

Suliko Liiv, Ene Alas
Université de Tallinn, Estonie



Synergies Pays Riverains de la Baltique
n°6 - 2009 pp. 241-247

Résumé: *L'Estonie s'est impliquée dans l'élaboration professionnelle de tests depuis plus d'une décennie, et ceci a donné au système éducatif estonien une grande force d'expérience. L'objectif du présent article est de donner une vue d'ensemble des progrès et des problèmes liés à l'amélioration de l'examen national d'anglais langue étrangère en Estonie.*

Mots-clés: *validité, fiabilité, descriptifs des épreuves*

Abstract: *Estonia has been involved in professional test-construction for over a decade and that has given the Estonian education system an enormous amount of experience. The article gives an overview of the development and problems related to the advancement of the national examination in English in Estonia.*

Keywords: *test validity, test reliability, test specifications*

I. Développement des tests

C'est en 1994 qu'on a commencé au niveau national en Estonie, dans le domaine de l'évaluation de la langue anglaise, à tenter systématiquement de suivre les principes d'évaluation des langues selon la tradition occidentale établie depuis quelques temps déjà. A cette époque, à l'Ouest, la recherche dans le domaine de l'évaluation des langues et du développement de tests adéquats étaient des composantes indépendantes et indispensables de l'enseignement et de l'évaluation des langues; l'Estonie, qui venait de retrouver une indépendance politique, économique et culturelle, désirait rapidement tirer parti de l'expérience des pays de l'Ouest et mettre en application les principes d'évaluation de l'anglais chez elle. Repensant à ce qu'était l'évaluation d'une langue étrangère avant l'établissement d'un examen national, S. Liiv affirme qu'« Il n'y avait pas d'examen final en anglais alors, et les professeurs disposaient d'une grande marge de manoeuvre pour compiler, gérer et noter les tests. Chaque école créait ses propres tests pour l'examen final et, par conséquent, les types d'exercices variaient énormément et on ne pouvait faire aucune comparaison des résultats d'une école à une autre : il y avait un manque de objectivité. » (Liiv, 2002 : 51-52)

La nécessité de créer un outil qu'on utiliserait pour évaluer la connaissance d'une langue dans toute l'Estonie amena le Ministre de l'Education de l'époque à former une équipe de travail qui commença à élaborer les premiers tests pilotes. La tentative de mettre en place un examen national pour les candidats aux lycées, à l'éducation supérieure et aux universités ne se limitait pas à l'anglais, mais incluait aussi toutes les langues enseignées en Estonie (l'estonien, le russe, l'anglais, l'allemand, le français), ainsi que les sciences. Les personnes chargées d'élaborer les tests avaient reçu à l'époque une formation générale commune, toutes matières confondues, mais de nos jours les groupes chargés du développement des examens nationaux travaillent de façon relativement indépendante, par spécialité. C'est en 1997 qu'a ouvert le centre national des examens et des qualifications, centre qui contrôle l'élaboration des examens nationaux, et dont une des missions principales est de garantir une gestion professionnelle et opportune des examens nationaux. Les limites de cet article ne nous permettront pas de faire des comparaisons avec d'autres matières et, par conséquent, nous nous bornerons à discuter de l'examen national de langue anglaise.

II. L'examen national d'anglais de nos jours.

L'élaboration et la mise en forme de l'examen national d'anglais suivent la directive du Ministère de l'Education et des Sciences de janvier 2001. Cette directive stipule les objectifs de l'examen national de la manière suivante :

- évaluer l'acquisition des objectifs pédagogiques tels qu'ils sont spécifiés dans le programme des écoles et des lycées ;
- donner aux écoles et aux professeurs la possibilité de comparer les résultats de leurs élèves avec ceux des autres élèves du pays ;
- mener la progression pédagogique dans le sens déterminé par le contenu et la forme de l'examen national ;
- établir des passerelles entre les différentes étapes et différents niveaux pédagogiques ;
- par un système de notation externe à l'examen, faire part de ses réactions auprès des responsables et permettre de prévoir ou même procéder à des changements dans le programme national, dans les manuels, dans la formation des professeurs en poste, et permettre ainsi le développement des domaines concernés.

Comme on peut le voir, l'objectif de l'examen national n'a pas différé, du moins dans les grandes lignes, de celui imaginé à l'origine. Par conséquent, ce dont les responsables de son élaboration doivent être conscients, c'est l'importance de l'effet de retour en ce qui concerne l'enseignement, les méthodes d'évaluation dans les écoles, ainsi que son impact sur les intéressés. « On entend par intéressés les concepteurs des tests, les professeurs, les étudiants, les personnes traitant les résultats, les gouvernements, et tout autre individu ou groupe qui a un quelconque intérêt à savoir de quelle façon utiliser les résultats, et s'ils sont utiles dans un contexte donné » (Fulcher, Davidson, 2007:14).

