
L'évaluation du niveau des compétences linguistiques à l'université : test standardisé ou test-maison ?

Réflexions sur le choix de l'université Tamkang de Taipei

David COURRON

Abstract

Confronted with the issues raised by the diversity and complexity of learner profiles, institutional promotion of multilingualism and industrialization of certifications under the influence of NICT, assessment in language didactics —especially within universities— is becoming a central focus among foreign language teaching scholars. The gap between curriculum contents and social purposes of language teaching, which the Common European Framework of Reference for Languages (CEFR) and its Reference Level Descriptions are expected to fill in, unrelentingly leads universities to offer tests and examinations able to meet students linguistic abilities via objective and reliable tests that guarantee fairness and a real comparability with one another. This concern has brought Taipei Tamkang University to introduce in 2002 an assessment examination of every coming graduate students' linguistic abilities. The test result is reproduced on the university official diploma in the form of a distinction. However choosing a home-made test conceived and realized on its own instead of a standardized test may lead to confusion in the actual aims of such an assessment. Either it aims at measuring the acquisition of points taught within the framework of a precise training program (e.g., subjective approach like regular term examinations) or at comparing the level acquired in various linguistic competences according to intangible criteria (e.g., objective approach like standardized tests such as French TCF and multilingual computer-based DIALANG). Beyond students' linguistic abilities, such an assessment unquestionably amounts to assessing the accuracy of teaching choices as well as quality and relevance of teaching in general, thus potentially placing the entire university in the hot seat. This is the reason why means of assessment whose psychometric qualities allow

comparisons between legitimately comparable performances should be given priority by universities.

Confrontée aux enjeux que représentent la variété et la complexité des profils d'apprenant, la valorisation institutionnelle du plurilinguisme et l'industrialisation des certifications sous l'effet de l'irruption des technologies de l'information et de la communication, la question de l'évaluation dans la didactique des langues s'impose avec une acuité nouvelle à la réflexion des praticiens de l'enseignement des langues étrangères en tête desquels se tiennent les établissements universitaires. Le renforcement de l'autonomisation de l'apprenant dans son acte d'apprentissage entreprend parallèlement de modifier son appréhension de l'évaluation qui « [...], rendue nécessaire, emprunte sa systématisation aux méthodes de mesure en psychométrie et en psychologie différentielle et constitue en se constituant comme domaine de connaissance, une ingénierie au service de fonctions sociales, aux premiers rangs desquelles, l'éducation, le recrutement et les formations au sein desquelles le marché des compétences linguistiques tient une place grandissante »¹.

Dans un tel contexte de décalage entre les cursus scolaires et les finalités sociales de l'apprentissage des langues, que le Cadre européen commun de référence pour les langues et ses référentiels² s'efforcent de réduire, les universités, qu'elles soient publiques ou privées, s'acheminent inexorablement vers une offre de tests et d'examens non plus seulement diplômants mais aussi et surtout valorisants, au sens où ils attribuent une valeur objective au niveau de compétence linguistique atteint par l'apprenant qui soit quantifiable pour être comparable. C'est le cas de l'université Tamkang³ à Taipei qui, à la demande de sa direction, a

¹ CHARDENET Patrick, « Les enjeux actuels de l'évaluation en didactique des langues », in *Le français à l'université*, Paris : Agence Universitaire de la Francophonie (AUF), 1er trimestre 2006, n°1, 11ème année, pp. 2-3 [en ligne]. Disponible sur : <<http://www.bulletin.auf.org/spip.php?rubrique6>> (consulté le 27.03.2007).

² Pour de plus amples précisions sur le CECR, se reporter à l'ouvrage préparé par la Division des Politiques Linguistiques du Conseil de l'Europe, *Cadre européen commun de référence pour les langues*, Paris : Didier, 2005, 192 p. ou consulter le site internet du Conseil de l'Europe <<http://www.coe.int/t/dg4/Linguistic/>> [en ligne] (consulté le 26.03.2007).

³ Établi en 1950 par le Dr. Clément C.-P. Chang, le collège professionnel d'anglais Tamkang n'a accédé au statut d'université qu'en 1958 avec l'attribution de diplômes de licence. Université à part entière en 1980, Tamkang est devenue une université généraliste qui se compose de 12

introduit, dès 2002, un examen d'évaluation des compétences linguistiques de ses étudiants en fin de 4^{ème} année. Les résultats de cet examen figurent sur le diplôme officiel de l'université Tamkang sous la forme d'une mention « assez bien », « bien » ou « très bien », l'absence d'une telle mention équivalant, sans toutefois prêter à conséquence sur la délivrance effective du diplôme, à un aveu de faiblesse de l'étudiant en langue étrangère.

Au-delà de la remise en question des qualités d'objectivité de l'évaluation autorisée par un tel examen, la décision de l'université Tamkang d'opter pour un test d'évaluation conçu et réalisé « avec les moyens du bord » selon les termes employés par les concepteurs de sa version française⁴ plutôt que pour un test standardisé comme le *Test de connaissance du français* (TCF) nous paraît trahir une confusion dans les objectifs réellement poursuivis par cet acte d'évaluation. En effet, soit celui-ci ambitionne de mesurer le suivi de la bonne acquisition de points enseignés dans le cadre d'une pédagogie particulière et la régularité de la progression dans le cadre d'un programme précis d'apprentissage (évaluation sommative, normative ou du savoir⁵), alors il ne peut prétendre qu'à une approche

facultés comprenant 53 départements et propose 17 programmes de maîtrise et 17 formations doctorales. Les membres du corps professoral et les personnels administratifs de Tamkang au nombre de 2 124 personnes encadrent plus de 27 000 étudiants. A présent, l'université Tamkang dispose de quatre campus situés à Tamsui, à Taipei, à Yi-lan (nouveau site de Lanyang) et... sur la Toile. En effet, le site internet de l'université permet non seulement de mettre à disposition de nombreux cours (éventuellement échangés avec ceux d'autres universités-partenaires) mais aussi de mettre au point des enseignements communs avec des universités étrangères. Ce réseau fait appel à des sommités internationales qui enseignent ainsi à distance [en ligne]. Disponible sur : <<http://foreign.tku.edu.tw/french/FR-Frameset.asp?topframe=up-1.html&mainframe=01-welcome-content01-p4.htm>> (consulté le 09.03.2007).

⁴ Les concepteurs du test d'évaluation de l'université Tamkang, dont M. Alain Monier, ont été rencontrés lors du séminaire de réflexion pédagogique consacré à l'évaluation et organisé, le 27 mai 2006, par le département de français de l'université catholique Fuijen de Taipei, auquel l'auteur a été convié en compagnie de Mme Nobuko Kurata, chef du département d'études françaises de l'université Nanzan. Les échanges nourris au cours de cette rencontre ont largement contribué à stimuler les réflexions proposées ici.

⁵ L'évaluation sommative « contrôle les acquis à la fin du cours et leur attribue une note ou un rang. Il ne s'agit pas forcément d'une évaluation de la compétence, [car elle] est souvent normative, ponctuelle et teste le savoir ». L'évaluation normative « classe les apprenants les uns par rapport aux autres » dans la classe ou dans le cadre d'une population donnée (les candidats

subjective à l'image des examens universitaires de fin de session dont l'horizon se borne à la classe. Soit, au contraire, il ambitionne non pas de mesurer mais plutôt de comparer le niveau acquis dans différentes compétences langagières, en vertu de critères intangibles et par là applicables par d'autres universités (évaluation formative, critériée ou de la capacité⁶), alors il doit prétendre à une approche objective à l'image des tests standardisés dont l'horizon s'étend en théorie à l'ensemble des apprenants de français du monde entier. Deux instruments d'évaluation proposent de résoudre cette tension entre subjectivité et objectivité, approximation et fiabilité, particularité et comparabilité par trop à l'oeuvre dans le test-maison de l'université Tamkang (I), d'une part, avec les garanties psychométriques de tests comme le TCF (II), et d'autre part, avec le test d'évaluation en ligne DIALANG qui allie objectivité et souplesse d'utilisation (III).

I–L'examen d'évaluation de 4ème année du département de français de l'université Tamkang : objectifs et analyse critique d'un test-maison

L'examen d'évaluation des compétences linguistiques des étudiants en fin de 4ème année du département de français de l'université Tamkang correspond à

d'un test). L'évaluation du savoir (ou du niveau) est celle de « l'atteinte d'objectifs spécifiques - elle porte sur ce qui a été enseigné - par voie de conséquence, elle est en relation au travail de la semaine ou du mois, au manuel, au programme. L'évaluation du savoir est centrée sur le cours. Elle correspond à une vue de l'intérieur ». *CECR...*, *op.cit.*, pp. 139–145.

⁶ L'évaluation formative est « un processus continu qui permet de recueillir des informations sur les points forts et les points faibles. L'enseignant peut alors les utiliser pour l'organisation de son cours et les renvoyer aussi aux apprenants. On utilise souvent l'évaluation formative au sens large afin d'y inclure l'information non quantifiable fournie par des interrogations et des entretiens. [...] La force de l'évaluation formative est de se donner pour but l'amélioration de l'apprentissage. » L'apprenant doit néanmoins être en mesure de tenir compte, de recevoir, d'interpréter et de s'approprier une telle information rétroactive. L'évaluation critériée « se veut une réaction contre la référence à la norme : on y évalue l'apprenant uniquement en fonction de sa capacité propre dans le domaine et quelle que soit celle de ses pairs ». L'évaluation de la capacité (ou mise en oeuvre de la compétence ou performance) est « l'évaluation de ce que l'on peut faire ou de ce que l'on sait en rapport avec son application au monde réel. Elle correspond à une vue de l'extérieur ». *Ibid.*

une commande passée auprès de tous ses départements de langues étrangères par la direction de cet établissement d'enseignement supérieur privé taïwanais de très grande renommée⁷. Cet examen, qui concerne donc tous les étudiants de la faculté des langues étrangères⁸, par l'octroi d'une mention sur le diplôme décerné par l'université à la fin des études semble viser outre le renforcement de la motivation des étudiants et l'élévation de leur niveau de compétences linguistiques, objectifs incontestablement louables, la confirmation de la reconnaissance sociale de l'image de marque de l'établissement. Et peut ainsi être assimilé à une pure mesure de prestige tant qu'aucune extension de cet acte d'évaluation dans les mêmes conditions de réalisation à d'autres établissements universitaires ne permettra la moindre comparaison entre les résultats⁹.

