



Speech-enabled Post-editing of Machine Translation

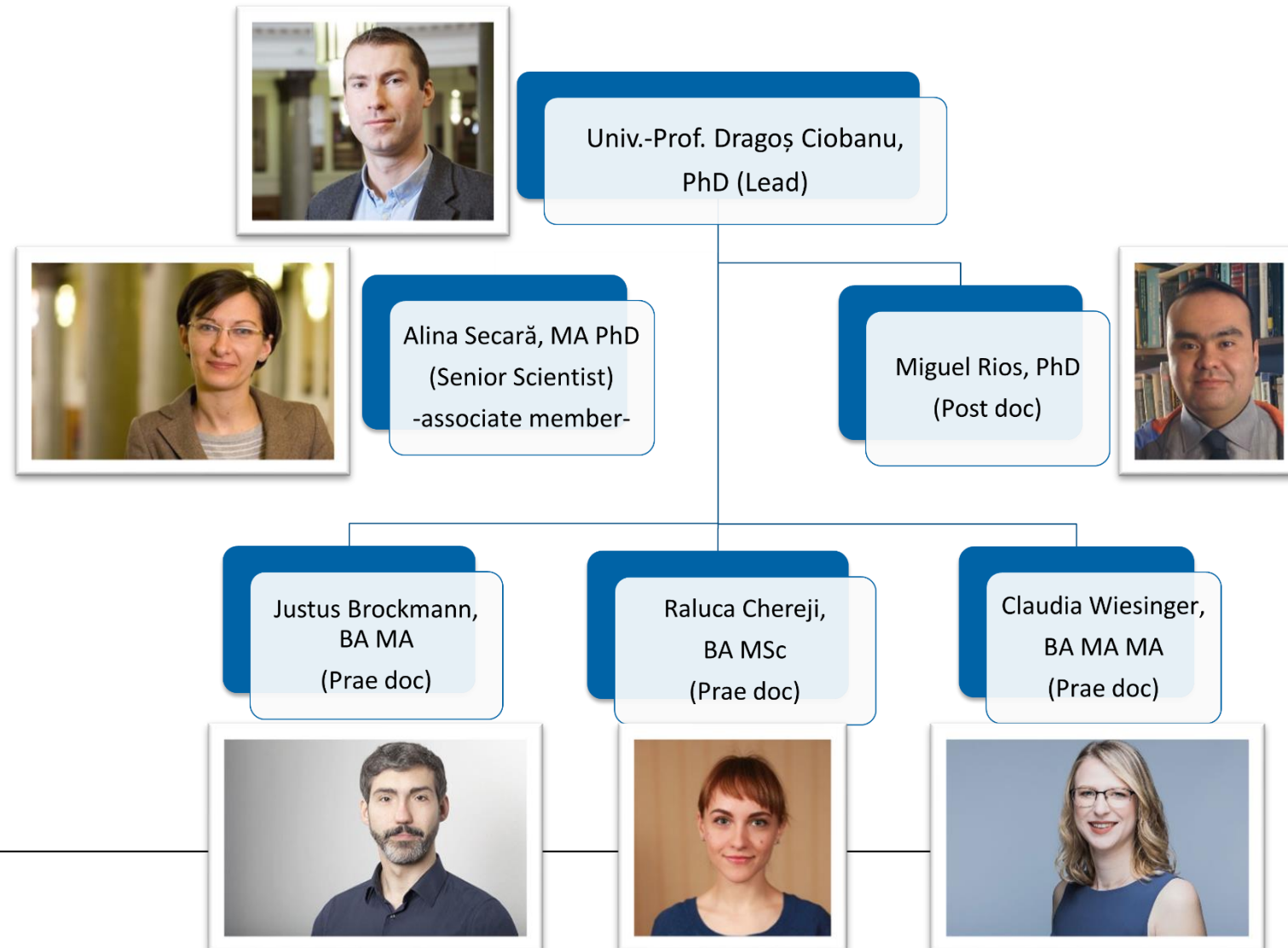
Univ.-Prof. Dragoş Ciobanu



HAITrans research group

Human and Artificial Intelligence
in Translation

- <https://haitrans.univie.ac.at/team/>





[Home](#) > [Issue08](#) > [Mossop article](#)



Empirical studies of revision: what we know and need to know

Brian Mossop, York University School of Translation & Government of Canada Translation Bureau

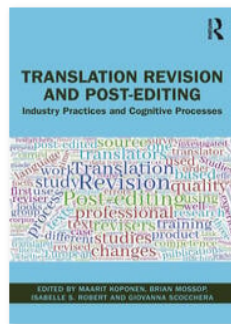
ABSTRACT

Translators and quality controllers generally acquire knowledge of how to revise their own or others' work by trial-and-error, by working under an experienced reviser, or by attending workshops. There are also one or two publications and in-house manuals that purvey advice for successful revising. Recently, however, Translation Studies scholars have begun to conduct empirical studies in which they observe the revision process through methods such as recording and playing back keystrokes, asking translators to think aloud into a microphone as they revise their own work, or comparing different revised versions of a given draft translation. This article reviews a selection of studies of revision in English, and concludes with some suggestions about questions that need attention.