Le nombre d'étudiants qui a choisi l'anglais comme matière d'examen national au cours des années et, par conséquent, qui considère cette langue comme importante pour leur choix de carrière, est exposé dans le tableau suivant :

Tableau 1. Nombre des participants à l'examen national au cours des années.

Année	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Candidats	9280	8769	9258	9461	8488	9311	9431	9099	9415	9590	9696

Partant du principe que le nombre moyen des candidats lycéens en Estonie est légèrement supérieur à 14000, on peut voir qu'environ 75% des lycéens choisissent l'anglais comme matière d'examen.

L'élaboration de l'examen national suit des critères qui proviennent, d'une part, du programme national et, d'autre part, du manuel *Year 12 Handbook*. Le niveau de langue qui doit être atteint à la fin des études de lycée en anglais est B2 dans toutes les compétences générales (compréhension écrite et orale, expression écrite et orale) comme stipulé par le CECR (Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer, 2001).

Le programme stipule donc le contenu de l'examen de façon très large. Le manuel *Year 12 Handbook* (Jõul 2005 et al.) donne des directives beaucoup plus concrètes. Le tout premier manuel date de 1995 et fut réédité de nombreuses fois, en même temps que l'élaboration de l'examen évoluait. Le manuel décrit la forme d'examen de chaque sous-compétence (expression écrite, compréhension orale, compréhension écrite, structures de langue), ainsi que le test d'expression orale, donne des exemples de styles de textes et de styles de consignes de test, fournit des réponses-types, des aides aux élèves et des échelles de notation pour chaque rubrique évaluée au cours de l'examen (écrit et oral).

En tenant compte de ces directives, la tâche de l'équipe d'élaboration de l'examen national est de constituer un test qui serait avant tout adéquat et fiable, c'est-à-dire pouvoir évaluer les compétences qu'il prétend évaluer, et ceci quelles que soient les conditions ou raisons du test. Chaque exemplaire d'examen national est un travail d'équipe, dont les consignes résultent du travail d'un certain nombre de concepteurs d'exercices qui ont reçu un entraînement pour concevoir des items qui permettent de tester chaque rubrique de compétence. C'est le responsable de l'équipe de chaque rubrique qui reçoit les items ou achève les consignes des concepteurs, assemble les items sous des consignes respectives, et enfin les soumet au spécialiste en chef responsable de la langue anglaise. Les items et les consignes sont discutés par des consultants indépendants, puis sont testés auprès des élèves de la onzième classe pour rendre compte de leur efficacité. Les écoles choisies sont représentatives de tous les types d'écoles dont les élèves présentent l'examen national, c'est-à-dire les écoles des villes, de la campagne, les écoles russe, les écoles estoniennes, etc. Le test des items d'examen amène à une évaluation statistique et qualitative. Selon Hughes, l'analyse statistique démontre la qualité du test (sa fiabilité, par exemple) en tant qu'entité et aussi dans ses divers items (leur difficulté par exemple, leur capacité à faire la différence entre les candidats forts et faibles) (Hughes, 2003 : 65). D'autre part, l'analyse qualitative est effectuée "afin de découvrir des erreurs d'interprétation, des réponses inattendues mais correctes, et autres signes montrant les faiblesses de tel ou tel item" (*ibid.*). Une fois que les tests sont considérés comme satisfaisants (ce qui signifie

le plus souvent apporter des modifications aux exercices, ou fréquemment abandonner tel ou tel exercice s'avérant inapproprié, et donc en refondre de nouveaux), le spécialiste en chef assemble deux versions finales de l'examen national (variante A et B); les deux étant, dans l'idéal, de qualité égale. Avant la finalisation, les versions de l'examen sont testées et relues par des natifs de langue anglaise. Ce procédé d'élaboration des examens a été suivi, avec très peu de changements, depuis 1997, année où le tout premier examen a été constitué.