Le test conçu par le corps professoral du département de français de l'université Tamkang a pour objet l'évaluation des compétences d'expression orale (sur 50 points), de compréhension orale (sur 50 points) et de compréhension écrite (sur 100 points), autour desquelles sont élaborées les trois parties¹⁰ qui le composent, à l'exception de celles d'expression écrite jugée trop complexe à mettre en oeuvre par crainte notamment de critères de correction subjectifs de nature à fausser l'évaluation. Ses concepteurs justifient toutefois cette exception au motif, d'une part, que les trois compétences soumises à évaluation permettraient d'obtenir une image suffisamment précise du niveau de compétence atteint en français par les étudiants, et d'autre part, que la difficulté d'élaborer des outils d'évaluation satisfaisants sur le fond (exercice proposé, critères objectifs d'évaluation) comme sur la forme (temps de correction) rendrait la tâche impossible à une petite

⁷ Pour illustrer le degré de renommée dont il est ici question, nous a été proposée l'image du « Berkeley de Taïwan ».

⁸ Environ 120 étudiants chaque année pour le seul département de français.

⁹ On notera qu'un seul établissement universitaire taïwanais, le Wenzao Ursuline College of Languages à Kaohsiung, a été déclaré par le Centre international d'études pédagogiques (CIEP) centre agréé pour le TCF dont il met en place deux sessions annuelles à l'attention de près de 300 candidats. Informations reçues à l'occasion d'un entretien avec M. Bruno Mège, chargé de programme au pôle certifications et évaluation du CIEP, rencontré aux 21ème Rencontres pédagogiques du Kansai organisées les 30 et 31 mars 2007 à Osaka par le Centre franco-japonais/Alliance française d'Osaka.

¹⁰ Voir en annexe des exemples de questions-réponses pour chacune des activités proposées dans les trois parties du test.

structure tel le département de français à Tamkang¹¹. Cette dernière réserve met ainsi en lumière l'obstacle principal à la conception d'un test-maison rigoureux : le respect de l'exigence d'objectivité et de son corollaire la garantie de comparabilité.

La compétence d'expression orale est évaluée en laboratoire de langue¹² à travers quatre épreuves proposées selon une progression de nature à ménager une gradation dans la difficulté tout en privilégiant la rapidité de réaction face à des énoncés simples¹³. La première consiste ainsi en la lecture d'un texte court¹⁴ assortie de réponses à 2 ou 3 questions relatives à son contenu, afin d'évaluer la fluidité et la compréhension de la lecture à travers la qualité de la prononciation, des liaisons, de l'intonation et de l'expression. La seconde tente de cerner les connaissances syntaxiques et lexicales en appréciant la rapidité de passage du mandarin au français au moyen de la traduction de 5 phrases. La troisième, organisée autour de dix questions, permet d'appréhender les capacités d'expression dans des domaines divers d'expression. La quatrième enfin juge de la richesse d'expression syntaxique et lexicale en situation de description d'un document iconographique, en général une photographie, guidée par quelques questions. Ce sont ainsi respectivement 10, 15, 20 et 5 points qui sont attribués à ces quatre épreuves, les seules que les enseignants du département de français soient contraints de corriger eux-mêmes.

¹¹ L'équipe enseignante du département de français de l'université Tamkang se compose de 29 professeurs dont 21 à temps complet et 8 à temps partiel. Parmi eux, 21 sont titulaires d'un doctorat. Le département compte 8 classes pour les quatre années universitaires à raison de 2 classes par niveau [en ligne]. Disponible sur : <<http://www2.tku.edu.tw/~tffx/verfrench/>> (consulté le 05.03.2007).

¹² L'ensemble des épreuves de compréhension orale nécessite entre 15 et 20 minutes en laboratoire de langue. Les réponses de chaque candidat peuvent alors y être enregistrées sur des cassettes audio remises ensuite aux différents correcteurs. Le recours aux équipements du laboratoire de langue relève d'un choix unanime contre un examen face à un jury de professeurs jugé trop « traumatisant ». Deux raisons ont justifié ce refus : l'énorme investissement en temps exigé des examinateurs et le facteur émotionnel avec un risque de distorsion des résultats entre les étudiants. En outre, ce mode de passation garantit une maîtrise égale du temps et l'absence d'élément extérieur favorisant ou perturbant le déroulement de l'examen.

¹³ Le vocabulaire, les structures grammaticales et les situations de communication employés dans le test ont été étudiés par le candidat durant les quatre années de son cursus à l'université.

¹⁴ Dans sa première version, la lecture portait sur des phrases hors contexte. Or, depuis 2006, les concepteurs du test privilégient les petits récits cohérents afin de faciliter le processus de lecture-compréhension.

Depuis l'édition 2005, à la suite des recommandations formulées par un comité d'évaluation de l'évaluation, un enseignant se voit désormais confier la charge de corriger l'ensemble d'une même épreuve, afin de supprimer tout risque d'écart de notation entre plusieurs correcteurs pour une même série d'items¹⁵.

L'évaluation de la compétence de compréhension orale pose aux concepteurs du test la question de l'extrême attention à porter à la création des items, afin de faire en sorte d'éviter l'écueil que représentent plusieurs réponses acceptables pour une même question. Ceci les a conduits à présenter les épreuves de cette partie du test sous la forme de questionnaires à choix multiples (QCM) avec l'avantage de pouvoir être traités informatiquement (par *mark sheets*) et ainsi de ne pas alourdir la tâche des enseignants chargés de la correction. A l'instar des épreuves de compétence d'expression orale, quatre épreuves sont proposées. La première, sur 15 points, vise à tester la connaissance des situations de communication et consiste en une série de 15 questions¹⁶ pour chacune desquelles quatre réponses sont suggérées au candidat. La seconde, sur 15 points également, cherche à évaluer ses connaissances syntaxiques et lexicales en lui proposant de sélectionner le meilleur équivalent d'une phrase simple parmi quatre choix. Les troisième et quatrième épreuves, dotées l'une et l'autre de 10 points, proposent de le faire réagir en contexte, à travers la compréhension, d'une part, de cinq textes courts (de 4 à 6 lignes) assortis d'une question pour chaque texte, et d'autre part, de cinq petits dialogues également assortis d'une question pour chaque dialogue. L'objectif recherché par les concepteurs du test est d'éprouver, dans un cas, la capacité à comprendre une suite d'énoncés et à en faire la synthèse, et dans un autre, la capacité à comprendre une situation de communication et à en extraire un élément important.

La troisième et dernière partie de l'examen d'évaluation des compétences

¹⁵ Lors des deux premières éditions de ce test, en 2003 et 2004, les correcteurs constitués en binôme recevaient un stock d'une trentaine de cassettes audio et évaluaient l'ensemble des épreuves d'expression orale, ce qui entraînait inévitablement, d'un binôme de correcteurs à l'autre, des problèmes d'objectivité dans l'appréciation et la notation des réponses.

¹⁶ Selon les concepteurs de ce test, le nombre important d'items (une quinzaine par épreuve) dans les parties relatives à la compréhension orale et écrite serait une garantie d'objectivité, dans la mesure notamment où, dans la masse des items, les étudiants courraient un moindre risque d'être pénalisés par une lacune en vocabulaire.

linguistiques des étudiants en fin de 4^{ème} année du département de français de l'université Tamkang, consacrée à la compréhension écrite et organisée autour d'une série de trois épreuves, a été introduite ultérieurement à la demande des étudiants qui se plaignaient de n'être uniquement évalués que sur leur point faible, l'oral. La première épreuve exige du candidat qu'il retrouve, parmi quatre propositions, l'élément manquant d'un énoncé. 15 phrases indépendantes les unes des autres lui sont ainsi proposées, afin d'apprécier ses connaissances grammaticales et lexicales. La seconde épreuve, formellement dérivée de la première, consiste à sélectionner un élément (verbe, nom, adverbe, adjectif, pronom, préposition etc.) qui pourrait se substituer à un élément de nature équivalente souligné dans un énoncé sans que le sens global de celui-ci n'en soit aucunement altéré. Dans ce cas, les 15 énoncés, ici aussi indépendants les uns des autres, doivent permettre d'estimer l'étendue (la variété ou la richesse pourrions écrire) des connaissances grammaticales et lexicales du candidat. La troisième épreuve vise, quant à elle, à évaluer la compréhension globale d'un texte grâce à une série de quatre questions. Cinq textes, de type article de presse, souvent argumentatifs composent cette dernière partie du test. Sur un total de 100 points, 30, 30 et 40 points sont respectivement attribués à ces épreuves de compréhension écrite.

Réponse à la demande formulée par la direction de l'université Tamkang de procéder à une évaluation des compétences linguistiques au seuil de l'obtention du diplôme, la décision du département de français d'opter pour un test conçu et réalisé en interne aura indubitablement contribué à sensibiliser l'équipe des professeurs aux questions que soulève l'évaluation, en particulier dans la phase d'élaboration d'un outil de connaissance des compétences des étudiants aussi discriminant et objectif que possible. Dans ce contexte, une des difficultés majeures rencontrées tient à la composition des épreuves, à la rédaction des consignes et à la confection de chacun des items du test. A titre d'exemple, afin d'éviter deux choix possibles dans les propositions de réponses ou l'emploi d'un vocabulaire trop technique ou spécialisé au détriment de celui de situations de communication ordinaires, les concepteurs ont évoqué le recours à l'élaboration d'une grille d'éléments grammaticaux et lexicaux à évaluer couplé à une réflexion sur le nombre précis d'items à placer dans le test. Parmi d'autres variables envisagées pour mieux objectiver le contenu du test, on soulignera, d'un côté, le

nombre de points accordés à chaque épreuve qui permet de moduler l'importance attribuée à chacune des compétences, et d'un autre côté, la mention de toutes les réponses proposées sur la feuille de réponses qui permet de jouer sur le degré de difficulté. Ces quelques préoccupations des auteurs du test du département de français de l'université Tamkang pointent ainsi à la fois la nécessité croissante de la mise en place d'une évaluation objective des compétences linguistiques à l'université et les limites d'un test-maison au regard des moyens qu'induit l'exigence d'une objectivité vérifiable, facteur essentiel de garantie de la comparabilité de l'évaluation menée. En s'appuyant davantage sur des actes de communication usuels qui exigent une assimilation complète des éléments de la langue, ce test-maison emprunte indiscutablement à la philosophie du TCF et au-delà s'inscrit dans la perspective actionnelle prônée par le *Cadre européen commun de référence pour les langues* (CECR)¹⁷. En revanche, et contrairement à ce qu'avancent ses concepteurs qui assurent s'être inspirés des deux types d'examen, il reste infiniment plus proche dans son esprit et ses partis-pris du Foreign Language Proficiency Test taïwanais que du TCF dont les très contraignants protocoles de conception des items intègrent des analyses psychométriques qui tendent, en réduisant le plus possible le poids de la subjectivité, à en augmenter la validité et la fiabilité dans le but de garantir la comparabilité des résultats du test à travers le monde quel que soit le candidat évalué.

