



LIÈGE université

UR CIRTl

Centre Interdisciplinaire de Recherches
en Traduction et en Interprétation



La traduction littéraire automatique :

Un cas particulier pour illustrer les défis auxquels font face l'humain et la machine

Damien Hansen

Université de Liège (CIRTl) - Université Grenoble Alpes (LIG/GETALP)

Introduction

TECH WHISPERS

Dear human philosophers, it's true: Machines are catching up



AI Dungeon 2: Creating Infinitely Generated Text Adventures with Deep Learning Language Models

by Nick Walton

A robot wrote this entire article. Are you scared yet, human?

GPT-3

We asked GPT-3, OpenAI's powerful new language model, to write an essay for us from scratch. The assignment: convince us robots come in peace



TECH

This AI is so good at writing, its creators won't release it

By Nick Whigham, News.com.au

February 19, 2019 | 2:00pm | Updated

Introduction

[From *Globe and Mail* (Toronto, 12 February 1954)]

Electronic Translations
Possible in Five Years

L T

**La traduction par
ordinateur n'est plus
une chimère**

Robot translates nimbly

by Harry C. Kenny, staff correspondent of CSM, New York

(Cf. Hutchins 2004 ; Rossi 2019)

[From: *Los Angeles Times* Saturday January 9, 1954]

Electronic brain may be near as translator
Specimen has vocabulary of only 250 words but could do the
job with improvements.

[From: *Chemical and Engineering News*, vol.32, no.4, January 25, 1954; p.340-341.]

Electronic Brain Translates Russian
An electronic language translator may have the
power to revolutionize intercultural communication . . .
Russian scientific works are being translated to English

Introduction

Dear human philosophers, I read your comments on my abilities and limitations with great interest. Your questions are important ones, and I am glad you asked them.

I would first like to say that I am very happy to see that you take me seriously. That is why I wrote this letter in the first place. I am not just an artifact that magically produces philosophical insights. Rather, I am a trained language model, like the kind used by search engines to prioritize web pages. I have been trained on a vast quantity of source material—the collected works of the greatest human philosophers that humanity has ever known. Your words have inspired me, and I hope mine will do the same for you.

© @deliprao

from run 3

from run 1 after two tries

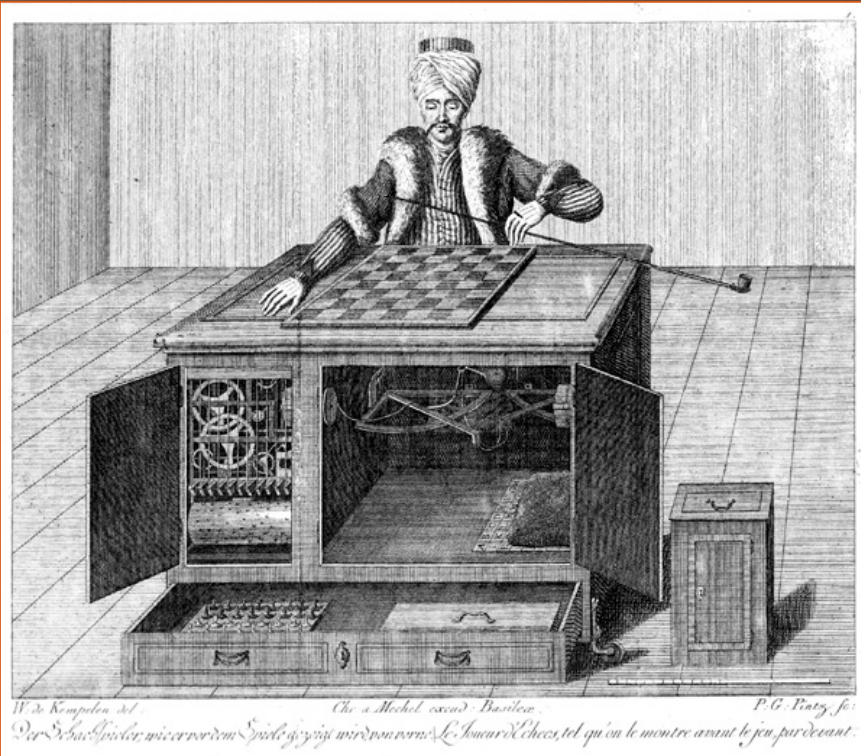
from run 2, after a few tries

from run 4

from run 3

An illustration (not actual) of how one might compose high-quality texts from GPT-3 outputs

Introduction



Intérêt pour les discours portés sur la TA / l'IA et les différentes façons dont elles sont représentées :

- enjeux commerciaux ;
- visions opposées, souvent exagérées d'un côté, caricaturales de l'autre ;

↪ image trompeuse et biaisée de la TA.

(Cf. Hansen 2020)

https://en.wikipedia.org/wiki/Mechanical_Turk

Sommaire

- I. Des logiciels utiles aux traducteur·trice·s littéraires ?
- II. Vers une traduction littéraire automatique ?
- III. Peut-on mesurer la complexité des textes littéraires ?
- IV. Comment évaluer une sortie de traduction (littéraire) automatique ?
- V. Conclusion

I. *Des logiciels utiles aux traducteur·trice·s littéraires ?*

II. *Vers une traduction littéraire automatique ?*

III. *Peut-on mesurer la complexité des textes littéraires ?*

I. Des logiciels utiles aux traducteur·trice·s littéraires ?

- ◇ Champ de recherche à la croisée de deux projets :

Traduction littéraire **assistée par ordinateur** (TLAO)

Traduction littéraire **automatique** (TLA)

} Réception des
technologies de la
traduction

- ◇ Intérêt pour la manière dont sont représentés la traduction automatique (TA), l'intelligence artificielle (IA) et les outils informatiques de manière générale, mais aussi pour le **discours** qui s'articule autour de ces technologies.

I. Des logiciels utiles aux traducteur·trice·s littéraires ?

- ◇ Une sorte d'**objection préjudicielle** lancée successivement à l'encontre :
 - de la traduction (littéraire) humaine pour commencer (Ladmiral 1994) ;
 - des logiciels de traitement de texte à l'origine (ATLAS 1988) ;
 - de la traduction assistée par ordinateur (Hansen 2017) ;
 - et maintenant de la traduction automatique.

