Personalized MT Systems for Literary Translators

Damien Hansen

University of Liège (CIRTI) – Grenoble Alpes University (LIG/GETALP)

Outline

I. Starting point

II. Literary corpora

III. Adapting a generic system

IV. Suggestion of a new paradigm for LMT

V. Conclusions

I. Starting point

II. Literary corpora

III. Adapting a generic system

I. Starting point

◈ Back in 2017:

–   We tried our hand at computer-assisted literary translation (CALT), but research within the corpus is always manual.

–   Could machine translation (MT) help by creating artificial matches and offering custom suggestions tailored to our domain and translator?

⇨   Adapt a system to the literary domain, by training it on relevant data (Besacier 2014, Toral & Way 2015).

⇨   Use the latest development of neural machine translation (Bahdanau et al. 2014).

I. Starting point

◈ Now that a few years have passed. Objective of the project as a whole:

 – train a system specialized to the literary field
 (as in Toral & Way 2018, Matusov 2019, Kuzman et al. 2019);

 – evaluate the resulting translation produced by the MT system
 (as in Tezcan et al. 2019, Macken et al. 2022, Castilho & Resende 2022);

 – address the issues that the inclusion of such a tool might bring to the field
 (as in Taivalkoski-Shilov 2019, Guerberof-Arenas & Toral 2020, Kenny & Winters 2020);

 – inform and comment on common discourses regarding literary texts and translation technology
 (as in Ruffo 2018, Slessor 2020, Daems 2021)

I. Starting point

◈ What this work suggests:

– a new experiment for the English–French pair that has not been re-attempted since SMT (Besacier 2014);

– an experiment in the fantasy genre, which present specific linguistic and textual challenges (Hansen et al. 2022);

– an adaptation procedure in which the system is adapted not just to the literary domain, but to a specific author, translator, genre and series;

– a new paradigm, or research avenue, for future experiments on LMT, and their inclusion into the larger area of CALT.

The corpus (*sep-only*)
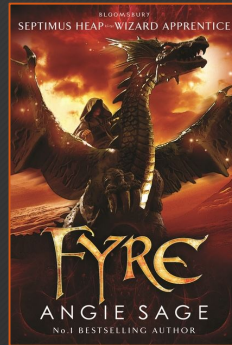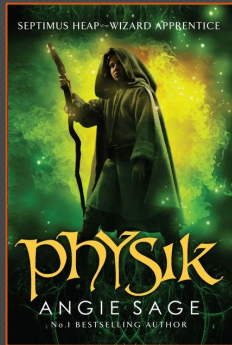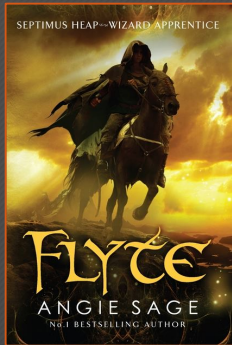
Sage, Angie. *Septimus Heap*. HarperCollins, 2005–2013. 7 vols.
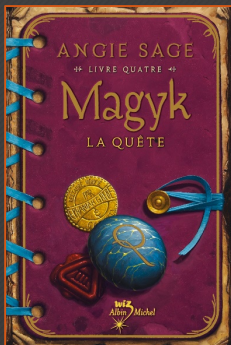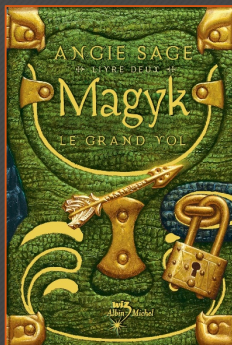
Serval, Nathalie, trans. *Magyk*. By Angie Sage, Albin Michel, 2005–2013. 6 vols.