¹⁷ La perspective actionnelle privilégiée par le CECR pour l'apprentissage, l'enseignement et l'évaluation des langues vivantes « [...] considère avant tout l'usager et l'apprenant d'une langue comme des acteurs sociaux ayant à accomplir des tâches (qui ne sont pas seulement langagières) dans des circonstances et un environnement donnés, à l'intérieur d'un domaine d'action particulier. Si les actes de parole se réalisent dans des activités langagières, celles-ci s'inscrivent elles-mêmes à l'intérieur d'actions en contexte social qui seules leur donnent leur pleine signification. Il y a "tâche" dans la mesure où l'action est le fait d'un (ou de plusieurs) sujet(s) qui y mobilise(nt) stratégiquement les compétences dont il(s) dispose(nt) en vue de parvenir à un résultat déterminé. La perspective actionnelle prend donc aussi en compte les ressources cognitives, affectives, volitives et l'ensemble des capacités que possède et met en oeuvre l'acteur social ». *CECR...*, *op.cit.*, p. 15.

II—Exigences et conditions d’objectivité dans l’évaluation des compétences linguistiques à l’université : les qualités psychométriques des tests standardisés comme garantie

Le *Foreign Language Proficiency Test* (FLPT)¹⁸ taïwanais est un test d’aptitude générale en langues étrangères, développé et administré par le Language Training and Testing Center (LTTC)¹⁹. Dispensé depuis 1965, à un rythme mensuel, en anglais, japonais, français, allemand et espagnol, il s’adresse aux candidats de niveau universitaire âgés d’au moins 18 ans. Le FLPT, qui se présente sous la forme d’un questionnaire à choix multiples (QCM), complété par une épreuve enregistrée d’expression orale à réponses libres de 15 minutes, comprend une épreuve dans chacun des domaines suivants : compréhension orale, structures grammaticales, vocabulaire et lecture et expression orale. De toute évidence, une description même succincte de ces épreuves suffit à désigner la source d’inspiration, tant dans

¹⁸ Le Foreign Language Proficiency Test (FLPT) est actuellement utilisé par plus de 400 établissements taïwanais, publics et privés. Il est déterminant pour la promotion et la sélection des candidats taïwanais aspirant à un recrutement professionnel ou universitaire à l’étranger. De nombreuses institutions taïwanaises telles que le Ministère de l’Éducation, le Bureau du Tourisme et l’Administration Centrale de la Fonction Publique y recourent régulièrement.

¹⁹ Ce centre de formation et d’évaluation linguistique a été établi en 1951 pour offrir la formation intensive en anglais dont avaient besoin les fonctionnaires que le gouvernement taïwanais s’app préparait à envoyer aux États-Unis dans le cadre des programmes d’assistance technique en oeuvre à ce moment-là. A partir de 1965, le LTTC a commencé à offrir des cours de japonais, de français, d’allemand et d’espagnol et à ouvrir des classes aux autres personnels gouvernementaux, ainsi qu’aux employés des organismes et entreprises privés et au grand public. C’est également à cette époque que le LTTC s’est lancé dans l’organisation de tests de compétence en langues étrangères. Doté du statut de fondation culturelle et éducative, octroyé en 1986 par le Ministère de l’Éducation, avec pour mission la satisfaction des besoins en développement économique et social par le biais de travaux sur la recherche et le développement en matière de formation linguistique et d’évaluation des compétences linguistiques, le LTTC est placé sous la surveillance d’un conseil d’administration, avec à sa tête le président de l’Université nationale de Taïwan, comprenant des représentants de l’Université nationale de Taïwan, du Conseil pour la planification économique et le développement, des Ministères de l’Éducation, des Affaires Économiques et des Affaires Étrangères et de la Fondation pour les échanges universitaires. Le LTTC dispose d’un site internet accessible à l’adresse <http://www.ltcc.ntu.edu.tw/E_LTTC/aboutthelttc.htm> (consulté le 05.03.2007).

L'esprit que dans la facture, de celles de l'examen d'évaluation des compétences linguistiques des étudiants en fin de 4^{ème} année du département de français de l'université Tamkang. L'épreuve de compréhension orale de français est composée de trois parties dans lesquelles chacune des questions non-écrites n'est posée qu'une seule fois. Dans la première partie (15 questions), le candidat doit, après avoir écouté chaque question, choisir la seule bonne réponse parmi les quatre proposées²⁰; dans la seconde (15 questions), il doit, après avoir écouté chaque courte phrase, lire les quatre correspondantes et trouver celle qui en donne la bonne explication²¹; et dans la troisième (10 dialogues et 10 histoires courtes), il doit, après avoir écouté chaque question, lire les quatre phrases correspondantes et trouver celle qui en donne la bonne explication²². L'épreuve de structures du français, organisée autour de 80 questions, propose quatre possibilités au candidat qui doit choisir celle qui est grammaticalement correcte²³. Quant à l'épreuve de vocabulaire et de lecture du français, qui rassemble 60 questions, elle se divise en une partie vocabulaire et expressions françaises et une partie compréhension écrite. Dans la première, les énoncés des questions 1 à 20 contiennent un blanc qu'il convient de remplir à l'aide de l'une des quatre réponses proposées²⁴ et ceux des questions 21 à 40 contiennent un élément souligné qu'il convient de remplacer par l'une des quatre expressions proposées²⁵. Dans la seconde partie, 20 questions à choix multiple permettent de tester les capacités en compréhension écrite du candidat. Enfin, l'épreuve orale de français est élaborée autour de cinq parties dans lesquelles il est demandé au candidat de lire des phrases en 45 secondes,

²⁰ A titre d'exemple : Où est ton frère ? (A- Il a trente ans/ B- Il est médecin/ C- Il est en France/ D- Il est très jeune). Voir les exemples d'épreuves du FLPT sur le site du LTTC [en ligne]. Disponible sur : <<http://www.lttc.ntu.edu.tw/flpt/FLPTf.htm>> (consulté le 05.03.2007).

²¹ A titre d'exemple : Jean a passé son enfance à Paris. (A- Il travaille à Paris/ B- Il a grandi à Paris/ C- Il habite à Paris/ D- Il a passé son examen à Paris). *Ibid.*

²² A titre d'exemple : (texte entendu) Après le printemps, nous avons l'été. Les enfants aiment bien cette saison car l'été est la saison des grandes vacances. La famille va passer quelques semaines à la campagne ou à la plage. (question posée) Quelle saison aiment les enfants ? (A- Le printemps/ B- L'été/ C- L'automne/ D- L'hiver). *Ibid.*

²³ A titre d'exemple : Qui...ton professeur ? (A- es/ B- est/ C- ai/ D- êtes). *Ibid.*

²⁴ A titre d'exemple : J'ai garé ma...au parking. (A- maison/ B- voiture/ C- chaise/ D- table). *Ibid.*

²⁵ A titre d'exemple : Jean voit tout en noir. (A- ne voit pas clairement/ B- vit dans un endroit sombre/ C- n'aime pas la lumière/ D- est pessimiste). *Ibid.*

de traduire un ensemble de phrases courtes en une minute, de répondre à des questions simples (15 secondes pour chacune des questions 1 à 5 puis 30 secondes pour chacune des questions 6 à 10), d'argumenter un point de vue en 90 secondes après un temps de réflexion d'une durée identique et de décrire une image dans les mêmes conditions de temps que précédemment. Cette dernière épreuve correspond d'ailleurs en tout point, à l'exception de la consigne d'argumentation d'un point de vue, à l'épreuve d'expression orale du test d'évaluation de l'université Tamkang, tandis que les deux épreuves de compréhension orale sont identiques jusque dans le nombre de questions posées dans leurs trois parties, et alors que l'épreuve de compréhension écrite de Tamkang s'inspire à la fois de celles de structures du français et de vocabulaire et de lecture du français du FLPT, respectivement pour l'évaluation des connaissances grammaticales et lexicales et pour l'évaluation de l'étendue de ces mêmes connaissances et de la compréhension globale d'un texte.

Toutefois, au-delà des références et des modèles qui sous-tendent et orientent la réflexion des concepteurs du test d'évaluation de l'université Tamkang, demeure l'hypothèque du poids de la subjectivité et par conséquent une interrogation sur la validité et la fiabilité de cette évaluation. Associés à la faisabilité²⁶, ces deux concepts sont à la base de tout processus d'évaluation. La validité de la procédure d'un test ou d'une évaluation est admise dès lors qu'il peut être « démontrer que ce qui est effectivement testé (le construct) est ce qui, dans le contexte donné, doit être évalué et que l'information recueillie donne une image exacte de la compétence des candidats en question »²⁷. Quant à la fiabilité, il s'agit de « la mesure selon laquelle on retrouvera le même classement des candidats dans deux passations (réelles ou simulées) des mêmes épreuves »²⁸. Trois exigences doivent alors présider au processus d'évaluation : premièrement, déterminer ce qui est évalué (spécification du contenu du test) ; deuxièmement, déterminer comment est interprétée la performance (formulation de critères déterminant

²⁶ La faisabilité d'une évaluation est un aspect essentiel dans la mesure où « [l]es examinateurs ne disposent que d'un temps limité. Ils ne voient qu'un échantillon limité de performances et il y a des limites au nombre et à la nature des catégories qu'ils peuvent manipuler comme critères. » *CECR...*, *op.cit.*, p. 136.