KEYWORDS

revision, empirical studies, quality control

Almost all talk and writing about revision tells us what *supposedly* happens ("all our translations are re-read by a second translator"), or how revisers *ought* to go about their jobs ("make no unnecessary changes"), or what techniques they *could* use ("try reading it aloud"). But what do revisers and self-revisers *actually* do? In human affairs, what we imagine people do, or what we think people ought to do, or what people claim that they do, may bear little relationship to what they really do.



Chapter

Post-Editing Human Translations and Revising Machine Translations

Impact on efficiency and quality

*By Joke Daems, Lieve Macken*Book [Translation Revision and Post-editing](#)

Edition	1st Edition
First Published	2020
Imprint	Routledge
Pages	21
eBook ISBN	9781003096962



Share

ABSTRACT

Post-editing of machine-translation output is generally considered to be a distinct process from the revision of a human-translated text. The main reasons for this assumption are the quality of machine-translation output and the fact that it might be easier to criticize the work of a machine than the work of a fellow human translator. With the global shift of statistical machine-translation systems to neural machine-translation systems, however, the quality of machine-translation output has improved. What was true about differences between revision and post-editing in 2010 might therefore no longer be true today. In addition, translators hired to revise a text are not always aware of the origin of the text. This chapter compares revision and post-editing products made by professional translation agencies when the actual origin of a text corresponded to the instructions they were given (revision of a human translation, post-editing of machine translation) and when the origin did not match the instructions (post-editing of a human translation and revision of machine translation). We look at the number of edits made, the quality of the revision and the optimality of the intervention.

Intento – Your MT Innovation Partner

MT Evaluation

Select the best-fit option among pre-trained models and custom models trained on your available data — optimal for your language pairs and domains.

2,909

Models evaluated by Intento

125k+

Language pairs available for evaluation

Evaluate best-fit MT for your data with [Intento MT Studio](#) or [ask our experts](#) for professional help.



MT Hub

Translate better, faster, and at scale. Keep your data secure and streamline your workflows.

70%

Less post-editing

97%

MT requires no human review

[Equip your team](#) with intelligent technology to create and translate content 4x faster, in real time.



MT Maintenance

Keep your engines at the forefront of cutting-edge technology. Stay up-to-date on new models and updates.

2,988

Crucial MT provider updates detected in 2021

347k+

Glossary terms added and checked for their impact on quality

Learn how to [evolve your MT program](#) over time.



Disclaimer

July 1—July 28, 2022

The MT systems used in this report were accessed from July 1 to July 28, 2022. Some of these systems may have changed since then.

Automatic scoring

This report demonstrates the performance of those systems exclusively on the datasets used for this report ([see slide 12](#)) using semantic similarity scores. The final MT decision requires Human LQA and depends on each specific use case.

Tailored Dataset

Data for all domains were collected in English from publicly available datasets and translated by e2f into 11 languages. The selected MT providers could not have had access to such data in the past for training their models.

* as defined in "Domain Adaptation and Multi-Domain Adaptation for Neural Machine Translation: A Survey" by Danielle Saunders

Plain Text Only

The evaluation was done on plain text data. We often see different results for tagged text (like those found in CAT/TMS systems) for some MT vendors and language pairs due to imperfect inline tag support.

Valid for a Specific Dataset

Normally, we run multiple evaluations for our clients using various language pairs and domains, and observe different MT system rankings than those provided in this report.


There's no "best" MT system

MT performance depends on how similar your data is to the data used to train the vendors' models, as well as their algorithms.

Trademarks

All third-party trademarks, registered trademarks, product names, and company names or logos mentioned in the Report are the property of their respective owners, and the use of such Third-Party Trademarks inures to the benefit of each owner. The use of such Third-Party Trademarks is intended to describe the third-party goods or services and does not constitute an affiliation by Intento and its licensors with such company or an endorsement or approval by such company of Intento or its licensors or their respective products or services.

Domains? What are these?

Domain is a corpus from a specific source that may differ from other domains in topic, genre, style, level of formality et cetera*. Basically, a combination of industry sector and content type 

Is NMT really that good?

“Compared to statistical machine translation, neural machine translation may be less adequate even if more fluent. In other words, the translation may diverge from the input in various ways, such as not translating part of the sentence or generating unrelated output words.”

(Koehn, 2020, p.165)

Impact of Domain-Adapted Multilingual Neural Machine Translation in the Medical Domain @TC44

Miguel Rios, Raluca-Maria Chereji, Alina Secară, Dragoş Ciobanu

HAITrans (Human and Artificial Intelligence in Translation) research group

Centre for Translation Studies

University of Vienna

<https://haitrans.univie.ac.at/>

Error Analysis (cont.)


Terminology-related errors: **98** MBart, and **64** fine-tuned MBart.

Error type	MBart ↓	Fine-tuned MBart ↓
Partial error	41	23
Source term copied	22	9
Inflectional error	2	4
Reorder error	1	3
Disambiguation issue in target	14	6
Incorrect lexical selection	9	6
Term drop	0	0
Other error	9	13



Hallucinations

Einladung Jahresgespräch

 Berger-Kuklik Barbara
To

[Inbox/Work/v-important/annual_review]

 Translate message to: English | Translation preferences

Liebe Kollegen und Kolleginnen!

