

## Traduction automatisée d'une oeuvre littéraire: une étude pilote

Laurent Besacier<sup>1</sup>

(1) LIG, Université de Grenoble, UJF - BP 53, 38041 Grenoble Cedex 9  
laurent.besacier@imag.fr

**Résumé.** Les techniques actuelles de traduction automatique (TA) permettent de produire des traductions dont la qualité ne cesse de croître. Dans des domaines spécifiques, la post-édition (PE) de traductions automatiques permet, par ailleurs, d'obtenir des traductions de qualité relativement rapidement. Mais un tel pipeline (TA+PE) est-il envisageable pour traduire une oeuvre littéraire ? Cet article propose une ébauche de réponse à cette question. Un essai de l'auteur américain Richard Powers, encore non disponible en français, est traduit automatiquement puis post-édité et révisé par des traducteurs non-professionnels. La plateforme de post-édition du LIG utilisée permet de lire et éditer l'oeuvre traduite en français continuellement, suggérant (pour le futur) une communauté de lecteurs-réviseurs qui améliorent en continu les traductions de leur auteur favori. En plus de la présentation des résultats d'évaluation expérimentale du pipeline TA+PE (système de TA utilisé, scores automatiques), nous discutons également la qualité de la traduction produite du point de vue d'un panel de lecteurs (ayant lu la traduction en français, puis répondu à une enquête). Enfin, quelques remarques du traducteur français de R. Powers, sollicité à cette occasion, sont présentées à la fin de cet article.

**Abstract.** Current machine translation (MT) techniques are continuously improving. In specific areas, post-editing (PE) allows to obtain high-quality translations relatively quickly. But is such a pipeline (MT+PE) usable to translate a literary work (fiction, short story) ? This paper tries to bring a preliminary answer to this question. A short story by American writer Richard Powers, still not available in French, is automatically translated and post-edited and then revised by non-professional translators. The LIG post-editing platform allows to read and edit the short story suggesting (for the future) a community of readers-editors that continuously improve the translations of their favorite author. In addition to presenting experimental evaluation results of the pipeline MT+PE (MT system used, automatic evaluation), we also discuss the quality of the translation output from the perspective of a panel of readers (who read the translated short story in French, and answered to a survey afterwards). Finally, some remarks of the official french translator of R. Powers, requested on this occasion, are given at the end of this article.

**Mots-clés :** traduction automatique, TA, oeuvre littéraire, post-édition.

**Keywords:** machine translation, MT, litterature, fiction, post-edition.

### 1 Introduction

La tâche de post-édition consiste à éditer la sortie textuelle produite, le plus souvent par une machine (système de traduction automatique, de reconnaissance optique de caractères, de transcription automatique, etc.) en vue de l'améliorer. Dans le domaine de la traduction de documents, où les systèmes automatiques sont utilisés, le processus suivant est généralement utilisé : le système de TA produit des traductions brutes qui sont manuellement post-édités par des traducteurs professionnels formés (post-éditeurs) corrigant les erreurs de traduction.

De nombreuses études ont montré les avantages de l'utilisation combinée traduction automatique+post-edition manuelle (TA+PE) dans un flux de diffusion. Par exemple, Garcia (2011) a montré que même si la post-édition de sorties brutes de TA ne conduit pas toujours à une amélioration en termes de productivité, elle permet de produire de meilleures traductions par rapport à des traductions manuelles directement issues du texte source<sup>1</sup>. Autodesk a aussi réalisé une expérience pour tester si l'utilisation de la TA permettrait d'améliorer la productivité des traducteurs. Dans cette expérience, les résultats<sup>2</sup> montrent que la post-édition des sorties issues d'un système automatique augmente significativement la productivité par rapport à une traduction partant "de zéro". Ceci est vérifié quelle que soit la paire de langues, l'expérience du traducteur et sa préférence exprimée concernant la méthode (post-éditer des sorties automatiques vs traduire à partir de zéro).