²⁷ *Ibid.*, p. 136.

²⁸ *Ibid.*

la réalisation d'un objectif d'apprentissage) ; et enfin, déterminer comment comparer (description des niveaux de compétences dans d'autres tests rendant possible la comparaison entre des systèmes différents de qualification). Afin de tendre vers une meilleure objectivité de leur test d'évaluation, les concepteurs de l'université Tamkang gagneraient incontestablement à essayer de le mettre en relation avec le CECR ainsi que le préconise le Conseil de l'Europe dans son avant-projet expérimental du Manuel « Relier les examens de langue au Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer (CECR) »²⁹. Ce document se présente tel un guide à l'attention des concepteurs de tests et d'examens soucieux de « mettre en oeuvre des démarches transparentes et concrètes pour situer leurs examens par rapport au CECR, à les appliquer et à en rendre compte dans un processus cumulatif de perfectionnement continu »³⁰. L'approche de ce manuel suggère de « décrire ce que recouvre l'examen ainsi que les démarches de passation et d'analyse, [de] mettre en relation les résultats de l'examen et les niveaux communs de référence présentés au chapitre 3 du CECR [et de] rendre compte des procédures suivies pour ce faire en apportant des preuves à l'appui »³¹, dans la mesure où la mise en relation d'un test ou d'un examen avec le CECR revient à mettre en oeuvre un processus complexe qui « relève d'une affirmation pour laquelle le producteur de l'examen doit apporter des preuves tant

29 Conseil de l'Europe, « Relier les examens de langue au Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer (CECR) », Strasbourg : Division des Politiques linguistiques, novembre 2003, DGIV/EDU/LANG (2003)10, 16 p. [en ligne]. Disponible sur : <<http://www.coe.int>> (consulté le 26.03.2007). Ce travail a été engagé au cours d'un séminaire organisé à Helsinki en juillet 2002 (DG IV/EDU/LANG-15-2002) afin d'apporter une réponse à la question complexe que se posent les organismes de certification et les centres d'examen : « Comment puis-je savoir que mon Niveau B1 est votre Niveau B1 ? » Il s'inscrit dans la continuation des travaux de la Division des Politiques linguistiques du Conseil de l'Europe notamment la publication du « Threshold level » (niveau seuil) dans les années 1970 et la conception du CECR dans les années 1990 : élaborer des outils de planification présentant des éléments de référence et des objectifs communs comme fondements d'une structure cohérente et transparente pour un enseignement/apprentissage efficace et une évaluation pertinente prenant en compte les besoins des apprenants et de la société afin de favoriser la mobilité des individus.

³⁰ *Ibid.*, p. 2.

³¹ *Ibid.*

au plan théorique qu'empirique »³². Le manuel présente quatre séries de démarches (la familiarisation, la spécification, la standardisation et la validation empirique) étroitement liées entre elles qui entrent chacune dans le processus de validation. La familiarisation permet de s'assurer que ceux qui participent à la mise en relation sont suffisamment initiés à l'approche prônée par le CECR. La spécification constitue un inventaire de ce que l'examen recouvre (contenu et types de tâches) par rapport aux catégories présentées dans le CECR. La standardisation englobe un ensemble de démarches « qui garantissent que les jugements émis au moment de l'évaluation des performances reflètent les concepts hypothétiques décrits dans le CECR et que les décisions relatives à la difficulté des tâches et des activités sont régies par des principes fondés sur les résultats de pré-tests ainsi que sur des avis d'experts »³³. Enfin, la validation empirique se définit comme une analyse de données de tests permettant d'apporter des preuves que l'examen lui-même et son lien au CECR sont valables. Ces deux dernières démarches « garantissent que l'on peut effectivement confirmer les hypothèses faites lors de la spécification et de la standardisation (« épreuves en cours de production ») par la passation réelle de l'examen (« épreuves mises en œuvre ») et par la mise à disposition des données sur le comportement des candidats appartenant à la population cible lors de la passation »³⁴. Toutefois, une validation empirique correcte exige un important savoir-faire en psychométrie³⁵ dont ne peuvent disposer de petites équipes de concepteurs comme celle du test d'évaluation de l'université Tamkang qui, pour l'heure, travaillent encore de manière isolée, ou toute structure universitaire désireuse de procéder à une évaluation objective des compétences linguistiques

³² *Ibid.*, p. 3.

³³ *Ibid.*, p. 4.

³⁴ *Ibid.*

³⁵ La psychométrie se définit comme la science qui étudie l'ensemble des techniques de mesures pratiquées en psychologie, ainsi que les techniques de validation de ces mesures. Ces techniques concernent tous les champs de la psychologie, ainsi que d'autres domaines de sciences connexes (comme la recherche en comportement du consommateur par exemple) mais elles ont été surtout développées à l'origine pour la mesure des performances intellectuelles (âge mental ou quotient intellectuel, quotients de développement dans la petite enfance) ou bien pour l'analyse des composantes de la personnalité (affectivité, émotions, relations avec les autres) [en ligne]. Disponible sur : <<http://fr.wikipedia.org/wiki/Psychom%C3%A9trie>> (consulté le 17.04.2007).

de ses étudiants. C'est la raison pour laquelle le recours au *Test de connaissance du français* (TCF) que le département d'études françaises de l'université Nanzan, centre de passation agréé par le Centre international d'études pédagogiques (CIEP) depuis 2006, s'apprête à faire passer à ses étudiants en fin de 2^{ème} année, nous paraît pouvoir garantir la meilleure comparabilité des notes obtenues de leurs compétences linguistiques.

Le TCF a pour objet l'évaluation des quatre compétences d'expression et de compréhension orale et écrite à travers cinq épreuves réparties entre épreuves obligatoires et épreuves facultatives. Les épreuves obligatoires du TCF, au nombre de trois, se présentent sous la forme d'un questionnaire à choix multiple (seule une des quatre réponses proposées pour chaque item est correcte) pour un total de 80 items présentés dans un ordre de difficulté progressif allant du niveau A1 au niveau C2 du CECR³⁶. L'épreuve de *compréhension orale* (30 items, 25 minutes) évalue ainsi la capacité du candidat à, en fonction de son niveau, soit repérer les informations essentielles transmises dans des situations de communication quotidienne, comprendre des énoncés simples portant sur des thèmes familiers (dans des messages, des annonces, des entretiens ou des discussions au téléphone), comprendre les points principaux d'une intervention sur des sujets familiers (rencontrés au travail, à l'université, pendant les loisirs), comprendre les idées principales d'interventions complexes sur un sujet concret ou abstrait, familier ou non rencontré dans la vie personnelle, sociale ou professionnelle, suivre une intervention sur des sujets abstraits et complexes et reconnaître une gamme étendue d'expressions idiomatiques ou enfin comprendre tout type de discours prononcé à un débit rapide. L'épreuve de *maîtrise des structures de la langue* (20 items, 20 minutes) s'attache à évaluer sa capacité à identifier et choisir les formulations correctes dans des structures syntaxiques et lexicales. Quant à l'épreuve de *compréhension écrite* (30 items, 45 minutes), elle apprécie sa capacité à, toujours en

³⁶ L'attestation de résultats du TCF mentionne un score global entre 100 et 699 points et une indication sur le niveau correspondant sur l'échelle en six niveaux du Conseil de l'Europe/CECR. Elle précise également le score détaillé par compétence assorti d'une indication de niveau. Le cas échéant, elle délivre une note sur 20 points pour les épreuves d'expression écrite et orale. A1 : élémentaire (100–199 points), A2 : élémentaire avancé (200–299 points), B1 : intermédiaire (300–399 points), B2 : intermédiaire avancé (400–499 points), C1 : supérieur (500–599 points) et C2 : supérieur avancé (600–699 points).

fonction de son niveau, soit repérer des informations globales dans des documents utilisés dans des situations de communication quotidienne (petits messages amicaux ou professionnels), repérer des informations globales et détaillées contenues dans des documents courants (petites annonces, prospectus, menus et horaires, etc.), repérer des informations détaillées portant sur des personnes, des faits ou des événements (lettres personnelles, etc.), repérer des informations fines et détaillées dans des textes en langue courante relatifs à la vie quotidienne ou au travail, comprendre les points de vue d'auteurs sur des sujets concrets ou abstraits dans des articles et des comptes rendus ou bien comprendre des textes abstraits ou complexes extraits d'ouvrages littéraires ou spécialisés. Deux épreuves facultatives, portant l'une sur l'expression orale, l'autre sur l'expression écrite, viennent alors compléter le dispositif d'évaluation mis en place à travers le TCF. Avec l'épreuve d'*expression orale*³⁷ (entretien de 15 minutes maximum), le candidat, évalué sur sa capacité à communiquer et à réagir aux sollicitations de l'examineur, se voit demander, en fonction de son niveau, soit de décrire des lieux (habitation, travail) et des personnes, de parler de ses conditions de vie, de son activité professionnelle ou scolaire, de raconter des expériences, des événements et des projets, d'expliquer un projet ou une idée, l'intrigue d'un livre ou d'un film et d'exprimer ses réactions, d'expliquer les avantages et inconvénients d'un projet, de donner son opinion, d'exprimer son accord et son désaccord, de présenter une argumentation claire et structurée dans un style approprié au contexte proposé ou enfin de présenter de façon détaillée et structurée des sujets complexes, de les développer et de conclure. Quant à l'épreuve d'*expression écrite* (d'une durée de 1h45), elle se compose de six exercices ordonnés par difficulté croissante qui correspondent à chacun des six niveaux du CECR³⁸. Parmi les exercices figurent la rédaction d'un message simple (environ 40 mots), la rédaction d'une lettre personnelle portant sur des situations de la vie quotidienne qui implique l'utilisation d'un registre courant (environ 60 mots), la rédaction d'un compte rendu d'expérience, d'un récit, l'exposé d'une opinion (environ 80 mots), la rédaction d'un texte explicatif, dans lequel

³⁷ La passation de l'épreuve d'expression orale se fait individuellement sous la forme d'un entretien en face à face avec un examinateur et ne nécessite aucun temps de préparation préalable.

