à savoir *les applications pratiques* 'practicality', *la fiabilité* 'reliability', *la validité* 'validity', *l'authenticité* 'authenticity' ainsi que *l'influence de l'épreuve sur l'enseignement* 'washback' (McNamara, 2000 : 9, 47-62, Brown & Abeywickrama, 2010 : 25-51, Weigle, 2013 : 48-56). Toutefois, d'après Weigle (2013 : 48) il est illusoire de croire que l'on peut maximaliser toutes ces composantes à la fois. Si l'on crée une épreuve dont la validité est très élevée, les applications pratiques seront affectées négativement ; et une épreuve dont la fiabilité est portée à son maximum risquera d'être peu valide. Il faudra donc créer un équilibre entre ces cinq paramètres.

2.1. Applications pratiques

Cette catégorie réfère à toutes les conditions pratiques relatives à l'épreuve, c'est-à-dire les frais liés à l'administration de l'épreuve, le temps employé à l'accomplir et à l'évaluer, le recours à des ressources technologiques disponibles, etc. (Brown & Abeywickrama, 2010 : 26-27, Weigle, 2013 : 48).

D'un point de vue administratif, cette épreuve n'est ni difficile ni coûteuse à gérer. Jusqu'en 2014, c'était une épreuve sur place surveillée ; désormais, c'est un devoir fait à la maison. Cependant, quant aux recours aux ressources technologiques disponibles, cette épreuve connaît certains défis. Comme c'est une épreuve, dont tous les outils de travail sont autorisés, il est permis aux étudiants de rendre leur travail en version électronique. Cependant, la Faculté des Lettres de l'Université de Copenhague a décidé que lors des épreuves sur place, les étudiants n'auront pas recours à Internet, afin d'éviter toute communication avec l'extérieur. Or, il est bien connu que le processus de traduction et de rédaction requiert le recours à Internet. Comme un résultat de cette critique, on a changé la forme d'examen de cette épreuve, de sorte qu'elle est désormais un devoir à faire à la maison. Ce changement a certes amélioré les applications pratiques de l'épreuve, mais a augmenté le risque de communication avec l'extérieur de façon considérable.

C'est néanmoins avec l'évaluation des copies que cette épreuve présente le plus grand nombre de défis. Contrairement à une épreuve indirecte, une épreuve directe⁷ comme celle-ci est exigeante, tant mentalement que temporellement, dans la mesure où chaque copie doit être évaluée individuellement. À cela s'ajoute le fait que les erreurs qui apparaissent vont de fautes grammaticales pures à des problèmes liés à l'acceptabilité des énoncés (Ellis & Barkhuizen, 2005 : 98). Il va de soi qu'il est beaucoup plus exigeant de corriger des énoncés dont l'acceptabilité peut être questionnée que des fautes grammaticales.

2.2. Fiabilité

Une épreuve fiable est censée mener au même résultat sous les mêmes conditions et des évaluations répétées (Brown & Abeywickrama, 2010 : 27, Weigle, 2013 : 49). La fiabilité peut être divisée en plusieurs composantes qui sont toutes « extérieures » à l'épreuve.

La fiabilité liée aux étudiants réfère à leur forme physique et mentale lors du jour de l'examen. Si l'étudiant est malade ou très nerveux, il est moins performant que d'habitude. La fiabilité relative à l'épreuve elle-même, réfère à ses conditions physiques, par exemple si le texte d'examen contient des erreurs typographiques ou s'il y a du bruit dans la salle d'examen. Lors de l'examen d'été en 2013, plusieurs étudiants se plaignaient du bruit provoqué par un grand projet de construction à l'extérieur de la salle d'examen. Il est pourtant difficile de tenir compte de ces deux types de fiabilité lors de l'évaluation, puisqu'elles sont difficilement mesurables.