1. le travail de Garcia (2011) est cependant sujet à controverses car la traduction manuelle sans post-édition semble avoir été faite sans autoriser l'utilisation d'aide numérique pour assister le traducteur (même pas un dictionnaire électronique)

2. amta2012.amtaweb.org/AMTA2012Files/html/5/5\_paper.pdf

Ces résultats mesurés en milieu académique (Garcia (2011)) ou industriel (Autodesk) sur des processus de traduction dans des domaines spécialisés, nous conduisent à poser la question suivante : quelle serait la valeur d'un tel processus (TA+PE) appliqué sur la traduction d'une oeuvre littéraire ? Combien de temps faut-il pour traduire une oeuvre de dix mille mots ? La traduction obtenue est-elle acceptable pour des lecteurs ? Qu'en penserait le traducteur attitré de l'auteur considéré ? Une traduction "low cost" produite par des communautés de fans (comme c'est le cas pour les séries TV) est-elle envisageable pour des romans ou des nouvelles ? Cet article tente d'apporter de premières réponses (certes très partielles) à ces questions. En plus de cette étude préliminaire, il restera au moins de ce travail la traduction inédite d'une nouvelle, ou plutôt d'un essai (*The Book of Me* de Richard Powers).

Le suite de l'article présente notre méthodologie générale, le choix de l'oeuvre et les systèmes de TA et de PE (tous deux issus du LIG) utilisés. Les données expérimentales collectées (et rendues disponibles à la communauté sur *github*) sont ensuite décrites dans la section 3. Dans la partie 4, nous tentons d'aller au delà d'un tâche de TA adaptée au domaine en sollicitant le point de vue d'un (petit) panel de lecteurs et l'avis du traducteur attitré de R. Powers. Pour finir, la partie 5 conclut ce travail préliminaire.

## 2 Traduction automatisée d'une oeuvre littéraire

### 2.1 Remarques préliminaires

Le format court de cet article ne permet pas de proposer une étude bibliographique exhaustive sur la traduction assise de documents littéraires ou, plus généralement, sur l'utilisation du TALN dans le domaine littéraire. Cependant, on peut noter que c'est un champ de recherche assez actif : un atelier y est consacré depuis 2012 : *ACL workshop on Computational Linguistics for Literature*<sup>3</sup>. On peut aussi mentionner la tenue récente, en France, de la *Journée d'étude sur la traduction de textes littéraires* organisée par l'ENS en 2013. Cependant, à notre connaissance, cette étude est la première sur la traduction automatisée d'une oeuvre littéraire. Il conviendrait également de définir ce qu'on appelle un *texte littéraire*. En ce qui nous concerne, nous incluons dans cette catégorie (notre définition est sans doute trop restrictive) toutes les oeuvres de fiction ou autobiographiques écrites sous la forme de romans ou de nouvelles. Dans ces textes, l'auteur exprime sa vision du monde, de son époque et de la vie en général tout en utilisant des *procédés littéraires* et une *technique d'écriture* (forme) lui permettant de créer des effets à l'aide du langage ou de faire passer des messages (explicites ou sous-entendus).

### 2.2 Choix de l'oeuvre

Le choix de l'oeuvre a été guidé par le fait que (a) nous avions un contact avec le traducteur français de l'auteur américain Richard Powers<sup>4</sup> (auteur notamment du roman *La chambre aux échos* qui a remporté le National Book Award et a été finaliste du Prix Pulitzer) (b) R. Powers publie des romans souvent en lien avec la science et, d'une certaine manière, ses textes contiennent quelques points communs avec des textes scientifiques ou techniques (ce qui réduit peut-être un peu le fossé entre traduction de textes scientifiques et littéraires ?).