³⁸ Du niveau A1 au niveau C2. Voir la définition des six niveaux communs de référence et leurs descripteurs in *CECR...*, *op.cit.*, pp. 23–38.

seront développés des points de vue et une argumentation (environ 100 mots), la comparaison de deux opinions et prise de position sur un thème général (100 à 125 mots) et la reformulation des idées essentielles contenues dans un document, exposé d'un point de vue argumenté (environ 100 mots). Le candidat est alors évalué sur sa capacité à respecter les consignes données, communiquer un message de façon claire, enchaîner des idées et faire preuve de cohérence dans son discours, exprimer son avis et l'argumenter, utiliser un éventail de vocabulaire en relation avec la tâche demandée, faire preuve de la maîtrise de structures complexes, faire preuve d'un esprit de synthèse et à reformuler avec ses propres mots.

Dans un environnement universitaire de plus en plus fortement concurrentiel, les exigences en terme d'évaluation des compétences linguistiques des étudiants ne s'accordent plus avec des tests n'offrant qu'une simple mesure absolue d'une capacité donnée. Elles requièrent, en revanche, des instruments d'évaluation qui autorisent la comparaison des performances d'un candidat avec celles d'autres candidats avec lesquels il est légitimement comparable. Les qualités psychométriques du TCF, de la standardisation à l'étalonnage, en passant par la sensibilité³⁹, la fidélité⁴⁰ et la validité, placent celui-ci au tout premier rang des instruments d'évaluation disponibles en français langue étrangère. Afin de balayer l'univers des possibles dans le domaine de compétence évalué, rendre comparables des tests composés d'items différents et localiser n'importe quel candidat sur une même échelle de compétence, le processus d'élaboration du TCF recourt à différents modèles de réponses à l'item, en vertu desquels « tous les items mesurant une compétence particulière peuvent être positionnés le long d'une échelle, leurs positions et leurs espacements étant déterminés par le niveau de difficulté auquel ils correspondent »⁴¹. Outre l'emploi d'un manuel du rédacteur, garant du respect de règles de rédaction strictes en conformité avec le CECR, et la construction de prétests soumis à plus de 250 candidats potentiels, les analyses psychométriques

³⁹ Également désignée finesse discriminative, c'est la qualité du test qui permet de mettre en évidence des différences, mêmes faibles, entre les individus, in « Les qualités psychométriques du TCF », Plaquette d'information aux centres de passation agréés, CIEP.

⁴⁰ La note obtenue à un test est composée d'une « note vraie » qui correspond réellement à une mesure de la compétence que l'on veut mesurer, et d'un « écart de mesure » aléatoire. Est fidèle un test qui comporte un faible écart de mesure. *Ibid.*

⁴¹ *Ibid.*

auxquelles est soumis le TCF, principalement l'application du modèle de mesure classique et le modèle de Rasch, conduisent à disqualifier près de 20% des items prétestés. L'analyse classique permet d'éliminer les items dont le fonctionnement est incompatible avec les qualités psychométriques requises, soit lorsque la réussite ou l'échec à l'item est quasi-systématique, soit lorsqu'il s'avère peu voire pas discriminant, soit lorsqu'il présente un problème de clé (bonne réponse) et/ou de distracteurs (autres réponses possibles)⁴². L'homogénéité et la fidélité du test, à savoir sa capacité à bien mesurer les compétences en français, sont appréciées au moyen d'une analyse factorielle et de l'indice du coefficient alpha de Cronbach⁴³. Le modèle de Rasch spécifie, pour sa part, la relation mathématique entre la performance observée et la variable latente et permet de calibrer les items⁴⁴. Pour chacun d'entre eux, les probabilités observées et les probabilités prévues par le modèle sont mises en comparaison et, afin de réduire l'écart de mesure, sont éliminés tous ceux pour lesquels les deux valeurs ne coïncident pas suffisamment. Ensuite, on s'assure que les difficultés estimées des items sont indépendantes de l'échantillon ayant servi à leur obtention ce qui entraîne le rejet de ceux qui introduiraient un avantage pour l'un ou l'autre sexe, le locuteur de telle langue ou les candidats de tel centre de passation. Le rapport annuel 2004 du CIEP

⁴² *Ibid.*

⁴³ Le coefficient alpha de Cronbach est un indice statistique variant entre 0 et 1 qui permet d'évaluer l'homogénéité (la consistance ou cohérence interne) d'un instrument d'évaluation ou de mesure composé par un ensemble d'items qui, tous, devraient contribuer à appréhender une même entité (ou dimension) "sous-jacente" : le niveau de connaissance ou de compétence sur un thème donné ; le niveau d'aptitude, d'attitude, de motivation, d'intérêt dans tel domaine ou par rapport à tel objet notamment. Cet indice traduit un degré d'homogénéité (une consistance interne) d'autant plus élevé(e) que sa valeur est proche de 1. Dans la pratique, on considère généralement que l'homogénéité de l'instrument est satisfaisante lorsque la valeur du coefficient est au moins égale à 0,80.

⁴⁴ L'analyse théorique de réponse aux items (*Item Response Theory* - IRT) « est utilisée pour déterminer la difficulté d'items isolés dans une banque d'items. Si vous êtes à un niveau avancé, vos chances de répondre à une question élémentaire sont élevées ; si vous êtes débutant, vos chances de répondre à un item avancé sont très faibles. [...] Avec le modèle de Rasch, on peut constituer une chaîne de tests et de questionnaires différents en utilisant des "items d'ancrage" communs à deux parties de la chaîne. [...] Ainsi, chaque partie peut avoir pour cible un groupe particulier d'apprenants et être néanmoins reliée à une échelle commune. » *CECR...*, *op.cit.*, p. 152.

préconise, par ailleurs, de recourir au modèle Many-facets de Rasch dans le but d'attribuer aux correcteurs d'expression écrite et orale un coefficient de sévérité destiné à corriger l'évaluation du candidat pour la rendre moins dépendante du correcteur⁴⁵.

Le TCF, pour la conception, le développement et la gestion commerciale duquel le CIEP a obtenu, en janvier 2003, une certification de norme de qualité ISO 9001, version 2000, reconduite en janvier 2004 à la suite de l'audit annuel de vérification, offre donc incontestablement toutes les garanties d'une évaluation objective des compétences linguistiques des étudiants à l'université, qu'elle permette d'apprécier le travail personnel fourni par chaque étudiant durant son cursus universitaire ou qu'elle fournisse indirectement aux équipes enseignantes matière à nourrir une réflexion sur la validité de leurs choix et de leurs pratiques pédagogiques. L'examen conçu par le département de français de l'université Tamkang nous semble donc bien en deçà des objectifs potentiels d'une solide évaluation des compétences linguistiques des étudiants en fin de cursus, toute considération relative à la participation financière au test exigée de la part des candidats mise à part. Toutefois, cet argument pécunier⁴⁶, dont on peut concevoir qu'il puisse devenir déterminant, ne devrait pas manquer de s'effondrer devant la gratuité et la souplesse d'utilisation caractéristiques du système d'auto-évaluation sur internet DIALANG.

III–Souplesse d'utilisation et garantie d'objectivité du test d'auto-évaluation informatisé en ligne DIALANG

Le système multilingue⁴⁷ d'évaluation DIALANG⁴⁸, destiné en priorité aux

⁴⁵ « Rapport annuel 2004 », CIEP, 1er janvier 2003 au 1er avril 2004, p. 28.

⁴⁶ En 2007, les frais pour l'ensemble des trois épreuves obligatoires s'élèvent à 40 euros et pour chacune des épreuves facultatives (expression écrite et orale) à 20 euros.

⁴⁷ « [...] les spécifications ont alors été traduites de l'anglais dans les treize autres langues. La démarche de traduction a fait l'objet d'un consensus. Il y a eu accord sur les orientations à suivre pour la traduction et la négociation, le premier critère de qualité étant l'intelligibilité pour l'utilisateur. Dans un premier temps, deux ou trois experts traduisaient indépendamment la spécification dans leur langue puis se concertaient pour se mettre d'accord sur une traduction unique. Les traductions étaient alors communiquées au Groupe de travail sur l'auto-évaluation

apprenants en langues désireux de recevoir une information analytique sur leurs compétences linguistiques, se présente comme une mise en application du *Cadre européen commun de référence pour les langues* à buts diagnostiques. Le cadre pour l'évaluation, les échelles descriptives utilisées pour rendre compte des résultats aux utilisateurs de même que les spécifications pour l'auto-évaluation s'appuient directement sur le CECR, ce qui, à l'instar du TCF, en garantit l'objectivité. DIALANG, qui se compose d'un volet d'auto-évaluation, de tests de langue et d'un programme d'analyse des résultats et de conseils (feed-back) disponibles en quatorze langues européennes⁴⁹, s'est rapidement imposé, depuis le lancement du projet en 1996 et en dépit de l'absence de développement d'une version pour les plateformes Macintosh, comme un outil d'évaluation d'une très grande souplesse d'utilisation auprès d'apprenants soucieux, d'une part, de connaître leur niveau de compétence en langue et de profiter d'une information sur les points forts et les points faibles de cette compétence, et d'autre part, d'obtenir des conseils

dont les membres avaient la compétence linguistique pour faire une ultime vérification en neuf langues. » *CECR...*, *op.cit.*, p. 163.

⁴⁸ Disponible sur : <<http://www.dialang.org/intro.htm>> (consulté le 17.01.2007), DIALANG est le fruit d'un projet initié par la Commission européenne et près de 25 institutions partenaires, principalement des universités telles Der Freien Universität Berlin, Jyväskylä Yliopisto (Centre for Applied Language Studies), CITO Arnhem, Lancaster University (Department of Linguistics and Modern English Language), à travers toute l'Union européenne. Le projet bénéficie du soutien financier de la Commission européenne, Direction générale pour l'Éducation et la Culture (DG XXII), dans le cadre du programme Socrates/Lingua Action D. L'action Lingua en faveur de la promotion de l'enseignement et de l'apprentissage des langues soutient les mesures qui visent à encourager et promouvoir la diversité linguistique au sein de l'UE, à contribuer à l'amélioration de la qualité de l'enseignement et de l'apprentissage des langues et à faciliter l'accès à des possibilités d'apprentissage tout au long de la vie répondant aux besoins de chacun. D'un côté, Lingua 1 s'attache, d'une part, à sensibiliser les citoyens à la richesse multilingue de l'UE et à les encourager à apprendre des langues tout au long de leur vie et faciliter l'accès à des ressources d'apprentissage en Europe, et d'autre part, à élaborer et diffuser des techniques novatrices et des bonnes pratiques dans le domaine de l'apprentissage des langues. D'un autre côté, Lingua 2 vise à garantir la mise à disposition d'une gamme suffisamment étoffée d'instruments d'apprentissage linguistique à l'usage des apprenants.

