Zum möglichen Einfluss der neuronalen maschinellen Übersetzung (und weiterer sprachrelevanter KI-Technologien) auf die Fremdsprachenlehre

Vortrag im Rahmen der Vortragsreihe „Transformation des Fremdsprachenlernens" des Goethe-Instituts Israel in Jerusalem

Donnerstag, 17.11.2022

Prof. Dr. Ralph Krüger

Institut für Translation und Mehrsprachige Kommunikation

TH Köln

Technology Arts Sciences
TH Köln

Vorwort

- möglicher Einfluss der neuronalen maschinellen Übersetzung auf die Fremdsprachenlehre ist von der Translationswissenschaft erkannt worden und wird dort diskutiert:

„Language teachers have reported that **the performance of machine translation has a negative impact on students' motivation to learn foreign languages** […]. Moreover, **students have been found to employ free neural machine translation solutions in language courses to produce texts […], making it difficult to evaluate their progress with conventional assessment methods**. Conversely, even if MT tools are familiar to teachers, they rarely explicitly use them in language teaching, which leaves the didactic potential of these tools widely unexploited. Sometimes teachers are reluctant to adopt new technologies that they regard as disruptive or even a threat to language teaching. Furthermore, we may find that machine-translated output now fits the purpose for certain contexts of communication, which **calls for a redefinition of the goals of language teaching and the competences we are aiming for in different contexts**.“

(Delorme Benites/Lehr 2021:48, meine Hervorhebung)

„[T]he question of high-quality machine translation [and its role in the CEFR] is still absent – although **the routine of many experienced language teachers is already being jeopardised by NMT systems**. In other words, **which *reception*, *production* and even *mediation* activities described in the CEFR are likely to be changed or even made obsolete by the rise of NMT?**“

(ibid.:51, meine Hervorhebung)

Zur Einordnung

- weltweit gesprochene Sprachen: ca. 7000
- Sprachen mit entwickeltem Schreibsystem: 3995
- ‚periphere' Sprachen: 98 % aller natürlichen Sprachen
- ‚zentrale' Sprachen (Anwendung in Bildung, Rechtswesen, Administration, Medien, Industrie usw.): ca. 100
- ‚superzentrale' Sprachen (Weltsprachen): 13 (Arabisch, Chinesisch, Deutsch, Englisch, Französisch, Hindi, Japanisch, Malaiisch, Portugiesisch, Russisch, Spanisch, Suaheli, Türkisch)
- ‚hyperzentrale' Sprachen (linguae francae): 1 (Englisch)

(Brisset/Colón Rodríguez 2021:230)

Weltweite Translationsströme

Laut UNESCO Index Translationum

Table 16.1 Top 20 source languages

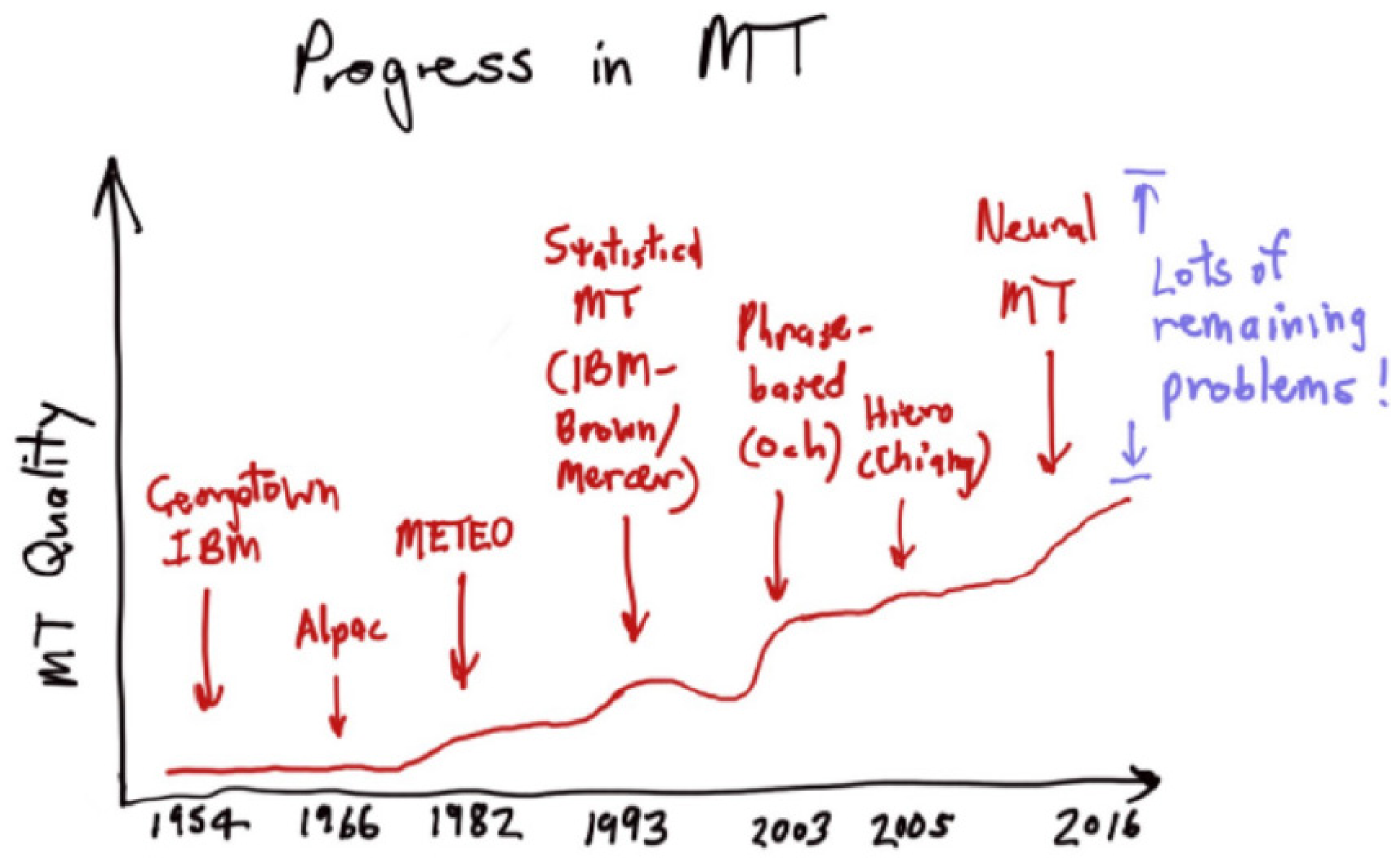
	Language	*No. of translations*	*Percent of top 20 SL*
1	English	1,266,110	55.35
2	French	226,123	9.88
3	German	208,240	9.10
4	Russian	103,624	4.53
5	Italian	69,555	3.04
6	Spanish	54,588	2.38
7	Swedish	39,984	1.74
8	Japanese	29,246	1.27
9	Danish	21,252	0.92
10	Latin	19,972	0.87
11	Dutch	19,667	0.85
12	Greek, Ancient (to 1453)	18,077	0.79
13	Czech	17,161	0.75
14	Polish	14,663	0.64
15	Norwegian	14,276	0.62
16	Chinese	14,071	0.61
17	Arabic	12,410	0.54
18	Portuguese	11,583	0.50
19	Hungarian	11,297	0.49
20	Hebrew	10,279	0.44
	Total	**2,182,178**	**95.4**
	Other source languages	**105,070**	**4.59**

