**Acceptance speech for the award of the academic title**

**DOCTOR HONORIS CAUSA, Timișoara, November 15, 2018**

**The State of Translation Studies in the World: Trends, Challenges, Perspectives**

**Prof. em. Dr. habil. Gyde Hansen**

**Introduction**

*Der technischen Universität von Timișoara danke ich ganz herzlich für die riesige Freude, die Sie mir durch die Verleihung der Ehrendoktorwürde bereiten.*

*Îi mulțumesc călduros Universității Tehnice din Timișoara pentru bucuria extraordinară pe care mi-ai dat-o, primind un doctorat onorific.*

1. **Trends**

The digital revolution with Artificial Intelligence is quickly changing our lives – also our field Translation Studies (TS).

Talking about the status of TS and recent trends implies talking about Machine Translation (MT). Today I will spare you the history of MT-systems. You can read it on the internet. What may be interesting is how we under the rapid changes can improve the quality of translations and adapt the translator training to the continually new challenges.

At the XXth FIT-World Congress in 2014 in Berlin, the theme was: “Man *vs.* Machine”? FIT is the International Federation of Translators for professional translators and associations as well as training institutes. They meet every three years.

Many of the participants, especially professional translators, asked about MT and expressed their concern that they could lose their jobs or that they would have to post-edit machine translations instead of translating themselves – perhaps for the rest of their working life. They also feared to earn less money for post-editing. In 2014, I was not sure what to answer “Que será, será …?”

Though there were lots of warnings against MT, I was interested in MT and my attitude was “*man* *and* machine” and not “*man versus machine*”. From 2013, my master students trained post-editing texts translated with *Statistical Machine Translation* (SMT). The linear SMT-systems are phrase- and example based using millions and millions of parallel texts and permanently learning from them. We used Google Statistical Machine Translation (GSMT) and Pons Machine Translation. Both systems could be downloaded for free. According to Schmitt (2014: 81), Google Translate was the most used machine translation software at that time.

We did many experiments and trained preparing texts for MT, for example, comparing our three languages and anticipating probable errors. The following example is from a brochure for Scandlines, a ferry between Denmark and Germany. It is an authentic text which shows the differences in pre- and post-modifying.



A few years later, at a summer academy in Innsbruck in 2016, I met colleagues and students from *all over the world* – look at the title of this speech - they were from both Europe, Russia, Ukraine and China. They asked the same questions as the participants two years earlier at the FIT congress. I showed them that GSMT already had improved – especially translation between English and French – but how should I be able to predict what would happen to translation and TS.

GSMT and MT with Pons still had a lot of problems a. o. with pragmatics, (e.g. deixis) with word-order (especially translating into German), long sentences, quotation marks, lack of words (e.g. prepositions and rare words), inflections (verbs, nouns, adjectives), reference (articles), the English -ing form, the German post-modifying genitive and compound words.

End of 2016, the capacity of computers had grown considerably, and via an artificial neural network capable of deep learning, self-learning of the devices had become possible. For translation, *Neural Machine Translation* (NMT) was launched in November 2016 with a few language pairs. Now NMT translates between more than 100 language pairs. Many platforms like Google Translate, Microsoft Translator and Systran apply and develop NMT – lately also Facebook and Amazon work on NMT.

There also appear some European translation services, Deepl Translator (Linguée.com) and Translate.2018.eu. Deeple is said to be more powerful than Google. Both systems still only work with six or seven languages. Applying Translate.2018.eu on some of my texts (22.10.2018) I realized that Google still provides better results.

Compared to the situation at the FIT congress in 2014, the situation has changed considerably. In 2009 and 2012, Franz Och, one oft the founders of phrase based Google Translate, had said in an interview with the German Bildzeitung: ”Ich werde erst zufrieden sein, wenn die Übersetzungen so gut sind wie professionelle menschliche Übersetzungen.” (My translation: I will not stop before the (machine) translations are as good as professional, human translations). With NMT, we get closer to this goal as it is improving in high speed combining knowledge and results from a. o.:

* + Inclusion of information from the context
	+ Word embeddings and word recognition
	+ Syntax-based neural translation models
	+ Linguistic annotation
	+ Traditional phrase-based SMT
	+ Training of the system and increasing the amounts of training data
	+ Several tool kits
	+ Quality checks via validation sets and back translation
	+ …
	1. **From Google Statistical Machine Translation (GSMT) to Google Neural Machine Translation (GNMT)**

Examples show the improvement. In 2014, we translated a text from Wikipedia – the description of “Full post-editing”. Here is the source text (ST):

“The expectation is that the outcome will be a text that is not only understandable but presented in some stylistically appropriate way, so it can be used for assimilation and even for dissemination, for inbound and for outbound purposes.”