⁴⁹ Il s'agit de l'allemand, de l'anglais, du danois, de l'espagnol, du finnois, du français, du grec, du gaélique (irlandais), de l'islandais, de l'italien, du néerlandais, du norvégien, du portugais et du suédois.

sur la façon d'améliorer leur habileté langagière tout en accroissant leur prise de conscience en ce qui concerne l'apprentissage des langues. A la différence de l'attestation du TCF, dont la validité certes limitée à deux ans ouvre néanmoins les portes de la plupart des établissements universitaires français aux non-francophones de niveau B2⁵⁰, DIALANG ne délivre aucune certification. Même si son public d'utilisateurs a principalement vocation à regrouper les apprenants isolés qui étudient les langues en autonomie ou dans le cadre d'un cursus formalisé, il ne nous paraît toutefois pas incongru de penser que les enseignants eux-mêmes pourraient y trouver un réel intérêt. Nous estimons raisonnablement pouvoir avancer, en effet, que l'équipe des concepteurs de l'examen d'évaluation de l'université Tamkang ne manquerait pas, en lui substituant le système d'auto-évaluation DIALANG, tout à la fois d'améliorer l'objectivité et la comparabilité de son test et des résultats qu'il délivre, de diminuer le fardeau que représente son élaboration annuelle par une équipe restreinte, d'introduire une réelle souplesse dans l'organisation pratique de sa passation et, peut-être est-ce là finalement l'avantage le plus intéressant, d'encourager ses étudiants à poursuivre leur apprentissage du français en autonomie au-delà de leurs études universitaires grâce aux commentaires et conseils fournis en fin de processus d'évaluation.

La démarche d'évaluation complète par le système DIALANG, qui requiert entre 30 et 45 minutes, mais seulement de 20 à 30 dans le cas où l'utilisateur ne choisit que le test de langue, s'ordonne autour de six étapes principales : un test de niveau sur l'étendue du vocabulaire, le choix de l'habileté (compréhension écrite/orale, expression écrite, vocabulaire, structures), une auto-évaluation (compréhension écrite/orale et expression écrite), une pré-estimation par le système de la capacité de l'utilisateur, l'administration d'un test de difficulté appropriée et des commentaires et conseils⁵¹. Après son inscription et la sélection

⁵⁰ Niveau B2 du CECR que le descripteur définit ainsi : Utilisateur indépendant « Peut comprendre le contenu essentiel de sujets concrets ou abstraits dans un texte complexe, y compris une discussion technique dans sa spécialité. Peut communiquer avec un degré de spontanéité et d'aisance tel qu'une conversation avec un locuteur natif ne comporte de tension ni pour l'un ni pour l'autre. Peut s'exprimer de façon claire et détaillée sur une grande gamme de sujets, émettre un avis sur un sujet d'actualité et exposer les avantages et les inconvénients de différentes possibilités. » *CECR...*, *op.cit.*, p. 25. Rappelons que le niveau B1 du CECR correspond au « niveau seuil » (Threshold Level) du Conseil de l'Europe élaboré au milieu des années 1970.

de la langue dans laquelle il souhaite prendre connaissance des instructions et des commentaires, l'utilisateur a la possibilité de se soumettre à un test de niveau, afin de permettre au système de délimiter l'étendue de son vocabulaire et « de décider à quel niveau de difficulté administrer les tests : facile, moyen, difficile »⁵². Une liste de 75 mots dont 25 n'existent pas lui est présentée, à charge pour lui de déterminer quels sont ceux qui appartiennent bien au lexique de la langue du test. L'habileté pour laquelle il souhaite être testé choisie, lui sont alors ensuite présentées des spécifications pour l'auto-évaluation qui couvrent l'habileté sélectionnée et par rapport auxquelles il doit se déterminer capable ou pas (par oui ou par non) de réaliser l'activité décrite dans chaque énoncé, avec le même objectif de détermination du niveau de difficulté des tests de langue proprement dits. Ce test d'auto-évaluation n'est cependant envisageable que si le choix porte sur un test de compréhension écrite, de compréhension orale ou d'expression écrite « mais pas pour les tests de vocabulaire et de structures, dans la mesure où le CECR lui-même ne propose aucun descripteur de ces deux aspects de la compétence langagière »⁵³. La décision de proposer, parallèlement au test d'auto-évaluation, un test de vocabulaire a été prise par les concepteurs de DIALANG en face des incertitudes qu'a fait naître en eux la faible fiabilité de la seule auto-évaluation employée comme instrument discriminant du niveau de difficulté adéquat : « Un test bref et objectif au cours duquel l'utilisateur aurait à démontrer sa maîtrise de certains aspects de la langue a été jugé nécessaire. Le test de vocabulaire a été choisi car la connaissance des mots fait partie intégrante de la lecture, de l'oral et de l'écrit et parce qu'il est possible de soumettre un nombre pertinent d'items de vocabulaire dans un laps de temps réduit, ce qui accroît d'autant la probabilité d'une pré-évaluation précise »⁵⁴. En outre, le choix d'un test offrant oui ou non pour réponse contribue à maximiser le nombre d'items et n'implique aucune traduction, si bien, qu'actuellement, le test de vocabulaire est légèrement plus favorisé que celui d'auto-

⁵¹ ALDERSON J. Charles, HUHTA Ari, « The development of a suite of computer-based diagnostic tests based on the Common European Framework », in *Language Testing*, July 1 2005, Vol. 22, n°3, pp. 301-320.

⁵² ALDERSON et HUHTA, « Computer-based diagnostic tests... », in *Language Testing*, *op. cit.*, p. 303. Nous traduisons de l'anglais.

⁵³ *Ibid.*

⁵⁴ *Ibid.*, p. 304.

évaluation. Ces deux tests « contribuent à rendre pour l'utilisateur l'expérience du test elle-même aussi significative que possible en terme de difficulté. Par ailleurs, ils servent aussi à renforcer la précision de l'évaluation des tests DIALANG : utiliser un seul test pour couvrir tous les niveaux du CECR de A1 à C2 reviendrait soit à une estimation moins fiable des compétences, soit à des tests si longs qu'ils rebuteraient les utilisateurs. Le recours à trois versions de tests qui se concentrent sur différentes parties du spectre des compétences fournit ainsi une solution fonctionnelle au problème posé »⁵⁵. Dès que le niveau de difficulté estimée par DIALANG a été déterminé, l'utilisateur peut commencer son évaluation proprement dite et lire une brève présentation du test avant que les items ne lui soient soumis. Ceux-ci se présentent sous quatre formes principales déclinées en diverses variantes dont il est possible de se faire une idée précise sur le site internet de DIALANG : choix multiple, menu déroulant, entrée de texte et questions à réponses courtes⁵⁶. Une fois les tests effectués, le système délivre à l'utilisateur quatre types de résultats pour qu'il puisse situer ses compétences linguistiques sur une échelle objective, que les concepteurs du test du département de français de l'université Tamkang pourraient aisément et en toute fiabilité exploiter à des fins docimologiques. Le premier résultat indique le niveau de l'utilisateur sur l'échelle des six niveaux du CECR et expose brièvement quelles compétences un apprenant de même niveau est supposé capable de mettre en oeuvre. Un second volet de résultats détaille l'ensemble des réponses de l'utilisateur sous la forme d'un tableau qui regroupe les items en fonction des sous-compétences qu'ils visent à évaluer, afin de lui fournir une image précise de ses forces et de ses faiblesses. DIALANG délivre également, sur une échelle de 1 à 1 000, le résultat obtenu au test de vocabulaire accompagné d'une brève explication de sa signification. Enfin, l'utilisateur est informé du degré de cohérence entre ses résultats après test et ceux

⁵⁵ *Ibid.*, p. 305.

⁵⁶ Pictorial Multiple Choice with Sound, Interactive Image with Sound, Video Clips in Listening, Drag and Drop Activity, Re-organisation, Highlighting/Underlining, Insertion, Deletion, Thematic Grouping, Multiple Response, Indirect Speaking with Audio Clips as alternatives, Benchmarking in Direct Writing, Multiple Benchmarks in Speaking, Drag and Drop Matching, Transformation, Combined Skills, Mapping and Flow Charting, Confidence in Response [en ligne]. Disponible sur : <<http://www.lanacs.ac.uk/fss/projects/linguistics/experimental/start.htm>> (consulté le 17.01.2007).

de son auto-évaluation. DIALANG fournit par ailleurs deux séries de conseils : la première consiste en une suite de tableaux qui exposent les différences entre le niveau assigné à l'utilisateur et ceux immédiatement supérieur et inférieur, et la seconde formule des recommandations pour l'aider à évoluer de son niveau actuel vers le niveau supérieur.

S'inscrivant dans le cadre des réflexions menées depuis plus d'une quinzaine d'années parmi les organismes certificateurs et les concepteurs de tests de langue, dont nombre sont désormais regroupés au sein d'ALTE⁵⁷, relatives en particulier à la nécessité de suivre scrupuleusement des standards professionnels à chacune des étapes de l'élaboration des tests d'évaluation des compétences linguistiques, les responsables du projet DIALANG se sont ainsi attachés à soumettre le leur à des procédures strictes de validation et de calibrage faisant appel aux ressources de l'analyse psychométrique. A titre d'exemple unique, nous noterons le soin apporté au développement des items dans chacune des 14 langues du système, confié à 14 équipes qui en auront produit près de 30 000, soit 2 000 par langue⁵⁸.

