

Sujet de recherche M2R sciences cognitives 2017-2018

Projet général : Evaluer la traduction automatique (TA), une étude des usages et perceptions

Sujet M2R : Acquisition et analyse de données sur les processus de traduction lors de la post-édition de TA statistique versus neuronale

Cadre général : L'objectif large du projet dans lequel s'inscrit ce stage est de proposer une analyse pluridisciplinaire des nouveaux modèles qui ont permis une amélioration extrêmement rapide de la qualité de la TA, et des nouveaux usages qui en découlent. Alors que les travaux analysant les bouleversements du champ de la traduction automatique suite aux progrès de la TA statistique, et leur impact sur la profession de traducteur, commencent tout juste à se développer (voir par ex. P. Koehn et al., 2013, A. Way, 2010, ou M. Ehrensberger-Dow et al. 2015), les nouveaux modèles basés sur des réseaux de neurones récurrents modifient encore considérablement la donne et laissent entrevoir une nouvelle évolution des usages (Knowles et Koehn, 2016).

Ce sont ces usages qu'il s'agit d'analyser dans le cadre du sujet proposé, à partir d'une simulation expérimentale du travail du traducteur lorsqu'il révise des sorties de TA (on parle alors de post-édition). Plusieurs analyses récentes de l'ergonomie cognitive de la post-édition évaluent la charge cognitive en se basant sur les caractéristiques de la tâche et des textes (voir par ex. Koponen, 2012 ou Čulo et al., 2014). Cependant, ces évaluations ne correspondent pas toujours à l'effort cognitif perçu (Martikainen et Kübler, 2016 : §8). Nous souhaitons combiner une évaluation quantitative basée sur l'enregistrement des processus au clavier et à l'écran (Lorenzo, 2002), et une évaluation plus qualitative de la charge cognitive, obtenue sous la forme de réponses à un bref questionnaire.

Question(s) : Les améliorations apportées par les modèles neuronaux modifient-elles la tâche de post-édition en réduisant la charge cognitive ? Ces modifications s'accompagnent-elles de perceptions plus favorables ?

Projet : Analyser les usages et perceptions de traducteurs professionnels et d'apprentis traducteurs (étudiants du Master de Traduction spécialisée multilingue de l'Université Grenoble Alpes), à partir de sorties de TA générées en utilisant les moteurs statistiques et neuronaux du LIG. Deux passages du même texte seront proposés, traduits par l'un puis l'autre moteur. La post-édition se fera dans un environnement de TAO connu des étudiants comme des traducteurs sollicités (SDL Trados Studio). Les données recueillies comprendront, outre le texte révisé, un enregistrement au clavier et à l'écran avec le logiciel InputLog. Les perceptions seront ensuite évaluées à partir d'un questionnaire.

Le travail de recherche de l'étudiant(e) M2R consistera à acquérir et analyser des données de plusieurs types afin de faire ressortir (le cas échéant) l'impact des nouveaux modèles de TA sur la charge cognitive.

Compétences requises : Bonne maîtrise de l'analyse statistique (régression multiple, modèles mixtes) et/ou très bonne maîtrise de l'anglais. Des compétences en psychologie expérimentale seront bienvenues.

Indemnisation (sous réserve de l'obtention du financement) : Un mois de stage indemnisé (400 euros) ou une mobilité de 4 jours, tous frais payés, à Dublin City University pour y effectuer une nouvelle prise de données, encadrée par deux grands spécialistes de ces questions avec lesquels nous collaborons Dorothy Kenny et Andy Way.

Lieu de stage : LIG ?

Responsables de stage : Caroline Rossi (ILCEA4) et Emmanuelle Esperança Rodier (LIG)

Co-encadrant : Laurent Besacier (LIG)

Contact mail : caroline.rossi@univ-grenoble-alpes.fr

Références bibliographiques :

- Čulo Oliver, Gutermuth Silke, Hansen-Schirra Silvia & Nitzke Jean (2014), « The Influence of Post-Editing on Translation Strategies », S. O'Brien, L. Winther-Balling, M. Carl, M. Simard & L. Specia (dir.), *Post-editing of Machine Translation: Processes and Applications*, Newcastle upon Tyne : Cambridge Scholars Publishing, 200-218
- Ehrensberger-Dow, M., Göpferich, S., & O'Brien, S. (2015). *Interdisciplinarity in Translation and Interpreting Process Research*. John Benjamins Publishing Company.
- Knowles, R., & Koehn, P. (2016). Neural interactive translation prediction. *AMTA 2016, Vol.*, 107.
- Koehn, P. et al. (2013). *Cognitive Analysis and Statistical Methods for Advanced Computer Aided Translation* (CASMACAT), Final Public Report. <http://www.casmacat.eu/uploads/Deliverables/final-public-report.pdf> [consulté le 26 janvier 2017].
- Lorenzo, M.P. (2002). "¿Es posible la traducción inversa?" In Gyde Hansen (ed.) *Empirical translation studies. Product and process*. Copenhagen: Samfundslitteratur, 85–124.
- Koponen, Maarit (2012) « comparing human perceptions of post-editing effort with post-editing operations », *Proceedings of the Seventh Workshop on Statistical Machine Translation*. Stroudsburg : Association for Computational Linguistics, 181-190.
- Martikainen, H., & Kübler, N. (2016). Ergonomie cognitive de la post-édition de traduction automatique : enjeux pour la qualité des traductions. *ILCEA. Revue de l'Institut des langues et cultures d'Europe, Amérique, Afrique, Asie et Australie*, (27). Consulté à l'adresse <https://ilcea.revues.org/3863>