Below are the translations into German of 2014 with GSMT and 2018 with GNMT.



Another example from the advertisement of a restaurant on Oahu/Hawaii:



Recently TS-students and university professors get more and more concerned: “Is it still worth to train translation? Will I ever get a job?” Will it be possible to earn a living with translation?” Professors of TS ask: “How long will I be allowed to train translators? Perhaps 15 years from now?”

1. **Challenges**

The most important challenge – perhaps even a threat – is that the whole translation sector will change rather quickly, and that professional translators and TS trainers – all of us – may not get much time to adapt to the new, rapidly changing situations. Translation management services may take over carrying out personality tests and push and stress translators.

Some questions appear:

* What kinds of texts will be left for human translation?
* Will it only be literature, creative websites, poetry, domain scientific texts, e.g. legal and technical texts?
* How much quality is needed? That may depend on the texts and their purpose. Special texts for special purposes or high-risk texts may need special treatment?
* Will all machine translated texts always have to be pre- and post-edited? There can be found an answer in the term “fitness for use”.

Philip Koehn (2009) talks about quality of SMT and defines it as “fitness for use” as follows:

The ultimate test for machine translation is its usefulness in performing tasks that involve translated material. (p. 10)

Machine translation is not an end itself; it is used to support some kind of task perhaps supporting the efforts of a human translator to more efficiently translate documents for publication, or perhaps helping someone to understand the contents of a document in an unknown foreign language. If these are the uses of machine translation, then machine translation will ultimately be evaluated in the marketplaceon how well it supports these tasks. (p. 237)

**2.1 Diversification of roles and skills**

The EU Parliament, one of the largest translation institutions in the world, uses already MT for 70% of their translations. They talk about a “language industry” (*LI*) and work with new professional profiles like “*Language Professionals*” or “*Intercultural Linguistic Mediators*”, who carry out different tasks like adapting texts to different purposes and cultures. They work with a diversification of roles and skills, e.g. drafting, adapting, editing, post-editing (Valter Mavrič, Direction générale de la traduction, European Parliament, Dubrovnik 2018). The same is the case for the EU Commission which disposes of one of the biggest archives of human translated texts in the world, “the EU’s collection of legislation and publications” (Kluvanec: 2014: 52)

Division of labor in translation is not new. Already in the Sovjet Union with its 100 different languages, translations were carried out in several phases of drafting, terminological research, translating, reviewing, comparison with the original, stylistic proposals, end-redaction – all phases by different assistants and experts (Dedecius 1986:115f).

The most interesting question demonstrating the present and future challenges of TS was asked this summer in an e-mail to me by the earlier chief of the Language Service at the United Nations Office in Geneva (UNOG), Ms. Marie-Josée de Saint Robert:

*“What competencies do revisers need that translators would not have – and what competencies do translators have that revisers would not have?”*

For the UNO, a huge global leader in translation, diversification of roles and skills is not new either. Translating a large variety of texts from and into languages from all over the world, the UNO always needs both, researchers, advisors regarding special knowledge, drafting assistants for text-preparation, terminologists, legal support, translators, reviewers and revisers for post-editing, and editors. *Many brains* were always at use at the UNO – and supporting them in their revision processes, I needed models, many brains.

My *Many brains* – I think I showed them to you some years ago – also can explain the present situation of language professionals. If revisers and translators need different qualifications and competences, our traditional translation training may not be sufficient anymore. Pre-editing and post-editing will require special training.

1. **Many Brains**

The following models are inspired by the theories of signs of the Danish linguist Louis Hjelmslev (1943), the German linguist Klaus Heger (1971), and Karl Bühler’s Organonmodell (1934).



Figure 1 shows the translator’s brain working on the source language text (SL text) in a communication situation. As much of the ST content as necessary in the situation is conveyed to the target text (TT content) – at the same time it is crucial to *abstract* from the ST form(shown with the red line). Whole texts can be regarded as signs that refer to phenomena in the world. “Necessary in the situation” will say that the TT must be adapted, e.g. regarding deixis or in relation to the TT receiver and its function.