En guise de conclusion

« Les politiques nationales, les traditions et les cultures de l'évaluation autant que les politiques, les cultures et les intérêts légitimes des organismes spécialisés dans les tests et les examens de langue sont des facteurs qui peuvent freiner l'intérêt qu'il y a à une reconnaissance mutuelle des qualifications »⁵⁹. Toutefois, il en va de l'intérêt de chacun que soient appliquées des procédures convenables en matière d'évaluation des compétences linguistiques dans les universités et tout particulièrement au sein des départements des facultés de langues étrangères. Car,

⁵⁷ L'Association des centres d'évaluation en langues en Europe (Association of Language Testers in Europe), créée en 1990, rassemble 27 membres institutionnels et associés qui représentent 24 langues européennes. Son objectif fondamental, l'établissement de standards professionnels à chaque étape des processus d'évaluation en langue afin de garantir aux apprenants l'accès à des certifications de haute qualité, s'est traduit en 1994 par la publication d'un Code de Bonne Pratique.

⁵⁸ Pour une étude détaillée des tests de calibrage de DIALANG, on se reportera à ALDERSON et HUHTA, « Computer-based diagnostic tests... », in *Language Testing, op. cit.*

⁵⁹ Conseil de l'Europe, « Relier les examens au CÉCR... », *op.cit.*, p. 6.

L'évaluation du niveau des compétences linguistiques à l'université : test standardisé ou test-maison ?

et l'exemple de l'examen d'évaluation des compétences linguistiques des étudiants en fin de 4^{ème} année du département de français de l'université Tamkang, qui correspond à une commande passée par la direction de cette université, le prouve à l'envi, au-delà de l'évaluation des seules compétences linguistiques des étudiants pour eux-mêmes, de tels actes d'évaluation conduisent déjà et conduiront incontestablement de plus en plus vers une évaluation critique des choix pédagogiques effectués par chaque institution. A travers l'évaluation des compétences linguistiques de leurs étudiants, ce sont la qualité et la pertinence de la réflexion pédagogique en matière d'enseignement des langues étrangères au sein des universités qui se trouvent ainsi mises sur la sellette. Dans un contexte accru de mobilité et de concurrence internationales, une telle évolution apparaît inéluctable, si bien qu'au regard des enjeux s'en remettre à un test-maison élaboré hors de tout système de référence de pratique et de qualité risquerait, par manque de fiabilité et de comparabilité des résultats obtenus, d'atteindre la crédibilité des établissements et de porter en soi le germe de leur propre disqualification.

Annexes

Annexe 1 : Examen d'évaluation de 4^{ème} année de l'université Tamkang, expression orale (exemples de questions)

1^{ère} partie : EXPRESSION ORALE (50 points)

1/1 Lisez le texte ci-dessous et répondez aux questions. Vous disposerez d'une minute trente (1'30) pour cette épreuve. (10 points)

Martine doit partir en vacances au Japon. Elle a un chien et elle ne peut pas l'emmener avec elle. Alors, elle se rend chez son amie Nicole et lui demande de bien vouloir le garder. Mais elle se heurte au refus de son amie. Elle insiste. Enfin, elle parvient à la convaincre de garder l'animal pendant les vacances.

Questions :

- 1 - Avez-vous un animal à la maison ? Si oui, lequel ? Si non, pourquoi ?
- 2 - Comment l'avez-vous obtenu ? En cadeau ? Acheté ?

1/2 Vous allez entendre 5 phrases en chinois. Chaque phrase ne sera dite qu'une seule

fois. Répétez la phrase en chinois et ensuite faites-en la traduction en français.
Vous aurez 30 secondes par phrase. (15 points)

1. 能否請您告訴我捷運站在哪裡？

1. Pourriez-vous me dire où se trouve la station de métro ?
2. 上星期天，吃過飯後，我們去散步。
2. Dimanche dernier, après avoir dîné, nous nous sommes promenés.

1/3 Vous allez répondre à 10 questions.

Chaque question ne sera posée qu'une seule fois.

Vous disposerez de 30 secondes après chaque question pour votre réponse. (20 points)

1. Où peut-on trouver des marchés de nuit intéressants à Taipei ?
2. Pour le nouvel an chinois, avez-vous reçu des enveloppes rouges ?
3. Quel jour de la semaine préférez-vous ? Pourquoi ?

1/4 Regardez cette photo et répondez aux questions. Vous disposerez de 1'30 (une minute trente secondes) pour y répondre. (5 points)

Questions :

- Où est-on ?
- Qui sont les personnages ?
- Que font-ils ?

Annexe 2 : Examen d'évaluation de 4^{ème} année de l'université Tamkang, compréhension orale (exemples de questions)

2^{ème} partie : COMPREHENSION ORALE (50 points)

2-1 Vous allez entendre 15 questions. Chaque question ne sera posée qu'une seule fois. Après avoir entendu la question, choisissez la bonne réponse parmi les 4 propositions. Vous disposerez de 15 secondes pour votre réponse. (15 points)

L'évaluation du niveau des compétences linguistiques à l'université : test standardisé ou test-maison ?

Ex : Comment allez-vous ?

- A - Je vais manger.
- B - Je vais bien, merci.
- C - Je vais à la bibliothèque.
- D - Nous venons de suite.

La seule réponse possible à la question est la proposition B : « Je vais bien, merci ». Vous devrez cocher la réponse B.

1 : Vous n'aimez pas aller à la piscine ?

- 1 A. Oui, je préfère la montagne.
- B. Si, j'y vais chaque semaine.
- C. Si, j'aime le poisson.
- D. Non, je déteste faire les courses.

2 : Il vous a rendu votre parapluie ?

- 2 A. Oui, demain.
- B. Non, il n'en a pas besoin.
- C. Oui, il me l'a rendu hier.
- D. Non, sans façons.

2/2 Vous allez entendre 15 phrases. Chacune ne sera dite qu'une seule fois. Ecoutez les 4 propositions et choisissez celle qui correspond le mieux à ce que vous avez entendu. Vous disposerez de 15 secondes pour votre réponse. (15 points)

Ex : Je n'aime pas beaucoup les romans policiers.

- A - Je déteste les romans policiers.
- B - J'adore les romans policiers.
- C - J'aime seulement un peu les romans policiers.
- D - Je préfère les romans policiers.

La proposition la plus proche de l'énoncé est la proposition C : « j'aime seulement un peu les romans policiers ». Vous devrez cocher la réponse C.

16 : Il n'est pas tellement favorable à ce projet.

- 16 A. Il aime beaucoup ce projet.

- B. Il n'aime pas beaucoup ce projet.
- C. Il adore ce projet.
- D. Il préfère ce projet.

17 : Je te verrai vers 21 heures ?

- A. Je vais te voir à partir de 21h.
- B. Je vais te voir à environ 21h.
- C. Je vais te voir bien après 21h.
- D. Je vais te voir bien avant 21h.

2/3 Vous allez entendre 5 textes ou histoires courtes. Vous ne les entendrez qu'une seule fois. Répondez ensuite à la question en choisissant la réponse la mieux appropriée. Vous disposerez de 15 secondes pour votre réponse.

Ex : Les industries traditionnelles principales de la Lorraine sont l'extraction du sel, des minerais de fer et de charbon. L'eau est une autre industrie d'extraction en Lorraine. L'eau minérale est mise en bouteille et vendue en France ainsi qu'à l'étranger.

Q : Quelles sont les extractions de la Lorraine ?

- A - L'eau et la quiche lorraine.
- B - Les minerais de fer et de charbon, le sel et l'eau minérale.
- C - Les roches minérales et l'eau.
- D - L'eau et les minéraux.

La bonne réponse est la B : « Les minerais de fer et de charbon, le sel et l'eau minérale ». Vous devez cocher la réponse B.

Texte 1 :

Quand il y a un typhon, on doit rester à la maison. C'est dangereux de sortir car le vent violent peut emporter des objets qui risquent de blesser ou même de tuer ceux qui s'aventurent à l'extérieur. En général, on n'aime pas les typhons, mais les enfants, eux, sont bien contents car quand il y a un typhon, ils ne vont pas à l'école.

Question : Pourquoi les enfants sont-ils contents quand il y a un typhon ?

- 31 A. Parce qu'ils peuvent jouer dans la rue.
- B. Parce qu'il n'y a pas d'école.

L'évaluation du niveau des compétences linguistiques à l'université : test standardisé ou test-maison ?

- C. Parce que le vent, c'est excitant.
- D. Parce que les enfants aiment le danger.

Texte 2 :

Société de production audiovisuelle recherche comédiens/chanteurs et comédiennes/chanteuses pour un spectacle musical qui commencera en septembre 2006. Les candidats sélectionnés devront être libres dès le mois d'avril pour des répétitions quotidiennes. Salaires en fonction des rôles joués. Contrats à durée déterminée. Se présenter le mercredi 8 mars, à 14 heures, à l'adresse suivante : 25 rue de la Paix - 75008 Paris.

Question : Quelle est la nature de ce texte ?

- 32
- A. Une annonce de rencontre
 - B. Une proposition de stage
 - C. Une offre de vente
 - D. Un casting

Texte 3 :

Pendant les vacances d'été, beaucoup d'étudiants taiwanais se rendent en France pour pratiquer leur français. Ils en profitent également pour visiter les châteaux et les musées dont ils ont entendu parler en classe. Mais ce qui les attire le plus reste la Provence. Est-ce à cause de son climat et de ses senteurs ou encore de ses paysages et des artistes qu'elle inspire ? Peut-être un peu de tout cela. N'est-elle pas aussi la région préférée des Français ?

Question : Qu'est-ce qui attire le plus les étudiants taiwanais en France ?

- 33
- A. La pratique du français.
 - B. La Provence.
 - C. Les musées.
 - D. La gastronomie.

2/4 Vous allez entendre 5 dialogues. Vous ne les entendrez qu'une seule fois.

Répondez ensuite à la question en choisissant la réponse la mieux appropriée.

Vous disposerez de 15 secondes pour votre réponse.

Le vendeur : Bonjour Madame, que désirez-vous ?

La dame : Je voudrais une écharpe.

Le vendeur : C'est pour offrir ?

La dame : Oui.

Le vendeur : Comment la trouvez-vous ?

La dame : Pas mal, mais c'est trop cher.

