

L'apport des technologies multimédias en évaluation didactique de la traduction

Philippe Gardy

Volume 60, numéro 3, décembre 2015

URI : <https://id.erudit.org/iderudit/1036136ar>

DOI : <https://doi.org/10.7202/1036136ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Les Presses de l'Université de Montréal

ISSN

0026-0452 (imprimé)

1492-1421 (numérique)

[Découvrir la revue](#)

Citer cet article

Gardy, P. (2015). L'apport des technologies multimédias en évaluation didactique de la traduction. *Meta*, 60(3), 406–430.
<https://doi.org/10.7202/1036136ar>

Résumé de l'article

Cet article présente les résultats d'une étude menée auprès de 88 étudiants en première année de traduction (anglais-français) quant aux apports potentiels de technologies multimédias au processus d'évaluation des travaux des étudiants. La comparaison des résultats obtenus par les apprenants ayant bénéficié d'une rétroaction vidéo et de ceux ayant reçu une rétroaction traditionnelle met en évidence une amélioration des résultats des premiers. Par ailleurs, une analyse qualitative met en exergue un net renforcement de leur perception des aspects positifs de l'évaluation. Ainsi, le recours à la vidéo peut être considéré comme un outil efficace dans le cadre d'un processus d'évaluation à visée formative des travaux des étudiants en traduction, et ce, indépendamment de la taille du groupe d'apprenants.

L'apport des technologies multimédias en évaluation didactique de la traduction

PHILIPPE GARDY

Université Laval, Québec, Canada

philippe.gardy@lli.ulaval.ca

RÉSUMÉ

Cet article présente les résultats d'une étude menée auprès de 88 étudiants en première année de traduction (anglais-français) quant aux apports potentiels de technologies multimédias au processus d'évaluation des travaux des étudiants. La comparaison des résultats obtenus par les apprenants ayant bénéficié d'une rétroaction vidéo et de ceux ayant reçu une rétroaction traditionnelle met en évidence une amélioration des résultats des premiers. Par ailleurs, une analyse qualitative met en exergue un net renforcement de leur perception des aspects positifs de l'évaluation. Ainsi, le recours à la vidéo peut être considéré comme un outil efficace dans le cadre d'un processus d'évaluation à visée formative des travaux des étudiants en traduction, et ce, indépendamment de la taille du groupe d'apprenants.

ABSTRACT

This article examines the results obtained in a study involving the use of video screen capture software as an innovative way of providing feedback to first-year university students in translation (English to French). The results of this study suggest that such technological tools can help students achieve better results during exams. Furthermore, qualitative analysis shows that the use of technology does enhance students' perception of the positive role that assessment can play during their training. Whether in small or larger groups, video screen capture software can be considered as an efficient tool within a formative assessment process.

MOTS CLÉS/ KEYWORDS

évaluation en traduction, outils technopédagogiques, rétroaction vidéo, évaluation formative, évaluation sommative
assessment in translation, technopedagogical tools, video screen capture, formative assessment, summative assessment

1. Introduction

Si l'on observe depuis une quinzaine d'années un regain d'intérêt pour l'évaluation en didactique de la traduction, comme en témoignent entre autres les travaux de Waddington (1999), Martínez Melis (2001), Gile (2005) ou Collombat (2009), les apports potentiels des outils multimédias au processus d'évaluation des travaux des étudiants n'ont en revanche, à notre connaissance, pas encore été étudiés. L'étude présentée ici s'inscrit dans la lignée des travaux de Waddington et de Stannard (2008), et s'appuie comme les leurs sur une démarche de recherche-action, qui, selon Liu (1992: 294), « naît de la rencontre entre une volonté de changement et une intention de recherche ». Il s'agit ici d'une recherche-action expérimentale, dans laquelle, selon Resweber (1995: 43), « la situation est construite et l'efficacité de l'intervention est

soumise à la mesure et la comparaison. » Sur la base des travaux de Paivio (1971, 1990) sur la théorie du double codage, de Stannard (2008) sur les apports de la rétroaction vidéo en didactique des langues, et de Froeliger (2004) sur les mécanismes de la confiance en soi, nous formulons l'hypothèse que l'utilisation de moyens multimédias en évaluation didactique de la traduction peut améliorer la qualité des apprentissages. L'objectif de cette recherche est donc de vérifier la validité de cette hypothèse.

En termes plus spécifiques, notre étude vise d'une part à estimer de façon concrète les effets de la rétroaction vidéo (RV) de leurs travaux sur l'apprentissage des apprenants, et d'autre part à recueillir l'opinion de ces derniers quant à l'utilisation de cet outil, notamment en ce qui a trait aux aspects psychologiques du processus d'évaluation. En effet, la perception des apprenants vis-à-vis de ce processus constitue à notre avis un élément fondamental de l'efficacité de l'évaluation didactique. Pour ce qui est des effets de la RV, ils seront déterminés en comparant l'évolution des résultats obtenus par les étudiants d'un groupe test, qui auront bénéficié de la vidéo, et de ceux des membres d'un groupe témoin, qui auront reçu une rétroaction traditionnelle.

Par ailleurs, l'une des caractéristiques de la recherche-action réside dans l'applicabilité de ses résultats. Selon Goyette et Lessard-Hébert (1987: 31), celle-ci est en effet « axée sur la dynamique du changement, portant à la fois sur sa production concrète et sur l'étude de son processus ou de ses effets ». Dans cette optique, il nous apparaît indispensable que les outils mis en œuvre dans notre étude puissent être intégrés par des enseignants dans leurs pratiques d'évaluation. Ce caractère d'utilisabilité se traduit par le respect de trois critères: l'absence de complexité technique des outils mis en œuvre; l'optimisation du temps nécessaire à leur mise en œuvre; et le caractère raisonnable des coûts engagés.

2. Le potentiel cognitif des technologies

2.1. Les modèles d'apprentissage

Selon Depover *et al.* (voir 2007: 35), il est nécessaire, pour cerner le potentiel cognitif des technologies de l'information et de la communication (TIC), de prendre en compte le modèle d'apprentissage dans lequel leur utilisation s'inscrit. Ainsi, le modèle behavioriste, qui est basé sur la transmission des connaissances, s'attache à rendre celle-ci la plus efficace possible tout en contrôlant ses effets sur l'apprenant. Selon Salomon *et al.* (1991), il s'agira alors d'apprendre de l'ordinateur plutôt que d'apprendre avec l'ordinateur. En revanche, dans l'approche cognitiviste, l'apprenant participe activement à son apprentissage et utilise ses structures cognitives pour s'appropriier des connaissances nouvelles. Cependant, d'après Depover *et al.* (2007: 36), « pour voir réellement reconnu le potentiel cognitif des TIC, il [a fallu] attendre le développement des approches interactionnistes, qui fondent l'apprentissage sur les échanges entre l'individu et son environnement matériel mais aussi humain ».

Dans cette perspective constructiviste, les dispositifs technologiques ne sont plus centrés sur la transmission de contenus, mais ont plutôt pour objectif d'engager l'apprenant dans des processus cognitifs. Ainsi, la dimension humaine reste primordiale et l'usage de la technologie, réfléchi. Selon Depover *et al.* (2007: 15), la mise en œuvre de ces modèles « s'est traduite par l'émergence d'approches nouvelles de

l'enseignement dans lesquelles le rôle prépondérant de l'apprenant était explicitement reconnu».

2.2. *Les liens entre apprentissage et technologies*

Depover *et al.* (voir 2007: 13) recensent trois approches d'analyse des liens entre l'apprentissage et les TIC. La première repose sur l'idée que l'usage qui est fait par les enseignants des technologies dépend de leur conception de l'apprentissage. La deuxième confère aux technologies une fonction de vecteur de changement: leur introduction crée un déséquilibre cognitif qui contribue à favoriser l'évolution des pratiques et le passage à des modèles d'apprentissage qui privilégient l'implication des apprenants. Enfin, la troisième approche conjugue et nuance les deux précédentes en reconnaissant aux technologies un potentiel de changement sur les pratiques, mais qui ne pourra s'exprimer que si le contexte est favorable. Pour ce faire, il est nécessaire d'agir sur le milieu en accompagnant le changement afin de lever les principales résistances. C'est cette troisième approche que nous avons adoptée pour la présente étude en reconnaissant explicitement que le potentiel cognitif des TIC ne peut s'exprimer pleinement qu'avec la participation d'acteurs humains prêts à s'engager.

2.3. *La théorie du double codage*

L'idée d'exploiter les ressources multimédias de l'ordinateur pour offrir à l'apprenant des possibilités multiples d'accès à l'information constitue une caractéristique évidente des logiciels à vocation didactique. L'utilisation du multimédia permettrait de transmettre l'information plus efficacement en vertu de la théorie selon laquelle il existe un lien entre la pluralité des modes de communication (visuels et auditifs: images, sons, texte, etc.) et l'efficacité de l'apprentissage. Cette théorie du double codage a été proposée par Paivio (1971 et 1990), qui confère à l'activité d'imagerie des effets bénéfiques sur la mémoire verbale. Pour expliquer ces effets facilitateurs sur la rétention des informations, Paivio avance que deux sortes de systèmes interviennent dans la mémorisation: d'une part un système d'imagerie mentale liée à notre perception de l'environnement, et d'autre part le système verbal qui résulte de notre expérience du langage. Ainsi, selon Depover *et al.* (2007: 26-27), «la présentation simultanée d'informations à partir d'un texte oral et d'une animation ou d'un film conduit à un approfondissement de l'apprentissage parce qu'elle favorise les liens entre les deux systèmes de codage».