The corpus (*sep-only*)

Measures of alignment (left) and language model (right) perplexity, inspired by Toral & Way (2015). Cf. Hansen et al. (2022).
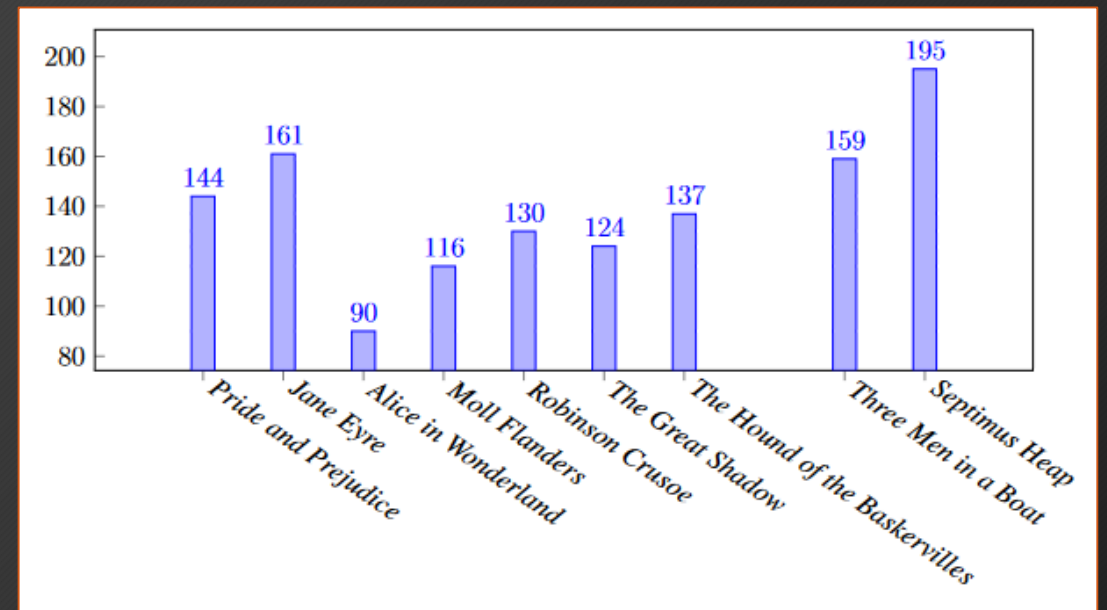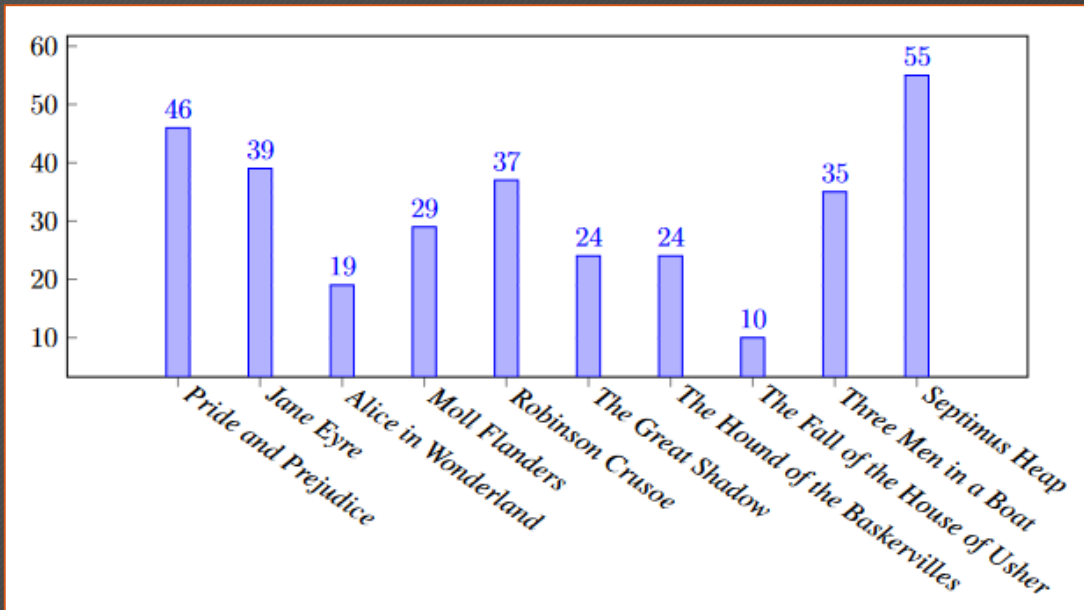
II. Literary corpora

The corpus (*sep-large*)

◈ Attempt to increase literary data, by creating corpora as in Toral & Way (2018).