La fiabilité liée au processus d'évaluation est celle qui pose le plus de défis pour cette épreuve. Ce genre de fiabilité se divise en deux catégories : *la fiabilité chez le même évaluateur* 'intra-rater reliability' et *la fiabilité entre plusieurs évaluateurs* 'inter-rater reliability' (Brown & Abeywickrama, 2010 : 28, Weigle, 2013 : 49). Pour ce qui est de la première catégorie, qui réfère à la stabilité d'un seul examinateur dans ses propres jugements, cette épreuve n'est pas sans poser problème, notamment dans la partie de l'expression écrite où les réponses peuvent varier considérablement d'une copie à l'autre, puisque elle repose sur un principe d'acceptabilité plutôt que de grammaticalité (voir 2.1.). Cependant, la deuxième catégorie pose davantage de défis, ce qui est dû premièrement à la construction de cette épreuve, qui repose sur deux conceptions de la langue opposées (voir section 2), et deuxièmement aux termes vagues du programme d'études. Si l'examineur et le co-examineur ne partagent pas la même conception de la langue, on peut avoir affaire à un véritable conflit entre les deux lors de l'attribution de la note. On peut fort bien s'imaginer que le co-examineur considère la compétence grammaticale comme la plus importante dans les deux sous-disciplines, alors que l'examineur tient les compétences socioculturelle, discursive et stratégique (Martinez, 1996 : 77, Canale, 1983) pour aussi importantes que la compétence grammaticale et *vice versa*.

2.3. Validité

La validité d'une épreuve peut être définie comme sa capacité à mesurer et à juger ce à quoi il est destiné (McNamara, 2000 : 47-54, Brown & Abeywickrama, 2010 : 29-36, Weigle, 2013 : 49-51). La validité est intimement liée à la fiabilité : si une épreuve est dépourvue de fiabilité, elle est en principe dépourvue de validité aussi. Si la fiabilité réfère à des conditions « extérieures » à l'épreuve, la validité a affaire aux conditions « internes » de celle-ci.

L'objectif de l'épreuve en question est double : traduire un texte danois en français et s'exprimer librement en français par écrit. Ces objectifs sont concrétisés dans les instructions du texte de l'examen d'été 2013 : « Traduisez en français le passage de *Eva kysse Tim* [...] jusqu'à [...] *vanvittig af den lyd* » et « Discutez le rôle des voyages touristiques dans la société moderne en faisant mention, à titre d'exemple, de leurs avantages et inconvénients » (voir annexe). En dépit de ces instructions plutôt claires, nous ne considérons pas que cette épreuve soit très valide, ce qui peut être imputé aux problèmes de fiabilité mentionnés. Il n'est pas évident de savoir si tous les éléments de la communication communicative sont testés ou seulement une partie d'entre eux. Et si tous ces éléments sont testés,

comment sont-ils évalués entre eux ? Bien qu'à première vue la première partie de l'examen (la traduction) paraisse plus valide que la deuxième (l'expression écrite), dans la mesure où la traduction d'un texte concerne notamment la transposition d'énoncés de la langue source à la langue cible en respectant les normes grammaticales et lexicales de la langue cible, force est de constater que des dimensions extra-linguistiques sont également tenues en compte lors de l'évaluation d'une traduction, ce qui ne ressort pas du texte d'examen.

2.4. Authenticité

L'authenticité d'une épreuve de langue reflète dans quelle mesure elle ressemble à une situation concrète de la langue cible (McNamara, 2000 : 7-9, Brown & Abeywickrama, 2010 : 36-37, Weigle, 2013 : 51-52). L'authenticité d'une épreuve se traduit notamment par un langage naturel, des sujets pertinents et actuels et le recours aux tâches reflétant celles que l'on observe dans le monde réel. Plusieurs de ces paramètres sont satisfaits dans l'épreuve étudiée ici (voir annexe) : le vocabulaire du texte d'examen 2013 appartient généralement au danois courant et standard (*middagsheden* 'la chaleur méridienne', *landskab* 'paysage', *bil* 'voiture', *kyssede* 'donna un baiser', *grinede* 'rit', etc.) avec quelques mots d'ordre poétique et spécialisé (*pinjetræer* 'piniers', *dinglede*, 'oscillait', *sørgende* 'attristé', *rampoverede* 'endommagés', *vidende smil* 'sourires de complicité', etc.). En outre, le texte est contextualisé, dans la mesure où il parle d'un jeune couple en voyage de noces en Grèce, et tous les étudiants en langue vont probablement connaître ce sujet, parce qu'ils sont censés avoir voyagé beaucoup, du moins pour explorer la culture de leur langue cible.