3. **Les débuts de l'utilisation de la rétroaction vidéo à des fins d'évaluation formative**

Il est possible de réaliser sur un ordinateur des captures d'écran statiques (sauvegarde d'une image figée de l'écran) ou dynamiques (enregistrement vidéo de l'ensemble des actions entreprises par l'utilisateur). L'utilisation des logiciels d'enregistrement vidéo à des fins de formation (principalement pour la prise en main d'outils informatiques, sous la forme de didacticiels) a connu un développement rapide avec l'essor d'Internet. Cependant, ce n'est qu'en 2007 que Stannard (2008) a réalisé un premier test de rétroaction vidéo (RV) à des fins d'évaluation formative. Le processus a été le suivant:

envoi électronique de leurs travaux par les étudiants; annotation des copies directement à l'écran; rétroaction audio-vidéo simultanée à l'annotation; puis envoi des fichiers vidéo aux étudiants en lieu et place des copies papier. Fondée sur les principes du double codage, la recherche-action de Stannard a suscité un réel engouement et la pratique de la RV a été reprise au Québec notamment par Lavoie (2010) et adaptée à l'enseignement de la traduction par Gardy (2011).

Selon Stannard (2008), les principaux enseignements de cette première expérience ont été, du point de vue de l'étudiant, l'appréciation des bénéfiques suivants :

- les aspects visuel et sonore de la rétroaction au lieu de l'écrit seul;
- la quantité de rétroaction fournie;
- la possibilité de réécouter à loisir la vidéo.

Du point de vue de l'enseignant, les conclusions portent sur les points suivants :

- la rétroaction très positive de la part des étudiants;
- l'importance de la préparation et la nécessité d'une prélecture des copies pour éviter les blancs lors de l'enregistrement vidéo;
- la quantité de l'information apportée. Selon Stannard (2008: 6), « You can speak around 200 words in a minute so a two minute video could be as much as 400 words »;
- le temps nécessaire au traitement des vidéos après l'enregistrement, variable en fonction de la durée de la vidéo et de la puissance de l'ordinateur;
- la nécessité d'une méthodologie quant au contenu de la vidéo et à la teneur de la rétroaction.

Les tests menés en 2010 auprès d'étudiants en traduction (Gardy, 2011 : 98-101) ont confirmé les conclusions de Stannard et mis en évidence la courbe d'expérience qui a permis de résoudre rapidement les écueils méthodologiques et techniques évoqués, notamment grâce au développement de la puissance des matériels et logiciels utilisés. C'est ce processus de correction que nous avons mis en œuvre pour les sujets du groupe test.

4. Description de la recherche

4.1. Présentation de la recherche et du corpus

Cette étude a été menée dans le cadre d'un cours universitaire d'introduction à la traduction, qui vise à permettre à l'étudiant de première année d'acquérir une méthode de travail adéquate fondée sur l'assimilation du métalangage de la traduction, l'exploitation efficace des ressources documentaires, le repérage et la caractérisation des principales difficultés récurrentes, ainsi que la justification des choix de traduction. Notre choix s'est porté sur cette population afin que les risques de biais liés à l'habitude de l'évaluation en milieu universitaire soient minimisés. Au cours de cette première session de la formation de ces futurs traducteurs, il s'agit de leur seul cours de traduction, les autres étant liés à des matières connexes. Ce cours compte 103 étudiants au début de la session, répartis en trois sections de taille équivalente: les sections A et B constituent le groupe témoin, et la section C le groupe test. L'enseignement est assuré par trois chargés de cours aux profils et à l'expérience similaires.

Sur le plan statistique, cet échantillon peut être considéré comme homogène. La technique d'échantillonnage est dite « aléatoire simple » (Depover, n.d. : 3), puisque

la répartition des étudiants entre les différentes sections résulte d'une procédure d'inscription qui s'apparente à un tirage aléatoire.

4.1.1. Cohérence et cohésion

La cohérence et la cohésion de l'enseignement constituent des paramètres essentiels pour la validité des résultats obtenus, au même titre que la représentativité de l'échantillon étudié. Ainsi, il aurait été possible de mener cette recherche au sein d'un seul et même groupe, pour lequel nous aurions effectué l'ensemble des évaluations. Cependant, chaque section comportant au plus 35 étudiants, les groupes test et témoin auraient compté de 15 à 20 sujets chacun, soit ce que Marien et Béaud (2003 : 14) considèrent comme des « micro-échantillons », dont la représentativité peut être remise en question.

Nous nous sommes assuré que le contexte d'enseignement était favorable à la minimisation des biais potentiels, tant dans sa dimension humaine que sur le plan matériel. Ainsi, notre connaissance des pratiques pédagogiques des enseignantes des sections A et B, ainsi que les liens de confiance qui nous unissent, ont permis la mise en place d'une atmosphère de coopération pleine et entière. Au cours de réunions préalables à l'étude, nous en avons exposé les objectifs et les paramètres, notamment la nécessité de dispenser de façon concertée un enseignement similaire dans les trois sections. Tous les matériels utilisés dans le cours ont été élaborés en commun et les aspects pratiques gérés à l'aide du portail numérique d'apprentissage de l'Université Laval, dont les fonctionnalités sont semblables à celles des plateformes d'apprentissage utilisées par la majorité des universités canadiennes. L'ensemble des apprenants a donc eu accès au site Internet du cours où figurait l'intégralité des informations et ressources afférentes (plan de cours, matériels de référence et d'apprentissage, forum, etc.). Ainsi, même si l'enseignement était assuré par trois personnes, la plateforme numérique était commune : tous les éléments utilisés en cours ou déposés sur le portail et toutes les informations fournies étaient identiques pour l'ensemble des étudiants des trois sections.

4.1.2. Modalités d'évaluation des étudiants

Dans une perspective constructiviste, il nous semble important d'évaluer les étudiants débutants sur leur maîtrise du processus de traduction et non sur leurs productions. Dans cette optique, les principaux modes d'évaluation sommative utilisés dans ce cours sont les suivants :

- 1) un exercice de repérage et de commentaire des difficultés présentées par un texte anglais de 244 mots ;
- 2) une première traduction de ce même texte, accompagnée de la justification des choix traductifs au regard des difficultés recensées lors de la première étape ;
- 3) une seconde traduction du texte, au regard des commentaires des enseignants sur la première ;
- 4) un examen de fin de session, qui se déroule en laboratoire, et qui consiste en la traduction d'un texte de 207 mots. Les étudiants ont accès à toutes les ressources de leur choix, la seule restriction concernant les moyens de communication.

Notre objectif étant d'étudier l'évolution des apprentissages des étudiants du groupe test par rapport à ceux du groupe témoin, notre corpus est composé des copies issues d'une part de la première traduction du devoir, et d'autre part de l'examen de

fin de session. Pour le devoir, les étudiants devaient traduire le texte ainsi que justifier leurs choix traductifs au regard des difficultés recensées précédemment. Les traductions ont été évaluées de façon critériée (respect du sens, qualité de la langue, adaptation au destinataire et qualité de la présentation) sur la base du barème en vigueur au sein du département. Cette méthode relève donc d'une analyse des erreurs. Quant aux commentaires rédigés par les étudiants, ce sont les critères de qualité de la langue, et de pertinence de la démarche et des sources, qui ont été utilisés. Quant au volet pratique de l'examen de fin de session (traduction d'un texte), il a été évalué selon les mêmes critères que ceux utilisés pour le devoir.

4.2. Méthodologie de la recherche

Selon Depover (n.d. : 3), l'utilisation de la recherche-action, qui consiste généralement à étudier un cas ou un nombre limité de cas particuliers sur une période de temps définie, conduit à mettre en évidence le concept de triangulation, qui « repose sur l'idée d'utiliser plusieurs méthodes de collecte de données, en vue d'étudier un phénomène » (Depover, n.d. : 7). Dans cette optique, nous avons choisi de recueillir des données quantitatives, mais aussi qualitatives, afin de pouvoir mesurer les effets de la RV sur certains paramètres comme la confiance, qui selon Froeliger (2004 : 50) constitue « un facteur essentiel de qualité en traduction ». Nous visons le respect des trois critères de qualité d'une recherche énoncés par De Ketele et Roegiers (2009 : 90) :

- la pertinence méthodologique (adéquation de la démarche, des instruments utilisés, choix des acteurs et du contexte) ;
- la validité méthodologique (adéquation suffisante entre, d'une part, les informations que nous déclarons récolter et traiter et, d'autre part, les informations effectivement recueillies et traitées) ;
- la fiabilité sur le plan méthodologique (les données récoltées et traitées sont-elles suffisamment indépendantes des auteurs de la recherche pour être compréhensibles et reproductibles par d'autres chercheurs ?).