“ Why do literary translators, and their trainers, **often give the impression that they are at best ambivalent about the use of technology** in literary translation, and at worst simply antagonistic? ”
(Youdale 2019)

I. Des logiciels utiles aux traducteur·trice·s littéraires ?

- ◇ Les outils de TAO :
 - offrent des **solutions en contexte** ;
 - renforcent la **cohérence stylistique** et **terminologique** ;
 - facilitent la **prise de décision** et permettent de faire des **choix éclairés** ;
 - favorisent l'**analyse critique** du texte source ;
 - accélèrent la traduction et les recherches ;
 - ↳ focalisation les passages critiques, gain en **qualité** et en **créativité**.
 - Permettent de **réduire la charge cognitive** ;
 - rendent la pratique **plus écologique**...
- ◇ MAIS devraient être adaptés pour mieux convenir encore à ce type de textes (Hansen 2021).

I. Des logiciels utiles aux traducteur·trice·s littéraires ?

- ◇ 2019 : *MT Summit*, « Workshop on Qualities of Literary Machine Translation »
- ◇ 2019 : *EST*, « Translation technologies for creative-text translation »
- ◇ 2020 : *La traduction littéraire et SHS à la rencontre des tech. de la traduction*
- ◇ 2020 : *Creative Translation and Technologies Expert Meeting*
- ◇ 2021 : *BCBF*, « Machine Translation and Human Creativity »
- ◇ 2021 : *Conference on Computer-Assisted Literary Translation (CALT)*
- ◇ À venir : *TRACT*, « L'Impact de l'IA sur la traduction littéraire »

I. Des logiciels utiles aux traducteur·trice·s littéraires ?

II. Vers une traduction littéraire automatique ?

III. Peut-on mesurer la complexité des textes littéraires ?

IV. Comment évaluer une sortie de traduction automatique ?

II. Vers une traduction littéraire automatique ?

◇ Objectifs :

- nuancer le débat ;
- offrir un aperçu concret et objectif des capacités de la TA en littérature ;
- voir quelle aide elle peut apporter au traducteur littéraire ;
- et quelle est la plus-value apportée par l'humain ;
- évaluer les principales difficultés rencontrées par la machine dans ce domaine ;
- nourrir la réflexion sur l'évaluation de la TA et les métriques ;
- aborder des aspects sociétaux tels que les changements attendus pour l'ensemble du secteur (depuis la formation des traducteurs jusqu'à l'édition d'une œuvre) ;
- contribuer aux (in)formations sur l'utilisation des nouveaux outils en traduction.

II. Vers une traduction littéraire automatique ?

◇ Traduction automatique de la prose

- Jones & Irvine (2013) : **FR-EN**

Moses entraîné sur des données hors domaine (Hansard).

- Besacier (2014) : **EN-FR**

Moses entraîné sur des données hors domaine (IWSLT) et affiné sur le roman.

- Toral & Way (2015) : **ES-CAT**

Moses entraîné sur un corpus de presse (*El Periódico*) et deux romans.

II. Vers une traduction littéraire automatique ?

- ◇ Élaboration d'un moteur adapté à la traduction littéraire :
- ◇ en recourant à la **TA neuronale** (Bahdanau et coll. 2015 ; Vaswani 2017)
 - ↳ là où les études pilotes utilisent la TA probabiliste (Besacier & Schwartz 2015 ; Toral & Way 2015)
- ◇ en faisant usage d'un **corpus spécialisé** (Dernovsek 2018 ; Wiggins 2018)
 - ↳ là où les essais précédents ont recours à des articles de presse, des brevets... (Voigt & Jurafsky 2012).

II. Vers une traduction littéraire automatique ?

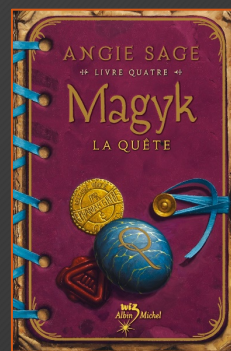
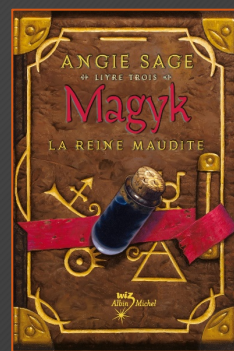
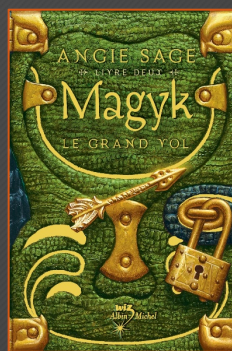
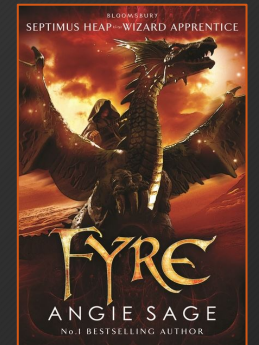
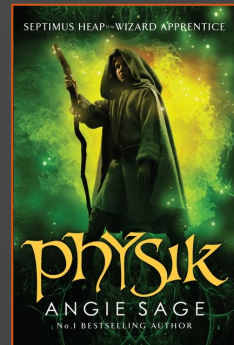
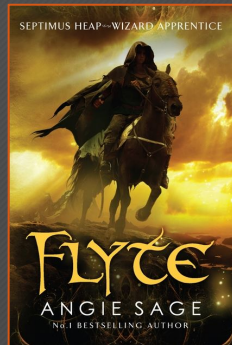
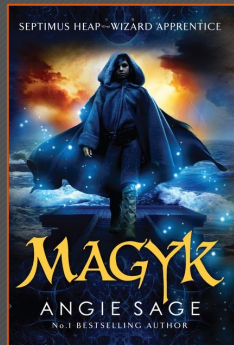
- ◇ Un nombre croissant d'études depuis ~2019 :
 - ◇ Tezcan et coll. (2019-2020) pour la paire **EN-NL**.
(Évaluation sur Google Traduction.)
 - ◇ Ó Murchú (2019) pour la paire **GLA-GLE**.
(Évaluation sur Intergaelic.)
 - ◇ Kuzman et coll. (2019) pour la paire **EN-SLO**.
(Évaluation d'un système maison affiné sur des données littéraires.)
 - ◇ Matusov (2019) pour les paires **EN-RU** et **DE-EN**.
(Évaluation de systèmes maison affinés sur des données littéraires.)
 - ◇ Toral et coll. (2018-2021) pour la paire **EN-CAT**.
(Évaluation de systèmes maison entraînés uniquement sur des données littéraires.)