Der vorherige Link ist leider falsch und Sie bekommen im Laufe des Tages einen neuen zugeschickt.


Mit freundlichen Grüßen

Invitation Annual Appraisal

 Berger-Kuklik Barbara
To

[Inbox/Work/v-important/annual_review]

 Translated | [Show Original](#) | [Translation preferences](#)

 Reply

Mr President, ladies and gentlemen, I should like to begin by thanking the rapporteur for his excellent

The previous link is unfortunately wrong and you will receive a new one during the day.

Yours sincerely ,



Some examples from DGT-RO (with thanks to Bogdan Marin & Ileana Busuioc)

the Ministry of Energy, Natural Resources, Environment and Mines (MiAmbiente), the Forest Conservation Institute (ICF), the Institute of Property (IP), the National Agrarian Institute (INA), the Revenue Administration Service (SAR), the Customs Revenue Subdirectorates (DARA), the Association of Honduran Municipalities (AMHON), the Honduran Social Security Institute, (IHSS), the Ministry of Labour and Social Security (STSS), the Ministry of Agriculture and Livestock (SAG), the National Directorate of Indigenous and Afro Honduran Peoples (DINAFROH), the Institute for Access to Public Information (IAIP), the Public Prosecutor's Office (MP), the Security Ministry (SSN, representing the National Preventive Police (PNP)), the National Defence Ministry (SEDENA, representing the armed forces of Honduras (FFAA)), the Attorney General's Office (PGR), the Central Bank of Honduras (BCH), the National School of Forest Sciences (ESNACIFOR) and other academic institutions concerned with environmental and forestry issues.

Ministerul Energiei, Resurselor Naturale, Mediului si Minelor (MICF), Institutul pentru Conservarea Naturii (ICF), Institutul National Agrare (IP), Serviciul de administrare fiscală (SAR), Subdirectia de Administrare Fiscală (AMHON), Ministerul Muncii și Securității Sociale (SAG), Ministerul Muncii și Securității Sociale (SAG), Ministerul Muncii și Securității Sociale (SAG), Ministerul Muncii și Securității Sociale (SAG), Ministerul Muncii și Securității Sociale (SAG), Ministerul Muncii și Securității Sociale (SSN), Ministerul Muncii și Securității Sociale (SSN, reprezentant al forțelor armate ale Honduras — FFAA), Biroul Procurorului General (PGR), Central Bank of Honduras (CSIB), Școala Națională de Științe Forestiere (ESNAF), Școala Națională de Științe Forestiere (ESNAF) si alte instituții academice care se ocupă de probleme de mediu și silvice.

Finally, I have asked for the branch manager's e-mail address a while back but I have not heard back – could you please send it to me as soon as possible? Many thanks!

Schließlich habe ich vor einiger Zeit um die E-Mail-Adresse des Niederlassungsleiters gebeten, aber ich habe noch keine Antwort erhalten. Herzlichen Dank!

Finally, I have asked for the branch manager's e-mail address a while back but I have not heard back; could you please send it to me as soon as possible? Many thanks!

Schließlich habe ich vor einiger Zeit um die E-Mail-Adresse des Niederlassungsleiters gebeten, aber noch keine Antwort erhalten. könnten Sie sie mir bitte so bald wie möglich zusenden? Herzlichen Dank!



EU Council Presidency translator

From:

Romanian English Bulgarian



To:

English

Provider:

Tilde

Translate

OBIECTIVUL 2: FOAMETE „ZERO”

Eradicarea foamei, asigurarea securității alimentare, îmbunătățirea nutriției și promovarea unei agriculturi durabile
Agricultura este importantă pentru întreaga societate. Dacă dorim o societate mai sănătoasă, trebuie să ajungem la o agricultură sustenabilă. România deține locul al șaselea în Europa din perspectiva dimensiunii terenurilor agricole. Pentru această țară este o prioritate eficiența agricolă și limitarea deșeurilor alimentare. Agenda 2030 țintește eradicarea foamei și securitatea alimentară prin creșterea productivității agricole și prin reducerea deșeurilor alimentare.



OBJECTIVE 2: “ZERO” HUNGER

Eradicating hunger, providing food security, improving nutrition and promoting sustainable farming
Agriculture is important for society as a whole. If we want a healthier society, we have to reach sustainable agriculture. Romania holds sixth place in Europe in terms of the size of agricultural land. Agricultural efficiency and limited food waste are a priority for this country. The 2030 Agenda aims to eradicate hunger and food security by increasing agricultural productivity and reducing food waste.



+ Upload or drag a file

Suggest translation

Machine Translation provided by Tilde
Rate this translation: ☆☆☆☆☆





In nominal terms, the VAT Gap decreased by €10.5 billion to €147.1 billion in 2016, a drop to 12.3% of total VAT revenues compared to 13.2% the year before.