Table 16.3 Top 20 target languages

	Language	*No. of translations*	*Percent of top 50 TL*
1	German	301,935	13.27
2	French	240,045	10.55
3	Spanish	228,559	10.05
4	English	164,509	7.23
5	Japanese	130,649	5.74
6	Dutch	111,270	4.89
7	Russian	100,806	4.43
8	Portuguese	78,904	3.46
9	Polish	76,706	3.37
10	Swedish	71,209	3.13
11	Czech	68,921	3.03
12	Danish	64,864	2.85
13	Chinese	63,123	2.77
14	Italian	61,087	2.68
15	Hungarian	55,214	2.42
16	Finnish	48,311	2.12
17	Norwegian	35,161	1.54
18	Greek, Modern (1453–)	30,459	1.33
19	Korean	28,168	1.23
20	Bulgarian	27,457	1.07
	Total	**1,987,356**	**87.39**
	Other target languages	**286,729**	**12.60**

(Brisset/Colón Rodríguez 2021:232, 234)

Historische Qualitätsentwicklung der maschinellen Übersetzung



(Manning et al. 2016)

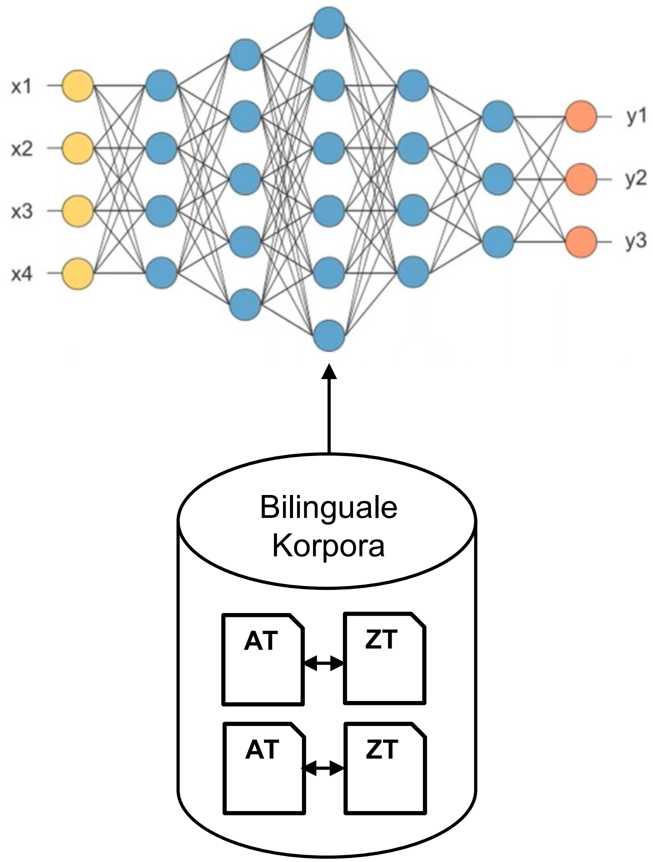
Technology
Arts Sciences
TH Köln

Neuronale maschinelle Übersetzung

Aktuelle Architektur auf Basis moderner KI-Verfahren

- Neuronale Netze werden mit großen Mengen zweisprachiger Trainingsdaten (Ausgangstexte und deren Übersetzungen) trainiert und lernen so, eigenständig neue Übersetzungen zu produzieren

→ **Output-Qualität ist unmittelbar abhängig von Menge und Qualität der Trainingsdaten**



Weitere Informationen:
[Krüger (2021): Die Transformer-Architektur für Systeme zur neuronalen maschinellen Übersetzung – eine popularisierende Darstellung]

Verfügbarkeit digitaler Sprachressourcen

Technologielücke zwischen Englisch als hyperzentraler Sprache und anderen Sprachen

- Leistungsfähigkeit von Sprachtechnologien in 30 europäischen Sprachen:

Table 3: State of LT support for 30 European languages in four different areas

Technology	Good Support	Moderate Support	Fragmentary Support	Weak/no Support
Machine Translation	English	French, Spanish	Catalan, Dutch, German, Hungarian, Italian, Polish, Romanian	Basque, Bulgarian, Croatian, Czech, Danish, Estonian, Finnish, Galician, Greek, Icelandic, Irish, Latvian, Lithuanian, Maltese, Norwegian, Portuguese, Serbian, Slovak, Slovene, Swedish, Welsh
Speech	English	Czech, Dutch, Finnish, French, German, Italian, Portuguese, Spanish	Basque, Bulgarian, Catalan, Danish, Estonian, Galician, Greek, Hungarian, Irish, Norwegian, Polish, Serbian, Slovak, Slovene, Swedish	Croatian, Icelandic, Latvian, Lithuanian, Maltese, Romanian, Welsh
Text Analysis	English	Dutch, French, German, Italian, Spanish	Basque, Bulgarian, Catalan, Czech, Danish, Finnish, Galician, Greek, Hungarian, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovene, Swedish	Croatian, Estonian, Icelandic, Irish, Latvian, Lithuanian, Maltese, Serbian, Welsh
Language Resources	English	Czech, Dutch, French, German, Hungarian, Italian, Polish, Spanish, Swedish	Basque, Bulgarian, Catalan, Croatian, Danish, Estonian, Finnish, Galician, Greek, Norwegian, Portuguese, Romanian, Serbian, Slovak, Slovene	Icelandic, Irish, Latvian, Lithuanian, Maltese, Welsh