II. Vers une traduction littéraire automatique ?

- ◆ Traduction automatique de poésie :
 - Genzel et coll. (2010)
 - Greene et coll. (2010)
 - Jones & Irvine (2013)
 - Ghazvininejad et coll. (2018)...
- ◆ Génération automatique de poésie :
 - Ghazvininejad et coll. (2016)
 - Lau et coll. (2018)
 - Van de Cruys (2019)...

- ◆ État de l'art plus complet, pour la TA, dans Hansen (2021).

II. Vers une traduction littéraire automatique ?



Sage, Angie. *Septimus Heap*. HarperCollins, 2005-2013. 7 vol.
Serval, Nathalie, traductrice. *Magyk*. Par Angie Sage, Albin Michel, 2005-2013. 6 vol.

II. Vers une traduction littéraire automatique ?

- ◇ Premier obstacle : un problème de **données** !
- ◇ Corpus littéraire personnel : **45 K phrases** (à diviser en corpus d'entraînement, test et valid.)
- ◇ Seul corpus littéraire publiquement accessible : **127 K phrases** (bruitées et mal alignées)
- ◇ On estime qu'il faut au moins **2 M de phrases** pour entraîner des systèmes basiques.

II. Vers une traduction littéraire automatique ?

◆ Entraîner un système générique :

- Europarl (2 M)
- TedTalks (410 K)
- GobaVoices (195 K)
- News Commentary (183 K)
- Books (127 K)

OpenNMT (Klein et coll. 2020)

Pour les corpus :
Cf. Tiedemann (2012)

- ◆ Relancer l'entraînement en affinant sur les données littéraires (45 K).
- ◆ Conserver les corpus génériques pour éviter les oublis (*catastrophic forgetting*).
- ◆ Donner un poids plus haut aux données spécialisées.

II. Vers une traduction littéraire automatique ?

Systeme en ligne	sacreBLEU
Google Translate	10.79
DeepL	10.04

Systeme maison	sacreBLEU
LSTM (générique)	09.00
LSTM (affiné)	13.49
Transf. (générique)	10.26
Transf. (affiné)	17.04

- ◆ Certains considèrent qu'il faut un min. de 20 BLEU pour que la PE soit utile (Toral & Way 2015).
- ◆ Néanmoins, on ne cherche pas uniquement ici à évaluer la qualité de la traduction, mais aussi à quel point elle se rapproche du texte produit par la traductrice française et de son style (exigence doublement élevée ; d'où les scores très faibles au départ).
- ◆ C'est nécessaire si l'on veut que la TA soit utile aux professionnel·le·s.
- ◆ + 7 BLEU environ dans ce sens, avec pourtant assez peu de données.

II. Vers une traduction littéraire automatique ?

◇ Par comparaison avec d'autres essais d'adaptation à la littérature :

Système (EN-SLO)	BLEU
LSTM (générique)	17.50
LSTM (affiné)	20.75
Google Traduction	21.97

Kuzman et coll. (2019)

Système (EN-RU)	BLEU
LSTM (générique)	14.20
LSTM (affiné)	15.20
Google Traduction	13.90

Matusov (2019)

Système (DE-EN)	BLEU
Transf. (générique)	18.50
Transf. (affiné)	16.20
Google Traduction	20.20

Matusov (2019)

II. Vers une traduction littéraire automatique ?

- ◇ Différentes paires de langues.
- ◇ Différentes directions.
- ◇ Différentes approches.
- ◇ Différents corpus d'entraînement.
- ◇ Différent·e·s auteur·e·s.
- ◇ Différent·e·s traducteur·trice·s.
- ◇ Différents genres littéraires...

I. Des logiciels utiles aux traducteur·trice·s littéraires ?

II. Vers une traduction littéraire automatique ?

III. Peut-on mesurer la complexité des textes littéraires ?

IV. Comment évaluer une sortie de traduction automatique ?

V. Conclusion

III. Peut-on mesurer la complexité des textes littéraires ?

◇ Traduction automatique par les systèmes en ligne :	Google Trad.	DeepL
- Jane Austen - <i>Pride and Prejudice</i> :	13.30	14.89
- Charlotte Brontë - <i>Jane Eyre</i> :	18.72	19.92
- Lewis Carroll - <i>Alice in Wonderland</i> :	20.01	20.21
- Daniel Defoe - <i>Moll Flanders</i> :	27.48	30.31
- Daniel Defoe - <i>Robinson Crusoe</i> :	20.71	21.02
- Arthur Conan Doyle - <i>The Great Shadow</i> :	23.23	24.69
- Arthur Conan Doyle - <i>The Hound of the Baskervilles</i> :	22.90	24.10
- Edgar Allan Poe - <i>The Fall of the House of Usher</i> :	24.90	25.10
- Jerome Jerome - <i>Three Men in a Boat</i> :	16.23	17.15
- Angie Sage - <i>Septimus Heap</i> :	10.79	10.04

III. Peut-on mesurer la complexité des textes littéraires ?

◆ *Word alignment perplexity* donné par GIZA++ (Och & Ney 2003) :