Toutefois, Weigle (2013 : 52) considère que les épreuves en expression écrite sont dépourvues d'authenticité pour plusieurs raisons. Premièrement, l'expression académique requiert que les étudiants aient une certaine connaissance sur le sujet à l'avance, obtenue au moyen de cours magistraux ou des articles spécialisés sur le sujet. Deuxièmement, la rédaction d'un texte académique ne se fait pas sous une contrainte de temps intense. Enfin, les étudiants savent à l'avance que le récepteur du texte est l'enseignant et suppose que celui-ci connaisse leur style d'écriture et qu'il soit capable de déduire le sens de leurs énoncés pas toujours réussis.

2.5. L'influence de l'épreuve sur l'enseignement

L'influence de l'épreuve sur l'enseignement et l'apprentissage des apprenants, aussi connu sous le terme de *washback* (McNamara, 2000 : 72, Brown & Abeywickrama, 2010 : 37-39, Weigle, 2013 : 54-55), se divise en un pôle positif

et un pôle négatif. On a affaire à du *washback* positif, lorsque les préparations à l'épreuve s'intègrent parfaitement avec les objectifs du cours. En revanche, le *washback* négatif concerne les cas où l'épreuve détermine l'enseignement d'une façon inadéquate, par exemple si elle ne reflète pas les principes didactiques et acquisitionnels du cours ou si l'enseignement amène seulement les élèves à faire passer l'examen avec un bon résultat.

Il est vrai que tous les examens universitaires dirigent dans une certaine mesure l'enseignement dispensé, car l'enseignement est destiné à préparer les étudiants à l'examen. Ce cours ne constitue pas une exception. Néanmoins, nous considérons que le *washback* peut dans une grande mesure être considéré comme positif dans notre cas, puisque l'enseignement aussi bien que l'épreuve est centré sur la dimension écrite au sens large, ce qui cadre parfaitement avec la conception de langue fonctionnelle et communicative (Canale, 1983, Lund, 2015 : 106-108). Ainsi, afin de développer la compétence grammaticale des étudiants, ils font différents exercices de grammaire sur des thèmes variés qui posent problèmes à des apprenants danois. De plus, les corrections apportées aux traductions des étudiants concernent notamment la grammaire et le lexique. Cependant, le cours s'ouvre aussi à d'autres domaines de la compétence communicative, par exemple la compétence textuelle et les structures argumentatives en focalisant sur l'apprentissage de différents connecteurs et marqueurs cohésifs sous la base du manuel de Garnier & Savage (2011).

3. Deux exemples de pratique

Dans la section précédente, nous avons étudié l'épreuve en traduction et en expression écrite selon cinq notions clés de la littérature scientifique sur l'évaluation qui se sont toutes avérées problématiques, à une exception près. Dans cette section, nous allons considérer deux exemples de pratique en analysant deux copies qui ont reçu la même note (la note moyenne 7 du barème de notation danois, note qui correspond à la note C dans le système européen) afin d'étudier si les problèmes liés à la fiabilité et à la validité de l'épreuve s'applique aussi à la pratique.

Selon le Ministère de l'Éducation nationale danoise, « la note 7 est attribuée à une bonne performance qui montre que les objectifs du cours sont acquis bien qu'il reste un certain nombre de manques »⁸ (Ministère de l'Éducation nationale danoise). Avec une telle définition ample, il est prévisible que cette note puisse être obtenue par de multiples voies. Les exigences de l'épreuve sont présentées dans la section 2. Comme nous l'avons déjà dit dans la section 1, notre analyse est fondée sur l'examen d'été en 2013 pour lequel un extrait du texte intitulé *Bryllupsrejse*

‘Voyage de noces’ de Naja Marie Aidt (2006) est utilisé. Cet extrait est présenté en annexe. Il en ressort que le texte d’examen en question comprend deux parties : un passage à traduire du danois en français (question 1) et un texte académique à rédiger sur les voyages touristiques dans la société moderne (question 2). Pour des questions d’espace, nous ne présenterons pas dans ce qui suit les réponses des étudiants dans leur intégralité, mais fournirons des exemples à titre illustratif.