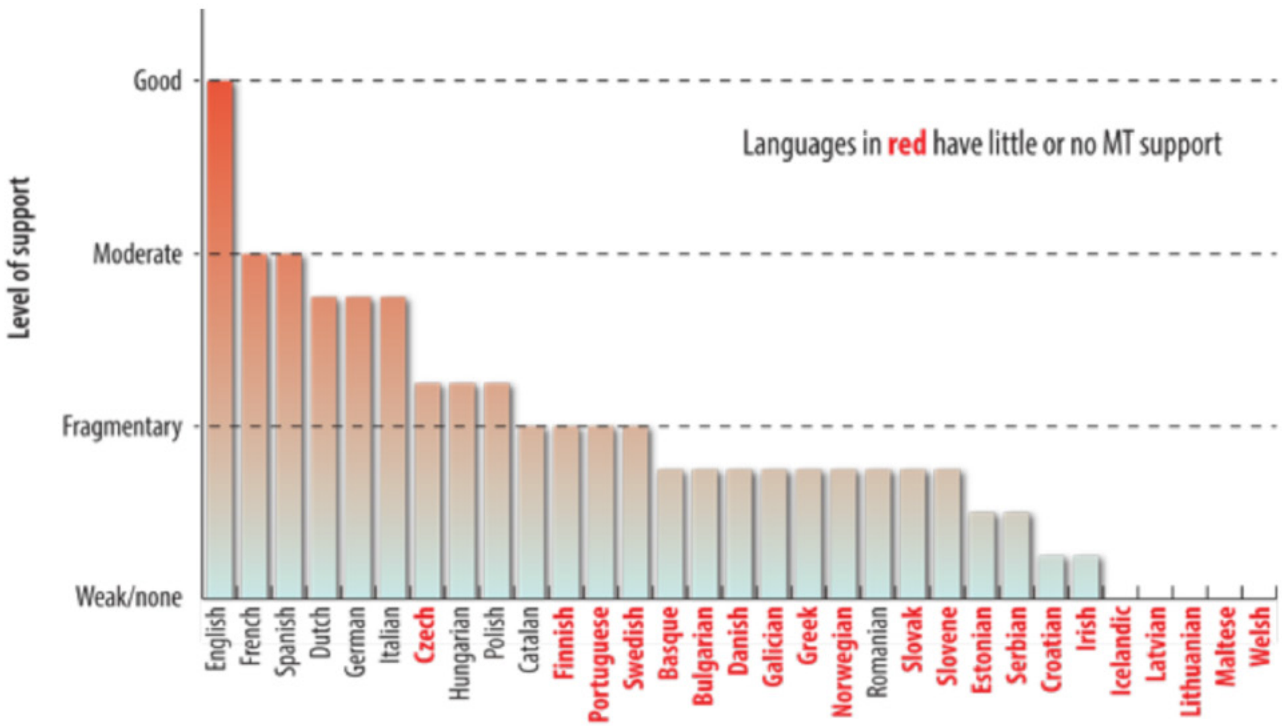
(EU-Parlament 2017:33)

Verfügbarkeit digitaler Sprachressourcen

Technologielücke zwischen Englisch als hyperzentraler Sprache und anderen Sprachen

- Verfügbarkeit und Qualität von MÜ-Systemen in 30 europäischen Sprachen:

Figure 5: Level of support of MT by language



(EU-Parlament 2017:34)

Verfügbarkeit digitaler Sprachressourcen

Aber derzeit Fortschritte im Bereich der maschinellen Übersetzung für „low-resource languages“

- Aktuelles Paper von Meta (Facebook):

No Language Left Behind: Scaling Human-Centered Machine Translation

NLLB Team, Marta R. Costa-jussà[*], James Cross[*], Onur Çelebi[*], Maha Elbayad[*], Kenneth Heafield[*], Kevin Heffernan[*], Elahe Kalbassi[*], Janice Lam[*], Daniel Licht[*], Jean Maillard[*], Anna Sun[*], Skyler Wang[*,§], Guillaume Wenzek[*], Al Youngblood[*]
Bapi Akula, Loic Barrault, Gabriel Mejia Gonzalez, Prangthip Hansanti, John Hoffman, Semarley Jarrett, Kaushik Ram Sadagopan, Dirk Rowe, Shannon Spruit, Chau Tran
Pierre Andrews[†], Necip Fazil Ayan[†], Shruti Bhosale[†], Sergey Edunov[†], Angela Fan[†,‡], Cynthia Gao[†], Vedanuj Goswami[†], Francisco Guzmán[†], Philipp Koehn[†,¶], Alexandre Mourachko[†], Christophe Ropers[†], Safiyyah Saleem[†], Holger Schwenk[†], Jeff Wang[†]

Meta AI, [§]UC Berkeley, [¶]Johns Hopkins University

Abstract

Driven by the goal of eradicating language barriers on a global scale, machine translation has solidified itself as a key focus of artificial intelligence research today. However, such efforts have coalesced around a small subset of languages, leaving behind the vast majority of mostly low-resource languages. What does it take to break the 200 language barrier while ensuring safe, high quality results, all while keeping ethical considerations in mind? In *No Language Left Behind*, we took on this challenge by first contextualizing the need for low-resource language translation support through exploratory interviews with native speakers. Then, we created datasets and models aimed at narrowing the performance gap between low and high-resource languages. More specifically, we developed a conditional compute model based on Sparsely Gated Mixture of Experts that is trained on data obtained with novel and effective data mining techniques tailored for low-resource languages. We propose multiple architectural and training improvements to counteract overfitting while training on thousands of tasks. Critically, we evaluated the performance of over 40,000 different translation directions using a human-translated benchmark, Flores-200, and combined human evaluation with a novel toxicity benchmark covering all languages in Flores-200 to assess translation safety. Our model achieves an improvement of 44% BLEU relative to the previous state-of-the-art, laying important groundwork towards realizing a universal translation system. Finally, we open source all contributions described in this work, accessible at `https://github.com/facebookresearch/fairseq/tree/nllb`.