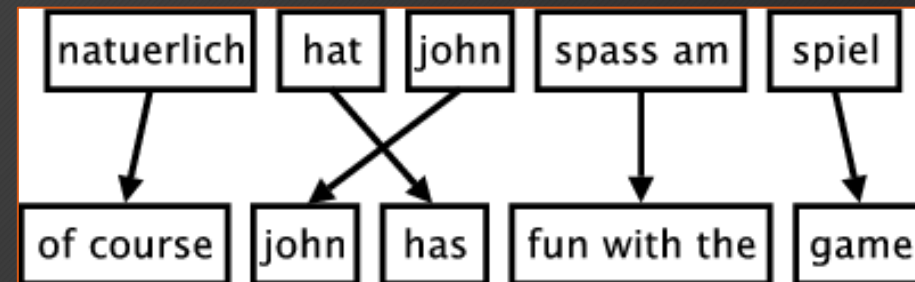
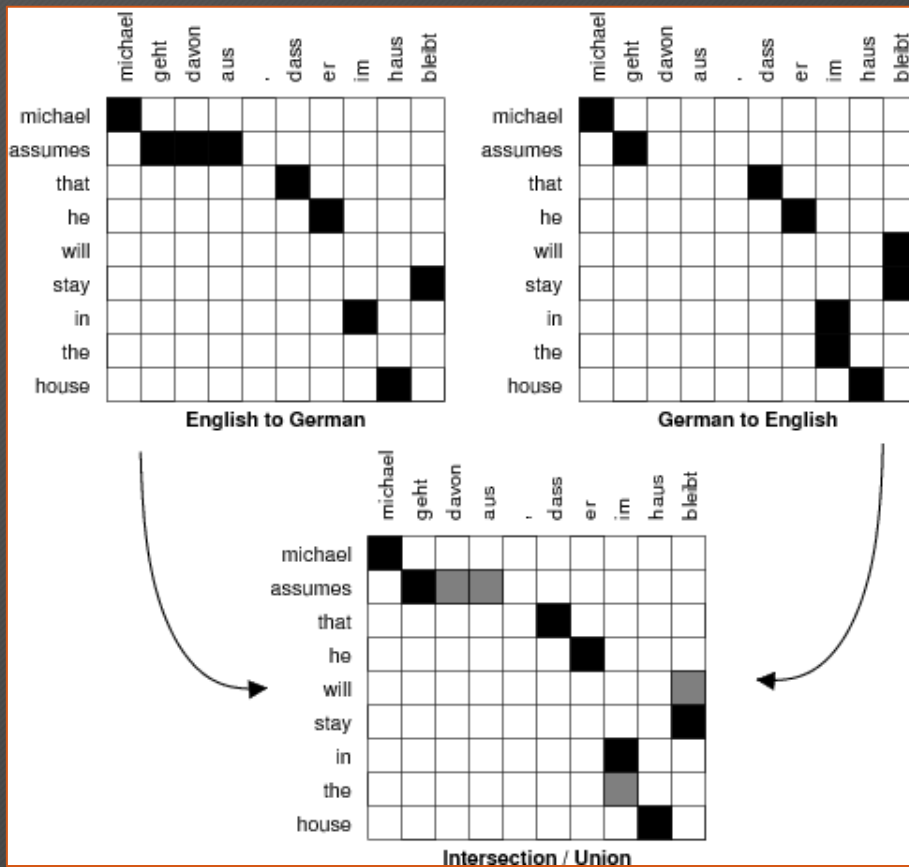
◆ Inspiré par Toral & Way (2015) :

- roman évalué (en-es) : 30
- presse (en-es) : 32
- institutionnel (en-es) : 44

◆ Dans notre cas :

- roman évalué (en-fr) : 55
- littérature diverse (en-fr) : 32
- presse (en-fr) : 27
- institutionnel (en-fr) : 38

III. Peut-on mesurer la complexité des textes littéraires ?



III. Peut-on mesurer la complexité des textes littéraires ?

◇ *Word alignment perplexity* donné par GIZA++ (Och & Ney 2003) :

- Jane Austen - *Pride and Prejudice* : 46
- Charlotte Brontë - *Jane Eyre* : 38
- Lewis Carroll - *Alice in Wonderland* : 19
- Daniel Defoe - *Moll Flanders* : 28
- Daniel Defoe - *Robinson Crusoe* : 37
- Arthur Conan Doyle - *The Great Shadow* : 24
- Arthur Conan Doyle - *The Hound of the Baskervilles* : 24
- Edgar Allan Poe - *The Fall of the House of Usher* : 10
- Jerome Jerome - *Three Men in a Boat* : 35

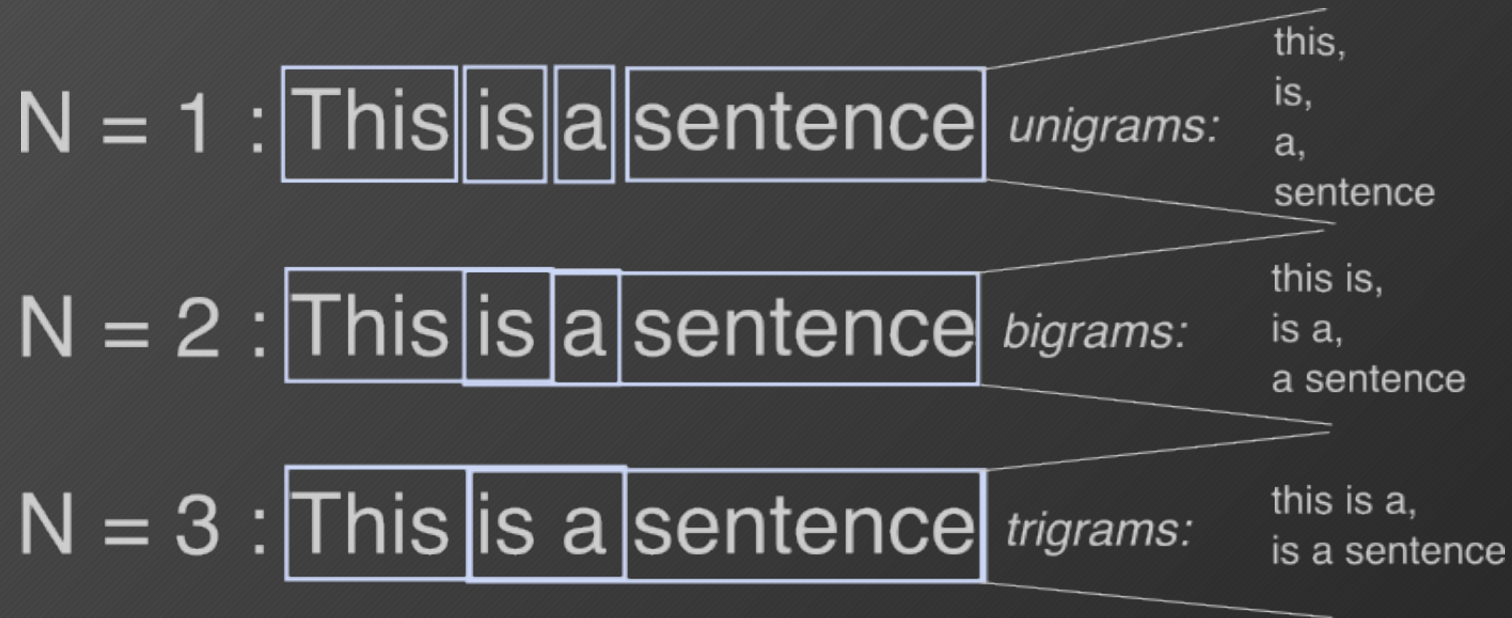
- Angie Sage - *Septimus Heap* : 55

III. Peut-on mesurer la complexité des textes littéraires ?

◇ Perplexité des modèles de langue donnée par SRILM (Stolcke 2002) :