Figure 2 illustrates self-revision. For the translator abstracting from the ST form is necessary, and from his/her own translation, the TT form – and at the same time to keep the ST content and the TT content, necessary in the situation, in the revised target language text (RTL text). Here the extra extended brain shall demonstrate that is difficult to abstract from one’s own translations. The reason is that we often fall in love with our own formulations and have the tendency to mobilize the same myelin again and again. Waiting some time also supports abstraction.



Figure 3 illustrates other-revision. It shows three brains. The first brain produces the ST. The second brain shows the translator’s brain. The third brain is that of the reviser. Both translator and reviser need the ability to abstract from the ST form. The reviser must also abstract from the translator’s TT form, and if he/she has the source text – also from the ST form. At the same time, they need to keep the ST content and the TT content – having the translation brief and the communication situation in mind.

The following Figure 4 is a model of SMT. It shows human – machine interaction with three human brains, SMT and automatic post-editing. The ST is drafted and pre-edited into a pre-edited source language text (PSL text), for example supported by controlled language. It is translated by SMT, and the machine translation language text (MTL text) is controlled automatically by automatic post-editing. At last the automatic post-edited target language (APMTL text) is revised by a human being and the result is the human revised machine target language text (HRMTL text). During all these stages, keeping and controlling the content is necessary – and at the same not to get lost in the forms. I have never seen how automatic post-editing is applied.



The following Figure 5 shows NMT in 2018. The automatic post-editing seems to be an included part of the NMT. According the terminology used in the field, the pre-edited source text is now called the INPUT and the MTTL text is the OUTPUT. This output is post-edited by a human being.



These models are flexible. We could add brains, showing brains of, for example drafting assistants or legal advisers. However, we could also delete some of them. As NMT learns extremely quickly, the first human brain to become dispensable could be the brain of the pre-editor.

 **3.1 Deep learning models**

For the moment, there are used several deep learning models, the Convolutional Neural Network (CNN), which is especially effective in the computer vision area, e.g. for image recognition, and secondly, the Recurrent Neural Network (RNN), which is effective in the natural language area. They are used by Google. For Multimodality Learning, i.e. creating simultaneously illustrations and texts, both models are applied together (Wei Di, Anurag Bhardwaj & Jianing Wei. 2018:71ff).

When NMT is improving so quickly it is a. o. due to these models. While Statistical Machine Translation still was caught in the form of words and sentences, NMT is based on semantic relationships. The translation appears via a process of encoding the meaning of the ST and decoding it through an input-layer and different hidden layers to an output-layer, the target text. The following figure 6 shows that NMT is not caught in the form of the source text. A human post-editor, however, must still avoid being caught in the MTTL texts form. (AIT in this model means that source texts also can be created artificially).



A closer look into the Neural MT perhaps can provide a simple explanation for an extremely complicated process, see figure 7:



Via a Thought vector, which represents a thought’s relations to other thoughts in a context, natural human reasoning is modelled (Geoffry Hinton, in a 2015 speech to the Royal Society in London). The ST input is encoded in a neural representation, i.e. the meaning of words, sentences, passages and texts and their embedding and semantic relationships in the context are defined – and with the RNN, the content of the text is decoded via many hidden layers, and thus the output data, the translation is created. This process is supported by an open-source software Seq2Seq and Tensor Flow, which is a kind of deep neural library, developed by Google Brain (2015/2017). In this process, the probability of possible words in their local and larger context is calculated.

1. **Pre-editing and Post-editing**

Post-editing of NMT has shown to be different from post-editing texts translated with SMT. NMT creates a new kind of errors and flaws. This will be shown in some of the following examples.

Are they different skills, and different from translation? If we zoom up the brain of the pre-editing person, he/she should have the basic qualifications and knowledge shown in the heart of the figure – and additionally most of the other qualifications and competences for pre-editing in front of the heart, see Figure 8:

**Figure 8**

Zooming up the brain of the post-editor, the qualifications and competences needed could look like in the following model, see Figure 9:



**Figure 9**

**4.1 Pre-editing**

Pre-editing a text requires the ability to anticipate the potential errors that may occur if the source text is translated. For MT, pre-editing can be supported by Controlled language (ASD-STE100: 1986/2004). Controlled language requires simplifying text a. o.

* using short sentences
* one sentence – one message
* a consistent terminology
* no use of pro-forms
* active voice
* one word – one meaning
* no (modal) particles
* only imperative in directive speech acts
* international illustrations and graphs
* ...

This controlled language I tried with my master students for the pre-editing of our source texts. We trained summaries of all kinds of texts. Pre-editing is not easy. Unexpectedly it showed that students who were good translators were poor pre-editors and vice versa.