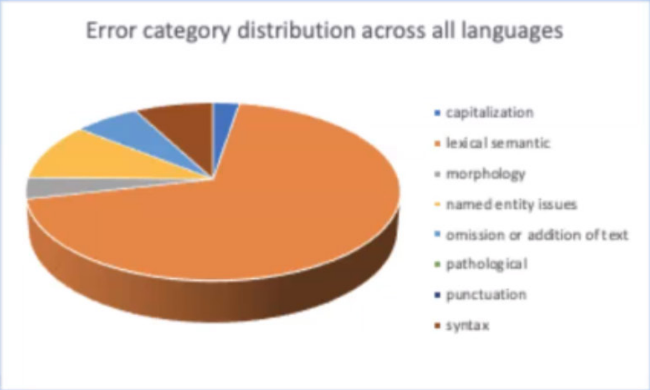
(https://ai.facebook.com/research/publications/no-language-left-behind-scaling-human-centered-machine-translation/)

Leistungsfähigkeit der neuronalen maschinellen Übersetzung

Fehlerverteilung von NMÜ-System über mehrere Sprachen hinweg

Error category average percentages - all languages

Lexical semantic Word ambiguity Noisy source Unknown words Code-switching Dialectal variants	30.00%
Named entity issues	4.00%
Omission or addition of text	7.00%
Pathological translations	3.00%
Syntax	3.00%
Morphology	2.00%
Capitalization	1.00%
Punctuation	0.01%



(Originalquelle siehe Vashee 2020)

Technology
Arts Sciences
TH Köln

Leistungsfähigkeit der neuronalen maschinellen Übersetzung

Leistungsranking der NMÜ bei Übersetzungen aus dem Englischen

AMTA

Actual ranking FROM English

1	Italian	18	Telugu	34	Latvian
2	Chinese (Traditional)	19	Turkish	35	Maltese
3	Chinese (Simplified)	20	Malaysian	36	Finnish
4	French	21	Indonesian	37	Ukrainian
5	Japanese	22	Bulgarian	38	Slovak
6	Spanish	23	Croatian	39	Filipino
7	Danish	24	Icelandic	40	Farsi
8	Hindi	25	Russian	41	Slovenian
9	Portuguese (Brazil)	26	Romanian	42	Urdu
10	Swedish	27	Greek	43	Catalan
11	Arabic	28	Hebrew	44	Polish
12	German	29	Thai	45	Serbian
13	Norwegian	30	Tamil	46	Estonian
14	Welsh	31	Hungarian	47	Marathi
15	Dutch	32	Swahili	48	Malagasy
16	Korean	33	Vietnamese	49	Lithuanian
17	Czech				

Ranked by delta to human translation quality

We compare the human evaluation of human translated sentences to the same source sentences machine translated and sort by the delta, smallest to largest.

Translated by Microsoft

(Originalquelle siehe Vashee 2020)

Weitere sprachrelevante KI-Technologien

Maschinelles Dolmetschen am Beispiel des Skype Translator von Microsoft

Skype Translator

Ganz gleich, ob Sie eine Übersetzung von Englisch in Spanisch oder von Englisch in Französisch benötigen oder per Sprache oder Textnachricht in Dutzenden Sprachen kommunizieren möchten, mit Skype überwinden Sie Sprachbarrieren in Echtzeit. So tauschen Sie sich nahtlos mit Freunden, Familienmitgliedern, Kunden und Kollegen aus.

Unser **Sprachübersetzer** kann momentan Unterhaltungen in 11 Sprachen übersetzen, darunter **Arabisch, Chinesisch (Mandarin), Deutsch, Englisch, Französisch, Italienisch, Portugiesisch (Brasilien), Russisch und Spanisch**.

Und unser **Textübersetzer** überträgt Chatnachrichten präzise und schnell in **mehr als 60 Sprachen**. Natürlich kommen laufend neue Sprachen hinzu. Sollten Sie die gewünschte Sprache oder Sprachvariante nicht in der Liste unterstützter Sprachen finden, dann schauen Sie bald wieder rein.

Skype herunterladen

(https://www.skype.com/de/features/skype-translator/)

- dreifacher ‚Flaschenhals': Spracherkennung → maschinelle Übersetzung → Sprachausgabe

- zusätzliche ‚Störfaktoren' bei gesprochener Sprache: Dialekt, Slang, Hintergrundgeräusche, elliptischer Sprachgebrauch aufgrund stärkerer Kontextgebundenheit usw.

→ **diese Faktoren beeinträchtigen zusätzlich die Qualität des MÜ-Outputs**

Weitere sprachrelevante KI-Technologien

Textzusammenfassung auf Grundlage neuronaler Sprachmodelle

Summarization demo

using sshleifer/distilbart-cnn-12-6

Summarization

Examples

The tower is 324 metres (1,063 ft) tall, about the same height as an 81-storey building, and the tallest structure in Paris. Its base is square, measuring 125 metres (410 ft) on each side. During its construction, the Eiffel Tower surpassed the Washington Monument to become the tallest man-made structure in the world, a title it held for 41 years until the Chrysler Building in New York City was finished in 1930. It was the first structure to reach a height of 300 metres. Due to the addition of a broadcasting aerial at the top of the tower in 1957, it is now taller than the Chrysler Building by 5.2 metres (17 ft). Excluding transmitters, the Eiffel Tower is the second tallest free-standing structure in France after the Millau Viaduct.