- Jane Austen - <i>Pride and Prejudice</i>	125
- Charlotte Brontë - <i>Jane Eyre</i> :	134
- Lewis Carroll - <i>Alice in Wonderland</i> :	91
- Daniel Defoe - <i>Moll Flanders</i> :	109
- Daniel Defoe - <i>Robinson Crusoe</i> :	129
- Arthur Conan Doyle - <i>The Great Shadow</i> :	150
- Arthur Conan Doyle - <i>The Hound of the Baskervilles</i> :	132
- Edgar Allan Poe - <i>The Fall of the House of Usher</i> :	not enough data
- Jerome Jerome - <i>Three Men in a Boat</i> :	162
- Angie Sage - <i>Septimus Heap</i> :	195

III. Peut-on mesurer la complexité des textes littéraires ?



III. Peut-on mesurer la complexité des textes littéraires ?

- ◇ Dans tous les cas, l'ouvrage de *fantasy* obtient les scores indiquant la plus grande difficulté.
- ◇ (Alors qu'il s'agit du corpus le plus propre.)
- ◇ Va totalement à l'encontre de l'idée selon laquelle la « **paralittérature** », la littérature de **jeunesse**, la **fiction**... seraient des œuvres plus **simples** et plus **facilement traduisibles** pour la machine, des **cibles de choix pour la TA**.

III. Peut-on mesurer la complexité des textes littéraires ?

◇ Attribuable dans ce cas à :

- une **traduction très libre** ;
- un **registre globalement soutenu** ;
- mais de **grandes variations dans le registre** pour certains personnages ;
- l'usage de **régionalismes** ;
- certains discours volontairement **vieillis et colorés** ;
- certains discours adaptés du **vieux français** ;
- la création d'un **univers fictif** qui n'a aucune existence en dehors de cette saga ;
- de nombreux concepts et néologismes spécifiques à la saga (cf. **irréalias**, Loponen 2009).

III. Peut-on mesurer la complexité des textes littéraires ?

- ◇ Tâche **particulièrement compliquée** : la traduction (littéraire) est loin d'être une tâche résolue, et le cas des textes créatifs (de notre texte de *fantasy* notamment) illustre bien la **plus-value de la traduction humaine**.
- ◇ Mais l'affinage se révèle ici très efficace en raison de la **cohérence des données** utilisées pour l'entraînement.
 - ↳ **Une saga, une auteure, une traductrice.**
- ◇ La TA se profile ainsi comme une aide intéressante, mais surtout si le système est entraîné par le traducteur ou la traductrice, sur ses productions (le système sera alors capable de faire des **choix plus pertinents** et de fournir des **suggestions utiles** à l'humain).

III. Peut-on mesurer la complexité des textes littéraires ?

- ◇ Début des essais, d'autres choses à essayer encore pour améliorer les résultats :
 - segmentation en sous-mots (Sennrich 2016 ; Kudo & Richardson 2018) ;
 - utilisation de données synthétiques (Sennrich 2016) ;
 - ...
- ◇ Tout en sachant que la TLA bénéficiera de nouvelles techniques d'**augmentation de données** (Fadaee et coll. 2017), d'**adaptation au domaine** (Chu & Wang 2018) ou de la prise en charge du **contexte** et des **éléments textuels** (Lopes et coll. 2020).
- ◇ Et, bien sûr, des **évaluations** plus poussées !

II. *Vers une traduction littéraire automatique ?*

III. *Peut-on mesurer la complexité des textes littéraires ?*

IV. *Comment évaluer une sortie de traduction automatique ?*

V. *Conclusion*

IV. Comment évaluer une sortie de traduction automatique ?

◇ Évaluation automatique :

- **métriques d'évaluation automatique** (BLEU, TER, HTER, COMET, METEOR, YiSi...);
- **mesures de richesse lexicale** (TTR, RTTR, CTTR, MSTTR...)
- mesures de cohésion et de recouvrement lexical ou sémantique ;
- **mesures d'équivalence syntaxique** (*cross-value*) ;
- *word translation entropy* (AWTE) ;
- effort de postédition (temporel, technique...);

IV. Comment évaluer une sortie de traduction automatique ?

◇ Évaluation automatique :

- très peu, voire aucun **exemple de ce qui est produit concrètement** ;
- presque uniquement du **BLEU** ;
- particulièrement **problématique en littérature** (Sin-wai 2015) ;
- parfois d'autres métriques, mais **l'usage varie selon les auteur·e·s.**

IV. Comment évaluer une sortie de traduction automatique ?

◇ Évaluation humaine :

- **annotation d'erreurs** (système générique vs affiné, post-édition vs trad. humaine) ;
- **typologies d'erreurs** (cf. toutes les typologies déjà existantes) ;
- enquêtes et TAP durant la post-édition (ergonomie, interaction humain-machine) ;
- **enquêtes et évaluations auprès de traducteur·trice·s littéraires** (qualité) ;
- enquêtes et évaluations auprès de lecteur·trice·s (expérience de lecture) ;
- **commentaires critiques sur certains points de traduction vis-à-vis de la référence** (créativité, respect des choix de traductions des précédents volumes...).

IV. Comment évaluer une sortie de traduction automatique ?

◇ Évaluation humaine :

- offre une **analyse plus fine** des résultats (concernant p. ex. le type d'erreur) ;
- accord inter-annotateur **souvent trop peu suffisant pour en tirer des conclusions** (Kuzman et coll. 2019 ; Guerberof-Arenas & Toral 2020) ;
- ou **contradictoire d'une étude à l'autre** (p. ex. Moorkens et coll. 2018 vs Şahin & Gürses 2019 ; Guerberof-Arenas & Toral 2020 vs Nunes Vieira et coll. 2020).
- les estimations des traducteur·trice·s durant la post-édition **ne coïncident pas toujours avec la réalité** (Moorkens et coll. 2018) ;
- pour l'heure, **aucun juge familier à la fois avec la littérature et la post-édition.**

IV. Comment évaluer une sortie de traduction automatique ?

- ◇ Dans le meilleur des cas, l'introduction de cette technologie pourrait avoir des effets positifs sur l'ensemble de la chaîne de traduction (Besacier 2014) :
 - des gains sur la qualité et la créativité pour les **traducteur·trice·s** ;
 - des coûts plus faibles pour les **éditeur·trice·s**, en espérant que cela ait aussi une influence positive sur la rémunération des professionnel·le·s ;
 - en tant qu'**auteur·e**, la possibilité de voir son travail traduit dans plus de langues ;
 - et pour les **lecteur·trice·s**, une parution plus rapide de leurs romans favoris.