În termeni nominali, deficitul de încasare a TVA a scăzut cu 10.5 miliarde EUR, ajungând la 147.1 miliarde EUR în 2016, ceea ce reprezintă o scădere a veniturilor totale din TVA cu 12,3 % față de 13,2 % în anul precedent.

Some examples from EN into DE (with thanks to my 2021W MT UE students)

Iryna Skliar

Iryna S**g**liar

James R Elliot

James R Elliott

He, H. and Ye, Z.

Er, H. and **Ya**, Z.

Roland W Force

Roland W **Macht**

Roland W **Kraft**

Roland W **Streitkräfte**;

Michael P Lesser

Michael **S.** Lesser

Tosepu, Ramadhan

Tosepu, **RAMADAN**

Fredrik Palm

Friedrich **von Schiller**

Very important!

What are the differences between **post-editing of machine translation hypotheses** and working **with a TM**?

TRUST

Project Settings Configuration | Copy Paste Batch Tasks Clipboard File Actions | Clear Formatting | Concordance Search | Select Previous Match | Show Translations | Add New Term | Quick Add New Term | Copy Source to Target | Merge Segments | Split Segments | Clear Target Segment | Change Segment Status | Navigation

Editor

Translation Results - marketing_en-de			
Project Settings			
Council & Committee Expenses			
1	Council & Committee Expenses	CM	Aufwendungen Vorstand & Ausschüsse
2	Council & Meeting Committee Expenses	75%	Vorstandssitzungen
3	Other Council & Committee expenses	72%	Sonstige Aufwendungen Ausschüsse

14	Membership	100%	Mitgliedsbeiträge	C+
15	Application Fees	100%	Aufnahmegebühren	C+
16	Administration Fees	100%	Verwaltungsgebühren	C+
17	Bank Interest	100%	Bankzinsen	C+
18	Examinations & Assessment	100%	Prüfungsgebühren	C+
19	-----	100%	-----	C+
20	-----	100%	-----	C+
21	-----	100%	-----	C+
22	Total Operating Income	100%	Summe betriebliche Erträge	C+
23	LESS OPERATING EXPENSES	100%	ABZÜGLICH BETRIEBLICHE AUFWENDUNGEN	C+
24	Advertising & PR	100%	Werbemaßnahmen	C+
25	Office Premises	100%	Bürräume	C+
26	Membership Services	100%	Leistungen für Mitglieder	C+
27	Council & Committee Expenses	100%	Aufwendungen Vorstand & Ausschüsse	C+
28	Administration			C+
29	I T Systems			C+
30	Finance			C+
31	Depreciation of Office Equipment			C+
32	-----			C+
33	-----			C+
34	-----			C+
35	Net Operating Surplus [Deficit]			C+
36	ADD OTHER INCOME			C+
37	Surplus [Defecit] on Directory			C+
38	Surplus on Publication Sales			C+

- Welcome
- Projects
- Files
- Reports
- Editor
- Translation Memories
- Qualityity

File Home Review Advanced View Add-Ins Help

Project Settings Configuration | Cut Copy Paste Batch Tasks | Clear Formatting | Concordance Search | Select Previous Match | Show Translations | Add New Term | Quick Add New Term | Copy Source to Target | Merge Segments | Split Segments | Clear Target Segment | Change Segment Status | Find | Replace | Select All

Editor

- 02_Accounts_1202eCoLoRe_en.xlsx.sdlxliff [Tra...]
- Worksheet
- Worksheet
- Worksheet
- EN_ecolore_01.docx.sdlxliff [Translation]*