4.2.1. Méthodologie de recueil des données quantitatives

Les copies des étudiants des sections A et B nous ont été transmises par leurs enseignantes avant qu'elles ne soient remises à leurs auteurs, afin que nous puissions les photocopier. Toutes les copies, aussi bien celles du groupe test que celles du groupe témoin, sont recorrigées pour assurer l'homogénéité du barème de correction entre les groupes. Une analyse de la typologie des erreurs relevées est ensuite effectuée pour les deux groupes afin de déterminer l'évolution de la nature et de la fréquence de ces erreurs. Par ailleurs, une liste de neuf risques spécifiques d'erreurs liées à des calques syntaxiques ou lexicaux potentiels a été préalablement définie, et leur occurrence vérifiée copie par copie. La majorité des copies corrigées se présentant sous forme papier, le dépouillement des données a été fait manuellement et les résultats reportés sur un tableau Excel qui permet la production automatique de données comparatives et statistiques.

4.2.2. Minimisation des biais statistiques

Selon Depover (n.d. : 3), l'élaboration du plan de collecte des données constitue « l'occasion d'une réflexion sur les erreurs susceptibles d'invalider les résultats ou du

moins de réduire leur validité». Certaines stratégies ont donc été mises en œuvre dans le but de minimiser ces risques, tant en amont du processus de collecte et de traitement des données que pendant son déroulement.

Les stratégies mises en œuvre en amont ont principalement été fondées sur des actions de communication visant à favoriser l'homogénéité des méthodes et des supports d'enseignement, ainsi que celle des processus d'évaluation. Ainsi, les trois enseignants se sont réunis pour convenir des modalités précises d'évaluation du devoir, en plus d'échanges fréquents par courriel. Un barème commun détaillant les principales erreurs rencontrées a été élaboré, puis progressivement enrichi tout au long du processus de correction, ce qui a facilité la mise en œuvre d'une approche homogène. Durant le processus de composition du corpus, nous nous sommes également assuré de l'homogénéité de la population étudiée. Pour ce faire, nous avons écarté de l'échantillon les étudiants allophones, ceux inscrits à la maîtrise, ainsi que ceux qui, pour une raison ou pour une autre, n'ont pas passé l'examen de fin de session. Ainsi, l'échantillon retraité comprend 88 sujets.

Par ailleurs, l'homogénéité des critères de correction et de l'application du barème constituant un élément crucial du processus, nous nous sommes efforcé de minimiser les risques liés au contexte de correction énoncés par Roberge (2006), notamment la fatigue. Pour éviter cet écueil susceptible de provoquer des variations dans l'évaluation des copies, nous nous sommes astreint à répartir le processus de recorection en y consacrant quotidiennement des plages horaires fixes de deux heures entrecoupées de pauses toutes les trente minutes. Selon Roberge, la fatigue et la surcharge cognitive qui en découle expliquent qu'un enseignant omette de corriger des erreurs. Par ailleurs, nous avons anonymisé les copies des étudiants du groupe test pour éviter l'effet Pygmalion, qui selon Roberge a pour effet qu'« un enseignant habitué à recevoir de bonnes copies d'un élève précis [...] considère que cet élève produit toujours de bons textes » (2006: 7).

Sur le plan informatique, nous avons eu recours aux fonctions de recherche avancée du logiciel Adobe Acrobat pour valider certaines des statistiques obtenues à la suite du relevé manuel des erreurs. Pour ce faire, nous avons téléchargé du portail l'original des copies constituant le corpus. Ces fichiers Word ont été enregistrés au format PDF afin d'éviter toute modification accidentelle de leur contenu, puis des recherches ciblées ont été menées sur ces fichiers PDF pour recenser de façon automatique et exhaustive toutes les occurrences d'un mot, d'un syntagme ou d'une caractéristique de présentation. Nous nous étions préalablement assuré lors de la phase de recorection de l'absence, pour les éléments étudiés, de fautes d'orthographe ou de présentation susceptibles d'affecter les résultats obtenus. L'utilisation de cet outil s'est avérée une aide précieuse pour la validation de l'exactitude des données saisies manuellement dans le tableau Excel. Pour ce qui est de l'enregistrement des RV, nous avons utilisé le logiciel Snagit¹.

4.2.3. Méthodologie de recueil des données qualitatives

La perception des étudiants du mode d'évaluation de leurs travaux constituant à notre avis un élément fondamental, nous avons soumis aux étudiants du groupe test un questionnaire d'enquête, passé à l'aide d'un site Internet spécialisé, préalablement, puis postérieurement à l'utilisation de la RV, afin de mesurer l'évolution de leur perception de l'évaluation et des apports de la RV. D'autres données ont été fournies

spontanément par les étudiants à l'occasion du processus d'évaluation du cours mené par la direction du département. Dans les deux cas, ces éléments qualitatifs ont été recueillis de façon anonyme.

5. Analyse des données quantitatives du corpus

Notre approche, suivie tant pour le devoir que pour l'examen, consiste d'une part à recenser de façon exhaustive les erreurs commises par les étudiants et à en dresser la typologie, et d'autre part à déterminer l'occurrence d'erreurs liées aux difficultés spécifiques mentionnées précédemment. La grille de correction habituellement utilisée par les enseignants en traduction de l'Université Laval comportant trois catégories (erreurs de transfert sémantique, erreurs de langue, erreurs de présentation) et 39 abréviations, nous avons volontairement, pour des raisons d'ordre pratique, restreint cette liste dans le cadre de notre recherche pour ne conserver que les éléments suivants :

- **sens** : toutes les erreurs de transmission du sens du texte source, du glissement au contresens ;
- **ajout ou omission** : tout ajout ou omission d'un mot ou d'un membre de phrase, sans pour autant que cela affecte le sens de façon significative. Si tel est le cas, l'erreur est comptabilisée dans la catégorie ci-dessus ;
- **calque** : toutes les erreurs résultant d'un calque du texte de départ, qu'il s'agisse d'un mot, d'un syntagme, d'une formulation peu idiomatique en français, d'un problème de structure ou de ponctuation, etc. ;
- **impropriété** : attribution à un mot d'un sens inexact ou contraire à l'usage ;
- **mal dit** : maladresse ou formulation boiteuse ;
- **orthographe et grammaire** : toutes les fautes liées à l'orthographe et à la grammaire ;
- **ponctuation** : toutes les erreurs de ponctuation qui ne résultent pas d'un calque ;
- **présentation** : toutes les erreurs de présentation qui ne résultent pas d'un calque.

Cette liste, volontairement limitative, ne reflète pas une quelconque pratique. Elle a été élaborée spécifiquement pour cette recherche dans le but de mettre en évidence les effets de la RV sur la récurrence des erreurs liées à des calques. Nous avons en effet constaté au fil des années que celles-ci, en raison du contexte linguistique nord-américain, constituent une source majeure d'erreurs chez les apprentis traducteurs canadiens francophones qui ont suivi nos cours.

5.1. Analyse des données quantitatives du devoir

Le texte du devoir, long de 244 mots, est extrait d'un communiqué de presse rédigé en anglais issu du site Internet de la société Nestlé, dans lequel il est question de l'inauguration d'une chaîne de production dans une usine du groupe. Ce site, principalement destiné aux médias et aux investisseurs du monde entier, reflète cependant la nationalité de cette société helvétique, puisque les sommes d'argent qui figurent dans le texte sont converties en francs suisses. Les étudiants ont reçu pour consigne de traduire pour le site Internet de Nestlé Canada.

5.1.1. Données statistiques relatives aux notes

Le Tableau 1 présente les données statistiques relatives aux notes brutes (c'est-à-dire celles attribuées par chaque enseignant) obtenues par les étudiants. On constate une forte homogénéité des deux sections qui composent le groupe témoin. En revanche, la moyenne du groupe test est supérieure de plus de trois points et demi à celle du groupe témoin.

TABLEAU 1
Statistiques des notes brutes du devoir

	groupe témoin			groupe test	écart test/ témoin
	section A	section B	moyenne	section C	
note la plus basse	55	53	54	44	-10
note la plus haute	87	89	88	88	0
moyenne	70,13	69,04	69,62	73,14	3,53
médiane	71,00	69,00	70,07	74,00	3,93
écart-type	8,33	9,16	8,72	10,95	2,23

Cependant, la phase de recorection des copies et d'harmonisation des méthodes d'évaluation a fourni des résultats légèrement différents résumés dans le Tableau 2.

TABLEAU 2
Statistiques des notes retraitées du devoir

	groupe témoin			groupe test	écart test/ témoin
	section A	section B	moyenne	section C	
note la plus basse	49	52	50	48	-2,5
note la plus haute	87	89	88	89	1
moyenne	69,41	68,43	68,95	71,75	2,80
médiane	70,00	69,00	69,53	72,00	2,47
écart-type	8,87	10,06	9,43	9,92	0,49

Ces variations ont entraîné une baisse de 0,67 point de la moyenne du groupe témoin, et de 1,39 point de celle du groupe test. Dans le premier cas, ces ajustements inférieurs à un point peuvent être considérés comme statistiquement négligeables. Dans le second cas, la baisse, qui est deux fois supérieure à celle du groupe témoin, est principalement imputable à l'attribution en plus grand nombre de points bonus pour des passages particulièrement réussis. La neutralisation de cette pratique sur les copies du groupe test explique l'évolution des notes. Le retraitement des copies a également entraîné un net resserrement de l'écart-type entre les groupes. Les différences de moyenne et de médiane, qui se sont également atténuées, s'expliquent par le fait que 21 % de la population du groupe test a obtenu une note supérieure ou égale à 80 %, contre seulement 12 % pour le groupe témoin.

