

TransCheck : un vérificateur automatique de traductions

Elliott MACKLOVITCH, Guy LAPALME
Laboratoire RALI – Université de Montréal
Montréal, Canada
{macklovi, lapalme}@iro.umontreal.ca

Résumé. Nous offrirons une démonstration de la dernière version de TransCheck, un vérificateur automatique de traductions que le RALI est en train de développer. TransCheck prend en entrée deux textes, un texte source dans une langue et sa traduction dans une autre, les aligne au niveau de la phrase et ensuite vérifie les régions alignées pour s'assurer de la présence de certains équivalents obligatoires (p. ex. la terminologie normalisée) et de l'absence de certaines interdictions de traduction (p. ex. des interférences de la langue source). Ainsi, TransCheck se veut un nouveau type d'outil d'aide à la traduction qui pourra à réduire le fardeau de la révision et diminuer le coût du contrôle de la qualité.

Abstract. We will present a demonstration of the latest version of *TransCheck*, an automatic translation checker that the RALI is currently developing. *TransCheck* takes as input two texts, a source text in one language and its translation in another, aligns them at the sentence level and then verifies the aligned regions to ensure that they contain certain obligatory equivalents (e.g. standardized terminology) and do not contain certain prohibited translations (e.g. source language interference). *TransCheck* is thus intended to be a new type of tool for assisting translators which has the potential to ease the burden of revision and diminish the costs of quality control.

Mots-clés : traduction assistée par ordinateur, vérification automatique de traductions, révision de traduction.

Keywords: machine-aided translation, automatic translation checking, translation revision.

En collaboration avec le Centre de recherche en technologies langagières à Ottawa, le laboratoire RALI est en train de développer un nouveau prototype de *TransCheck*, un vérificateur automatique de traductions. Comme son nom l'indique, *TransCheck* est un outil qui vise à aider un réviseur humain à déceler automatiquement certaines erreurs de traduction et à faire respecter certaines normes linguistiques au sein d'un grand service de traduction. Le système ressemble un peu à un vérificateur d'orthographe, dans le sens qu'il sert à repérer des erreurs potentielles dans un texte. Mais contrairement à un vérificateur d'orthographe, qui décèle des erreurs de forme monolingues, *TransCheck* cible des erreurs de correspondances entre deux textes, un texte source et sa traduction. Le mot « librairie », par exemple, est une forme correcte en français, mais il constitue néanmoins une erreur en tant que traduction du mot anglais « library ». Un vérificateur monolingue serait muet devant de tels cas, alors qu'un vérificateur bi-textuel comme *TransCheck* sera en mesure de le signaler.

Conceptuellement, comment fonctionne ce vérificateur automatique de traductions ? *TransCheck* prend en entrée deux textes, un texte source dans une langue et une ébauche de traduction dans une autre. Dans un premier temps, le système segmente chaque texte en unités linguistiques (c.-à-d. des phrases et des mots), pour ensuite les aligner au niveau des phrases. Une fois qu'il a établi cette correspondance entre les phrases qui sont des traductions mutuelles, *TransCheck* y applique les différents modules de détection d'erreurs. Certains de ces modules sont basés sur des grammaires internes, comme les grammaires des différents types d'expressions numériques (p. ex. les dates ou les expressions monétaires) ; d'autres modules exigent des données externes fournies normalement par l'utilisateur (p. ex. la terminologie). *TransCheck* signale une erreur potentielle lorsque, dans une région alignée, il repère un item source sans trouver son correspondant obligatoire, ou lorsque le système repère un item source et trouve aussi son correspondant interdit dans la région cible.

Nous pouvons illustrer ces différents cas de figures à l'aide de la capture d'écran à la page suivante. Comme on peut y voir, le prototype actuel est pourvu d'une interface graphique dans laquelle le texte source apparaît à gauche et le texte cible à droite. Les cases à cocher au dessus des deux textes permettent à l'utilisateur d'activer ou de désactiver les différents modules de détection d'erreurs. Les régions alignées sont démarquées par une ligne horizontale bleue et, au sein de ces régions, les erreurs décelées sont signalées par un soulignement rouge. Dans la troisième région, par exemple, où l'utilisateur a pointé sur le mot souligné « sniper », le système signale une erreur potentielle de terminologie ; et, dans le cadre inférieur de la fenêtre, il indique les termes français qu'il cherchait mais n'a pas trouvés du côté cible. Dans la région suivante, il s'agit d'un autre type de correspondance obligatoire : une expression numérique cette fois, qui se trouve à être traduit par la paraphrase « un tour complet ». Dans la dernière région alignée, nous voyons un exemple d'une interdiction de traduction : normalement, le nom « deception » en anglais et son homographe français sont considérés comme étant des faux-amis.