Welcome
 Projects
 Files
 Reports
Editor
 Translation Memories
 Qualityity

Translation Results - marketing_en-de,OPUS-CAT

Project Settings

Council & Committee Expenses

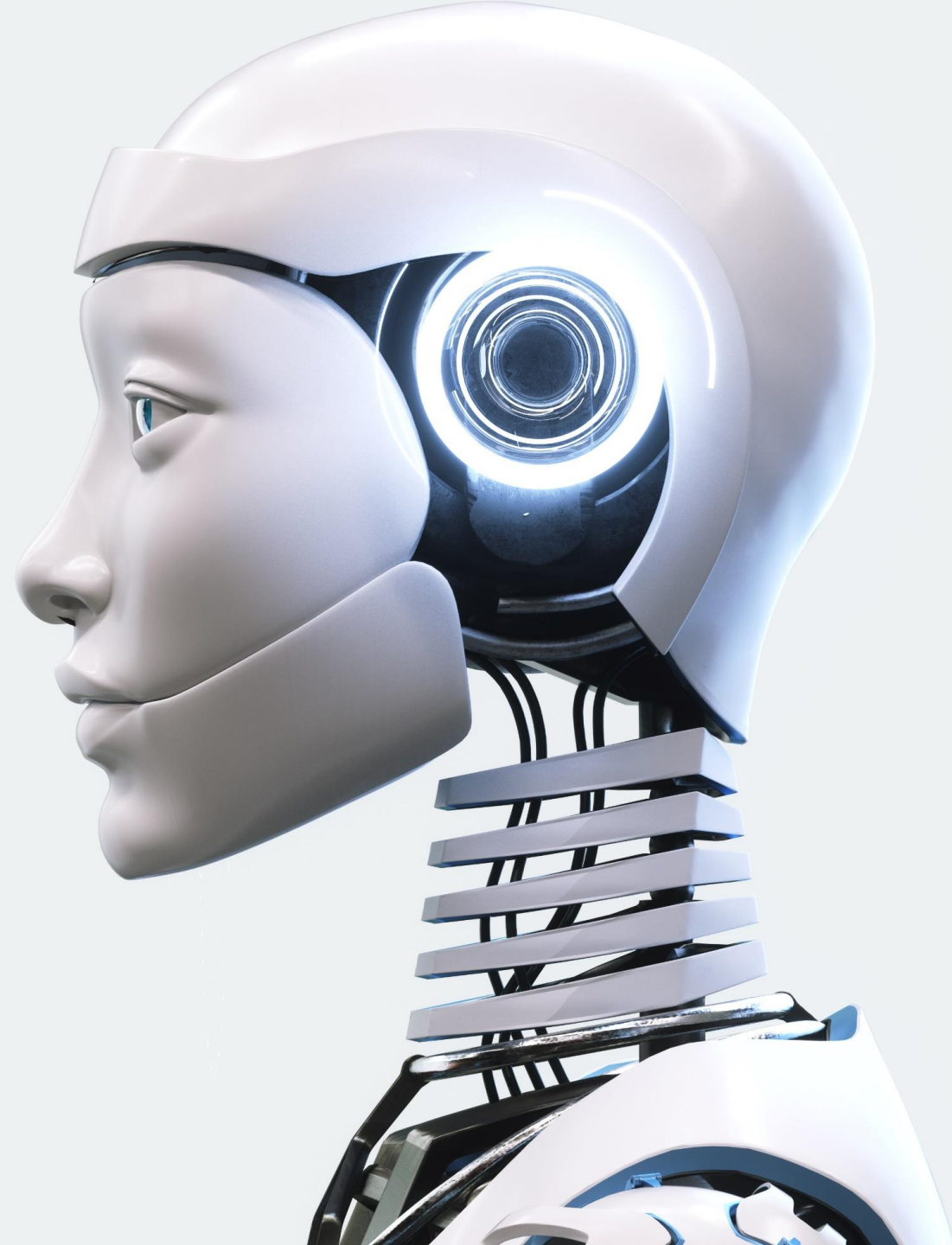
1	Council & Committee Expenses	CM	Aufwendungen Vorstand & Ausschüsse	C
2	Council & Meeting Committee Expenses	75%	Vorstandssitzungen	
3	Other Council & Committee expenses Expenses	72%	Sonstige Aufwendungen Ausschüsse	
4	Council & Committee Expenses	NMT	Auslagen des Rates und des Ausschusses	mtSystem opus-cat

marketing_en-de 20/10/2022 17:42:20 khfneu

Translation Results - marketing_en-de,OPUS-CAT
 Fragment Matches - marketing_en-de,OPUS-CAT
 Concordance Search - marketing_en-de,OPUS-CAT
 Comments TQAs (0) Messages (1)

14	Membership	100%	Mitgliedsbeiträge	C+
15	Application Fees	100%	Aufnahmegebühren	C+
16	Administration Fees	100%	Verwaltungsgebühren	C+
17	Bank Interest	100%	Bankzinsen	C+
18	Examinations & Assessment	100%	Prüfungsgebühren	C+
19	-----	100%	-----	C+
20	-----	100%	-----	C+
21	-----	100%	-----	C+
22	Total Operating Income	100%	Summe betriebliche Erträge	C+
23	LESS OPERATING EXPENSES	100%	ABZÜGLICH BETRIEBLICHE AUFWENDUNGEN	C+
24	Advertising & PR	100%	Werbemaßnahmen	C+
25	Office Premises	100%	Bürräume	C+
26	Membership Services	100%	Leistungen für Mitglieder	C+
27	Council & Committee Expenses	100%	Aufwendungen Vorstand & Ausschüsse	C+
28	Administration	100%		C+
29	I T Systems	100%		C+
30	Finance	100%		C+
31	Depreciation of Office Equipment	100%		C+
32	-----	100%		C+
33	-----			C+
34	-----			C+
35	Net Operating Surplus [Deficit]			C+
36	ADD OTHER INCOME			C+
37	Surplus [Defecit] on Directory			C+
38	Surplus on Publication Sales			C+







Search for Articles:

Title / Keyword

Author / Affiliation

Informatics

All Article Types

Search

Adv

Journals / Informatics / Volume 6 / Issue 4 / 10.3390/informatics6040051



informatics

Submit to this Journal

Review for this Journal

Edit a Special Issue

Article Menu

Article Overview

- Abstract
- Share and Cite
- Article Metrics
- Related Articles
- Order Article Reprints

Article Versions

Open Access Article

Speech Synthesis in the Translation Revision Process: Evidence from Error Analysis, Questionnaire, and Eye-Tracking

by Dragoș Ciobanu ^{1,*} Valentina Ragni ² and Alina Secară ¹

¹ Centre for Translation Studies, University of Leeds, Leeds, LS2 9JT, UK

² School of Modern Languages, University of Bristol, Bristol, BS8 1TE, UK

* Author to whom correspondence should be addressed.