Par l'intermédiaire de son traducteur français (J-Y Pellegrin), R. Powers a été informé, par e-mail, de notre démarche et il a donné son accord. Nous avons ensuite choisi une oeuvre, non encore traduite en français, intitulée *The Book of Me* et publiée pour la première fois dans les colonnes du magazine GQ<sup>5</sup>. C'est un texte narratif, raconté à la première personne, où l'auteur est le personnage principal de l'histoire. Il y raconte l'année 2008 au cours de laquelle il est devenu la neuvième personne au monde à voir son génome entièrement séquencé. Bien que le thème soit la génétique et malgré le style simple et clinique employé par l'auteur, *The Book of Me* est bien une oeuvre littéraire où l'auteur, qui enseigne la technique narrative à l'université, ne met jamais de côté son ambition poétique, son humour et sa fascination pour l'irrationnel.

### 2.3 Méthodologie générale

La traduction de textes par cascade TA+PE (traduction automatique puis post-édition) a déjà été évaluée au LIG sur des données journalistiques. En 2012, la collecte de post-éditions manuelles de 12 000 énoncés (équivalente à un livre de 500 pages environ) a été réalisée par Potet *et al.* (2012). Ce corpus collecté via *crowdsourcing* et disponible en ligne<sup>6</sup>, est un des plus gros corpus existant concernant la post-édition de données issues d'un système de traduction automatique appliquée sur des données libres de droit. Il est par exemple trois fois plus important que celui collecté par Specia *et al.* (2010) qui fait référence dans le domaine. La même méthodologie a été appliquée sur l'oeuvre littéraire choisie mais

3. [sites.google.com/site/clfl2014a/](http://sites.google.com/site/clfl2014a/)

4. [fr.wikipedia.org/wiki/Richard\\_Powers](http://fr.wikipedia.org/wiki/Richard_Powers)

5. [www.gq.com/news-politics/big-issues/200810/richard-powers-genome-sequence](http://www.gq.com/news-politics/big-issues/200810/richard-powers-genome-sequence)

6. [www-clips.imag.fr/geod/User/marion.potet/index.php?page=download](http://www-clips.imag.fr/geod/User/marion.potet/index.php?page=download)

(différence importante) il n'y avait qu'un seul post-éditeur (pas de *crowdsourcing*) et le texte postédité a été ensuite révisé.

Plus précisément, le corpus a été divisé en trois parties égales. Une boucle *traduction-postédition-adaptation* est appliquée sur les trois blocs de texte selon le processus suivant :

- Après traduction automatique du premier tiers (par un système de TA probabiliste initial non adapté à la tâche), les sorties traduites sont révisées par une étudiante en Master de traduction,
- Ce premier tiers post-édité sert à adapter le système de TA Anglais-Français. Vu la faible taille des données post-éditées, seuls les poids du modèle log-linéaire sont adaptés en utilisant le premier tiers corrigé comme corpus de développement. Une méthode similaire est suggérée par Pecina *et al.* (2012) pour l'adaptation au domaine avec quantité de données limitées (nous sommes conscients que d'autres techniques d'adaptation plus avancées auraient pu être utilisées mais ceci n'était pas le thème central de notre contribution),
- Ensuite, le second tiers de l'oeuvre est traduit avec le système de TA adapté, puis la sortie est post-éditée et un second système de TA adapté est obtenu à partir de ces nouvelles données. Ce second système est utilisé pour traduire la troisième et dernière partie de l'oeuvre,
- Une fois la PE terminée, le texte final est révisé : d'abord par le post-éditeur puis par un autre réviseur (non professionnel, francophone, ayant cependant une bonne connaissance de la langue anglaise),
- Les temps de post-édition et de révision sont mesurés.