En résumé, la faiblesse des écarts constatés entre notes brutes et notes retraitées s'explique à notre avis par la qualité du travail préparatoire au processus d'évaluation des copies, elle-même liée à la collaboration étroite des trois enseignants.

5.1.2. Recensement des erreurs commises lors du devoir

Le recensement des erreurs commises par les étudiants fait apparaître une homogénéité certaine des trois sections, ce qui semble logique au regard de leur composition aléatoire. Comme le montre le Tableau 3, le nombre moyen d'erreurs commises par les sujets du groupe témoin s'élève à 25,7, soit 6 % de plus que pour ceux du groupe test (moyenne de 24,2). La répartition par catégorie ne met pas en évidence de différence significative entre les deux groupes. Par ailleurs, les calques fautifs constituent dans les deux cas la principale source d'erreurs, soit respectivement 30 % et 34 % du nombre total².

TABLEAU 3

Nombre moyen et typologie des erreurs par copie (devoir)

	groupe témoin			groupe test	écart test/ témoin
	section A	section B	moyenne	section C	
sens	2,7	3,6	3,1	3,0	-0,1
ajout ou omission	0,8	1,2	1,0	1,0	-0,0
calque	8,0	7,6	7,8	8,3	0,4
impropriété	4,9	5,0	5,0	4,4	-0,6
mal dit	3,2	2,9	3,0	2,4	-0,6
orthographe et grammaire	2,7	2,5	2,6	2,2	-0,4
ponctuation	1,7	1,5	1,6	1,4	-0,2
présentation	1,9	1,4	1,7	1,6	-0,1
total	25,8	25,8	25,7	24,2	-1,5
% de calques	31 %	30 %	30 %	34 %	
% calques + impropriétés	50 %	49 %	50 %	52 %	

Cela confirme les observations de Larose (1998: 17) au sujet de traducteurs en exercice au Canada, pour lesquels « les impropriétés et les calques sont les [erreurs les] plus courantes. On associe ces erreurs à la trop grande littéralité et au manque de connaissance de la LA [langue d'arrivée] ». Ces résultats confirment également notre hypothèse énoncée précédemment. Par ailleurs, les impropriétés représentent bien la deuxième source d'erreurs, et ce, pour les deux groupes.

5.2. Analyse des données quantitatives de l'examen

Le texte d'une longueur de 207 mots utilisé pour l'examen est issu du site Internet du *Globe and Mail*. Il porte sur un sujet d'actualité internationale de 2012, à savoir l'approche des élections présidentielles en France. Les étudiants ont reçu pour consigne de traduire à destination d'un public canadien. Un lien menant au texte original leur a été fourni, ce qui leur permet de prendre connaissance de l'intégralité du texte d'origine, long de 669 mots, dont la lecture rapide peut fournir si nécessaire des précisions quant aux éléments à traduire. Ce texte présente des difficultés sémantiques moyennes et il y est principalement question des effets de la crise de l'euro sur les gouvernements européens. Cette situation, abondamment commentée dans les médias canadiens, avait déjà été abordée au cours de la session à l'occasion d'une traduction à vocation purement formative.

5.2.1. Données statistiques relatives aux notes

Le Tableau 4 présente les données statistiques relatives aux notes brutes obtenues à l'examen par les étudiants. On constate que l'écart entre les résultats des sections A et B s'est légèrement accentué, mais que ceux-ci restent cependant assez homogènes. Quant à l'écart entre les groupes test et témoin, il est légèrement inférieur à celui constaté pour le devoir, mais s'élève néanmoins à près de trois points.

TABLEAU 4
Statistiques des notes brutes de l'examen

	groupe témoin			groupe test	écart test/ témoin
	section A	section B	moyenne	section C	
note la plus basse	54	58	56	53	-3
note la plus haute	84	90	87	93	6
moyenne	68,72	70,71	69,65	72,54	2,89
médiane	68,50	71,50	69,90	73,00	3,10
écart-type	9,38	8,16	8,81	11,08	2,27

Contrairement à ce qui s'est passé pour le devoir, la phase de recorection et d'harmonisation a cette fois donné lieu à des écarts relativement importants, présentés dans le Tableau 5.

TABLEAU 5
Statistiques des notes retraitées de l'examen

	groupe témoin			groupe test	écart test/ témoin
	section A	section B	moyenne	section C	
note la plus basse	45	53	49	52	3
note la plus haute	83	88	85,5	93	7,5
moyenne	62,97	66,71	64,72	70,75	6,03
médiane	63,50	66,00	64,67	72,50	7,83
écart-type	10,44	9,92	10,20	11,30	1,10

Ainsi, la moyenne des notes du groupe témoin est en recul de près de cinq points, alors que celle du groupe test ne baisse que de moins de deux points. La dispersion des notes dans les deux sections du groupe témoin s'est accentuée, tandis que celle du groupe test est restée quasiment stable. L'analyse des écarts observés après la phase de recorection nous a permis de mettre en évidence trois facteurs permettant d'expliquer cette évolution :

- en raison de la date de l'examen, placé en toute fin de session, les enseignants n'ont pas été en mesure de se réunir pour préparer la correction comme cela avait été le cas pour le devoir. Même si le courriel facilite la communication, nous sommes d'avis que la dynamique issue d'une rencontre « physique » est propice à la fécondité des échanges ;
- corollaire de ce qui précède, certaines erreurs n'ont pas été recensées dans le barème et d'autres n'ont pas donné lieu à un consensus. Ainsi, le calque lexical *coupures budgétaires* a été accepté pour les étudiants du groupe témoin, mais sanctionné pour ceux du groupe test. Cela a également été le cas pour l'anglicisme *critiques*, employé au sens de *détracteurs*. À elle seule, l'harmonisation de ces deux erreurs explique un écart de 1,45 point entre la moyenne brute et la moyenne retraitée, soit presque la

- moitié de la variation constatée entre groupe test et groupe témoin ;
- enfin, il est possible que des facteurs physiologiques et psychologiques mis en avant par Roberge, par exemple la fatigue liée à la fin de session, aient eu des effets négatifs sur l'homogénéité de l'application des éléments du barème.

5.2.2. Recensement des erreurs commises lors de l'examen

Le recensement du nombre d'erreurs commises par les étudiants fait apparaître, comme c'était le cas pour le devoir, une grande homogénéité des deux sections qui composent le groupe témoin. Pour ces étudiants, les calques fautifs représentent toujours la première source d'erreurs, dans des proportions comparables à celles du devoir.

TABLEAU 6
Nombre moyen et typologie des erreurs par copie (examen)

	groupe témoin			groupe test	écart test/ témoin
	section A	section B	moyenne	section C	
sens	6,3	6,0	6,2	4,7	-0,1
ajout ou omission	1,4	1,3	1,4	2,4	0,0
calque	7,0	7,8	7,4	5,8	0,4
impropriété	4,0	3,2	3,6	4,3	-0,6
mal dit	2,5	2,9	2,7	2,5	-0,6
orthographe et grammaire	1,9	2,3	2,1	2,4	-0,4
ponctuation	0,9	1,0	1,0	0,5	-0,2
présentation	1,6	1,2	1,4	1,1	-0,1
total	25,7	25,6	25,6	23,8	-1,9
% de calques	27 %	30 %	29 %	24 %	
% calques + impropriétés	43 %	43 %	43 %	42 %	

Si l'on constate également une relative stabilité du nombre d'erreurs commises par les étudiants du groupe test, leur typologie a en revanche évolué (Tableau 7). Ainsi, le nombre de calques est en recul sensible, passant de 8,3 à 5,8 par copie³, ce qui équivaut à une baisse de près de 30 %. Ainsi, pour le groupe test, alors que 34 % des erreurs commises lors du devoir étaient imputables à des calques, cette proportion a chuté à 24 % pour l'examen.

TABLEAU 7
Évolution de la typologie des erreurs examen/devoir

	groupe témoin			groupe test	écart test/ témoin
	section A	section B	moyenne	section C	
sens	3,7	2,4	3,1	1,7	-1,4
ajout ou omission	0,7	0,1	0,4	1,4	1,0
calque	-1,0	0,1	-0,5	-2,5	-2,0
impropriété	-0,9	-1,8	-1,3	-0,1	1,2
mal dit	-0,7	0,0	-0,3	0,1	0,5
orthographe et grammaire	-0,8	-0,2	-0,5	0,2	0,7
ponctuation	-0,8	-0,5	-0,6	-0,9	-0,2
présentation	-0,3	-0,3	-0,3	-0,5	-0,3
total	-0,1	0,0	-0,1	-0,4	-0,4

Les erreurs d'ordre sémantique sont également moins nombreuses dans les copies du groupe test (4,7 en moyenne contre 6,2 pour le groupe témoin). En revanche, celles liées à des ajouts ou omissions et à des impropriétés affichent une tendance inverse. Alors que pour le devoir, groupe témoin et groupe test présentaient un pourcentage identique d'occurrence des erreurs sélectionnées, les étudiants de ce dernier en ont commis à l'examen 21 % de moins que leurs homologues (voir Tableau 9 *infra*). Certaines de ces difficultés ont été traitées correctement par les deux groupes exactement dans les mêmes proportions. Il s'agit de deux cas de participe présent et de deux risques de calque pouvant amener des anglicismes lexicaux. En revanche, les cas de comparatifs elliptiques ont été moins fréquents dans le groupe test (13 % moins d'occurrences en moyenne), ainsi que ceux de pléonasmes. Enfin, aucun calque fautif de *with* par *avec* n'a été recensé dans les copies du groupe test, alors qu'un quart des étudiants du groupe témoin n'ont pas évité cette difficulté.