3.1. Première copie

D’un point de vue grammatical, la deuxième partie de la première copie (l’expression écrite) est meilleure que la première (la traduction). C’est une tendance récurrente pour la plupart des copies, ce qui s’explique par le fait que la traduction est beaucoup plus contraignante que l’expression écrite, puisque l’étudiant doit transposer les structures linguistiques de la langue source à la langue cible aussi soigneusement que possible, alors qu’il peut choisir les structures dont il a déjà la maîtrise pour l’expression écrite.

Les erreurs de cette copie sont en particulier limitées à la syntaxe verbale, notamment à l’aspect verbal et au mode, par exemple le recours à l’indicatif pour le subjonctif après la conjonction *pour que* et l’emploi du présent plutôt que l’imparfait dans la locution verbale *il y a longtemps* employée en contexte de passé. En outre, certaines erreurs sont observées pour l’emploi des articles, surtout le recours à l’article partitif plein (*des*) à la place de l’article partitif réduit (*de*). Ensuite, ce travail présente certains problèmes au niveau du lexique, à titre d’exemple le substantif *caillots* à la place de *cailloux*, et au niveau des homonymes : **il l’a regarda* pour *il la regarda*. Enfin, cette copie contient certaines erreurs orthographiques.

Cela dit, cette copie témoigne d’une compétence grammaticale remarquable à divers égards. À quelques exceptions près, l’étudiant maîtrise la différence aspectuelle entre le passé simple et l’imparfait. Il emploie correctement le conditionnel pour traduire le futur dans le passé et le gérondif pour indiquer la simultanéité. En outre, l’étudiant maîtrise généralement la morphologie et emploie un vocabulaire précis et varié. Dans la partie consacrée à l’expression écrite, on observe également une syntaxe complexe avec un usage fréquent de structures subordonnées et de syntagmes verbaux infinis, de même que l’emploi d’un vaste échantillon de connecteurs de différents types : causals, adversatifs, consécutifs, énumératifs, etc., ce qui établit une chaîne argumentative cohérente et logique.

Si cette copie possède toutes ces qualités grammaticales, pourquoi n'a-t-elle obtenu qu'une note moyenne ? Cela est probablement dû à la façon dont est présenté le contenu dans la partie sur l'expression écrite. De temps à autre, le texte semble construit et la langue peu authentique et élaborée. On a l'impression que cet étudiant est un apprenant ayant appris le français par les yeux (Ferris, 2012 : 54), c'est-à-dire notamment par la lecture et les règles grammaticales. Cette impression est renforcée par le fait que le texte suit un modèle : premièrement, tous les aspects positifs relatifs aux voyages touristiques sont présentés, suivis de tous les aspects négatifs, ce qui crée une certaine monotonie dans l'argumentation. En outre, l'étudiant a de temps à autres recours à des constructions très formelles et peu courantes qui témoignent d'une connaissance grammaticale subtile. On cite le recours fréquent à des connecteurs tels que *peut-être* et *aussi* qui, lorsque situés en tête de phrase, sont suivis de l'inversion complexe : *peut-être peut-il ; aussi les déchets peuvent-ils*. Par le dernier exemple, l'étudiant a fort probablement voulu montrer qu'il maîtrise la règle selon laquelle *aussi* prend un sens de causalité lorsque suivi de l'inversion complexe. Une solution plus courante aurait sans doute été *par conséquent* ou *c'est pourquoi*.

3.2. Deuxième copie

La deuxième copie a aussi obtenu la note moyenne 7, mais pour d'autres raisons que celle qui précède. Ici nous avons affaire à un étudiant qui est grammaticalement moins performant que la précédente. Cela s'observe notamment pour ce qui est de la syntaxe verbale, en particulier pour l'aspect verbal, où l'étudiant à plusieurs reprises emploie l'imparfait pour des actions clairement délimitées. En outre, on observe un seul cas où le subjonctif est employé à la place de l'indicatif et plusieurs cas où les formes des pronoms personnels sont mélangés, par exemple *lui* à la place de *le* ou *la* et *vice versa*. De plus, cette copie présente plusieurs problèmes liés à l'accord en genre et en nombre, notamment entre le sujet et le participe passé : **elle [...] étant accoupi ; *la voiture noire couvert de poussière ; *les gens qui sont parti*. S'y ajoutent l'antéposition d'adjectifs normalement postposés : **un cher inconvenient économique* et une ponctuation insuffisante.