## 2.4 Système de TA utilisé

Le système LIG est entraîné à partir des données fournies lors de la campagne d'évaluation IWSLT et optimisé pour la tâche de TA anglais-français de la campagne 2012. Les corpus suivants sont utilisés pour construire le modèle de traduction (total cumulé de 25M phrases environ) : *news-c*, *europarl*, *un*, *ted*, *eu-const*, *dgt-tm*, *pct* et une extraction de 5M de phrases du corpus *gigaword*. La partie en français des mêmes corpus est utilisée pour apprendre le modèle de langue, avec l'ajout du corpus *news-shuffle* fourni au cours de la campagne WMT 2012. Le système LIG est un système à base de segments (*phrase-based*) qui s'appuie sur la boîte à outils *Moses* Hieu *et al.* (2007). Trois modèles de traduction appris à partir de différents corpus (*ted*; *news-c+europarl+un+eu-const+dgt-tm+pct*; *gigaword5M*) sont utilisés. Un modèle de langue 5-gramme est appris séparément sur chaque corpus à l'aide de la boîte à outils SRILM Stolcke (2002) avec un modèle de repli Kneser-Ney modifié, puis les modèles sont interpolés en optimisant la perplexité sur le corpus *dev2010* de la campagne IWSLT. Le système du LIG obtient des scores BLEU de 36,88 et 37,58 sur les corpus *tst2011* et *test2012* de IWSLT respectivement (BLEU évalué avec casse et ponctuation). Plus de détails sur ce système (classé honorablement lors de IWSLT 2012) se trouvent dans Besacier *et al.* (2012). Il est clair que ce système n'est pas optimal pour traduire des textes littéraires et il serait souhaitable, pour de futurs travaux, de rassembler au moins des textes littéraires en français pour adapter le modèle de langue cible (ou même envisager d'avoir accès aux autres œuvres et traductions de l'auteur).

## 2.5 Post-edition

Nous utilisons l'interface de post-édition du LIG proposée initialement par Huynh *et al.* (2008). Celle-ci a été utilisée pour de nombreux projets : traduction d'articles de l'encyclopédie EOLLS, accès multilingue à des dizaines de sites Web. La figure 1 montre l'interface de post-édition en mode avancé qui permet, pour chaque segment source, de charger plusieurs traductions automatiques (issues de Google, Moses<sup>7</sup>, etc.) et de les corriger. Le temps de post-édition est mesuré et l'historique des corrections est stocké dans une base de données. Cet outil est la colonne vertébrale ayant donné lieu au concept d'iMAG proposé par le LIG (*interactive Multilingual Access Gateway*) qui permet la navigation dans un site ou sous-site Web dans plusieurs langues d'accès, avec amélioration incrémentale et contrôlée de la qualité des traductions (voir le lien en note de bas de page<sup>8</sup> pour plus de détails).

# 3 Données expérimentales

## 3.1 Corpus et statistiques de post-édition

L'oeuvre, composée de 545 segments et 10731 mots est divisée en trois blocs identiques. La Table 1 résume le nombre de mots des données source et cible (TA ou PE<sup>9</sup>). Sans surprises, un ratio supérieur à 1,2 est observé entre cible française (TA) et source anglaise. On constate cependant que ce ratio tend à diminuer après post-édition de la sortie française.

7. ici, la sortie de notre système était chargée en priorité par rapport à Google

8. aximag.fr/AXiMAG-homePage.html

9. La post-édition utilisée ici est obtenue après chaque itération du processus ; la dernière étape de révision n'est donc pas prise en compte à ce stade.



FIGURE 1 – Interface de post-édition en mode avancé

TABLE 1 – Corpus source, cible traduite et cible corrigée

Itération (nb. seg)	Anglais (nb. mots)	TA Français (nb. mots)	PE Français (nb. mots)
It.1 (184)	3593	4295	4013
It.2 (185)	3729	4593	4202
It.3 (176)	3409	4429	3912
<b>Total (545)</b>	<b>10731</b>	<b>13317</b>	<b>12127</b>

### 3.2 Performances du système de TA

La Table 2 résume les performances de TA (mesurées avec BLEU) calculées sur le corpus complet avec les systèmes issus de chaque itération. Le temps de post-édition requis pour chaque bloc est également indiqué. Les scores BLEU, qui sont directement comparables, ne montrent pas de véritable amélioration du système. Il semble donc que l'adaptation des poids seuls (qui donne lieu à des améliorations dans Pecina *et al.* (2012)) soit peu efficace dans notre cas. Le temps de post-édition diminue cependant légèrement à chaque itération (mais là encore, les différences sont faibles et il est difficile de dire si la diminution du temps de PE est dûe à l'adaptation du système de TA ou à une productivité croissante du post-editeur qui s'adapte à la tâche). Au final, le temps total de PE est estimé à 15h environ.