◈ A synthetic corpus in the manner of Caswell et al. (2019): 150 novels, from many French-speaking countries, various time periods, in varying genres. Back-translated by DeepL.

◈ A parallel corpus of 40 novels translated by Nathalie Serval, in the fantasy, *fantastique* or science-fiction genres.

◈ A parallel corpus of 30 novels typical of the fantasy genre.

◈ No more than 2 works per author or translator.

II. Literary corpora

	Segments	Tokens EN	Tokens FR
Europarl	2,007,723	49,867,465	54,553,979
Video Game	1,370,431	21,041,902	22,804,380
TED	410,443	7,041,745	7,464,033
GlobalVoices	195,387	3,503,600	3,980,602
News	183,251	4,055,180	4,952,704
Books	127,021	2,737,133	2,770,418
Total	4,294,256	88,247,025	96,526,116

Generic corpora

	Segments	Tokens EN	Tokens FR
Synthetic	338,233	14,339,224	15,130,086
Translator	111,322	3,571,242	3,569,595
Parallel	100,055	4,014,409	4,365,486
Sep. (trn)	37,348	550,536	541,779
Sep. (val)	7,225	109,859	106,621
Sep. (tst)	704	10,181	10,073
Total	594,887	22,595,451	23,723,640

Specialized corpora

(Tiedeman 2012; Hansen & Houlmont, forthcoming)

III. Adapting a generic system

1
Un paquet dans la neige
Silas Heap s'enveloppa dans sa cape pour se protéger de la neige.
Sa longue marche à travers la Forêt l'avait glacé jusqu'aux os.

La construction de nouvelles maisons les avait amenés a p
Le Château n'avait pas tardé à attirer des artisans d
Il avait tant prospéré que la place avait fini p
Si on y vivait à l'étroit, l'atmosphère y éta
Comme le soleil d'hiver plongeait derrière les murs du Château,
Il fallait qu'il atteigne la porte Nord avant la tombée de la nui
Une présence vivante, mais à peine.
Un petit cœur humain battait près de lui.
Silas s'immobilisa.
En tant que magicien ordinaire, il percevait certaines
La neige tombait dru autour de lui ; déjà, elle reco
Puis il entendit un bruit. Était-ce un pleur, un
Il n'aurait su le dire, mais cela lui suffit.
Il trouva le paquet dans un buisson sur le
En le ramassant, il eut la surprise de
Silas la cala sur son bras, se dema
Bien qu'emmaillotée dans une ér
Silas eut le sentiment péni
Il songea à Sarah qui l'attendait bien au chaud aup
Il enveloppa soigneusement le bébé dans sa cape verte
Il atteignit le Château juste comme Gringe, le gardie
— C'était moins une, marmonna Gringe.
Vous autres magiciens, vous êtes une drôle d's
Me demande ce que vous pouvez fabriquer d
Silas brûlait de fausser compagnie à
Il piocha un penny en argent dans une de ses poches et le
— Merci, Gringe.
Bonne nuit.
Gringe regarda le penny comme s'il allait le mordre
— Marcia Overstrand, elle, m'a donné une dem
Mais cette femme-là, elle a de la classe

Pre-processing

◈ Septimus corpus aligned manually.

◈ Larger literary corpus aligned automatically, paragraph by paragraph, with *Logiterm* (Terminotix 2018).

◈ 16K vocabulary with SentencePiece unigram encoding (Kudo 2018).