Informatics **2019**, *6*(4), 51; <https://doi.org/10.3390/informatics6040051> (registering DOI)

Received: 1 July 2019 / Revised: 1 November 2019 / Accepted: 5 November 2019 / Published: 11 November 2019

(This article belongs to the Special Issue [Advances in Computer-Aided Translation Technology](#))

View Full-Text

Download PDF

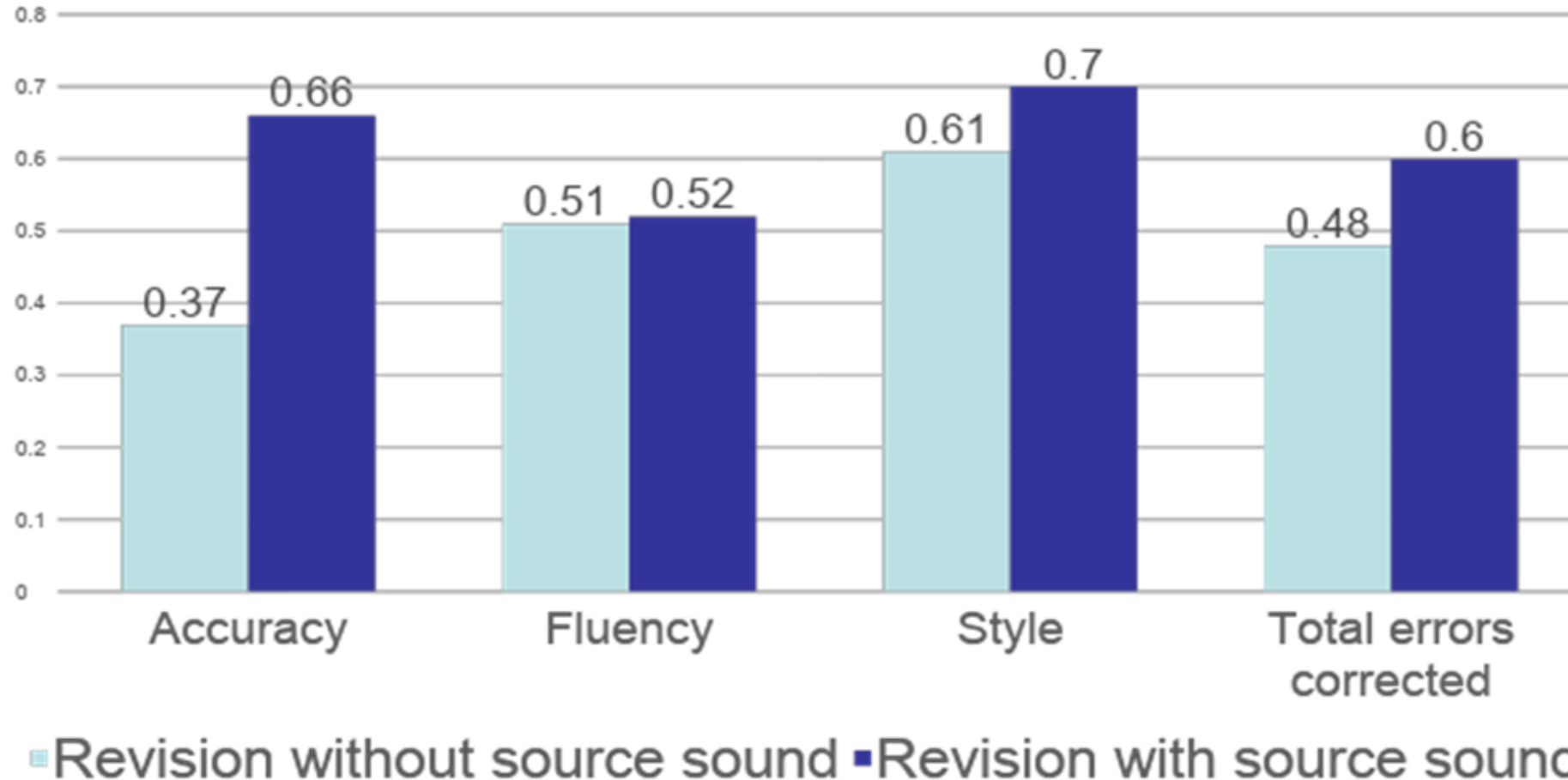
Browse Figures

Review Reports



Why?