III. Les consignes de notation

Les notations objective et subjective ont été utilisées dès le départ dans l'examen national de langue anglaise. Les tests de compréhension orale, de compréhension écrite et ceux de structures de langue ont toujours été notés de façon objective, puisque leur notation dépend de réponses-types pour chaque item. Ces réponses-types sont fournies en même temps que l'élaboration des items, et sont examinées à nouveau lors de la période d'essai des tests, ce qui aboutit invariablement à des réponses quelque fois acceptables, mais qu'on a souvent négligé d'envisager. Une fois les réponses-types établies, le correcteur n'a le droit à aucun jugement. Il n'en va pas de même pour le test de compréhension orale, pour lequel les élèves doivent remplir des blancs ou donner des réponses courtes, et par conséquent les problèmes d'orthographe entrent en compte. On ne peut donc là pas préétablir des réponses-types complètes avant que le test ne soit accompli. Afin d'assurer une notation uniforme, une réunion de concertation est organisée après l'exécution du test et l'on prend un échantillon d'une centaine de copies pour déterminer l'éventail de fautes acceptées. En principe, on n'enlève pas de « points pour des fautes de grammaire ou d'orthographe, à condition que la réponse donnée soit effectivement la bonne » (Hughes, 2003 : 170). Il est cependant nécessaire de faire preuve de clarté. Une fois les décisions requises prises, on procèdera à la notation à proprement parler selon les modèles de réponses établis.

Les parties d'expression écrite et d'expression orale de l'examen national sont notées de manière subjective, c'est-à-dire que des équipes de correcteurs sont entraînées à noter les copies des candidats ou leur performance à l'oral. Pour ce qui est de l'écrit, les correcteurs se sont basés généralement sur deux échelles de notation: l'une pour les lettres, et l'autre pour les rapports et essais. Le total maximum de points pour un test en particulier est de 20 points pour les deux exercices, le souci principal ayant toujours été, lors de l'élaboration des échelles de notation, de savoir à quoi attribuer les points dans cette compétence. L'échelle de notation pour les lettres est passée de l'attribution de points pour l'exécution de l'exercice, la mise en forme de la lettre et la qualité de langue (ceci en 1999), à celle de l'attribution de points pour l'exécution de l'exercice, le vocabulaire et registre de langue, la grammaire et l'orthographe (en 2001), à l'exécution de l'exercice, la mise en forme et qualité de langue (jusqu'à 2006), et finalement à celle de l'exécution de l'exercice et la qualité de langue (en 2007). Il est aussi intéressant de remarquer que, jusqu'à l'échelle de notation de 2007, les sous-compétences ont été appréciées de façons différentes. Donnons par exemple l'échelle de

1999, selon laquelle le candidat pouvait obtenir un maximum de 2 points pour l'exécution de l'exercice, mais un maximum de 3 points pour le vocabulaire et registre, ainsi que pour la grammaire et l'orthographe. Dans l'échelle de notation de la lettre en 2006, le candidat ne pouvait obtenir qu'un maximum de 2 points pour l'exécution de l'exercice et la mise en forme de la lettre, alors que les critères de langue étaient évalués sur une échelle de 0 à 4 points. Ce type de notation peut amener l'enseignement d'une classe à se concentrer sur la langue en elle-même (c'est-à-dire la grammaire et le vocabulaire), et à négliger d'autres facettes de l'écriture, tel que le contenu ou l'organisation du travail, ceci étant au désavantage de l'étudiant, s'il devait composer sur des thèmes où ces mêmes facettes étaient nécessaires à l'exécution de l'exercice. Tous les tests écrits sont notés par deux correcteurs, et dans le cas d'un désaccord notable de note, on fait appel à un troisième correcteur pour évaluer la copie.

La notation de l'oral a également subi quelques changements importants. Les difficultés de l'élaboration d'une échelle de notation sont identiques à celles rencontrées lors de l'établissement d'une échelle d'évaluation de l'écrit, à savoir quel critère privilégier pour la notation. Ici aussi, l'échelle est passée d'une échelle tenant compte de tous les critères de 1999, à un nombre de points variable selon les différents critères (en 2001), pour revenir à l'échelle initiale (en 2007). L'échelle de notation actuelle permet d'évaluer les compétences du candidat sur une base de quatre critères : communication, vocabulaire, grammaire, prononciation et élocution. Pour un exposé complet de l'échelle de l'oral de 2007, voir Alas 2007. La performance orale des candidats est évaluée par un examinateur indépendant pendant l'examen oral. L'examinateur ne prend pas part à l'entrevue, mais ne fait qu'évaluer la performance du candidat selon l'échelle de notation établie.

IV. Evaluation de l'examen et les problèmes

Tous les examens nationaux d'anglais sont validés ensuite par toute une série de données statistiques faisant état d'une analyse des items, qui contrôle (1) que les items permettent bien de faire la différence entre les individus selon leurs compétences (paramètre de discrimination), (2) le niveau de difficulté de chaque item (paramètre de difficulté), (3) et la probabilité qu'un individu de niveau médiocre puisse répondre correctement à l'item (paramètre de chance, ou de "pile ou face") (Bachman, 1990: 204). Bien que l'on puisse avoir accès aux données de l'examen national d'anglais depuis toutes ces années, nous ne traiterons pas de ce sujet dans cet article. Etant donné que la validation de l'examen contribue à poursuivre l'élaboration des tests, elle est d'une extrême importance car elle permet aux auteurs des items d'avoir un retour sur la qualité de leur travail et sur la fiabilité des résultats des tests.