Le vendeur : Et celle-là ?

La dame : Pas mal, mais c'est encore trop cher.

Le vendeur : Je suis désolé, Madame, je ne connais pas vos goûts. En tout cas, le prix et la qualité vont de pair.

Question : Que dit le vendeur à propos de la qualité et du prix ?

A - La qualité est plus importante que le prix.

B - Le prix est aussi important que la qualité.

C - La qualité est moins importante que le prix.

D - Il ne dit pas grand chose.

La bonne réponse à la question est la réponse B : « Le prix est aussi important que la qualité ». Vous devrez cocher la réponse B.

Dialogue 1 :

Homme : Excusez-moi madame, je cherche la poste.

Femme : Ah oui, elle est dans l'avenue de Bercy.

Homme : Par là ?

Femme : Non, par là. Vous voyez la banque, au coin de la rue ? C'est par là.

Homme : Ah, oui. Merci beaucoup.

Q : Où l'homme veut-il aller ?

36 A. À la poste.

B. Au Café de Bercy.

C. Au bureau de tabac.

D. À la banque.

Dialogue 2 :

Homme : Bonjour, je voudrais une baguette s'il vous plaît.

Femme : Vous désirez autre chose ?

Homme : Oui, deux croissants au beurre.

Femme : Voilà. Et avec ça ?

Homme : Ce sera tout, merci. Combien je vous dois ?

Femme : Ça vous fera trois euros et 20 centimes.

Q : Qu'est-ce que l'homme achète ?

37 A. Deux baguettes et deux croissants.

L'évaluation du niveau des compétences linguistiques à l'université : test standardisé ou test-maison ?

- B. Une baguette et un croissant.
- C. Une baguette et un pain au chocolat.
- D. Une baguette et deux croissants.

Annexe 3 : Examen d'évaluation de 4^{ème} année de l'université Tamkang, compréhension écrite (exemples de questions)

3^{ème} partie : Épreuve de compréhension écrite (100 points)

3-1 Vous devez remplir les blancs de chacune des 15 phrases avec l'une des quatre propositions. (30 points)

Ex : Nous serons _____ en retard.

- A - jamais
- B - probablement
- C - pas
- D - personne

La seule possibilité est B : « probablement ». Vous cochez la case B sur la feuille des réponses.

1. Il aime plaisanter et prendre la vie du bon côté, c'est quelqu'un d' _____ .
 - A. optimiste
 - B. inconvenant
 - C. indépendant
 - D. inquiétant
2. Quand j'étais au lycée, je jouais du piano chaque jour, mais maintenant je n' _____ joue plus.
 - A. y
 - B. à
 - C. au
 - D. en

3-2 Vous devez cocher la proposition la plus proche de l'élément souligné pour chacune des 15 phrases. (30 points)

Ex : Nous apprenons le français depuis quatre ans.

- A - pensons
- B - étudions
- C - comprenons
- D - écoutons

La proposition la plus proche de l'élément souligné est B : « étudions ». Vous cochez la réponse B sur la feuille des réponses.

16. Nous ne pouvons pas donner suite à votre demande de remboursement.
- A. comprendre.
 - B. retrouver.
 - C. satisfaire.
 - D. décider.
17. Martin est au chômage depuis trois mois.
- A. sans emploi.
 - B. en vacances.
 - C. fiancé.
 - D. en stage.

3-3 Vous devez cocher la bonne réponse aux 4 questions des 5 textes. (40 points)

Texte 1

Les épaves de notre vie quotidienne ne sont pas destinées à terminer leur vie sous la surface des mers et des océans et pourtant, des scientifiques n'hésitent pas à couler des carcasses de navires, des voitures, des blocs de béton, ou encore des paquets de pneus liés entre eux. On trouve encore des bus, des camions et même des frigos. Loin de nuire à la qualité de la vie sous-marine, ils contribuent à sa bonne santé.

Mais on ne coule pas n'importe quoi, n'importe comment et n'importe où. En 1980, Taiwan passait par le fond 31 bateaux et de nouvelles structures en béton, véritable « hôtels » pour poissons, commencèrent à être immergées à des profondeurs allant de 15 à 65 mètres. Elles ont permis de multiplier la biomasse par cent et le nombre des espèces par quatre. Dans ces récifs artificiels, chaque espèce peut trouver une place et un abri bien adaptés.

Aujourd'hui encore, des épaves variées et des récifs modulaires sont immergés dans des zones où la faune et la flore ont du mal à se fixer, faute de niches écologiques accueillantes. Leurs zones d'immersion ne ressemblent en rien à des cimetières marins mais bien plutôt à des jardins luxuriants.

(d'après Télé 7 jours du 27 mai au 2 juin 1995)

Questions :

31. De quoi parle cet article ?
- A. d'un accident.
 - B. d'environnement.
 - C. de pollution.
 - D. d'un fait divers.
32. Pourquoi immerge-t-on des épaves ?
- A. pour s'en débarrasser.
 - B. parce qu'il y a beaucoup de place dans la mer.
 - C. parce que l'eau salée va les détruire.
 - D. pour améliorer la biomasse.
33. Grâce à ces immersions, les espèces ont pu être multipliées par...
- A. 2.
 - B. 3.
 - C. 4.
 - D. 10.
34. Pourquoi ces espèces ont-elles pu se développer ?
- A. parce que ces récifs sont adaptés à leurs besoins.
 - B. parce que la température de l'eau est plus chaude.
 - C. grâce à l'aquaculture.
 - D. grâce à la pureté de l'eau.

Texte 2

La "Bibliothèque rose" fête ses 150 ans

La célèbre collection fête ses 150 ans. Une histoire qui a commencé en 1852 quand Louis Hachette crée la "Bibliothèque des chemins de fer", afin d'occuper les voyageurs pendant leurs trajets. Il achète alors les kiosques implantés dans les gares, dont le succès ne se démentira pas... Cent cinquante ans plus tard, ils sont rouge et blanc et portent l'enseigne Relay.

Ce n'est que quatre ans plus tard que la couleur rose est attribuée à cette collection pour enfants - pour la différencier des autres - avec le premier roman de la comtesse de Ségur, *Les Nouveaux Contes de fées*, illustré par Gustave Doré. Suivront une vingtaine d'autres récits de cette grand-mère en mal d'occupation : *Les Petites Filles modèles*, *Le Général Dourakine*, *Mémoires d'un âne...* Autant d'histoires qui sont devenues des classiques.

En 1924, la "Bibliothèque verte", pendant "garçons" de la "Rose", voit le jour. Mais avec l'après-guerre, la mixité rattrape les collections "Rose" et "Verte" : l'une s'adresse aux plus petits, l'autre aux plus grands. En 1955, à la "Bibliothèque rose", c'est l'explosion d'Enid Blyton : *Le Club des cinq*, *Le Clan des sept*, la série des *Oui-Oui...* Dans la foulée, d'autres séries-cultes ont vu le jour, comme *Alice* de Caroline Quine ou *Fantômette* de Georges Chaulet.

C'est d'ailleurs lui qui offre son plus beau cadeau à la "Bibliothèque rose" avec, pour cet anniversaire, un livre inédit : *Le Retour de Fantômette*. Le 15 mars, les lecteurs pourront, après dix-huit ans d'absence, découvrir le 50^e épisode de la petite justicière. Révolution technologique oblige, Fantômette a maintenant un téléphone portable et sa complice Ficelle passe ses journées sur Internet... (d'après Le Monde du 25/02/06)

Questions :

35. À l'origine, la collection a été créée pour...
 - A. être lue dans les trains.
 - B. l'éducation des enfants.
 - C. promouvoir de nouveaux auteurs.
 - D. proposer des livres bon marché.
36. Actuellement, les couleurs de la collection sont le...
 - A. rose et vert.
 - B. noir et blanc.
 - C. rouge et blanc.
 - D. bleu et rouge.
37. En 1924, on crée la « Bibliothèque verte » pour ...
 - A. les parents.
 - B. les garçons.
 - C. les filles.
 - D. les bébés.
38. Le dernier livre de « Fantômette » a été publié il y a...
 - A. 1 an.
 - B. 4 ans.
 - C. 12 ans.
 - D. 18 ans.

Quelques références bibliographiques

- CÁRDENAS Abel, « La Evaluación de la Proficiencia en Español como Lengua Extranjera: Descripción y Crítica de la Entrevista de Proficiencia Oral », in *Academia Literature and Language*, Nagoya, January 2006, Vol. 79, pp. 53–72.
- CÁRDENAS Abel, « La Evaluación de la Proficiencia en Español como Lengua Extranjera: Introducción y Contextos », in *Academia Literature and Language*, Nagoya, January 2005, Vol. 77, pp. 191–208.
- CÁRDENAS Abel, « La Evaluación de la Expresión Oral en los Exámenes de los Diplomas de Español como Lengua Extranjera: Una Guía para Profesores y Estudiantes », in *Academia Literature and Language*, Nagoya, January 2001, Vol. 69, pp. 25–67.

L'évaluation du niveau des compétences linguistiques à l'université : test standardisé ou test-maison ?

DAVIES Alan, «Introduction: Language Testing and the Golden Rules», in *Language Assessment Quarterly: An International Journal*, 2004, Vol. 1, n°2&3, pp. 97–107.

HACKETT Ed, «The development of a computer-based version of PET», in *Research Notes*, Cambridge: Cambridge ESOL, November 2005, issue 22, pp. 9–13 [en ligne]. Disponible sur : <http://www.cambridgeesol.org/rs_notes/> (consulté le 07.03.2007).

TAGLIANTE Christine, *L'évaluation et le Cadre européen commun*, Paris : CLE International/ SEJER, nouvelle édition 2005, 207 p.

VAN AVERMAET Piet, KUIJPER Henk, SAVILLE Nick, «A Code of Practice and Quality Management System for International Language Examinations», in *Language Assessment Quarterly: An International Journal*, 2004, Vol. 1, n° 2&3, pp. 137–150.

WILSON Juliet, «Ethical issues in the testing of young learners», in *Research Notes*, Cambridge: Cambridge ESOL, November 2005, issue 22, pp. 6–7 [en ligne]. Disponible sur : <http://www.cambridgeesol.org/rs_notes/> (consulté le 07.03.2007).