5.3. Analyse des écarts entre devoir et examen

Les étudiants des groupes test et témoin ont obtenu des résultats homogènes lors du devoir : comme le montre le Tableau 8, la proportion d'erreurs liées aux difficultés sélectionnées est identique entre les deux groupes, à savoir 48 %, et l'écart des nombres moyens d'erreurs ne dépasse pas 6 %. Seule la pondération des calques au sein de la typologie des erreurs diffère, mais ces valeurs (31 % et 34 %) restent cependant relativement proches.

TABLEAU 8

Comparaison statistique résumée (devoir)

	groupe témoin	groupe test	écart test/témoin en %
	sections A et B	section C	
nombre moyen d'erreurs	25,66	24,21	-6 %
% de calques	30,5 %	33,8 %	11 %
% d'erreurs de sens	12,1 %	12,4 %	2 %
% erreurs sélectionnées	48,3 %	47,6 %	-1 %
moyenne retraitée	68,95	71,75	4 %

En revanche, la typologie des erreurs commises à l'examen présente des évolutions plus marquées entre les deux groupes (Tableau 9). Ainsi, alors que les étudiants du groupe test avaient commis au devoir 11 % de calques de plus que leurs homologues, cette tendance a fait plus que s'inverser à l'examen. Le différentiel d'évolution atteint ainsi 26 %. Parallèlement, le nombre d'erreurs de sens non imputables à des calques est également en baisse marquée dans le groupe test.

TABLEAU 9

Comparaison statistique résumée (examen)

	groupe témoin	groupe test	écart test/témoin en %
	sections A et B	section C	
nombre moyen d'erreurs	25,64	23,78	-7 %
% de calques	28,7 %	24,4 %	-15 %
% d'erreurs de sens	24,0 %	19,8 %	-18 %
% erreurs sélectionnées	40,2 %	31,8 %	-21 %
moyenne retraitée	64,72	70,75	9 %

Malgré les précautions prises lors de la préparation de cette étude, et notamment l'accent mis sur l'homogénéisation de l'enseignement dans les trois sections, ainsi que le soin apporté aux procédures de recorection et d'harmonisation des copies, plusieurs facteurs peuvent expliquer les évolutions constatées, par exemple :

- le choix des textes : même si tous les étudiants sont logés à la même enseigne, il se peut que ceux férus d'actualité internationale aient été avantagés lors de l'examen ;
- le stress lié à l'examen : ce facteur, clairement identifié par les étudiants, peut expliquer une contre-performance lors d'une épreuve en temps limité ;
- les styles d'enseignement : même si la matière présentée et les supports informatiques utilisés sont identiques, la personnalité et la sensibilité de chaque enseignant peuvent influencer, de façon perceptible ou non, sur ses méthodes d'enseignement.

Cependant, la probabilité que les deux premiers de ces facteurs aient exercé une influence plus grande sur les étudiants du groupe témoin que sur ceux du groupe test n'est pas fondée statistiquement. L'analyse des résultats du devoir infirme par ailleurs l'hypothèse d'une hétérogénéité des groupes. Ainsi, l'utilisation de la RV, qui représente dans notre étude la principale variable du contexte d'enseignement, peut constituer une explication des variations constatées. Ses effets seront détaillés lors de l'analyse des données qualitatives du corpus.

6. Analyse des données qualitatives du corpus

6.1. Analyse de la perception de l'évaluation

Afin de déterminer les effets de l'utilisation de la RV sur la perception que les étudiants ont de l'évaluation, nous les avons interrogés par questionnaire en début de session, puis après la fin de celle-ci.

6.1.1. Perception préalable de l'évaluation

Le groupe test était composé de 33 personnes au début de la session. Selon De Ketele et Roegiers (2009 : 145), « dans le cadre d'une *enquête par questionnaire*, c'est surtout la représentativité de l'échantillon qui déterminera le caractère systématique du recueil d'informations ». C'est pourquoi nous avons sensibilisé les étudiants à l'importance d'un taux de réponse aussi élevé que possible pour assurer cette représentativité et éviter le besoin de méthode de redressement de l'échantillon. Ce questionnaire a été passé de façon anonyme sur Internet et a recueilli 29 réponses (Tableau 10), ce qui correspond à 88 % des étudiants du groupe test. Ils devaient répondre à la question

« Pour moi, l'évaluation (les évaluations que je reçois en tant qu'étudiant), c'est : » en indiquant leur degré d'accord avec cinq assertions.

TABLEAU 10

Résultats du questionnaire pré-test

	pas d'accord du tout	plutôt pas d'accord	d'accord	tout à fait d'accord	Total
un mal nécessaire	6,90% 2	20,69% 6	37,93% 11	34,48% 10	29
une source de stress	0,00% 0	13,79% 4	65,52% 19	20,69% 6	29
un moyen de savoir ou j'en suis	0,00% 0	3,45% 1	48,28% 14	48,28% 14	29
quelque chose d'utile	3,45% 1	3,45% 1	62,07% 18	31,03% 9	29
un outil d'apprentissage à part entière	0,00% 0	27,59% 8	41,38% 12	31,03% 9	29

On constate que les étudiants ont une vision positive de l'évaluation : pour 93 % d'entre eux, l'évaluation est utile ou très utile, et elle constitue pour 72 % un outil d'apprentissage à part entière. Elle représente également pour la quasi-totalité des répondants (97 %) un bon moyen de se situer dans ses apprentissages. Cependant, elle est vécue comme un mal nécessaire par 72 % des apprenants, principalement en raison du stress occasionné par ce processus. Ainsi, plus de huit étudiants sur dix indiquent s'en ressentir modérément, et plus d'un sur cinq de façon intense. Les étudiants avaient également la possibilité de laisser des commentaires, qui sont retranscrits intégralement ici :

Bien qu'une évaluation devrait m'aider à apprendre de mes erreurs en bonne partie, le stress qui vient avec [*sic*] et surtout la note sur le relevé m'empêchent de voir les évaluations comme des outils d'apprentissage très efficaces...

L'évaluation me permet de savoir si j'ai bien compris la matière, et de connaître ce que je maîtrise moins bien. C'est aussi une façon de savoir si mes efforts valent la peine.

Je crois qu'il est important de faire plusieurs évaluations au cours d'une session ; même si cela peut représenter une grande source de stress, elle permet de faire le point sur une partie de matière... mais bon, ce n'est pas un test qui peut nous dire exactement ou [*sic*] on en est rendu et une note ne peut à mon avis pas tout le temps nous éclairer adéquatement sur la réalité, bien que c'est forcément ce que l'on a tendance à regarder en premier...

Les évaluations me stressent énormément, mais au moins ça me permet de voir si j'ai bien compris et quelles sont mes difficultés.

C'est quelque chose d'utile, mais en même temps il y a beaucoup de stress.

Il se dégage de ces commentaires à la fois la compréhension de la nécessité de l'évaluation et la mise en avant du facteur stress. En effet, il est évoqué dans quatre des cinq commentaires. L'ambivalence des étudiants envers l'évaluation est donc nette.

6.1.2. Analyse de la perception postérieure à la recherche

En raison du retraitement de l'échantillon, le questionnaire postérieur à l'utilisation de la RV a été soumis à 28 étudiants après la fin de la session et a obtenu un taux de réponse de 86 %. Le nombre de questions a été volontairement limité à dix de façon à réduire le temps nécessaire pour y répondre et donc à maximiser les chances d'une participation élevée. En effet, Marien et Béaud (2003: 14) considèrent qu'un « micro-échantillon », soit moins de 20 participants, peut susciter des doutes quant à sa représentativité. La première question est identique à celle posée préalablement à la recherche, et les réponses du Tableau 11 peuvent être directement comparées à celles du tableau 10.

TABLEAU 11

Résultats question 1 du questionnaire post-test

	pas d'accord du tout	plutôt pas d'accord	d'accord	tout à fait d'accord	Total
un mal nécessaire	20,83% 5	20,83% 5	37,50% 9	20,83% 5	24
une source de stress	4,17% 1	25,00% 6	54,17% 13	16,67% 4	24
un moyen de savoir ou j'en suis	0,00% 0	0,00% 0	33,33% 8	66,67% 16	24
quelque chose d'utile	0,00% 0	0,00% 0	20,83% 5	79,17% 19	24
un outil d'apprentissage à part entière	0,00% 0	0,00% 0	29,17% 7	70,83% 17	24

Ainsi, si la perception des étudiants quant à l'utilité de l'évaluation et son intérêt en tant que repère dans l'apprentissage n'évolue pas de façon significative lorsqu'on additionne les réponses positives (*d'accord*) et très positives (*tout à fait d'accord*), un net transfert des premières vers les secondes peut cependant être observé, notamment pour ce qui relève de l'utilité de l'évaluation. L'évolution la plus nette concerne le critère de l'évaluation en tant qu'outil d'apprentissage à part entière. En effet, les perceptions positives ou très positives à cet égard passent de 72 à 100 %, dont 40 % de hausse pour la seconde catégorie. Quant aux deux critères liés aux aspects négatifs de l'évaluation, leur évolution est moins marquée, néanmoins réelle, puisque les répondants ne sont plus que 58 % à considérer l'évaluation comme un mal nécessaire (contre 72 % auparavant) et 70 % à y voir une source de stress. Même si ce pourcentage reste élevé, il est en baisse de 16 % par rapport au premier questionnaire. Cependant, plusieurs commentaires des étudiants relativisent l'importance attribuée au stress, ce qui n'était pas le cas lors du questionnaire pré-test :

Pour moi, une bonne évaluation, même si c'est une source de stress, me renseigne sur mes forces et mes faiblesses.