III. Adapting a generic system

```
# Training parameters:
batch_type: "tokens"
batch_size: 4096
valid_batch_size: 16

bucket_size: 32768
train_from: ./out/septimus/models/septimus-train_st
reset_optim: all

# Optimization
model_dtype: "fp32"
optim: "adam"
learning_rate: 2
warmup_steps: 8000
decay_method: "noam"
average_decay: 0.0005
adam_beta2: 0.998
max_grad_norm: 0
label_smoothing: 0.1
param_init: 0
param_init_glorot: true
normalization: "tokens"

# Model
encoder_type: transformer
decoder_type: transformer
enc_layers: 6
dec_layers: 6
heads: 8
rnn_size: 512
word_vec_size: 512
transformer_ff: 2048
dropout_steps: [0]
dropout: [0.1]
attention_dropout: [0.1]
position_encoding: true
```

Training & tuning

◈ OpenNMT-py (Klein et al. 2017)

◈ Transformer architecture (Vaswani et al. 2017)

◈ Parameters of the *base* model

◈ 200 000 steps (*generic*) ⇨ 50 000 (*sep-only*)

⇨ 50 000 (*sep-large*)

NMT

```
# Translation
#############

for checkpoint in ./out/${OUT}/models/*.pt; do

done

for FILE in ./out/${OUT}/translations/*_bpe.txt ;
    filename=$(basename $FILE _bpe.txt)
    spm_decode \
        --model=./data/${DIR}/subword/unigram_16_fr.model \
        --input_format=piece \
        < ./out/${OUT}/translations/${filename%.*}_bpe.txt \
        > ./out/${OUT}/translations/${filename%.*}_tok.txt
done

for FILE in ./out/${OUT}/translations/*_tok.txt ; do
    filename=$(basename $FILE _tok.txt)
    perl detokenizer.perl \
        -l fr \
        -lc \
        < ./out/${OUT}/translation
        > ./out/${OUT}/transla
done

rm ./out/${OUT}/translations/*{tok,bpe}.txt

# Evaluation
############

sacrebleu ./data/${DIR}/tra.fr \
    --input ./out/${OUT}/translations/*.txt \
    --language-pair en-fr \
    --metrics bleu chrf ter \
    --chrf-word-order 2 \
    --tokenize 13a \
    --width 2 \
```

Evaluation

◈ 704 segments (3 chapters from the 6th volume).

◈ Three metrics provided by sacreBLEU (Post 2018)

- BLEU   (formal similarity w/ the ref., compared by n-grams);

- chrF2++   (same, comparison by characters, words and bigrams);

- TER   (# of modifications necessary to produce the ref.).

◈ And COMET (Rei et al. 2020)

- (comparison of embeddings to measure semantic similarity).

III. Adapting a generic system

Evaluation metrics

System	BLEU ⇧	chrF2++ ⇧	TER ⇩	COMET ⇧
Google	10.79	35.20	91.08	-0.240
DeepL	10.04	34.88	92.81	-0.248
Generic	09.93	33.14	92.24	-0.388
Sep-only	18.56	40.43	76.06	-0.126
Sep-large	**19.08**	**41.44**	**75.98**	**-0.066**

◈ Online systems tested on 25/11/2020

◈ BLEU   #:1|c:mixed|e:no|tok:13a|s:exp|v:2.0.0
     chrF2++   #:1|c:mixed|e:yes|nc:6|nw:2|s:no|v:2.0.0
     TER   #:1|c:lc|t:tercom|nr:no|pn:yes|a:no|v:2.0.0

◈ wmt20-comet-da model

◈ Comparison with other adaptation procedures (Matusov 2019, Kuzman et al. 2019):

◈ The difference between generic and adapted is less marked, but generally confirmed.

◈ However, online tools scores are generally well above 10 and above adapted models in terms of BLEU.

◈ The findings are consistent with Kuzman et al. (2019), who notice that data from a specific author-translator couple is better than a lot of literary data.

Further analyses

◈ Other evaluations and error annotation.

◈ More details on how the system has adapted.

◈ Presentation during the main conference on 05/07.

IV. Suggestion of a new paradigm for LMT

◈ System adapted to the work of an individual translator (as alluded to already by Besacier 2014 and Toral & Way 2015).

◈ Although still far from human production, much better performance of the adapted MT (as noted by also by Kuzman et al. 2019), and much closer to the human reference.

◈ This approach thus actively rejects technological determinism, as defined by Ruffo (2018), "whereby technology acts as a subject in shaping society and culture. On the contrary, humans regain their active role of agents in determining, accepting, rejecting and interpreting technological artefacts."

◈ What is more, having a system that is tailored to individual human productions could play an important role in the emotional response and therefore acceptance of such a tool (cf. Koskinen & Ruokonen 2017, Daems 2021).