Attention while translating / revising / PEMT



Ciobanu, D., Ragni, V., & Secară, A. (2019). Speech Synthesis in the Translation Revision Process: Evidence from Error Analysis, Questionnaire, and Eye-Tracking. *Informatics*, 6(4)(51), Article 4. <https://doi.org/10.3390/informatics6040051>



imminent Vienna's
Research Center

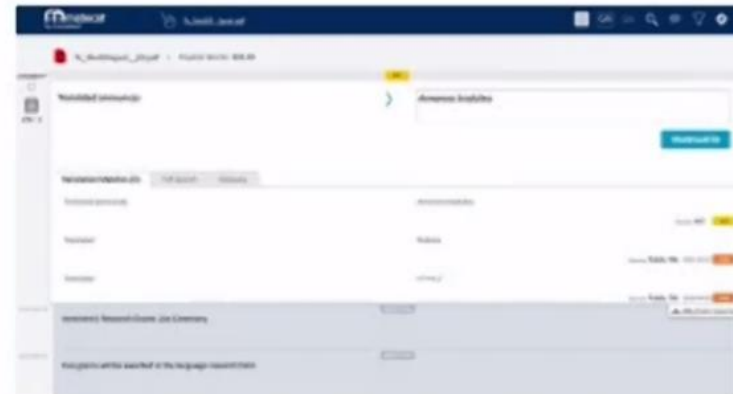
Project 1 - Human Computer Interaction

Dragoș Ciobanu

Professor of Computational Terminology and Machine Translation
University of Vienna, Austria

The impact of Speech Synthesis on cognitive load, productivity
quality during post-editing machine translation (PEMT)

and



Imminent



YouTube AT

Search

Volume Analysis - MateCat x Translate - GALA_Demo - x Translate - GALA_23_Demo_Fi x Translate - GALA_23_Demo_E x Translate - GALA_23_Demo_Ri x +

staging.matecat.kromin.it/translate/gala-demo/en-GB-de-DE/211-6742e14ce9e0#3487

matecat by translated

1 GALA_Demo

CEA_COVID-19_Part1.docx • Payable Words: 299.75

3487

MT

COVID-19 Stigma & Discrimination: Should we be scared of people who have had COVID-19?

COVID-19 Stigma & Diskriminierung: Sollten wir Angst vor Menschen haben, die COVID-19 hatten?

TRANSLATED

Translation Matches (1) TM Search Glossary

COVID-19 Stigma & Discrimination: Should we be scared of people who have had COVID-19? COVID-19 Stigma & Diskriminierung: Sollten wir Angst vor Menschen haben, die COVID-19 hatten?

Source: MT German MT

Add private resources

3488 This factsheet is designed to help those working with communities to provide the facts to those who express fear and stigma of people who have recovered from COVID-19.

3489 Can you recover from COVID-19?

3490 Yes, people can fully recover from COVID-19.

3491 COVID-19 positive status IS NOT permanent or a death sentence.

1:08 / 1:38

Job ID: 211 English → German -% Weighted words: 300 To-do: 300

Imminent Matecat demo for GALA 2023

Unlisted



Dragoș Ciobanu
58 subscribers

Analytics

Edit video

Like

Dislike

Share

Download

Save

All

Listenable

Recently uploaded

Watched

Talks



The Growth Mindset | Carol Dweck | Talks at Google



THE LANGUAGE OF BUSINESS
THE BUSINESS OF LANGUAGE
12-15 March 2023

Speech-enabled Machine Translation Post-editing (MTPE): Effective or Distracting?

**Claudia Wiesinger, Justus Brockmann, Raluca Chereji,
Alina Secară, Miguel Rios, Dragoş Ciobanu (Principal Investigator)**

Human and Artificial Intelligence in Translation (HAITrans)
University of Vienna

#GALA2023

Impact on cognitive load

- Increase in MFD between silence and sound conditions

= **NOT statistically significant** (neither in source text nor in target text)

Mean Fixation Duration (MFD)		
	Source text	Target text
Silence	287 ms	318 ms
Sound	384 ms	393 ms

Impact on productivity

Methods

- Time tracking
- Questionnaires

Key performance indicators

- Post-edited words per hour
- Self-perception

Experiment duration per text part (words/h)				
	Text 1	Text 2	Text 3	Text 4
P1	724	691	585	678
P2	844	789	674	787
P3	933	946	827	791
P4	1226	1617	1289	1121
P5	704	671	836	678
P6	1089	1391	1047	1021

Average productivity (words/h)		
P1	654	684
P2	817	730
P3	880	869
P4	1421	1205
P5	770	674
P6	1240	1034