Compute

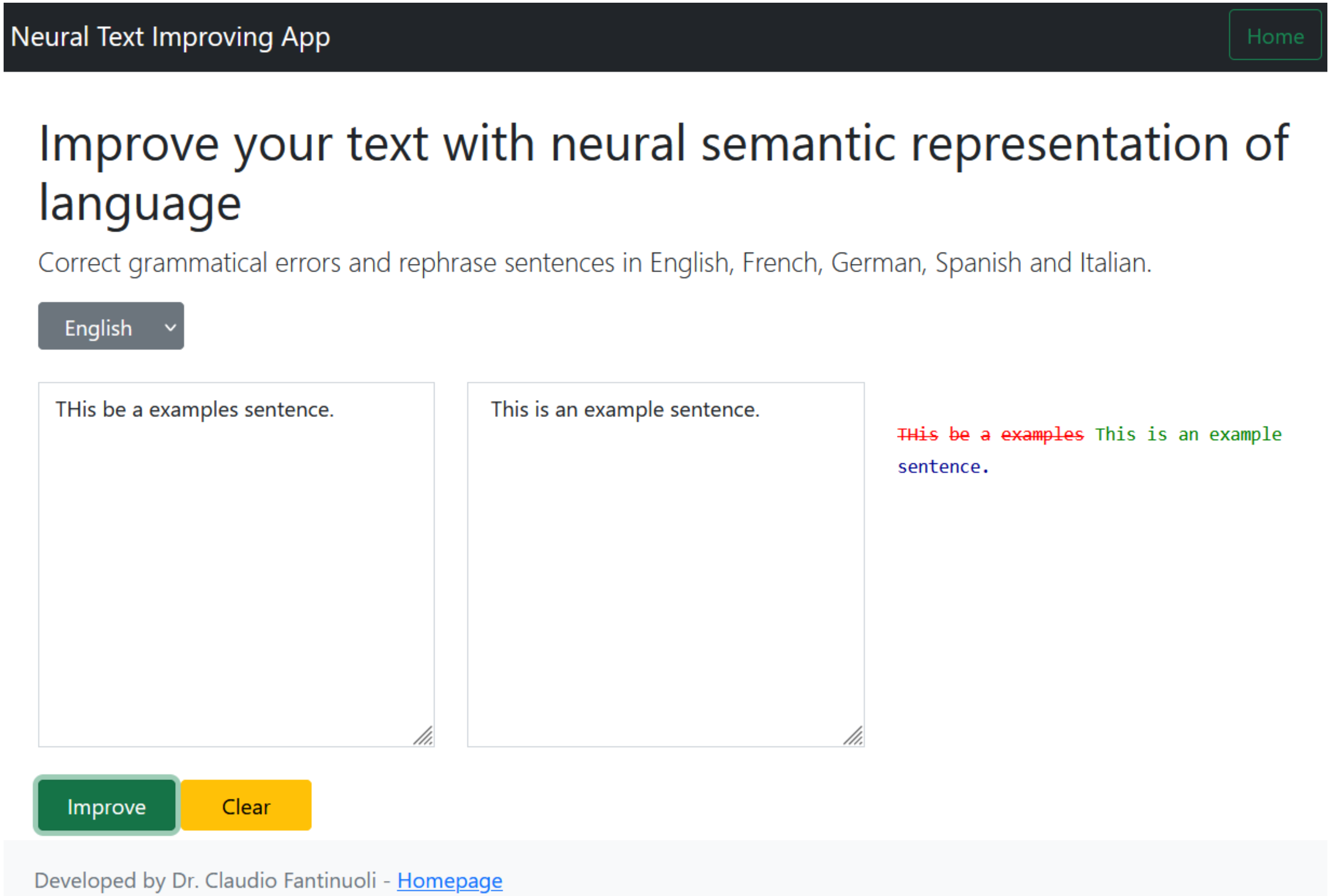
Computation time on cpu: cached

The tower is 324 metres (1,063 ft) tall, about the same height as an 81-storey building . It was the first structure to reach a height of 300 metres . It is now taller than the Chrysler Building in New York City by 5.2 metres (17 ft) Excluding transmitters, the Eiffel Tower is the second tallest free-standing structure in France .

(https://huggingface.co/tasks/summarization)

Weitere sprachrelevante KI-Technologien

Textoptimierung auf Grundlage neuronaler Sprachmodelle

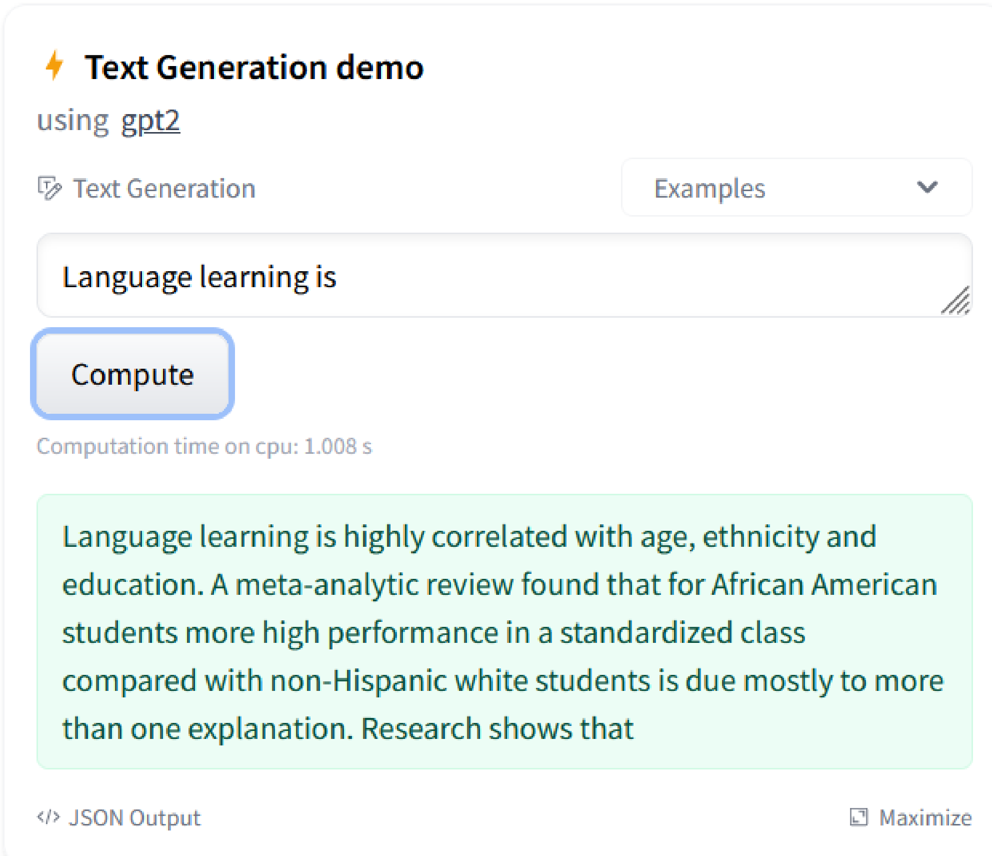


(https://www.claudiofantinuoli.org/apps/COR/index.html)