Cela dit, plusieurs éléments grammaticaux sont maîtrisés, par exemple le conditionnel pour exprimer le futur dans le passé et le gérondif pour indiquer la simultanéité. En outre, l'étudiant maîtrise l'emploi du subjonctif dans différents types de subordonnées, par exemple dans les subordonnées complétives précédées par le verbe *falloir* et dans les subordonnées relatives après un antécédent nié (*il n'y a personne qui sache*). Ainsi, on peut considérer le recours au subjonctif à la

place de l'indicatif mentionné dans le paragraphe précédent comme une *faute* 'mistake', c'est-à-dire une méprise linguistique commise par l'apprenant, puisque la règle linguistique n'est pas entièrement acquise, et non comme une *erreur* 'error', à savoir une lacune fondamentale dans l'interlangue de l'apprenant (Ellis & Barkhuizen, 2005 : 104). De plus, l'étudiant possède une compétence stylistique, parce qu'il utilise la forme formelle *l'on*, appropriée dans un essai académique, à la place de *on*, afin d'éviter l'élosion de l'e-caduc en *que: il faut que l'on travaille*. Enfin, l'étudiant a recours à un vocabulaire précis et nuancé (*avantageux, inconvenient, menace, etc.*) et à des verbes sémantiquement complexes (*inspirent, soulève, etc.*).

Bien que cette copie possède plus d'erreurs grammaticales que la précédente, elle a obtenu la même note, ce qui s'explique par la façon dont est organisé le contenu. Ici, il n'est pas présenté sous forme d'un modèle fixe et imité comme dans le cas précédent, mais de façon organique : l'étudiant présente les avantages et inconvénients liés aux voyages touristiques, mais les mélange de façon dynamique tout au long du travail et y intègre également des réflexions personnelles et des éléments inattendus, par exemple en faisant une comparaison entre un voyage concret et imaginaire. En outre, on a l'impression d'avoir affaire à un apprenant ayant appris le français par l'oreille (Ferris, 2012 : 54), c'est-à-dire par de longs séjours dans la culture cible. En effet, l'étudiant emploie des expressions et collocations naturelles en français : *avoir pour simple but, il est avantageux, soulever un paradoxe, etc.* et n'a pas recours à d'expressions qui illustrent la maîtrise de règles grammaticales subtiles et rares.

En conclusion, l'analyse de ces deux copies a montré que la même note peut s'obtenir par différentes voies, soit grâce aux qualités grammaticales aux dépens de l'organisation textuelle, soit grâce à l'organisation textuelle aux dépens de la grammaire. En d'autres termes, même en pratique, cette épreuve de langue pose des défis au sujet de la fiabilité et de la validité. Il importe de souligner que l'argumentation exposée ici repose sur notre évaluation en tant qu'examineur. Il se peut bien que le co-examineur ait valorisé d'autres aspects des copies, et pourtant il est parvenu au même résultat que nous. Dans beaucoup de cas, les deux examinateurs⁹ attribuent la note sans même consulter les copies ensemble, sauf en cas de désaccord. En outre, on peut bien s'imaginer que si les deux examinateurs ne partageaient pas la même conception de la langue, le résultat serait bien différent. Si l'examineur avait eu une conception formelle de la langue, la première copie aurait fort probablement eu une note au-dessus de la moyenne (10 ou B dans le système européen), et la deuxième une note au-dessous de la moyenne (4 ou D dans le système européen). Tous ces éléments ne renforcent guère la fiabilité et la validité de cette épreuve.