TABLE 2 – Evaluation automatique (BLEU) sur corpus complet et mesure du temps de PE pour chaque bloc de l'itération

Système TA utilisé	score BLEU (corpus complet)	temps PE (bloc it.)
It.1 (non adapté)	38,95	5h37mn
It.2 (tuning sur bloc 1)	37,51	4h45mn
It.3 (tuning sur bloc 1+2)	39,38	4h35mn

La lecture de l'oeuvre traduite à ce stade (après PE) est peu satisfaisante. En effet, la post-édition est faite "segment par segment", sans vue d'ensemble sur le texte. Il en résulte une manque d'homogénéité très gênant pour un texte littéraire. Pour cette raison, deux révisions du texte traduit sont également effectuées : une par le post-éditeur lui-même (4h) et une par l'auteur de cet article (6h). La version finale de l'oeuvre traduite (qui aura donc été obtenue en  $15+4+6=25$ h) sert de base aux évaluations, plus qualitatives, présentées dans la partie suivante. Les différences entre la version post-éditée brute (15h de travail) et la version révisée (25h de travail) ne sont pas analysées ici, mais pourront l'être dans le futur puisque les deux versions sont rendues disponibles sous *github*.

## 4 Aller au delà d'une tâche de TA adaptée au domaine

### 4.1 Le point de vue des lecteurs sur la traduction post-éditée

Neuf lecteurs ont accepté de lire l'oeuvre traduite et ont répondu à un questionnaire, toujours ouvert sur [fluidsurveys.com](http://fluidsurveys.com)<sup>10</sup>. La version *pdf* de l'essai traduit ainsi que fichier tableur rassemblant les résultats du sondage sont également rendus disponibles dans *github*. Après trois questions permettant de mieux cerner le profil du lecteur, une première partie (5 questions) interroge les lecteurs sur la lisibilité et la qualité du texte littéraire traduit. Une seconde partie (7 questions) vérifie que certaines subtilités du texte ont été bien comprises.

Le texte est jugé globalement lisible (5 TBien et 3 Bien), compréhensible (8 oui, 1 non) et contenant peu de fautes (8 rarement, 1 souvent). Les questions de compréhension les plus faciles sont bien traitées par les lecteurs qui répondent tous correctement (4 questions). Cependant, trois questions donnent lieu à des réponses différentes selon les lecteurs :

- 2 lecteurs ont mal répondu à une question apparemment simple (*qui finance le séquençage du génome de Powers ?*),
- la question *Lors de l'écriture de la nouvelle, combien de personnes se sont déjà fait séquencer le génome ?* était ambigu puisqu'on pouvait répondre 8 ou 9 (en comptant ou non Powers) et a donné lieu à des réponses différentes des lecteurs
- seuls 4 lecteurs sur 9 ont su donner une chronologie correcte des étapes du processus de séquençage d'un génome ; le texte traduit ne contient cependant pas de contresens ; ce résultat mitigé indique peut-être un désintérêt de certains lecteurs pour les aspects les plus techniques de l'oeuvre.

En résumé, nous pouvons dire que ce sondage, qui reste très limité, montre toutefois que le texte (produit selon notre méthodologie) a été jugé acceptable et assez lisible par nos lecteurs (dont 3 indiquent lire très souvent, 4 assez souvent et 2 rarement). Nous citons également quelques remarques mises dans les commentaires libres : "J'ai remarqué peu de fautes, quelques néologismes (j'ai considéré que c'étaient des néologismes et pas des erreurs de traduction car ça avait du sens)" - "Texte très fluide et lecture très facile malgré des termes scientifiques précis" - "J'ai trouvé le texte un peu difficile parce qu'il contient des mots complexes et qu'il traite d'un domaine que je ne connais pas du tout".