Oui, les évaluations peuvent être une source de stress, mais sans elles, nous ne pourrions pas apprendre de nos erreurs. La théorie ne dit pas tout. Il faut essayer par nous-mêmes et faire des choix, et ensuite voir si c'étaient de bons ou mauvais choix.

En conclusion, il apparaît que l'utilisation de la RV joue un rôle positif dans l'évolution de la perception que les étudiants ont de l'évaluation, principalement en renforçant le sentiment de son utilité et de son statut d'outil d'apprentissage à part entière.

6.2. Analyse de la perception des apports de la rétroaction vidéo

Le questionnaire proposé aux étudiants comportait neuf questions relatives à l'utilisation de la RV et il leur était possible d'ajouter des commentaires après chaque question. La première d'entre elles consistait à indiquer leur degré d'accord au regard de quatre assertions (tableau 12), chaque degré renvoyant à une valeur numérique croissante comprise entre un et cinq. L'évaluation moyenne (colonne de droite du tableau) correspond donc à un coefficient d'appréciation mesuré sur une échelle de un à cinq.

TABLEAU 12

Perception des apports de la RV

	pas d'accord du tout	plutôt pas d'accord	ni ni	d'accord	tout à fait d'accord	Total	Évaluation moyenne
la RV m'incite à analyser mes erreurs	0% 0	4,17% 1	4,17% 1	25% 6	66,67% 16	24	4,54
la RV me permet de mieux comprendre mes erreurs	0% 0	0% 0	0% 0	33,33% 8	66,67% 16	24	4,67
la RV me permet d'éviter certains pièges récurrents	0% 0	0% 0	4,17% 1	37,50% 9	58,33% 14	24	4,54
la RV me permet d'apprendre de mes erreurs	0% 0	0% 0	0% 0	20,83% 5	79,17% 19	24	4,79

Ces quatre assertions visent ainsi à appréhender, outre la perception des étudiants, leurs attitudes et comportements postérieurement à la réception de leur copie corrigée, notamment au regard de certains des travers du processus d'évaluation souvent mentionnés par la documentation, à savoir le sentiment d'arbitraire ressenti par les étudiants, leur fréquente incompréhension des commentaires qui leur sont adressés et leur passivité relativement à la rétroaction. Il ressort des réponses des étudiants que tous sont d'avis que la RV leur permet de mieux comprendre leurs erreurs que dans un contexte de rétroaction traditionnelle. Pour 92 % d'entre eux, la RV les incite également à analyser leurs erreurs, ce qui selon Robinson *et al.* (2013: 261) constitue un facteur prédictif de la satisfaction des étudiants: «if feedback is to be effective it must engage the student and develop self-regulated learning [...]. Poor satisfaction with feedback is likely to occur if students see the feedback as an end in itself and do not work independently with the feedback provided to improving their performance.» Par ailleurs, 96 % des étudiants considèrent que la RV leur permet d'éviter certains pièges récurrents. Cette perception confirme les résultats de l'analyse quantitative qui a mis en évidence une nette diminution des erreurs liées aux calques. Enfin, la totalité des étudiants sont d'avis que la RV leur permet d'apprendre de leurs erreurs, dont 79 % qui sont tout à fait d'accord avec cette assertion. Ces résultats confirment les travaux de Nicol et Macfarlane-Dick (2006: 204), pour qui «there is considerable research evidence to show that effective feedback leads to learning gains».

D'autres aspects ont été évoqués par les étudiants dans leurs commentaires, notamment la démythification, du point de vue de l'étudiant, de l'acte d'évaluation:

Plus interactive qu'une correction manuelle, elle [la RV] donne un bon aperçu de la démarche de correction entreprise par les professeurs. Les commentaires oraux sont argumentés et plus compréhensibles que les erreurs de traduction étiquetées dans une correction manuelle.

La notion de lisibilité exprimée dans ce commentaire rejoint une des préoccupations exprimées par Robinson *et al.* dans l'étude qu'ils ont menée auprès d'étudiants universitaires du premier cycle. Ils ont en effet mis en évidence que 30 % d'entre eux éprouvaient parfois des difficultés à déchiffrer les commentaires des enseignants. Selon ces auteurs, « staff should be aware that increasing reliance on technology amongst students means that they are less experienced at deciphering handwriting » (2013 : 267). Au vu des différentes contraintes qu'ils évoquent, Robinson *et al.* suggèrent la mise en œuvre et l'évaluation de nouvelles méthodes de rétroaction, position à laquelle nous ne pouvons que souscrire :

Evaluation of new methods of giving feedback which allow students to receive help on how to use their feedback without experiencing the emotional or logistical problems associated with organising a face to face meeting with their tutors is needed. In addition, new technologies such as video and podcasts allow staff to give fuller comments, in a shorter time, without raising problems of legibility [...] and so should be a good method of increasing student satisfaction with feedback in a time effective manner.

(Robinson *et al.*, 2013 : 269)

6.2.1. Aspects comportementaux

Selon Veslin et Veslin (1992 : 23), « la correction des copies, pour un certain nombre d'élèves, ne représente pas un moment particulièrement lumineux de la situation pédagogique [...]. Il arrive que le devoir ne sera même pas relu. » Ainsi, les résultats d'un sondage⁴ que nous avons mené en 2011 auprès des enseignants de notre département ont mis en évidence que lorsqu'un examen de fin de session se tient sur papier et que les copies sont mises à disposition des étudiants à un guichet accessible plus de 37 heures par semaine, seulement 40 % d'entre eux vont y récupérer la leur. Cette pratique apparaît comme un corollaire aussi logique que révélateur de l'importance attachée par les étudiants à la seule note. Si le développement progressif de la remise des notes en ligne fera évoluer ces statistiques, il n'est cependant pas acquis qu'il modifie de façon perceptible le comportement des étudiants. En revanche, il apparaît que l'utilisation de la RV engendre une telle modification. Ainsi, le Tableau 13 résume les réponses fournies par les étudiants du groupe test à la question : « Lorsque je reçois ma rétroaction, quels sont les fichiers que j'ouvre ? »

TABLEAU 13

Comportement des étudiants à la réception des fichiers

Choix de réponses	Réponses	
les deux fichiers	87,50%	21
le fichier Word seulement	0%	0
le fichier vidéo seulement	12,50%	3
d'abord le fichier Word	58,33%	14
d'abord le fichier vidéo	41,67%	10
Nombre total de répondants : 24		

À la réception des fichiers, 42 % des répondants déclarent ouvrir d'abord le fichier vidéo. Il ressort également des commentaires reçus que le centre d'intérêt d'une proportion non négligeable des étudiants s'est déplacé de la note vers les éléments de la rétroaction, puisque la note n'apparaît qu'à la fin de la vidéo, alors qu'elle figure en lecture directe et immédiate sur le fichier Word. *A contrario*, selon une nouvelle étude (Gardy, 2015) dans laquelle 357 étudiants de dix universités canadiennes ont été interrogés, 90 % des répondants ont indiqué s'intéresser d'abord à la note.

TABLEAU 14
Utilisation ultérieure des vidéos

Choix de réponses	Réponses	
jamais	8,33%	2
parfois	50%	12
souvent	25%	6
toujours	16,67%	4
Total		24

Par ailleurs, alors que les enseignants déplorent souvent que la note monopolise l'attention des étudiants et que la copie remise à l'apprenant finisse rapidement au fond d'un tiroir – voire de la poubelle –, dans notre étude, 92 % des répondants ont affirmé réécouter parfois ou souvent la vidéo (tableau 14). Le format numérique de la rétroaction en favorise donc l'accessibilité et le partage. Cette caractéristique s'avère très intéressante pour les étudiants qui résident loin de leur lieu d'études et qui sont déjà rentrés chez eux au moment de la remise des examens finaux, et *a fortiori* pour ceux qui suivent des cours à distance.

6.2.2. Autres aspects de la rétroaction vidéo

Selon Gardy (2015), dans près de 90 % des cas, les étudiants en traduction reçoivent leur copie pendant le cours, le plus souvent au début de celui-ci. Outre le stress généré par la découverte de la note, cette pratique ne facilite pas l'attention des apprenants. En effet, comme nous avons pu le constater en observant à quatre occasions le comportement d'étudiants lors de la remise de leurs travaux, il s'écoule en moyenne plus de huit minutes avant que l'enseignant n'arrive à capter l'attention de tous. Par ailleurs, le temps nécessaire à la distribution des copies diminue d'autant la durée de la rétroaction.