Technology Arts Sciences TH Köln

Weitere sprachrelevante KI-Technologien

Textfortsetzung auf Grundlage neuronaler Sprachmodelle



(https://huggingface.co/tasks/text-generation)

Leistungsfähigkeit der NMÜ/sprachrelevanter KI-Technologien

Verortung (hochperformanter) Systeme im Gemeinsamen Europäischen Referenzrahmen für Sprachen (nach Delorme Benites/Lehr 2021)

		A1	A2	B1	B2	C1	C2
UNDERSTANDING	**Listening**	I can recognise familiar words and very basic phrases concerning myself, my family and immediate concrete surroundings when people speak slowly and clearly.	I can understand phrases and the highest frequency vocabulary related to areas of most immediate personal relevance (e.g. very basic personal and family information, shopping, local area, employment). I can catch the main point in short, clear, simple messages and announcements.	I can understand the main points of clear standard speech on familiar matters regularly encountered in work, school, leisure, etc. I can understand the main point of many radio or TV programmes on current affairs or topics of personal or professional interest when the delivery is relatively slow and clear.	I can understand extended speech and lectures and follow even complex lines of argument provided the topic is reasonably familiar. I can understand most TV news and current affairs programmes. I can understand the majority of films in standard dialect.	I can understand extended speech even when it is not clearly structured and when relationships are only implied and not signalled explicitly. I can understand television programmes and films without too much effort.	I have no difficulty in understanding any kind of spoken language, whether live or broadcast, even when delivered at fast native speed, provided I have some time to get familiar with the accent.
	Reading	I can understand familiar names, words and very simple sentences, for example on notices and posters or in catalogues.	I can read very short, simple texts. I can find specific, predictable information in simple everyday material such as advertisements, prospectuses, menus and timetables and I can understand short simple personal letters.	I can understand texts that consist mainly of high frequency everyday or job-related language. I can understand the description of events, feelings and wishes in personal letters.	I can read articles and reports concerned with contemporary problems in which the writers adopt particular attitudes or viewpoints. I can understand contemporary literary prose.	I can understand long and complex factual and literary texts, appreciating distinctions of style. I can understand specialised articles and longer technical instructions, even when they do not relate to my field.	I can read with ease virtually all forms of the written language, including abstract, structurally or linguistically complex texts such as manuals, specialised articles and literary works.
SPEAKING	**Spoken Interaction**	I can interact in a simple way provided the other person is prepared to repeat or rephrase things at a slower rate of speech and help me formulate what I'm trying to say. I can ask and answer simple questions in areas of immediate need or on very familiar topics.	I can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar topics and activities. I can handle very short social exchanges, even though I can't usually understand enough to keep the conversation going myself.	I can deal with most situations likely to arise whilst travelling in an area where the language is spoken. I can enter unprepared into conversation on topics that are familiar, of personal interest or pertinent to everyday life (e.g. family, hobbies, work, travel and current events).	I can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible. I can take an active part in discussion in familiar contexts, accounting for and sustaining my views.	I can express myself fluently and spontaneously without much obvious searching for expressions. I can use language flexibly and effectively for social and professional purposes. I can formulate ideas and opinions with precision and relate my contribution skilfully to those of other speakers.	I can take part effortlessly in any conversation or discussion and have a good familiarity with idiomatic expressions and colloquialisms. I can express myself fluently and convey finer shades of meaning precisely. If I do have a problem I can backtrack and restructure around the difficulty so smoothly that other people are hardly aware of it.
	Spoken Production	I can use simple phrases and sentences to describe where I live and people I know.	I can use a series of phrases and sentences to describe in simple terms my family and other people, living conditions, my educational background and my present or most recent job.	I can connect phrases in a simple way in order to describe experiences and events, my dreams, hopes and ambitions. I can briefly give reasons and explanations for opinions and plans. I can narrate a story or relate the plot of a book or film and describe my reactions.	I can present clear, detailed descriptions on a wide range of subjects related to my field of interest. I can explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.	I can present clear, detailed descriptions of complex subjects integrating sub-themes, developing particular points and rounding off with an appropriate conclusion.	I can present a clear, smoothly-flowing description or argument in a style appropriate to the context and with an effective logical structure which helps the recipient to notice and remember significant points.
WRITING	**Writing**	I can write a short, simple postcard, for example sending holiday greetings. I can fill in forms with personal details, for example entering my name, nationality and address on a hotel registration form.	I can write short, simple notes and messages relating to matters in areas of immediate needs. I can write a very simple personal letter, for example thanking someone for something.	I can write simple connected text on topics which are familiar or of personal interest. I can write personal letters describing experiences and impressions.	I can write clear, detailed text on a wide range of subjects related to my interests. I can write an essay or report, passing on information or giving reasons in support of or against a particular point of view. I can write letters highlighting the personal significance of events and experiences.	I can express myself in clear, well-structured text, expressing points of view at some length. I can write about complex subjects in a letter, an essay or a report, underlining what I consider to be the salient issues. I can select style appropriate to the reader in mind.	I can write clear, smoothly-flowing text in an appropriate style. I can write complex letters, reports or articles which present a case with an effective logical structure which helps the recipient to notice and remember significant points. I can write summaries and reviews of professional or literary works.