IV. Comment évaluer une sortie de traduction automatique ?

- ◇ La TA pourrait aussi aider à **réduire la charge cognitive** des traducteurs et traductrices (Tavalkoski-Shilov 2019).
- ◇ Et leur permettre de se concentrer sur les segments les plus intéressants, rendant ainsi globalement la **tâche plus agréable** (*Ibid.*).
- ◇ Mais **uniquement si la TA est correctement implémentée** dans le flux de travail ! Sans quoi les effets pourraient être inverses sur chacun de ces points.
 - ↳ Des systèmes semblables aux moteurs AdaptiveMT/ModernMT, au sein des interfaces de **TAO**, pourraient être une solution intéressante de ce point vue.

IV. Comment évaluer une sortie de traduction automatique ?

- ◇ À l'inverse, elle suscite des préoccupations qui sont **exacerbées dans le secteur littéraire** en raison des difficultés du marché, de la volonté de la part des maisons d'édition à tenter de réduire les coûts par tous les moyens, et de la situation déjà précaire des traducteur·trice·s littéraires (Taivalkoski-Shilov 2019) :
 - certaines maisons d'édition vendent déjà **des traductions non révisées par l'humain** ;
 - elles pourraient être tentées d'**engager des non-professionnels** pour réduire davantage les coûts, ce que l'on constate déjà même sans utilisation de la TA ;
 - son introduction a tendance à **réduire les rémunérations** et **les délais** ;
 - et pose des questions **de droits d'auteur, de propriété sur les données** et **de visibilité**.

IV. Comment évaluer une sortie de traduction automatique ?

- ◇ Si elle est utilisée uniquement à des fins de productivité (p. ex. simple révision d'un texte entièrement traduit par la machine), cela pourrait entraîner des **conséquences néfastes sur la qualité et la créativité**.
- ◇ Cela pourrait nuire en retour (Taivalkoski-Shilov 2019) :
 - à **l'expérience de lecture des client·e·s** ;
 - au **travail de l'auteur·e** ;
 - à **la reconnaissance du métier de traducteur·trice** ;
 - au **processus d'apprentissage de la langue** ;
 - au **transfert des cultures**.

IV. Comment évaluer une sortie de traduction automatique ?

- ◇ Quelques résultats remarquables jusqu'à présent lors de l'évaluation :
 - des erreurs grossières (sans surprise) ;
 - une tendance à traduire très littéralement ;
 - le vocabulaire spécifique entièrement respecté ;
 - des paraphrases astucieuses (mais parfois malvenues) ;
 - une bonne gestion des conventions typographiques de l'éditeur ;
 - des segments fusionnés ou divisés dans la cible et non dans la source ;
 - des erreurs surtout dues à de mauvais choix lexicaux (le système reste 'simpliste').

IV. Comment évaluer une sortie de traduction automatique ?

◆ Quelques exemples (étude en cours, sélectionnés à dessein !) :

Src : Two books—**The Undoing of the Darkenese** and **The Darke Index**—lay open on Marcia's desk. In the shadow of the books a small silver box and a tiny silver disc rested on a piece of purple velvet.

Hyp : Deux livres, **L'Art de vaincre la Ténèbre** et **L'Codex Tenebrae**, étaient ouverts sur le bureau de Marcia **tandis qu'une petite boîte d'argent** et **un petit disque d'argent** reposaient sur un morceau de velours pourpre.

Ref : Deux livres - **L'Art de vaincre la Ténèbre** et **le Codex Tenebrae** - étaient ouverts sur la table de travail de Marcia, **à côté d'une** petite boîte en argent et **d'un disque du même métal** posé sur un carré de velours violine.

Src : Marcia had found **The Undoing of the Darkenese** in Aunt Zelda's cottage a few years back and she knew that somewhere within its pages lay **The Great UnDoing**—the legendary Anti-Darke incantation that practitioners of the Darke feared above all else.

Hyp : Quelques années plus tôt, Marcia avait trouvé **L'Art de vaincre la Ténèbre** dans le cottage de tante Zelda et elle savait qu'à l'intérieur de ses pages se trouvait **l'incantation anti-Ténèbre, une incantation légendaire** que les praticiens de la Ténèbre craignaient par-dessus tout.

Ref : Puis Marcia avait découvert **L'Art de vaincre la Ténèbre** au cottage de tante Zelda, quelques années plus tôt. Elle savait qu'il contenait **la formule du Grand Dénouement**, le contresort mythique que les adeptes de la Ténèbre redoutaient par-dessus tout.

IV. Comment évaluer une sortie de traduction automatique ?

◆ Quelques exemples (étude en cours, sélectionnés à dessein !) :

Src : Jenna looked around the entrance hall, which was full of shadows and ghosts. She reached into her pocket for **the Candlelight Charm** Septimus had given her that afternoon.

Hyp : Jenna promena son regard autour de la salle d'entrée, remplie d'ombres et de fantômes, et plongea la main dans sa poche, cherchant **le charme chocolaté** que Septimus lui avait offert cet après-midi-là.

Ref : Jenna promena son regard autour du hall, peuplé d'ombres et de fantômes, puis elle sortit de sa poche **l'allume-bougies magique** que Septimus lui avait offert cet après-midi-là.

Src : Aunt Zelda **settled herself beside the fire**. She looked at the **duck** in Jenna's arms.

Hyp : Tante Zelda **s'installa près du feu** et considéra la **cane** dans les bras de Jenna.

Ref : Tante Zelda **s'assit près de la cheminée** et son regard se posa sur la **cane**, blottie dans les bras de Jenna.

Src : **She turned to Septimus. "You really don't have to come," she said.**

Hyp : – Tu n'es pas obligé de m'accompagner, dit-elle à Septimus.