◈ Human-adapted systems as a way to mitigate some of the threats posed by LMT:

- noise and muffling of translators' voice (Kenny & Winters 2020);

- creativity (Guerberof-Arenas & Toral 2020),

- cognitive load/friction and work conditions (O'Brien 2012);

- plagiarism (Şahin & Gürses 2019);

- authorship, intellectual property, translators' rights, (Larsonneur 2020);

- temptation from mercenary editors to hire non-professionals, to drastically reduce remuneration and deadlines, to sell unrevised translations… (Taivalkoski-Shilov 2019)

  ✍ All of which would ultimately affect quality, reader experience, recognition of the work…

◈ Translators are not against translation technologies *per se*, but rather against the tools that do not account for the specific challenges of their work and the "human aspects" of it (Ruffo 2018, Koskinen & Ruokonen 2017).

IV. Suggestion of a new paradigm for LMT

◈ Lacour (2019):

| Human-Assisted Machine Translation | vs | Computer-Assisted Human Translation |
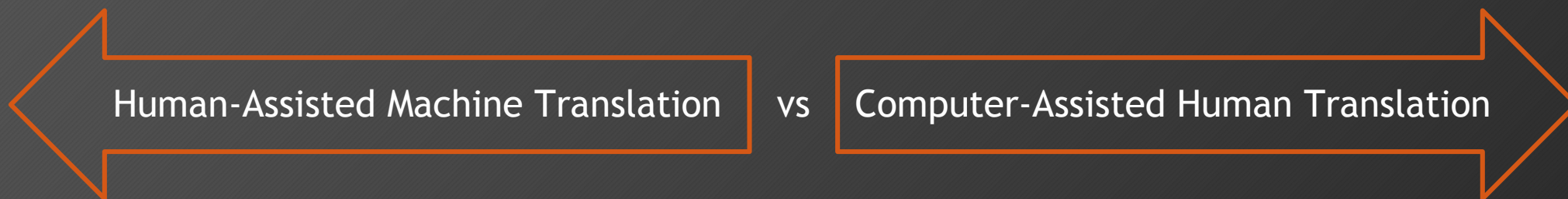
◈ What this means for LMT:

Raw PE	MT included in the larger picture of CALT
Standard PE interface	Dedicated work environment
Focus on productivity and profit	Focus on analysis, reflexion and creativity

A different scenario

◈ Interactive system offering sug-gestions for the current segment;

◈ in a tailored interface (e.g. CAT, although we could do better);

◈ combined with other corpus tools:

– corpus search

– translation memories

– machine translation

– termbases

– edit pane free of pre-translated text
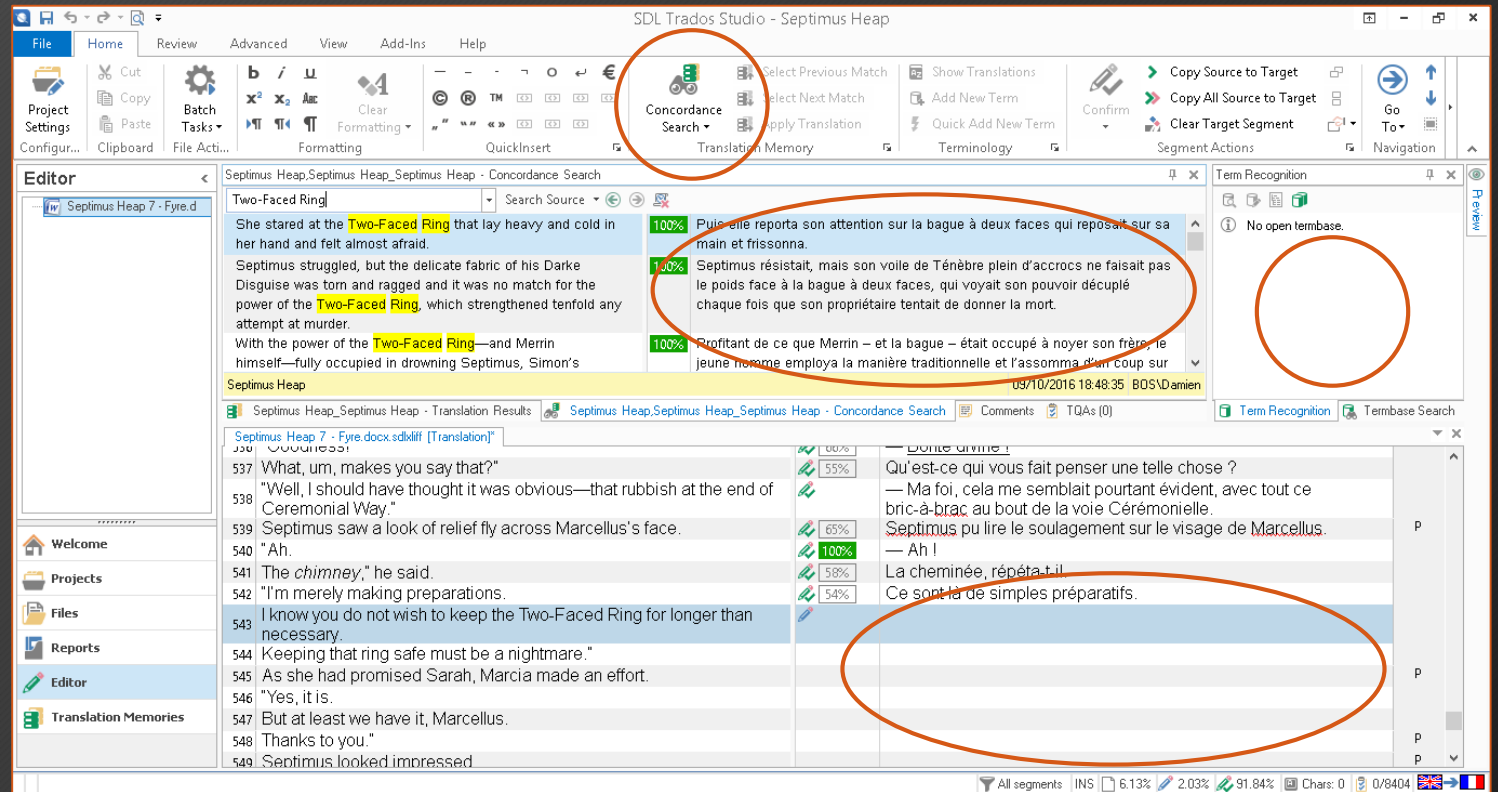
– …

A different scenario

◈ Ideally, one could choose to merge segments and the MT would offer a new translation, allowing one to work with paragraphs if desired (our system can partially handle this although not specifically designed to).

◈ Translators would be able to train their own individual systems through translation memories (this is the idea behind ModernMT, AdaptiveMT or Lilt, although I do not know how efficient they are).

◈ This use is also closer the practice of the very few that already use MT (Slessor 2020), and could be expected to rise with more useful and personalized suggestions.

◈ There is an actual demand from translators, but mostly for tools that support their needs (i.e. not productivity), while cost and lack of training and awareness are other strong factors against the use of technologies (*Ibid.*, Daems 2021).

V. Conclusions

◈ Will MT replace humans in the literary field?

  ↳ Not by a long shot. Machines do not "think" or "create" anything. But they learn well, and training them on relevant data allow them to make more useful suggestions.

◈ So, could it be of help to literary translators?

  ↳ More so if they are trained on individual productions and implemented in a way that does not constrain the translation process. Then, professionals would be able to focus on what matters (creative segments, genre-specific strategies, personal style...).

◈ Are "fiction genres" the trojan horse of literary machine translation?

  ↳ Not necessarily for fantasy it seems (due to its how far it is from standard language), but it further illustrates the need to fine-tune on more than just literary data.

V. Conclusions

- ◈ We now know that literary translators are looking for and benefit from corpus exploration tools that already exist in CAT tools (although they do not always realize that).

- ◈ Individualized MT takes this one step further by creating *ad-hoc* suggestions, based on the same corpus of personal translations.