Silence Sound

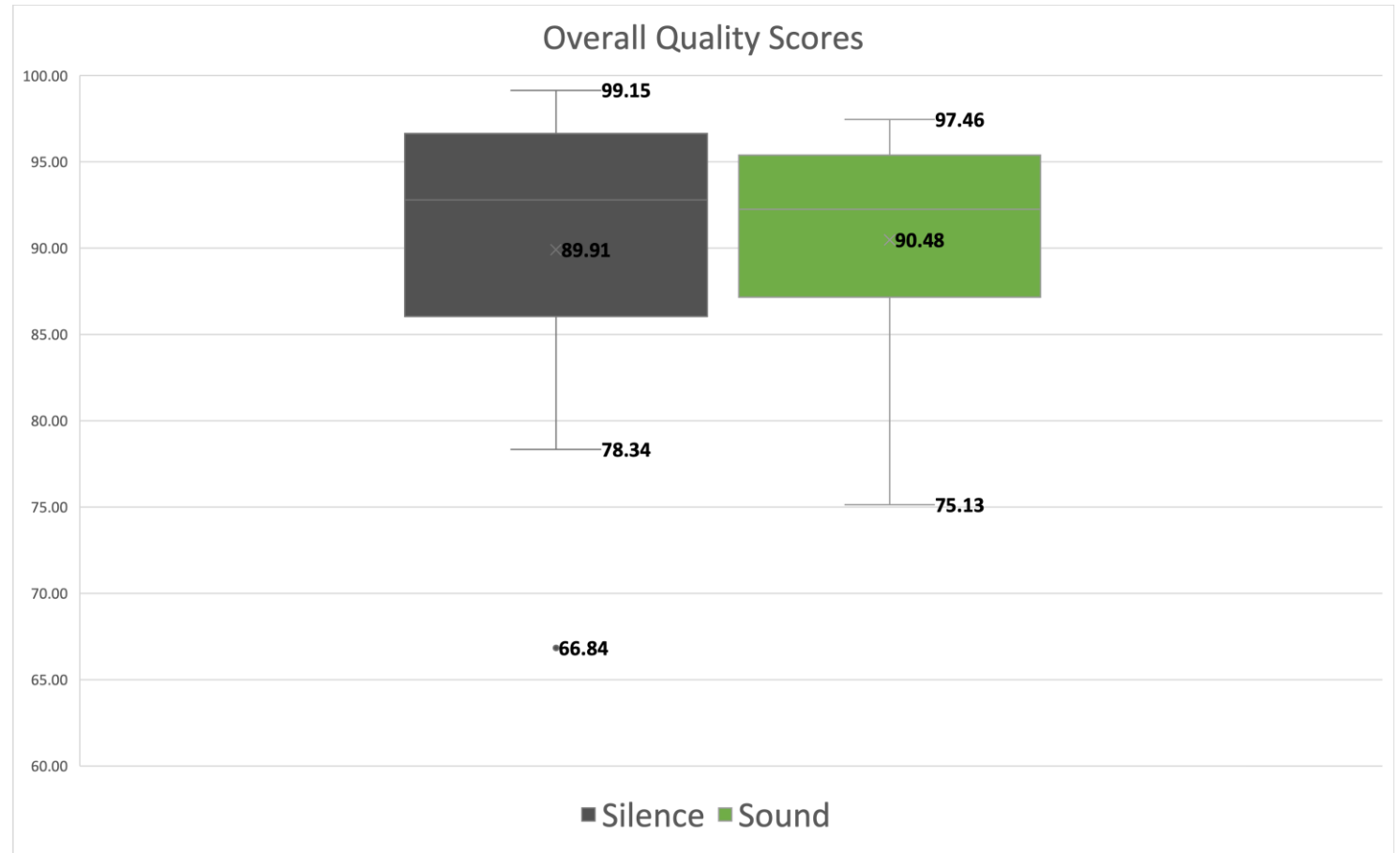
Impact on quality

Methods

- Error correction analysis
(Multidimensional Quality Metrics)

Key performance indicators

- Overall Quality Score
- Error correction rate



Impact on quality



Final thoughts

It was good to get **other sensory input** when I started to lose my concentration.

P1

I find it **rather irritating** and I feel it makes me **work more slowly**.

P6

Short and simple: **it helped me**.

P5

I think it could make post-editing **more dynamic**.

P4

I **was not very keen** on using it, because I expected it to be too unnatural and thus disturbing, but [...] I **felt better** when using it.

P3

Yes, I **would use it** [...] in cases I am not sure my translation is well understandable.

P2

Conclusions

- Potential of **speech technologies** for translation, revision, PEMT
 - Supported by research results
 - Especially relevant given the growth of NMT deployment

- Investigation of **contexts in which these tools could be integrated** in translators' workflows (doctoral projects)

Thank you!

HAITrans - Human and Artificial Intelligence in Translation

HAITrans - Human and Artificial Intelligence in Translation - is a research group based in the University of Vienna Centre for Translation Studies. It investigates the behavioural and cognitive effects which technologies such as machine translation and automatic speech recognition and synthesis have on translators, as well as their impact on the profession, practice, training and society at large.

At present, the core research areas of the Vienna HAITrans Group are:

- 1 Effects of speech technologies (Speech to Text and Text to Speech) on translation, revision and post-editing machine translation (PEMT) tasks (R1)
- 2 Technology-supported translation, revision, and PEMT practices (R2)
- 3 (Translation) technology for accessibility (R3)
- 4 Translation technology didactics (R4)

In our qualitative and quantitative investigations we use data gathered via eye-tracking, questionnaires, focus groups, corpora, and translation environment tool metrics. We also collaborate with academic partners, international organisations, language service providers, dedicated professional associations and cultural-sector partners.

Contact

Centre for Translation Studies (ZTW)
Kolingasse 14-16
1090 Vienna
T: +43-1-4277-58080
haitrans@univie.ac.at



Dragoş Ciobanu

dragos.ioan.ciobanu@univie.ac.at

HAITrans research group

<https://haitrans.univie.ac.at/>