



HAL
open science

Une gageure : tester l'oral

Anne Péchou

► **To cite this version:**

Anne Péchou. Une gageure : tester l'oral. Les Après-midi de LAIRDIL, 1994, 01, pp.51-53. hal-04058366

HAL Id: hal-04058366

<https://ut3-toulouseinp.hal.science/hal-04058366>

Submitted on 4 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Une gageure : tester l'oral

La manière d'évaluer l'anglais en tant que langue étrangère a beaucoup évolué au cours des dernières années, comme en témoigne l'évolution des pratiques à l'INSA de Toulouse. L'évaluation de la performance orale, en particulier, a pris une place accrue dans la notation globale. Inspiré par l'une des épreuves de l'examen de Cambridge First Certificate, l'examen oral en première et deuxième année - 400 étudiants environ - prend la forme d'un jeu de rôles ou d'une discussion où deux étudiants doivent résoudre un problème. L'examineur reste en retrait et n'intervient pas.

Plusieurs années de ce type d'évaluation ont inspiré les réflexions suivantes.

Choix du type de test

- *Un test doit refléter ce qu'on enseigne.* Hughes use le terme de *backwash* pour designer l'effet d'un type de test sur l'apprentissage et sur la façon d'enseigner. Il serait vain d'insister sur la pratique de la langue de communication en cours si on persistait à évaluer par une épreuve écrite. L'apprenant est peu enclin à fournir un effort à l'oral s'il sait qu'il ne sera pas évalué à l'oral.
- *Il doit être discriminant*, c'est-à-dire permettre d'utiliser une vaste gamme de notes.
- *Il doit tenir compte des conditions matérielles*: temps, locaux et personnel disponible.

Durée de l'épreuve

Quinze minutes en moyenne pour deux étudiants avec un seul examinateur, faute de moyens pour en avoir deux, soit huit étudiants à l'heure. Le choix des "paires" est laissé aux étudiants. Chaque groupe est convoqué un quart d'heure à l'avance et se prépare dans la salle où l'épreuve est soutenue.

Problèmes rencontrés

La difficulté principale est d'assurer l'harmonisation de la notation entre une dizaine d'examineurs ou plus. On sait que certains notent beaucoup plus sévèrement que d'autres et qu'un étudiant, évalué par un autre enseignant, lorsqu'on procède à un échange de groupes à l'occasion d'un test oral peut obtenir une note très éloignée de celle qu'il attendait. De gros écarts subsistent, même à l'issue d'une session de formation où l'équipe d'examineurs s'était entraînée à partir de documents vidéos montrant des étudiants de divers niveaux en situation d'examen.

Des écarts se manifestent également suivant que l'évaluateur est de langue maternelle anglaise ou non. Habituellement, les francophones ont tendance à noter plus sévèrement que leurs collègues anglophones - qui ont une plus grande tolérance des erreurs, à condition qu'elles n'entravent pas la compréhension.

Recommandations

Afin de corriger les écarts, il semble essentiel d'établir des critères.

Ceux-ci peuvent être traditionnels : grammaire, vocabulaire, prononciation, fluidité et contenu. Ils peuvent aussi inclure les notions de registre et prendre en compte l'interaction.

Les critères serviront de guide à l'examineur pendant l'épreuve, mais il n'est pas souhaitable d'affecter rigidelement d'une note chacun des critères.

Il faut, tout en tenant compte des critères sélectionnés, attribuer la note finale sur une impression d'ensemble, ce qu'Underhill appelle *impression making*. Cette impression ne peut échapper à la subjectivité. Mais nous sommes toujours évalués subjectivement lorsque nous participons à une interaction, que ce soit dans notre langue maternelle ou en langue étrangère.

Françoise Cormon, méthodologue chevronnée aux Etudes Pédagogiques de Genève, confirme ce point de vue. Après avoir passé des années à peaufiner des grilles d'évaluation pour l'oral, elle les délaisse à présent en faveur d'un jugement global basé sur une liste de critères. Elle a observé que les écarts de notation entre examinateurs, à partir d'une impression d'ensemble par rapport à un niveau moyen requis, étaient moindres ainsi que lorsque les enseignants effectuaient un total de points pour tous les critères.

Conclusions

En dépit des résistances rencontrées (difficultés de mise en place, perte de temps, manque de fiabilité) un examen oral est la seule façon d'évaluer l'oral - si le développement des capacités à s'exprimer à l'oral est l'objectif d'un cours de langue. La notation par deux examinateurs - qui entraîne des frais supplémentaires - semble un moyen très sûr de pallier les disparités d'évaluation mentionnées ci-dessus.

Références bibliographiques

- Underhill, N. 1987. *Testing Spoken Language*. Cambridge: Cambridge University Press.

Ce livre offre une description très complète des différents types de tests. Son but est d'encourager les examinateurs à expérimenter diverses techniques et à en changer si elles ne conviennent pas, plutôt que de suivre automatiquement une façon de faire. Les opérations statistiques pour assurer la fiabilité des notes d'un même examen et entre plusieurs épreuves conviennent pour des épreuves écrites mais ne sont pas adaptées à l'oral, par essence plus fluide et spontané. L'auteur insiste sur l'aspect humain de la rencontre entre deux personnes que constitue un test oral - à distinguer de "*la machine à saucisses*" des résultats statistiques. Il faut rétablir l'équilibre entre traitement statistique et évaluation subjective.

- Hughes, A. 1989. *Testing for Language Teachers*. Cambridge: Cambridge University Press.

Dans ce livre plus complet, le test oral n'est qu'un chapitre parmi les autres sortes de tests, pour l'écrit, la lecture, l'écoute, la grammaire et le vocabulaire. L'auteur insiste sur la nécessité de cohérence entre test et enseignement. On trouve un chapitre sur la *fiabilité* (les résultats obtenus avec un groupe de niveau donné sont similaires à ceux qu'on obtiendra à une autre occasion), un autre sur la *validité* (un test mesure avec précision ce qu'il est destiné à mesurer) et une annexe sur l'analyse statistique des résultats de tests.

Anne Péchou