4. Conclusions

Cet article a étudié une épreuve de langue en traduction et en expression écrite en français en deuxième année à l'université de Copenhague selon cinq notions clés de la littérature scientifique sur l'évaluation. L'analyse a montré que cette épreuve pose des défis par rapport à tous ces paramètres, sauf pour ce qui est du *washback*. Les défis étaient plus prononcés pour la fiabilité et la validité de l'épreuve. Les problèmes relatifs à la fiabilité, en particulier celle liée à l'évaluateur, sont dus à la construction de l'épreuve qui repose implicitement sur deux conceptions de la langue divergentes et aux objectifs imprécis du programme d'études. Ces problèmes sont surtout prononcés dans les cas où les deux examinateurs ne partagent pas la même conception de la langue. L'analyse a également montré que les problèmes liés à la fiabilité affectent aussi la validité de l'épreuve. L'analyse des deux copies choisies a confirmé les problèmes de fiabilité et de validité de l'épreuve, car elle a montré que la même note peut s'obtenir par des voies différentes.

Il importe de se demander si un devoir sur l'expression écrite libre est une tâche pertinente pour évaluer la compétence écrite des étudiants dans la langue cible. Cette tâche est inauthentique, du moins comme genre académique, car les étudiants n'ont souvent pas une connaissance profonde sur le sujet et ont par conséquent des difficultés à rédiger un texte réfléchi, nuancé - et académique. Pour cette raison, on peut envisager d'utiliser les sujets abordés dans les cours de littérature ou d'actualités comme thèmes pour l'examen en expression écrite et de demander aux étudiants de rédiger leur texte en fonction d'un récepteur bien défini, par exemple un associé intéressé ou un adversaire. En outre, on peut se demander si ce type d'examen reflète les compétences requises par les employeurs. La traduction et l'argumentation académique sont certes des compétences pertinentes pour les chercheurs, les traducteurs et les futurs professeurs de lycée, mais puisque les candidats en langue s'orientent de plus en plus vers les entreprises et les organisations (internationales), il faut aussi envisager de les initier à d'autres genres textuels tels que les lettres commerciales et les sites web.

Afin de résoudre les problèmes liés à la fiabilité et à la validité, les objectifs des programmes d'études doivent tout d'abord être révisés et précisés tant par rapport à la notion de langue employée qu'aux éléments de la compétence communicative en jeu. En outre, il faut envisager de proposer des cours d'entraînement des nouveaux membres des jurys d'examineurs nationaux. Ainsi l'évaluation des épreuves de langue ne restera pas une activité aléatoire, ce qui n'est ni de l'intérêt des institutions éducatives ni de celui des enseignants ni de celui des étudiants.

Bibliographie

- Aidt, N.M. 2006. *Bavian*. Copenhagen : Gyldendal.
- Brown, H.D., Abeywickrama, P. 2010. *Language Assessment. Principles and Classroom Practices*. New York : Longman.
- Canale, M. 1983. From communicative competence to communicative language pedagogy. In : *Language and Communication*. Londres : Longman.
- Ellis, R., Barkhuizen, G. 2005. *Analysing Learner Language*. Oxford : Oxford University Press.
- Ferris, D.R. 2012 [2002]. *Treatment of Error in Second Language Student Writing*. Ann Arbor : University of Michigan Press.
- Garnier, S., Savage, A.D. 2011. *Rédiger un texte académique en français*. Paris : Ophrys.
- Lund, K. 2015 [2009]. Fokus på sprog. In: *Sprogfag i forandring - pædagogik og praksis*. Frederiksberg : Samfundslitteratur.
- Martinez, P. 1996. *La didactique des langues étrangères*. Paris : PUF.
- McNamara, T. 2000. *Language Testing*. Oxford : Oxford University Press.
- Ministère de l'Éducation nationale danoise. <http://uvm.dk/Uddannelser/Paa-tvaers-af-uddannelserne/7-trins-skalaen/Karakterer-paa-7-trins-skalaen> [consulté le 30 mai 2015].
- Programme d'Études 2012. Studieordning for det centrale fag på bachelorniveau i Fransk sprog, litteratur og kultur, 2012-ordningen. Université de Copenhague, Institut d'Études anglaises, germaniques et romanes.
- http://hum.ku.dk/uddannelser/aktuelle_studieordninger/fransk/fransk_ba12_rev13_just_14_15.pdf [consulté le 27 mai 2015].
- Weigle, S.C. 2013. *Assessing Writing*. Cambridge : Cambridge University Press.