- ◈ The main problem, at the moment, is the still limited quality of MT and the very unintuitive as well as the constraining aspect of PE interfaces and the task itself, whereas translators require tools that are "as invisible as possible" (Daems 2021).

- ◈ But this avenue is all the more relevant from an ethical standpoint as some translation agencies are already introducing literary PE (cf. Macken et al. 2022), and as websites sharing NMT-translated novels are popping up on the Web.

Bibliography

◈ Besacier, Laurent. "Traduction automatisée d'une oeuvre littéraire : une étude pilote". *Actes de la 21e conférence sur le Traitement Automatique des Langues Naturelles*, edited by Philippe Blache et al., ATALA, 2014, pp. 389-394, https://aclanthology.org/F14-2001/.

◈ Caswell, Isaac, et al. "Tagged Back-Translation". *Proceedings of the Fourth Conference on Machine Translation*, vol. 1, ACL, 2019, pp. 53-63, doi: 10.18653/v1/W19-5206.

◈ Daems, Joke. "Wat denken literaire vertalers echt over technologie?" *Webfilter*, 2021, https://biblio.ugent.be/publication/8701185.

◈ Guerberof-Arenas, Ana, and Antonio Toral. "The Impact of Post-Editing and Machine Translation on Creativity and Reading Experience". *Translation Spaces*, vol. 9, no. 2, 2020, pp. 255-282, doi: 10.1075/ts.20035.gue.

◈ Hansen, Damien, and Pierre-Yves Houlmont. "A Snapshot into the Possibility of Video Game Machine Translation". Forthcoming.

◈ Hansen, Damien, et al. "La traduction littéraire automatique : Adapter la machine à la traduction humaine individualisée". *Journal of Data Mining and Digital Humanities*, special issue *Vers une robotique du traduire*, 2022, HAL: 03583562.

◈ Kenny, Dorothy, and Marion Winters. "Machine translation, ethics and the literary translator's voice". *Translation Spaces*, vol. 9, no. 1, 2020, pp. 123-149, doi: 10.1075/ts.00024.ken.

◈ Klein, Guillaume, et al. "OpenNMT: Open-Source Toolkit for Neural Machine Translation". *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, edited by Mohit Bansal and Heng Ji, ACL, 2017, pp. 67-72, https://aclanthology.org/P17-4012.

Bibliography

◈ Koskinen, Kaisa, and Minna Ruokonen. "Love letters or hate mail? Translators' technology acceptance in the light of their emotional narratives". *Human Issues in Translation Technology*, edited by Dorothy Kenny, Routledge, pp. 7-25, https://usuaris.tinet.cat/apym/TT/koskinen.pdf.

◈ Kudo, Taku. "Subword Regularization: Improving Neural Network Translation Models with Multiple Subword Candidates". *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, edited by Iryna Gurevych and Yusuke Miyao, vol. 1, ACL, 2018, pp. 66-75, doi: 10.18653/v1/P18-1007.

◈ Kuzman, Taja, et al. "Neural Machine Translation of Literary Texts from English to Slovene". *Proceedings of the Qualities of Literary Machine Translation*, edited by James Hadley et al., EAMT, 2019, pp. 1-9, https://www.aclweb.org/anthology/W19-7301.pdf.

◈ Macken, Lieve, et al. "Literary Translation as a Three-Stage Process: Machine Translation, Post-Editing and Revision". *Proceedings of the 23rd Annual Conference of the European Association for Machine Translation*, EAMT, 2022, pp. 101-110, https://aclanthology.org/2022.eamt-1.13.

◈ Matusov, Evgeny. "The Challenges of Using Neural Machine Translation for Literature". *Proceedings of the Qualities of Literary Machine Translation*, edited by James Hadley et al., EAMT, 2019, pp. 10-19, https://www.aclweb.org/anthology/W19-7302.pdf.

◈ Larsonneur, Claire. "Neural Machine Translation: From Commodity to Commons?" When Translation Goes Digital Case Studies and Critical Reflections, edited by Renée Desjardins et al., Palgrave Macmillan, 2021, pp. 257-280, doi: 10.1007/978-3-030-51761-8_11.