Il ne faut pas s'imaginer qu'un outil comme *TransCheck* pourra déceler toutes les erreurs de traduction, ce qui représente un problème qui est probablement plus difficile que la traduction automatique elle-même. Par contre, un système qui serait en mesure d'alléger le fardeau des réviseurs en automatisant les aspects les plus mécaniques et fastidieux de leur travail serait certainement bien accueilli. Ainsi, le prototype actuel traite un sous-ensemble d'erreurs que l'on peut détecter par des techniques simples et formelles, c.-à-d. sans 'compréhension profonde' des textes. Ceci dit, *TransCheck* n'est pas du tout un système fermé ; au fur et à mesure que notre compréhension des relations traductionnelles avancera, nous pourrons y intégrer d'autres types de vérifications.

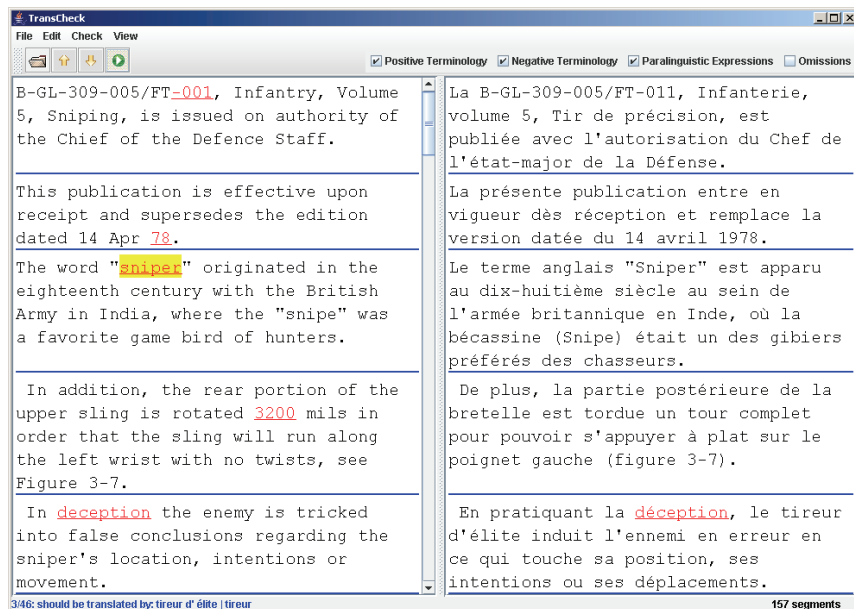


Figure 1 : Une capture d'écran du système *TransCheck* où les erreurs signalées par le système sont soulignées en rouge.

Finalement, il est important de noter que *TransCheck* est un outil d'aide à la traduction qui vise surtout à appuyer le traducteur et non pas à le remplacer. Ainsi, le système signale des erreurs potentielles, mais c'est à l'utilisateur de décider lesquelles doivent réellement être corrigées et lesquelles peuvent rester telles quelles. Autre élément interactif dans l'emploi de *TransCheck*, l'utilisateur peut aisément modifier ou ajouter des entrées au glossaire de la terminologie et à l'anti-dictionnaire des interdictions dont *TransCheck* se sert pour effectuer ces vérifications.

Références

- JUTRAS J-M. (2000). An Automatic Reviser : The TransCheck System. *Sixth Applied Natural Language Processing Conference*, 127-134.
- MACKLOVITCH E. (1995). TransCheck - or the Automatic Validation of Human Translations. *Proceedings of MT Summit V (Luxembourg)*.
- RUSSELL, G. (2005). Automatic Detection of Translation Errors: the TransCheck System. *Translating and the Computer 27 (ASLIB Conference)*.