Si l'utilisation de la RV permet d'optimiser la gestion du temps lors du cours et d'éviter les pertes d'attention, elle rend également possible une modification du cours, tant dans son contenu que dans sa dynamique. En effet, les informations recueillies auprès des enseignants (Gardy, 2015) mettent en évidence que la rétroaction des devoirs et examens faite en cours se limite souvent à un commentaire général des difficultés rencontrées. La RV quant à elle permet aux étudiants de prendre connaissance de ces commentaires préalablement au cours, ce qui libère du temps qui peut être utilisé dans une vision autre que rétrospective. Ainsi, au lieu de consister en une énumération des erreurs rencontrées, le cours peut être l'occasion d'aborder différentes méthodes de résolution de problèmes et de faire réfléchir les étudiants sur leur

approche méthodologique. Il en résulte une participation accrue de leur part, d'autant plus que ceux qui n'ont pas obtenu une note satisfaisante ont eu le temps nécessaire à l'évacuation des sentiments négatifs qui auraient pu s'ensuivre.

6.3. *Évolution de la perception de l'évaluation*

La dernière question posée aux étudiants visait à déterminer si l'utilisation de la RV avait modifié leur perception de l'évaluation. Il s'agit de la seule question ouverte du questionnaire, dans le but d'encourager les étudiants à préciser leur pensée au-delà d'une simple réponse binaire. Sur les 24 étudiants, 19 (soit 79%) ont répondu par l'affirmative, dont 3 par un simple *oui*, et 5 par la négative (dont 2 par un simple *non*). L'analyse des commentaires liés aux réponses négatives fait apparaître qu'un des répondants considérait déjà l'évaluation comme un outil d'apprentissage :

Je voyais déjà les corrections comme une façon de confirmer mes connaissances ou de corriger le tir.

Par ailleurs, les deux autres commentaires reconnaissent explicitement les avantages de la RV :

Non, mais elle permet de mieux comprendre les erreurs commises et de s'y référer dans le futur pour éviter les mêmes erreurs.

Non, en ce qu'elle joue le même rôle qu'une rétroaction papier. La RV est tout de même de loin supérieure à la correction classique.

L'analyse des réponses fournies aux questions précédentes par les deux répondants qui se sont limités à un *non* à celle-ci fait apparaître qu'ils avaient répondu *tout à fait d'accord* aux questions portant sur leur perception des apports de la RV. Par ailleurs, les commentaires écrits en complément à d'autres de leurs réponses indiquent qu'eux aussi considéraient déjà antérieurement l'évaluation comme un outil d'apprentissage :

Ce n'est pas le résultat chiffré qui m'intéresse, mais le fait de voir et de comprendre mes erreurs afin de ne plus les refaire et, ainsi, de m'améliorer.

Plus interactive qu'une correction manuelle, elle donne un bon aperçu de la démarche de correction entreprise par les professeurs. Les commentaires oraux sont argumentés et plus compréhensibles que les erreurs de traduction étiquetées dans une correction manuelle.

L'analyse des commentaires liés aux réponses positives à la question posée met en évidence quatre principaux éléments de satisfaction : une meilleure compréhension des erreurs, la personnalisation de la rétroaction, l'accent mis sur la méthodologie et la dédramatisation de la situation d'évaluation, les deux premiers étant étroitement liés.

Certains des types d'erreurs sont parfois un peu nébuleux donc, quand elles nous sont expliquées, c'est très formateur. De plus, je ne suis pas le genre d'étudiante à consulter les professeurs quand je ne comprends pas une évaluation alors qu'ici, j'ai toute l'information nécessaire.

Cette méthode me donne vraiment l'impression que l'évaluation me sert d'apprentissage, contrairement aux évaluations pour lesquelles nous ne recevons qu'une note sur la première page sans le moindre commentaire. Parce qu'elle est personnalisée, la rétroaction vidéo m'apparaît encore plus utile que le retour sur l'évaluation effectué en classe.

Oui, car elle permet de comprendre la manière correcte d'aborder une phrase qui nous a posé problème.

Oui. Cette rétroaction représente une sorte de rencontre un à un [*sic*] avec le chargé de cours.

Par ailleurs, l'idée d'une baisse du stress lié à réception de la note a été évoquée, ainsi que la proximité avec l'enseignant que crée la vidéo et qui tend à dédramatiser la situation d'évaluation. La perception des étudiants de l'évaluation s'en trouve donc modifiée et ils la vivent moins comme une sanction que comme un outil.

De cette façon j'ai moins l'impression d'être évaluée. J'ai plus l'impression d'apprendre. C'est un peu comme si je prenais [*sic*] un cours privé [*sic*] de traduction. Je reçois des commentaires pertinents pour mon travail à moi et non pour l'ensemble du travail des étudiants.

Oui. Quand on reçoit un devoir corrigé et qu'on y voit des erreurs, on se sent pas bons. Mais grâce au [*sic*] vidéo, on comprend nos erreurs. Je vois mon évaluation comme un apprentissage et non une critique.

Enfin, l'arbitraire lié à la correction, qui dans une étude de Veslin et Veslin (voir 1992: 25) était mis en avant par plus de la moitié des apprenants, ou la subjectivité, évoquée par près de 43 % des étudiants en traduction (Gardy, 2015), ne sont plus mentionnés en tant que source de frustration.

Cela montre que les marques de correction sont toutes justifiables et moins arbitraires qu'on peut parfois le penser. Généralement, quand on a les explications, on accepte mieux la note au bas de la page.

En conclusion, il apparaît que pour une très large majorité des répondants, l'utilisation de la RV a modifié leur perception du processus d'évaluation et de son utilité en tant qu'outil d'apprentissage, en jouant sur des aspects aussi bien techniques (compréhension des erreurs et lisibilité des éventuels commentaires) que psychologiques (dédramatisation de l'acte d'évaluation et diminution du sentiment d'arbitraire). L'utilisation de moyens multimédias, qui permet une diversité d'approches propres à soutenir l'attention, est également valorisée par les apprenants, comme le rappelle Hamon :

L'apport majeur des TICE semble devoir être conçu comme allant bien [*sic*] delà de la seule dimension pratique, puisque ces instruments impliquent non seulement une incrémentation de la motivation mais tendent aussi à agir sur l'appropriation de connaissances et de compétences. (Hamon, 2013: 559)

7. Conclusion

7.1. La rétroaction vidéo: des apports mesurables

L'utilisation d'outils technopédagogiques aux fins d'évaluation des travaux d'étudiants en traduction, si elle tend à se développer, reste encore minoritaire et souvent limitée au seul usage de la fonction *Commentaire* de Word. Indépendamment des procédures d'évaluation utilisées, il apparaît que l'usage raisonné d'outils multimédias, en l'occurrence la rétroaction vidéo, est susceptible d'améliorer l'efficacité de l'évaluation envisagée dans une approche socioconstructiviste. Ainsi, la comparaison

de l'évolution des résultats obtenus par les sujets d'un groupe test par rapport à ceux d'un groupe témoin met en évidence une amélioration perceptible des résultats obtenus par les premiers. Cependant, d'autres études à plus grande échelle seraient nécessaires pour confirmer la validité des résultats quantitatifs, qui peuvent avoir été affectés par des biais inhérents au protocole adopté. En revanche, ces derniers n'ont pas d'influence sur les apports qualitatifs de l'utilisation de la RV mis en avant par les étudiants. Ainsi, il apparaît que la RV répond aux besoins et aux attentes des étudiants mieux que ne peuvent le faire les méthodes traditionnelles. Elle répond également à plusieurs des critères de l'évaluation formative, à savoir «la création d'un climat de confiance, l'individualisation de l'aide apportée, l'auto-évaluation, la responsabilisation de l'apprenant et la régulation des apprentissages» (Gardy, 2011 : 100).

La rétroaction vidéo entraîne également des modifications sur le déroulement des cours en classe. En effet, lors de la remise des copies papier de devoir ou d'examen aux étudiants, il est fréquent que des éléments de correction soient longuement apportés par l'enseignant et que des questions, voire des réclamations, soient soulevées par les étudiants. Au Canada, où la session universitaire est de quinze semaines, le temps consacré à l'évaluation sommative peut représenter l'équivalent de trois cours (deux cours consacrés aux épreuves d'examen et un, en cumulé, à la correction des travaux). La RV, qui évite d'avoir à revenir en classe sur la litanie des erreurs rencontrées lors de la correction, permet de libérer du temps pour des activités plus formatives. Ainsi, les séances de correction, où il n'y a plus de remise physique des copies ni de baisse générale de l'attention qui va souvent de pair, se concentrent désormais sur les approches méthodologiques et des échanges apaisés, même s'ils peuvent être passionnés, entre étudiants et enseignant.

7.2. Les contraintes liées à l'utilisation de la rétroaction vidéo

Si les avantages liés à la rétroaction vidéo sont manifestes, le développement de son utilisation se heurte à des contraintes tant matérielles qu'humaines. Les aspects matériels n'ont cependant rien d'insurmontable, puisque l'équipement requis se limite à un micro de qualité correcte et à un logiciel d'enregistrement vidéo, ce qui représente un investissement total d'environ 50 dollars. L'installation et le fonctionnement du logiciel ne nécessitent aucune compétence informatique particulière.