We tried, for example, to translate the directive speech acts into many languages – apart from German and Danish also into French, Italian and Russian. German and Danish dispose of special directive forms, the *directive infinitive* in German and in Danish a form which is called *S-passive*.



With human translation that is no problem. An imperative in English is translated with the respective forms into German and Danish. With GSMT and GNMT, we met unexpected challenges.



GSMT in 2014 shows what can happen even when the directive in the ST is an imperative. The Danish TT is nearly perfect, but the German TT shows a kind of machine confusion. As we have an imperative in the ST, the error in the German text where also imperative can be used seems to be unnecessary.

In August 2018, I tried to translate the same texts again with GNMT. The result was surprising. As expected, the German TT is perfect. Imperative is used even applying the polite “Sie” – but what about the Danish TT? With the old GSMT there had been only one pronoun lacking and now the Danish TT does not give any meaning?



This example shows a sneaky kind of errors. The translation of the imperative “chop” is wrong, as “hæld” means “pour”. The words “løgløg” for “springløg” and “fødevareprocessor” (grocery processor) for “food processor” do not exist in Danish. The system simply made up words and guessed.

**4.2 Post-editing**

Going back to the important question of Mme. Marie-Josée de Saint Robert: *“What competencies do revisers need that translators would not have – and what competencies do translators have that revisers would not have?”*

Post-editing of HT is a skill we have trained for many years. We trained fair minimal invasive corrections. Many students, especially bilingual students, tended to over-editing – they inserted unnecessary changes. The same was a problem at translation agencies and organizations where translators usually corrected each other’s translations. Over-editing created conflicts and friction between colleagues. Many of the students and later professional translators (my earlier test persons from translation process research) said that they would prefer to write a new translation instead of post-editing.

As can be seen from the examples with *hallucination*, in this article, it must be quite demanding to post-edit translations with NMT, as the system invents words and formulations that sound nice and seem to fit well, but don’t really make sense.

We trained post-editing with Googli. Googli was our virtual student. All our translations during the last years of our master study were also translated by Googli, and the students had to post-edit Googli’s translations. We defined Googli’s talents as follows:

* Googli hat keine Seele/kein Bewusstsein
* Googli hat kein Verständnis
* Googli hat kein Einfühlungsvermögen
* Googli hat weder Gewissen noch Verantwortungsgefühl
* Googli kann auch keine Verantwortung übernehmen
* Googli interessiert sich nicht für den Auftrag
* Googli weiβ nichts über Anpassung an die Kommunikationssituation und an die ZT-Empfänger
* Googli kann Vieles nicht
* …

Translated with Google NMT into English, September 2018

* Googli has no Googli has no soul / no consciousness
* Googli has no understanding
* Googli has no empathy
* Googli has no conscience or sense of responsibility
* Googli can not take responsibility either
* Googli is not interested in the job
* Googli knows nothing about adapting to the communication situation and to the ZT recipients
* Googli can not do much
* …

Translated with Google NMT into Rumanian, September 2018

* Google nu are suflet / conștiință
* Google nu are nicio înțelegere
* Google nu are empatie
* Google nu are conștiință sau sentiment de responsabilitate
* Google nu poate să-și asume nici o responsabilitate
* Google nu este interesat de acest post
* Google nu știe nimic despre adaptarea la situația de comunicare și la destinatarii ZT
* Google nu poate face prea mult
* …

**5. Challenges for NMT**

We saw what happened when GNMT is “hallucinating the output”. Confronted by infrequent or rare words, or metaphors, the system invents an output that is fluent and sounds good, but it does not make sense. This could be seen in the Danish example, where the words in bold don’t exist. Koehn and Knowles (2018:30) say about this weakness: “This is of particular concern when MT is used for information gisting — the user will be misled by hallucinated content in the NMT output.”

Another example may show what can happen:



The reason for this error may be that it is only one single sentence. As shown above, NMT depends on word-embedding in a context. And it is metaphoric use.

Apart from the problems with rare words and metaphors, other challenges for NMT are according to Koehn and Knowles (2018) domain specific texts, difficulties with long sentences (here pre-editing might help). Furthermore, the training samples are still too small. NMT is based on probability and there is still a need to explore and perhaps define the optimal beam size, which is the amount of predicted probable possible translations. Word alignment and phrase alignment does not always work. However, NMT is a self-learning system and gets better in rapid pace.