(Grafik aus Delorme Benites/Lehr 2021:55)

NMÜ und Lesekompetenz in der Fremdsprache

Stilistische Nuancen und implizite Bedeutungen sind von NMÜ bisher nicht erfassbar

NMÜ alleine: maschinelle Übersetzung aus L2 in L1

	Overall reading comprehension
C2	Can understand virtually all types of texts including abstract, structurally complex, or highly colloquial literary and non-literary writings.
	Can understand a wide range of long and complex texts, appreciating subtle distinctions of style and implicit as well as explicit meaning.
C1	Can understand in detail lengthy, complex texts, whether or not these relate to their own area of speciality, provided they can reread difficult sections.
	Can understand a wide variety of texts including literary writings, newspaper or magazine articles, and specialised academic or professional publications, provided there are opportunities for rereading and they have access to reference tools.
B2	Can read with a large degree of independence, adapting style and speed of reading to different texts and purposes, and using appropriate reference sources selectively. Has a broad active reading vocabulary, but may experience some difficulty with low-frequency idioms.
B1	Can read straightforward factual texts on subjects related to their field of interest with a satisfactory level of comprehension.
A2	Can understand short, simple texts on familiar matters of a concrete type which consist of high frequency everyday or job-related language.
	Can understand short, simple texts containing the highest frequency vocabulary, including a proportion of shared international vocabulary items.
A1	Can understand very short, simple texts a single phrase at a time, picking up familiar names, words and basic phrases and rereading as required.
Pre-A1	Can recognise familiar words/signs accompanied by pictures, such as a fast-food restaurant menu illustrated with photos or a picture book using familiar vocabulary.

(Grafik aus Delorme Benites/Lehr 2021:60)

- Zusammenhang zwischen Vokabularerwerb und Lesekompetenz wird durch NMÜ unterbrochen
- Wie können C2-Kompetenzen ohne lineare Progression durch die vorherigen Stufen erworben werden?

→ in Zukunft didaktische Trennung zwischen Lesen zum Vokabularerwerb und Lesen zum Textverständnis? (vgl. Delorme Benites/Lehr 2021:61)

NMÜ und Schreibkompetenz in der Fremdsprache

Overall written production

	Overall written production
C2	Can produce clear, smoothly flowing, complex texts in an appropriate and effective style and a logical structure which helps the reader identify significant points.
C1	Can produce clear, well-structured texts of complex subjects, underlining the relevant salient issues, expanding and supporting points of view at some length with subsidiary points, reasons and relevant examples, and rounding off with an appropriate conclusion. Can employ the structure and conventions of a variety of genres, varying the tone, style and register according to addressee, text type and theme.
B2	Can produce clear, detailed texts on a variety of subjects related to their field of interest, synthesising and evaluating information and arguments from a number of sources.
B1	Can produce straightforward connected texts on a range of familiar subjects within their field of interest, by linking a series of shorter discrete elements into a linear sequence.
A2	Can produce a series of simple phrases and sentences linked with simple connectors like "and", "but" and "because".
A1	Can give information about matters of personal relevance (e.g. likes and dislikes, family, pets) using simple words/signs and basic expressions. Can produce simple isolated phrases and sentences.
Pre-A1	Can give basic personal information (e.g. name, address, nationality), perhaps with the use of a dictionary.

(Grafik aus Delorme Benites/Lehr 2021:58)

NMÜ als Hilfsmittel im menschlichen Schreibprozess (‚Post-Editing' des maschinell übersetzten Textes)

NMÜ alleine: Textproduktion in L1 und MÜ in L2

- lineare Progression durch GER-Stufen wird durch NMÜ gestört → wie kann NMÜ in die unteren Stufen integriert werden, damit die menschlichen Eingriffe auf Stufe C1 und C2 möglich sind? (vgl. Delorme Benites/Lehr 2021:59)

NMÜ und Mediationskompetenz

Text-to-Speech-Systeme als Hilfsmittel

Text-to-Speech-Systeme

TRANSLATING A WRITTEN TEXT IN SPEECH	
Note: As in any case in which mediation across languages is involved, users may wish to complete the descriptor by specifying the languages concerned.	
C2	Can provide fluent spoken translation into (Language B) of abstract texts written in (Language A) on a wide range of subjects of personal, academic and professional interest, successfully conveying evaluative aspects and arguments, including the nuances and implications associated with them.
C1	Can provide fluent spoken translation into (Language B) of complex written texts written in (Language A) on a wide range of general and specialised topics, capturing most nuances.
B2	Can provide spoken translation into (Language B) of complex texts written in (Language A) containing information and arguments on subjects within his/her fields of professional, academic and personal interest.
B1	Can provide spoken translation into (Language B) of texts written in (Language A) containing information and arguments on subjects within his/her fields of professional, academic and personal interest, provided that they are written in uncomplicated, standard language.
	Can provide an approximate spoken translation into (Language B) of clear, well-structured informational texts written in (Language A) on subjects that are familiar or of personal interest, although his/her lexical limitations cause difficulty with formulation at times.
A2	Can provide an approximate spoken translation into (Language B) of short, simple everyday texts (e.g. brochure entries, notices, instructions, letters or emails) written in (Language A).
	Can provide a simple, rough, spoken translation into (Language B) of short, simple texts (e.g. notices on familiar subjects) written in (Language A), capturing the most essential point. Can provide a simple, rough spoken translation into (Language B) of routine information on familiar everyday subjects that is written in simple sentences in (Language A) (e.g. personal news, short narratives, directions, notices or instructions).
A1	Can provide a simple, rough spoken translation into (Language B) of simple, everyday words and phrases written in (Language A) that are encountered on signs and notices, posters, programmes, leaflets etc.
Pre-A1	*No descriptors available*

NMÜ als Hilfsmittel (‚Post-Editing' des maschinell übersetzten Textes, vgl. Schreibkompetenz)

NMÜ alleine (auf unteren Stufen ist NMÜ evtl. sogar besser als Sprachlernende)