Malgré les énormes progrès faits dans le domaine de l'évaluation de la compétence en langue étrangère, il reste des problèmes qui n'ont pas été résolus.

L'équipe nationale de développement des tests cherche toujours à employer de nouveaux auteurs d'items, qui pourraient considérablement apporter leur contribution à la création d'items plus variés et de meilleure qualité. Cela permettrait aussi d'accélérer le processus d'élaboration, et d'éviter des prises de décision de dernière minute.

Il y a aussi le problème de créer deux versions du même examen, de même type et de même fiabilité, pour l'examen national de chaque année, où la grande majorité des candidats choisit la première version, alors que la seconde version n'est que peu choisie (par exemple : en 2006, 9552 candidats ont choisi la version A, alors que 38 seulement ont pris la version B). Les ressources utilisées dans l'élaboration des deux versions, cependant, sont identiques, et semblent gaspillées avec la version B.

Le troisième problème concerne la confidentialité du test : il s'agit d'informer les professeurs et les élèves du déroulement de l'examen national en veillant à ne pas révéler les items en particuliers, les consignes et les questions. La procédure d'élaboration des tests actuelle, qui est fondée sur un grand nombre d'auteurs d'items supervisés par des chefs d'équipe par compétence, qui à leur tour laissent le travail à des consultants indépendants et à des spécialistes par matière, aussi responsables de la formation des professeurs quant aux techniques d'évaluation, à la mise en application des échelles de notation et des méthodes d'évaluation, garantit, on peut l'espérer, d'un côté des tests qui restent confidentiels (donc valides et fiables), ainsi qu'une anxiété réduite de la part des professeurs concernés par l'examen national, quoiqu'il faille encore y veiller.

V. Conclusion

L'examen national d'anglais est bien établi désormais. C'est l'examen le plus communément présenté, un examen de compétence d'une langue étrangère au niveau national estonien qui est comparable à tout autre examen national de langue étrangère dans les autres pays d'Europe. Les auteurs des examens suivent les directives de l'Union Européenne, telles qu'elles sont décrites par le CECR (Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer). L'Estonie est devenue membre d'organisations internationales d'examens, comme l'ALTE (*Association of Language Testers in Europe*), l'EALTA (*European Association for Language Testing and Assessment*), etc., en élaborant ses propres tests selon les pratiques de ces organisations.

Nous savons comment élaborer des tests valides et fiables, et comment les gérer professionnellement. La création des tests suit des directives, une mise en pratique de caractéristiques particulières dans les tests, un modèle de rédaction des items, une période d'essai, une gestion des examens, et une analyse statistique comme il est spécifié au niveau international. L'élaboration de tests a eu un effet de retour sur les méthodes d'enseignement des langues dans les écoles où les professeurs sont entraînés à enseigner et à tester une compétence particulière en employant les meilleurs moyens, à choisir un manuel et à le doubler d'un apport d'éléments extérieurs afin que l'élève en bénéficie pleinement. Tous les tests déjà effectués sont sur dossiers et à la

disposition des élèves et des professeurs pour leur information sur la page du site des examens et des diplômes nationaux.

Les éducateurs sont beaucoup plus conscients de l'importance de la façon d'évaluer les élèves. L'évaluation fait maintenant partie du programme de formation des enseignants à l'Université de Tallinn. On entraîne systématiquement les correcteurs à prendre des décisions expertes sur les performances orales ou écrites des étudiants. On organise des ateliers et des conférences semestrielles afin de familiariser les professeurs avec les problèmes d'élaboration des examens nationaux et leur faire part des expériences passées. S'ils ont appris des techniques européennes d'évaluation des langues, les experts en évaluation estoniens partagent maintenant leur savoir-faire au niveau local lors de conférences internationales.

Bibliographie

- Alas, E. 2007. «Developing the national examination in the English language», in *Open!*, n° 32, pp. 2-4.
- Bachman, L. 1990. *Fundamental Considerations in Language Testing*, Oxford University Press.
- Fulcher, G., Davidson F. 2007. *Language Testing and Assessment*, London, New York : Routledge Applied Linguistics, 403 p.
- Hughes, A. 2003. *Testing for Language Teachers*, Cambridge University Press, 172 p.
- Jöul, M., Lätt, V., Mere, K., Sass, E., Türk, Ü., Vilu, M. 2005. *Year 12 Handbook*, Tallinn, 160 p.
- Liiv, S. 2002. « Foreign Language Competence and Testing » in *Perspectives on English and American Language and Literature*, Tallinn, pp.51-59.