Annexe

Texte pour l'examen d'été 2013

”Det var helt sikkert et udslag af Tims iver og grænseløse spontanitet, som fik dem til at begive sig op ad bjerget midt i middagsheden. Dette græske landskab, som Eva aldrig havde brudt sig særligt om, virkede mere fjendtligt og tørt end nogensinde. Pinjetræer og vild oliven dinglede som hjælpeløse sørgende ud over skrænterne, og den allestedsnærværende duft af timian gav hende kvalme. Men Tim ville se kvindernes by, Olimpos, som lå på toppen af bjerget. Og de kørte opad i den gamle, ramponerede bil han havde lejet af en amerikansk kvinde, som med sine løse gevandter og rynkede, solhærgede hud, mindede ham om hans egen mor. Udstødningsrøret raslede hen over den grusede vej. Eva kyssede Tim på halsen. Han så på hende. Deres ansigter lyste op i strålende, vidende smil. [...] Men da Eva lidt efter insisterede på, at de gjorde holdt, tog Tim et billede af hende, mens hun sad på hug og tisede med blottet underkrop, og hun fór op og løb efter ham, hun forsøgte at rive kameraet fra ham, hun var virkelig vred, men han grinede bare, løb op ad vejen, og her lykkedes det ham at tage endnu et billede: Hun står med skrævende ben og råber med åben mund, mens hun peger truende på ham. Bag hende ser man det sølvgrønne vildnis og den støvede sorte bil. Den venstre side af hendes ansigt er stærkt oplyst af solen. [...]

Hun satte sig ind i bilen, smækkede døren og bedyrede, at fra nu af ville hun ikke tale med ham i mindst en halv time. Han rystede på hovedet og gassede op. Han grinede og sagde, at hun var en furie. Han sagde, at han elskede hende. Men Eva gav sig ikke. De var begge tørstige, men de havde for længst drukket det sidste vand. Småsten fra vejen sprang op og ramte ustandselig bilen mens de kørte, og efter nogen tid var hun ved at blive vanvittig af den lyd.”

Questions

1. Traduisez en français le passage de *Eva kyssede Tim* [...] jusqu'à [...] *vanvittig af den lyd* (1094 unités).

2. Dans le texte, on peut lire que le couple Eva et Tim passe leurs vacances en Grèce. Discutez du rôle des voyages touristiques dans la société moderne en faisant mention, à titre d'exemple, de leurs avantages et inconvénients (Écrivez 200 mots au minimum).

Notes

1. « oversættelsesmæssige problemer mellem dansk og fransk syntaks samt [...] akademisk argumentation ».
2. « at udnytte sin sproglige viden i forbindelse med oversættelsesopgaver fra dansk til fransk ».
3. « at udnytte sin sproglige viden til at fremstille et synspunkt overbevisende i en fransk tekst ».
4. « oversætte en ikke-teknisk, middelsvær tekst til et dækkende og korrekt fransk ».
5. « opbygge en akademisk argumentation på skriftligt fransk ».
6. « udtrykke sig nuanceret, grammatisk korrekt og kommunikativt hensigtsmæssigt på fransk ».
7. Une épreuve directe est une épreuve dans laquelle les apprenants accomplissent la tâche ciblée, alors que dans une épreuve indirecte les apprenants n'accomplissent pas la tâche ciblée, mais une tâche qui y est reliée (Brown & Abeywickrama, 2010 : 31). A titre d'exemple, une épreuve directe en expression écrite peut être une rédaction dans laquelle les apprenants emploient un vaste éventail de structures linguistiques de la langue cible, tandis qu'une épreuve indirecte pourrait être une tâche à choix multiples où les apprenants doivent cocher la bonne réponse sans produire d'énoncés linguistiques eux-mêmes.
8. « Karakteren 7 gives for den gode præstation, der demonstrerer opfyldelse af fagets mål, med en del mangler ».
9. Il importe de souligner que contrairement à bien d'autres pays, on a deux examinateurs pour la plupart des épreuves au Danemark : celui qui a enseigné le cours (le premier examinateur) et le censeur (le deuxième examinateur), ce dernier étant très souvent externe à l'établissement.