La contrainte de temps représente en revanche un obstacle majeur au développement de la RV. En effet, le temps nécessaire à la réalisation des vidéos s'ajoute à celui de la correction des copies à proprement parler. Ainsi, pour un groupe de 30 étudiants dont l'évaluation se ferait sur la base de deux travaux à la maison et de deux examens, la surcharge de travail pour l'enseignant, après quelques tâtonnements, s'élèverait à environ douze heures⁵ par session. Cependant, l'expérience nous a montré que l'utilisation de la RV entraîne des gains de temps sur certaines tâches annexes. Ainsi, les étudiants sont beaucoup moins nombreux, voire très rares, à solliciter l'enseignant au sujet de leurs travaux, que ce soit par courriel ou en demandant une rencontre. Même si n'avons pas trouvé de données relatives au temps passé par les enseignants à répondre aux demandes des étudiants concernant l'évaluation de leurs travaux, il apparaît que la surcharge temporelle liée à l'utilisation de la RV pour des groupes de taille réduite ou moyenne (c'est-à-dire jusqu'à 30 étudiants) est en réalité inférieure au nombre d'heures fourni précédemment.

Il est évident que la tendance à l'augmentation de la taille des groupes que connaissent la plupart des universités ne joue pas en la faveur de la RV. En effet, selon Hamon (2013 : 560), « [e]n ce qui concerne la didactique de la traduction, nous avons pu constater que les volumes importants d'effectif conditionnaient de façon significative les pratiques de classe en présentiel », et ces effets peuvent être étendus aux processus d'évaluation. Ainsi, il n'est pas matériellement envisageable de pratiquer des RV personnalisées dans un groupe de 50 ou 60 apprenants. Si la taille croissante des groupes ne permet pas d'envisager cette pratique, il existe cependant la possibilité d'adapter le processus en enregistrant alors des vidéos thématiques, où l'enseignant peut par exemple commenter les principales erreurs rencontrées lors de la correction ou mettre en évidence des aspects méthodologiques ou des pratiques exemplaires. Cette solution, que nous adoptons pour les grands groupes, recueille l'approbation des étudiants, comme le montrent ces commentaires formulés à l'occasion de l'évaluation anonyme d'un cours suivi par 55 étudiants :

Les techniques d'évaluation permettent aux élèves de s'améliorer réellement et de développer une méthode de travail rapide et efficace.

J'ai aimé la façon d'analyser les traductions hebdomadaires, les analyses vidéo des examens et devoirs étaient une excellente façon d'apprendre de nos erreurs.

Cette approche, si elle ne permet pas d'individualiser le suivi des apprenants, constitue cependant à notre avis un compromis acceptable, puisque les autres avantages liés à l'utilisation de la rétroaction vidéo sont conservés.

NOTES

1. Une version d'essai du logiciel peut être téléchargée sur le site de l'éditeur : www.techsmith.fr/snagit.html
2. Les pourcentages ont été arrondis pour faciliter la lecture, mais une décimale a été conservée pour les nombres relatifs de façon à respecter l'intégrité des calculs qui en découlent.
3. Afin de faciliter les comparaisons entre les résultats de la présente étude et ceux obtenus par d'autres chercheurs, les statistiques d'erreurs sont présentées en annexe en fréquence relative (nombres d'erreurs par 100 mots).
4. GARDY, Philippe (octobre 2011) : étude menée par courriel auprès de 12 enseignants du département Langue, linguistique et traduction de l'Université Laval, Québec.
5. Calculs fondés sur les hypothèses suivantes : vidéos de cinq minutes et marge de sécurité de 20 %.

RÉFÉRENCES

- AUDET, Lucie (2008) : *Recherche sur les facteurs qui influencent la persévérance et la réussite scolaire en formation à distance*. Montréal : REFAD.
- COLLOMBAT, Isabelle (2009) : La didactique de l'erreur dans l'apprentissage de la traduction. *JoSTrans*. 12:37-54.
- DE KETELE, Jean-Marie et ROEGIER, Xavier (2009) : *Méthodologie du recueil d'informations : Fondements des méthodes d'observation, de questionnaires, d'interviews et d'études de documents*. Bruxelles : De Boeck.
- DEPOVER, Christian, KARSENTI, Thierry, et al. (2007) : *Enseigner avec les technologies : Favoriser les apprentissages, développer des compétences*. Québec : Presses de l'Université du Québec.
- DEPOVER, Christian (n.d.) : *Méthodes et outils de recherche en sciences de l'éducation*. Consultée le 2 février 2014, <http://ute.umh.ac.be/methodes>
- FROELIGER, Nicolas (2004) : Les mécanismes de la confiance en traduction – aspects relationnels. *JoSTrans*. 2:50-62.

- GARDY, Philippe (2011) : La rétroaction vidéo comme outil d'apprentissage. *Synergies*. 6:91-104.
- GARDY, Philippe (2015) : *L'évaluation en didactique de la traduction et l'intégration des outils technopédagogiques : étude qualitative et expérimentation*. Thèse de doctorat non publiée. Québec: Université Laval.
- GILE, Daniel (2005) : *La traduction: la comprendre, l'apprendre*. Paris: Presses universitaires de France.
- GOYETTE, Gabriel et LESSARD-HÉBERT, Michelle (1987) : *La recherche-action: Ses fonctions, ses fondements et son instrumentation*. Québec: Les Presses de l'Université du Québec.
- HAMON, Yannick (2013) : *Les TICE pour la production écrite et la traduction de l'italien vers le français: le cas de la SSLMIT de Forlì*. Thèse de doctorat non publiée. Bologne: Università di Bologna.
- LAROSE, Robert (1998) : Méthodologie de l'évaluation des traductions. *Meta*. 43(2):153-186.
- LAVOIE, Sylvie (2010) : *Une méthode d'évaluation alliant rapidité, précision et chaleur humaine*. Consultée le 22 février 2014, <www.profweb.qc.ca/fr/publications/recits/une-methode-devaluation-alliant-rapidite-precision-et-chaleur-humaine/index.html>.
- LIU, Michel (1992) : Présentation de la recherche-action: définition, déroulement et résultats. *Revue internationale de systémique*. 6(4):293-311.
- MARIEN, Bruno et BÉAUD, Jean-Pierre (2003) : *Guide pratique pour l'utilisation de la statistique en recherche: le cas des petits échantillons*. Montréal: Agence universitaire de la francophonie.
- MARTÍNEZ MELIS, Nicole (2001) : *Évaluation et didactique de la traduction: le cas de la traduction dans la langue étrangère*. Thèse de doctorat non publiée. Barcelone: Universitat Autònoma de Barcelona.
- NICOL, David et Macfarlane-Dick, Debra (2006) : Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*. 31(2):199-218.
- PAIVIO, Allan (1971) : *Imagery and Verbal Processes*. New York: Holt, Rinehart and Winston.
- PAIVIO, Allan (1990) : *Mental Representations: A Dual Coding Approach*. New York: Oxford University Press.
- RESWEBER, Jean-Paul (1995) : *La recherche-action*. Paris: Presses universitaires de France.
- ROBERGE, Julie (2006) : *Corriger les textes de vos élèves: Précisions et stratégies*. Montréal: Chenelière Éducation.
- ROBINSON, Sarita, POPE, Debbie, et al. (2013) : Can we meet their expectations? Experiences and perceptions of feedback in first year undergraduate students. *Assessment & Evaluation in Higher Education*. 38(3):260-272.
- SALOMON, Gavriel, PERKINS, David, et al. (1991) : Partners in cognition: Extending human intelligence with intelligent technologies. *Educational Researcher*. 20(3):2-9.
- STANNARD, Russel (2008) : A new direction in feedback. *Humanising Language Teaching*. 10(6). Consulté le 11 décembre 2013. <www.hltmag.co.uk/dec08/mart04.htm>.
- VESLIN, Jean et VESLIN, Odile (1992) : *Corriger des copies, évaluer pour former*. Paris: Hachette Éducation.
- WADDINGTON, Christopher (2001) : Different methods of evaluating student translations: The question of validity. *Meta*. 46(2):311-325.

ANNEXES

ANNEXE 1

Typologie des erreurs commises lors du devoir (nombre d'erreurs pour 100 mots)

DEVOIR	groupe témoin			groupe test
	section A	section B	moyenne	section C
sens	1,1	1,5	1,3	1,2
ajout ou omission	0,3	0,5	0,4	0,4
calque	3,3	3,1	3,2	3,4
impropriété	2,0	2,0	2,0	1,8
mal dit	1,3	1,2	1,2	1,0
orthographe et grammaire	1,1	1,0	1,1	0,9
punctuation	0,7	0,6	0,7	0,6
présentation	0,8	0,6	0,7	0,7
nombre d'erreurs/100 mots	10,6	10,5	10,5	9,9

ANNEXE 2

Typologie des erreurs commises lors de l'examen (nombre d'erreurs pour 100 mots)

EXAMEN	groupe témoin			groupe test
	section A	section B	moyenne	section C
sens	3,1	2,9	3,0	2,3
ajout ou omission	0,7	0,6	0,7	1,2
calque	3,4	3,7	3,6	2,8
impropriété	1,9	1,5	1,8	2,1
mal dit	1,2	1,4	1,3	1,2
orthographe et grammaire	0,9	1,1	1,0	1,2
punctuation	0,5	0,5	0,5	0,3
présentation	0,8	0,6	0,7	0,5
nombre d'erreurs/100 mots	12,4	12,4	12,4	11,5