### 4.2 Le point de vue du traducteur de R. Powers

Pour finir cette étude pilote, un dixième lecteur a été sollicité : le traducteur français de l'auteur, J-Y Pellegrin, enseignant chercheur à Paris-Sorbonne. Son avis est résumé ici sous la forme de questions réponses. Le manque de place ne nous permet pas de commenter ces remarques mais nous pensons qu'elles sont assez explicites pour être délivrées en l'état.

**Lisibilité ?** "Le texte auquel vous êtes parvenu restitue une image fidèle du contenu de l'article de Powers. Le pari de la lisibilité est gagné et certains passages (notamment ceux qui portent sur les aspects scientifiques de l'expérience décrite) sont très convaincants."

**Imperfections ?** "Il reste bien sûr des imperfections, des lourdeurs, voire des erreurs ponctuelles, qui appellent une correction"

#### Principales erreurs ?

- "Le défaut le plus répétitif, celui dont souffre d'ailleurs le travail de tout traducteur débutant, est le calque syntaxique, là où le français structure différemment la phrase .... On comprend, mais ça ne sonne pas vraiment français"
- "Autre défaut assez fréquent, la perte des idiomatismes du français au profit d'anglicismes. Parfois ces anglicismes peuvent être plus dérangeants lorsqu'ils flirtent avec le franglais comme dans « connaissances actionnables » (p. 18) au lieu de « connaissances pratiques / utilisables »."
- "Un troisième défaut tient à la non prise en compte de certains repères culturels .... Par exemple, Powers fait plusieurs références à la topographie de Boston qui donnent lieu à des inexactitudes dans la traduction : « la rivière Charles » par exemple (p. 12) qui n'est pas une rivière mais plutôt un fleuve ; c'est pourquoi on traduira par « la Charles River » ou simplement « la Charles »"

**Ce texte pourrait-il servir de base de départ à un traducteur littéraire professionnel ?** "Instinctivement, je serais tenté de répondre non pour l'instant, parce que, dès son premier jet, le traducteur possède des réflexes qui lui permettent de produire un texte plus « propre » que celui auquel vous êtes parvenu .... Cependant, ce traducteur passera plus de 25 heures à produire les 42 feuillets de 1500 signes correspondant au texte de Powers. À raison de 7 feuillets par jour en moyenne, il faut 6 journées de 8h pour venir à bout du texte .... Si, en revanche, je pouvais ne travailler que sur votre texte, en oubliant complètement celui de Powers parce que j'aurais la garantie que votre traduction ne comporte aucune erreur, ni oubli, ni aplatissement par rapport à l'original, mais qu'elle demande simplement à être améliorée, rendue plus fluide, dans un français plus authentique, les choses seraient différentes et le gain de temps sans doute considérable. ".

10. [https://fluidsurveys.com/surveys/manuela-cristina/un-livre-sur-moi-qualite-de-la-traduction/?TEST\\_DATA=](https://fluidsurveys.com/surveys/manuela-cristina/un-livre-sur-moi-qualite-de-la-traduction/?TEST_DATA=)

## 5 Conclusion

### 5.1 Données rassemblées et mise en ligne

Les données de cet article sont disponibles sur le lien suivant [github.com/powersmachinetranslation/DATA](https://github.com/powersmachinetranslation/DATA). On y retrouve notamment :

- les 545 segments source, cible (TA et PE) mentionnés table 1,
- l'oeuvre traduite et révisée en français, ayant été lue par un panel de 9 lecteurs,
- les résultats du questionnaire (9 lecteurs) compilés dans un tableur.