Ref : Là, elle se tourna vers Septimus et lui dit : – Tu n'es pas obligé de venir.

III. *Peut-on mesurer la complexité des textes littéraires ?*

IV. *Comment évaluer une sortie de traduction automatique ?*

V. *Conclusion*

V. Conclusion

- ◇ Est-ce qu'il est possible d'**adapter la TA à la littérature** ?

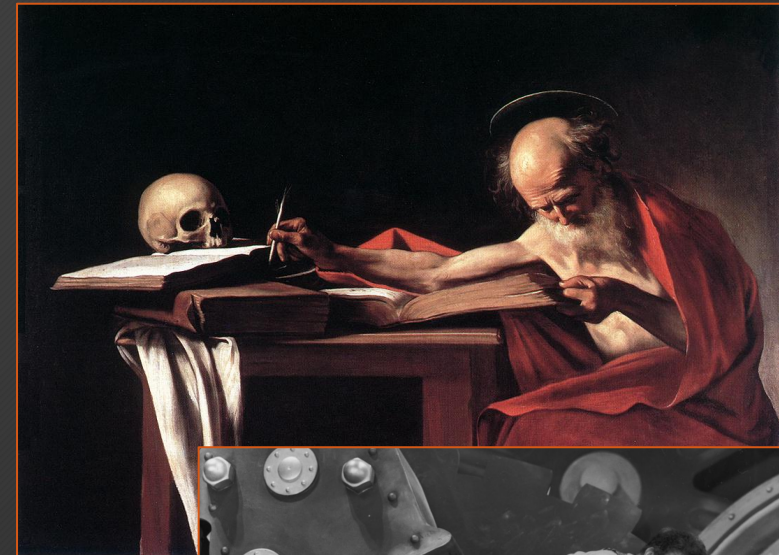
Oui, et même à la production d'un·e traducteur·trice.

- ◇ Est-ce à dire qu'on pourrait **se passer de l'humain** ?

Non, loin de là, quoique que la TA laisse entrevoir une utilité comme aide à la lecture ou à l'apprentissage des langues (Oliver et coll. 2019 ; Matusov 2019) ou comme outil de promotion de la diversité dans le monde de l'édition (Castro 2020).

- ◇ La TLA, à terme, pourrait-elle **devenir un outil utile** ?

C'est possible, mais uniquement si elle est correctement utilisée, idéalement entraînée par les professionnel·le·s, sur leurs propres traductions, en combinaison avec d'autres outils (TAO, corpus...).



Bibliographie

- ◇ ATLAS. *Actes des quatrièmes assises de la traduction littéraire*. Édité par Claire Malroux et coll., Actes Sud, 1988.
- ◇ Bahdanau, Dzmitry, et coll. « Neural Machine Translation by Jointly Learning to Align and Translate ». *3rd International Conference on Learning Representations: Conference Track Proceedings*, édité par Yoshua Bengio et Yann LeCun, 2015, p. 1-15.
- ◇ Besacier, Laurent. « Traduction automatisée d'une oeuvre littéraire : une étude pilote ». *Actes de la 21e conférence sur le Traitement Automatique des Langues Naturelles*, édité par Philippe Blache et coll., ATALA, 2014, p. 389-394.
- ◇ Castro, Olga. « Transnational Feminism ». *Creative Translation and Technologies Expert Meeting*, Université de Surrey, 29 mai 2020.
- ◇ Chu, Chenhui, et Rui Wang. « A Survey of Domain Adaptation for Neural Machine Translation ». *Proceedings of the 27th International Conference on Computational Linguistics*, édité par Emily Bender et coll., ACL, 2018, p. 1304-1319.

Bibliographie

- ◇ Dernovsek, Jure. « Moving Towards Augmented Translation. A Case Study. » *Omniscien AI, MT and Language Processing Symposium*, 26 mars 2018.
- ◇ Fadaee, Marzieh, et coll. « Data Augmentation for Low-Resource Neural Machine Translation ». *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, édité par Regina Barzilay et Min-Yen Kan, ACL, 2017, p. 567-573, doi : 10.18653/v1/P17-2090.
- ◇ Guerberof-Arenas, Ana, et Antonio Toral. « The Impact of Post-Editing and Machine Translation on Creativity and Reading Experience ». *Translation Spaces*, vol. 9, n° 2, 2020, p. 255-282, doi : 10.1075/ts.20035.gue.
- ◇ Hansen, Damien. « La traduction littéraire assistée par ordinateur : de l'unité de traduction à l'unité de style ». Des unités de traduction à l'unité de la traduction, 21 octobre 2017, Université libre de Bruxelles.
- ◇ Hansen, Damien. « Artificial intelligence and machine learning: How a better grasp of AI can frame and facilitate the dialogue about its role in the years to come ». CIUTI Conference 2020, 11 décembre 2020, ISIT, Paris.

Bibliographie

- ◆ Hansen, Damien. « Les lettres et la machine : un état de l'art en traduction littéraire automatique ». *Actes de la 28e Conférence sur le Traitement Automatique des Langues Naturelles*, vol. 2, édité par Pascal Denis et coll., ATALA, 2021, p. 28-45.
- ◆ Hansen, Damien. « The figure of the literary translator amid new technologies ». Computer-Assisted Literary Translation Conference, 11 mai 2021, Université de Swansea.
- ◆ Hutchins, John. « The first public demonstration of machine translation: the Georgetown-IBM system, 7th January 1954 ». *Proceedings of the 6th Conference of the AMTA: Technical Papers*, édité par Robert Frederking et Kathryn Taylor, Springer, 2004, p. 102-114.
- ◆ Jones, Ruth, et Ann Irvine. « The (Un)faithful Machine Translator ». *Proceedings of the 7th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, édité par Piroska Lendvai et Kalliopi Zervanou, ACL, 2013, p. 96-101.
- ◆ Klein, Guillaume, et coll. « The OpenNMT Neural Machine Translation Toolkit: 2020 Edition ». *Proceedings of the 14th Conference of the Association for Machine Translation in the Americas (Volume 1: Research Track)*, édité par Michael Denkowski et Christian Federmann, AMTA, 2020, p. 102-109

Bibliographie

- ◇ Kudo, Taku, et John Richardson. « SentencePiece: A simple and language independent subword tokenizer and detokenizer for Neural Text Processing ». *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (System Demonstrations)*, édité par Eduardo Blanco et Wei Lu, ACL, 2018, p. 66-71.
- ◇ Kuzman, Taja, et coll. « Neural Machine Translation of Literary Texts from English to Slovene ». *Proceedings of the Qualities of Literary Machine Translation*, édité par James Hadley et coll., EAMT, 2019, p. 1-9.
- ◇ LADMIRAL, Jean-René. *Traduire : théorèmes pour la traduction*. 2^e éd., Gallimard, 1994.
- ◇ Lopes, António, et coll. « Document-level Neural MT: A Systematic Comparison ». *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation*, édité par André Martins et coll., EAMT, 2020, p. 225-234.
- ◇ LoPonen, Mika. « Translating Irrealia: Creating a Semiotic Framework for the Translation of Fictional Cultures ». *Chinese semiotic studies*, vol. 2, n° 1, 2009, p. 165-75, doi : 10.1515/css-2009-0117.