 **6. The State of TS in the World – Trends, Challenges and future Perspectives**

Going back to questions about jobs and skills and to the many brains, the trend will be more and more automatization, a change from *personal to digital*. Many of the traditional jobs will disappear. New kinds of jobs may come up – like for example professional pre- and post-editor instead of translator. Still for some time, human pre-editing will be important to check the input regarding potential cultural, pragmatic or stylistic challenges like, for example, rare words, special terminology and metaphors. As machines cannot be made liable, human post-editing will still be crucial. High risk texts will perhaps still be translated by HT if clients are willing to pay for human quality. Future professional translators will have to be flexible linguistic experts with the capacity to work in many different roles. Regrettably, in most countries the translation profession as well as post-editing lack a clear professional profile, recognition and prestige – and a title which would be in accordance with the knowledge and expertise needed for linguistic mediation.

I cited Franz Och who in 2009 and 2012 said that he would not stop before machine translations are as good as professional, human translations”. He may have been too modest. Perspectives are for the moment that future NMT or even another upcoming artificial intelligent MT system will probably translate better than any human translator. And they will be able to translate around the clock. This will cause new challenges.

We will have to live with rapid *CHANGE* – and the future problem may be adapting to the digital development with all its advantages and disadvantages. However, there are also new *CHANCES*.

See a GNMT of a Danish text from the newspaper Politiken from October 22 into your language:

EU-Parlamentet har netop vedtaget en resolution om, at sprogteknologi skal styrkes for alle europæiske sprog. Parlamentet anbefaler, at EU afsætter midler til forskning og udvikling for at sikre, at alle europæiske sprog får adgang til den sprogteknologi, der er afgørende for at kunne søge og bruge viden digitalt.

Parlamentul European tocmai a adoptat o rezoluție care consolidează tehnologia lingvistică pentru toate limbile europene. Parlamentul recomandă ca UE să aloce fonduri pentru cercetare și dezvoltare pentru a se asigura că toate limbile europene au acces la tehnologia lingvistică esențială pentru căutarea și utilizarea digitală a cunoștințelor.

And a machine at work:



1. **References**

ASD-STE100: 1986. Controlled Language: Aerospace and Defense: Simplified Technical English (2004): <http://www.asd-ste100.org>. (accessed October 2018)

Braun, Otto.1955. “Fragen der literarischen Übersetzung.” In: Von der Verantwortung des Übersetzers. *Beiträge zur deutschen Gegenwartsliteratur Heft 4. Diskussionsmaterial zur Vorbereitung des IV. Deutschen Schriftstellerkongresses*. Berlin: Deutscher Schriftstellerverband der DDR.

Bühler, Karl. 1934. *Sprachtheorie*. 3. Aufl. 1978. Stuttgart: Fischer.

De Saint Robert, Marie-Josée. 2018. Citation from e-mail.

Dedecius, Karl. 1986. *Vom Übersetzen*. Frankfurt am Main: Suhrkamp.

Heger, Klaus. 1971. *Monem, Wort und Satz*. Tübingen: Niemeyer.

Hjelmslev, Louis. 1943/19662. *Omkring Sprogteoriens Grundlæggelse*. København: Akademisk Forlag.

Kluvanec, Daniel. 2014. “Getting the right mix – Approaches to machine translation in the European Commission.” In: Baur, W. et al. *Man vs. Machine?* Volume I. Berlin. BDÜ Fachverlag. 51-57.

Koehn, Philipp. 2009. *Statistical Machine Translation*. Cambridge: Cambridge University Press.

Koehn, Philipp & Knowles, Rebecca. 2017.“Six Challenges for Neural Machine Translation”. In: *Proceedings of the First Workshop on Neural Machine Translation.* Vancouver: Association for Computational Linguistics. 28-39.

Mavrič,Valter. 2018. Direction générale de la traduction, European Parliament, mündlich in Dubrovnik 2018.

Och, Franz. 2009. So funktioniert ”Google Translate”. Interview in der Bildzeitung.

Schmitt, Peter.A. (2014). “Wer hat Angst vor Google MÜ”? In Baur, W. et al. *Man vs. Machine*? Vol.I. Baur, Berlin: BDÜ Fachverlag, 77-93.

Wei Di, Anurag Bhardwaj & Jianing Wei. 2018. *Deep Learning Essentials.* Packtpub.com.

Yinyin, Liu. (2017). <https://ai.intel.com/building-skip-thought-vectors-document-understanding/> (accessed November 6, 2018)