◈ Lacour, Phillipe. "TraduXio : un environnement numérique pour la traduction collaborative et multilingue de textes culturels." *Ateliers Digit-Hum: Les humanités numériques en langue*, 17 oct. 2019, Paris. https://hdl.handle.net/10670/1.kg77a8.

Bibliography

◈ O'Brien, Sharon. "Translation as Human–Computer Interaction". *Translation Spaces*, vol. 1, no. 1, 2012, pp. 101–122, doi: 10.1075/ts.1.05obr.

◈ Post, Matt. "A Call for Clarity in Reporting BLEU Scores". *Proceedings of the Third Conference on Machine Translation: Research Papers*, edited by Ondřej Bojar et al., ACL, 2018, pp. 186–191, doi: 10.18653/v1/W18-6319.

◈ Rei, Ricardo, et al. "COMET: A Neural Framework for MT Evaluation". *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, edited by Bonnie Webber et al., ACL, 2020, pp. 2685–2702, doi: 10.18653/v1/2020.emnlp-main.213.

◈ Ruffo, Paola. "Human-Computer Interaction in Translation: Literary Translators on Technology and Their Roles". *Proceedings of the 40th Conference Translating and the Computer*, edited by David Chambers et al., Éditions Tradulex, 2018, pp. 127–131, https://www.asling.org/tc40/wp-content/uploads/TC40-Proceedings.pdf.

◈ Şahin, Mehmet, and Sabri Gürses. "Would MT Kill Creativity in Literary Retranslation?" *Proceedings of the Qualities of Literary Machine Translation*, edited by James Hadley et al., EAMT, 2019, pp. 26–34, https://aclanthology.org/W19-7304/.

◈ Sage, Angie. *Septimus Heap*. HarperCollins, 2005–2013. 7 vols.

◈ Serval, Nathalie. *Magyk*. By Angie Sage, Albin Michel, 2005–2013. 6 vols.

◈ Slessor, Stephen. "Tenacious technophobes or nascent technophiles? A survey of the technological practices and needs of literary translators". *Perspectives*, vol. 28, no. 2, pp. 238-252, 2020, doi: 10.1080/0907676X.2019.1645189.

Bibliography

◈ Taivalkoski-Shilov, Kristiina. "Ethical Issues Regarding Machine(-Assisted) Translation of Literary Texts". *Perspectives: Studies in Translation Theory and Practice*, vol. 27, no. 5, 2019, pp. 689–703, doi: 10.1080/0907676X.2018.1520907.

◈ Terminotix Inc. *Logiterm*. V. 5.8.0.18042, 2018.

◈ Tezcan, Arda, et al. "When a 'Sport' Is a Person and Other Issues for NMT of Novels". *Proceedings of the Qualities of Literary Machine Translation*, edited by James Hadley et al., EAMT, 2019, pp. 40–49, https://www.aclweb.org/anthology/W19-7306.pdf.

◈ Tiedemann, Jörg. "Parallel Data, Tools and Interfaces in OPUS". *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, edited by Nicoletta Calzolari et al., ELRA, 2012, pp. 2214-2218, http://www.lrec-conf.org/proceedings/lrec2012/pdf/463_Paper.pdf.

◈ Toral, Antonio, and Andy Way. "Translating Literary Text between Related Languages using SMT". *Proceedings of the Fourth Workshop on Computational Linguistics for Literature*, edited by Anna Feldman et al., ACL, 2015, pp. 123–132, doi: 10.3115/v1/W15-0714.

◈ Toral, Antonio, and Andy Way. "What Level of Quality can Neural Machine Translation Attain on Literary Text?" *Translation Quality Assessment: From Principles to Practice*, edited by Joss Moorkens et al., Springer, 2018, pp. 263-287, https://arxiv.org/pdf/1801.04962.pdf.

◈ Vaswani, Ashish, et al. "Attention is All you Need". *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*, edited by Ulrike von Luxburg et al., Curran Associates Inc., 2017, pp. 6000-6010, https://papers.nips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html.

Don't hesitate to get in touch:
damien.hansen@uliege.be
(LiteraryLudeme)

Bibliography and slides can easily
be found in my repository
(ORBi)

Thank you for the attention