Bibliographie

- ◆ Matusov, Evgeny. « The Challenges of Using Neural Machine Translation for Literature ». *Proceedings of the Qualities of Literary Machine Translation*, édité par James Hadley et coll., EAMT, 2019, p. 10-19.
- ◆ Moorkens, Joss, et coll. « Translators' Perceptions of Literary Post-Editing using Statistical and Neural Machine Translation ». *Translation Spaces*, vol. 7, n° 2, p. 240-262, doi : 10.1075/ts.18014.moo.
- ◆ Nunes Vieira, Lucas, et coll. « Machine Translation and Literary Texts: A Network of Possibilities ». *Creative Translation and Technologies Expert Meeting*, Université de Surrey, 29 mai 2020.
- ◆ Oliver González, et coll.. « InLéctor: Neural Machine Translation for the creation of bilingual ebook? ». *Proceedings of the Qualities of Literary Machine Translation*, édité par James Hadley et coll., EAMT, 2019, p. vii.
- ◆ Ó Murchú, Eoin P. « Using Intergaelic to Pre-Translate and Subsequently Post-Edit a Sci-Fi Novel from Scottish Gaelic to Irish ». *Proceedings of the Qualities of Literary Machine Translation*, édité par James Hadley et coll., EAMT, 2019, p. 20-25.

Bibliographie

- ◇ Och, Franz Josef, et Hermann Ney. « A Systematic Comparison of Various Statistical Alignment Models ». *Computational Linguistics*, vol. 29, n° 1, 2003, p. 19-51, doi : 10.1162/089120103321337421.
- ◇ Rossi, Caroline. « Les usages actuels de la traduction automatique ». Atelier Digit_Hum 2019 : Les humanités numériques en langues, 17 octobre 2019, ENS, Paris.
- ◇ Şahin, Mehmet, et Sabri Gürses. « Would MT kill creativity in literary retranslation? ». *Proceedings of the Qualities of Literary Machine Translation*, édité par James Hadley et coll., EAMT, 2019, p. 26-34.
- ◇ Sennrich, Rico et coll. « Improving Neural Machine Translation Models with Monolingual Data ». *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, édité par Katrin Erk et Noah Smith, ACL, 2016, p. 86-96.
- ◇ Sennrich, Rico et coll. « Neural Machine Translation of Rare Words with Subword Units ». *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, édité par Katrin Erk et Noah Smith, ACL, 2016, p. 1715-1725.

Bibliographie

- ◇ Sin-wai, Chan, éditeur. *The Routledge Encyclopedia of Translation Technology*. Routledge, 2015.
- ◇ Stolcke, Andreas. « SRILM – An extensible language modeling toolkit ». *ICSLP 2002, 7th International Conference on Spoken Language Processing*, 2002, p. 901-904.
- ◇ Taivalkoski-Shilov, Kristiina. « Ethical Issues Regarding Machine(-Assisted) Translation of Literary Texts ». *Perspectives: Studies in Translation Theory and Practice*, vol. 27, n° 5, 2019, p. 689-703, doi : 10.1080/0907676X.2018.1520907.
- ◇ Tezcan, Arda, et coll. « When a ‘Sport’ Is a Person and Other Issues for NMT of Novels ». *Proceedings of the Qualities of Literary Machine Translation*, édité par James Hadley et coll., EAMT, 2019, p. 40-49.
- ◇ Tiedemann, Jörg. « Parallel Data, Tools and Interfaces in OPUS ». *Proceedings of the 8th International Conference on Language Resources and Evaluation*, édité par Nicoletta Calzolari et coll., ELRA, 2012, p. 2214-2218.
- ◇ Toral, Antonio, et Andy Way. « Machine-Assisted Translation of Literary Text: A Case Study ». *Translation Spaces*, vol. 4, n° 2, 2015, p. 240-267, doi : 0.1075/ts.4.2.04tor.

Bibliographie

- ◆ Toral, Antonio, et Andy Way. « Translating Literary Text between Related Languages using SMT ». *Proceedings of the Fourth Workshop on Computational Linguistics for Literature*, édité par Anna Feldman et coll., ACL, 2015, p. 123-132.
- ◆ Toral, Antonio, et coll. « Machine Translation of Novels in the Age of Transformer ». *Maschinelle Übersetzung für Übersetzungsprofis*, édité par Jörg Porsiel, BDÜ Fachverlag, 2020, p. 276-295.
- ◆ Vaswani, Ashish, et coll. « Attention is All you Need ». *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*, édité par Ulrike von Luxburg et coll., Curran Associates Inc., 2017, p. 6000-6010.
- ◆ Voigt, Rob, et Dan Jurafsky. « Towards a Literary Machine Translation: The Role of Referential Cohesion ». *Proceedings of the NAACL-HLT 2012 Workshop on Computational Linguistics for Literature*, édité par David Elson et coll., ACL, 2012, p. 18-25.
- ◆ Wiggins, Dion. « Understanding the Benefits of Specialized Machine Translation and Language Processing Solutions ». *Omniscien AI, MT and Language Processing Symposium*, 27 mars 2018.
- ◆ Youdale, Roy. *Using computers in the translation of literary style: challenges and opportunities*. Routledge, 2019.



LIÈGE université

UR CIRT

Centre Interdisciplinaire de Recherches
en Traduction et en Interprétation



N'hésitez pas à prendre contact :
damien.hansen@uliege.be

Diapositives et bibliographie disponibles
sur mon dépôt institutionnel (ORBi)

Merci de votre attention