### 5.2 Commentaires et questions ouvertes

Nous avons présenté une première expérience de traduction automatisée d'une oeuvre littéraire (essai en anglais d'une vingtaine de pages). Les résultats issus d'un pipeline TA+PE ont été présentés et, pour aller au delà, les avis d'un panel de lecteurs et d'un traducteur ont été sollicités. Le texte traduit, obtenu après 25h de travail humain, est jugé acceptable par les lecteurs mais l'avis du traducteur professionnel reste mitigé. Cette approche suggère une méthodologie de traduction rapide et "low cost", analogue aux traductions de sous-titres de séries TV trouvées sur le Web. Pour l'auteur, c'est la possibilité d'avoir son oeuvre traduite dans un plus grand nombre de langues (plusieurs dizaines au lieu d'une poignée - cet essai de R. Powers a d'ailleurs aussi été traduit en roumain avec la même méthodologie). Mais celui-ci est-il prêt à sacrifier la qualité de traduction (et son contrôle sur celle-ci) au prix d'une diffusion plus large de ses œuvres ?

Pour le lecteur qui ne peut lire l'auteur en langue source, c'est la possibilité d'avoir accès plus rapidement à une traduction (certe imparfaite) de son auteur favori. Pour le lecteur non natif mais capable de lire l'oeuvre en langue source, c'est la possibilité d'avoir une aide sur les parties qu'il a du mal à comprendre. Une dernière chose : le titre de l'oeuvre *The Book of Me* est resté inchangé sur la version française car aucune traduction satisfaisante n'a été trouvée pour illustrer le fait que *book* fait référence ici à un livre mais aussi à l'ADN de l'auteur ; ce paradoxe illustre bien toute la difficulté de la traduction d'une œuvre littéraire.

## Remerciements

Merci à Manuela Barcan qui a assuré la première phase de post-édition des traductions automatiques en français et en roumain au cours de l'été 2013. Merci à J-Y Pellegrin, traducteur français de Richard Powers, pour son aide et son ouverture d'esprit. Merci à V. Bellynck et C. Boitet pour leur aide avec l'outil de post-édition Sectra\_W.

## Références

- BESACIER L., LECOUTEUX B., AZOUIZ M. & LUONG NGOC Q. (2012). The LIG English to French Machine Translation System for IWSLT 2012. In *In proceedings of the 9th International Workshop on Spoken Language Translation (IWSLT)*.
- GARCIA I. (2011). Translating by post-editing : is it the way forward ? *Journal of Machine Translation*, **25**(3), 217–237.
- HIEU H., BIRCH A., CALLISON-BURCH C., ZENS R., AACHEN R., CONSTANTIN A., FEDERICO M., BERTOLDI N., DYER C., COWAN B., SHEN W., MORAN C. & BOJAR O. (2007). Moses : Open source toolkit for statistical machine translation. In *ACL'07, Annual Meeting of the Association for Computational Linguistics*, p. 177–180, Prague, Czech Republic.
- HUYNH C.-P., BOITET C. & BLANCHON H. (2008). Sectra\_w.1 : an online collaborative system for evaluating, post-editing and presenting mt translation corpora. In *LREC'08, Sixth International Conference on Language Resources and Evaluation*, p. 28–30, Marrakech, Morocco.
- PECINA P., TORAL A. & VAN GENABITH J. (2012). Simple and effective parameter tuning for domain adaptation of statistical machine translation. In *Proceedings of the 24th International Conference on Computational Linguistics*, p. 2209–2224, Mumbai, India : Coling 2012 Organizing Committee.
- POTET M., EMMANUELLE E R., BESACIER L. & BLANCHON H. (2012). Collection of a large database of french-english smt output corrections. In *Proceedings of the eighth international conference on Language Resources and Evaluation (LREC)*, Istanbul, Turkey.
- SPECIA L., CANCEDDA N. & DYMETMAN M. (2010). A dataset for assessing machine translation evaluation metrics. In *7th Conference on International Language Resources and Evaluation (LREC-2010)*, p. 3375–3378, Valetta, Malta.
- STOLCKE A. (2002). Srilm : An extensible language modeling toolkit. In *ICSLP'02, 7th International Conference on Spoken Language Processing*, p. 901–904, Denver, USA.