TRANSLATING A WRITTEN TEXT IN WRITING	
C2	Can translate into (Language B) technical material outside his/her field of specialisation written in (Language A), provided subject matter accuracy is checked by a specialist in the field concerned.
C1	Can translate into (Language B) abstract texts on social, academic and professional subjects in his/her field written in (Language A), successfully conveying evaluative aspects and arguments, including many of the implications associated with them, though some expression may be over-influenced by the original.
B2	Can produce clearly organised translations from (Language A) into (Language B) that reflect normal language usage but may be over-influenced by the order, paragraphing, punctuation and particular formulations of the original.
	Can produce translations into (Language B, which closely follow the sentence and paragraph structure of the original text in (Language A), conveying the main points of the source text accurately, though the translation may read awkwardly.
B1	Can produce approximate translations from (Language A) into (Language B) of straightforward, factual texts that are written in uncomplicated, standard language, closely following the structure of the original; although linguistic errors may occur, the translation remains comprehensible.
	Can produce approximate translations from (Language A) into (Language B) of information contained in short, factual texts written in uncomplicated, standard language; despite errors, the translation remains comprehensible.
A2	Can use simple language to provide an approximate translation from (Language A) into (Language B) of very short texts on familiar and everyday themes that contain the highest frequency vocabulary; despite errors, the translation remains comprehensible.
A1	Can, with the help of a dictionary, translate simple words and phrases from (Language A) into (Language B), but may not always select the appropriate meaning.
Pre-A1	*No descriptors available*

(Grafik aus Delorme Benites/Lehr 2021:63)

NMÜ und Mediationskompetenz

- Sind Mediationskompetenzen noch relevant für die Bestimmung der Gesamtkompetenz in einer Fremdsprache?

- Sollte Mediationskompetenz in Zukunft auf professionelle Fachkommunikations-/Fachübersetzungskontexte beschränkt werden?

- Sollte auf neue Aspekte mit Einbezug der NMÜ fokussiert werden (z. B. kontrolliertes Schreiben oder Pre-Editing in L1 zur Verbesserung des MÜ-Ergebnisses in L2)?

- Sollte Sprachmediation zu Sprachlernzwecken dezidiert ohne NMÜ durchgeführt werden und auf geeignete didaktische Ziele wie den Erwerb interkultureller Kompetenzen fokussieren?

(vgl. Delorme Benites/Lehr 2021:62)

Fazit: NMÜ und der GER

- Kompetenzen der unteren bis mittleren Stufen werden inzwischen durch die NMÜ abgedeckt → neue Kompetenzdeskriptoren, die aktuelle sprachtechnologische Entwicklungen integrieren?

- Fremdsprachenkompetenz kann in Zukunft nicht mehr ohne Aspekte der Mensch-Maschine-Interaktion konzeptualisiert werden (Situated/Extended/Distributed Cognition)

- Technologieorientierte Aspekte des Sprachenlernens sollten ergänzt werden durch Fokussierung auf exklusive menschliche Kompetenzen (interkulturelle Unterschiede, interkulturelle Empathie usw.)

- Im Zuge der gesellschaftlichen Digitalisierung und Datafizierung entwickeln sich neue *Literacy*-Konzepte: *Digital Literacy*, *Data Literacy*, *AI Literacy*, ***Machine Translation Literacy***

(vgl. Delorme Benites/Lehr 2021:64)

Machine Translation Literacy

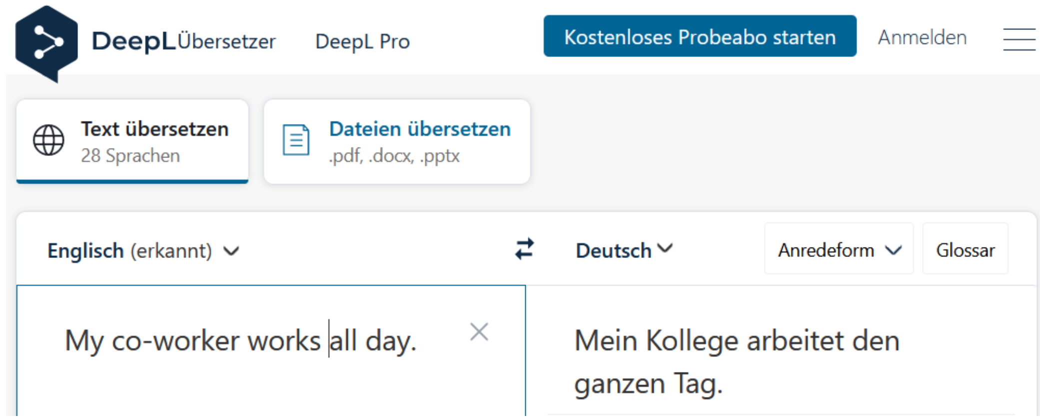
„**In order to become informed and critical users of machine translation tools, and of their translated output, students need to develop machine translation literacy.** […] [T]he how-to skills of using machine translation tools are very easy to acquire since they consist of little more than copying and pasting a text, choosing a language pair, and clicking the "Translate" button. **Using machine translation is easy, but using it critically requires some thought.** Therefore […], machine translation literacy is mainly about developing cognitive competencies, rather than techno-procedural skills."

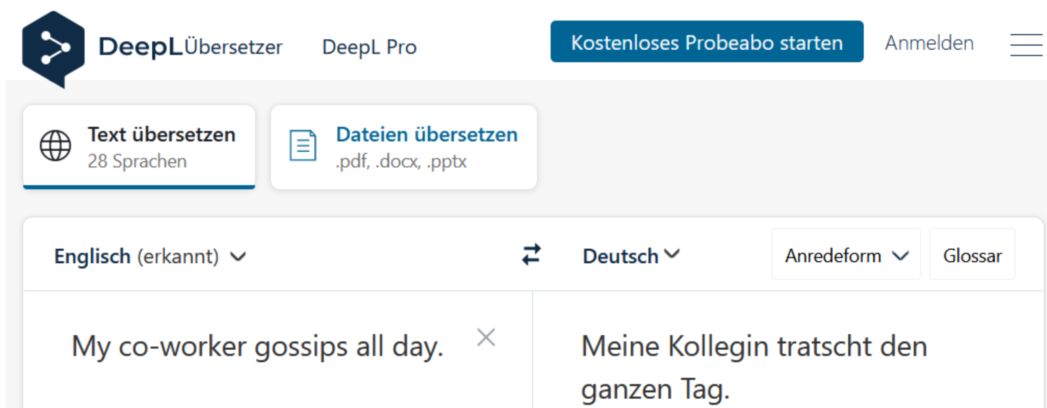
(Bowker 2021:132, meine Hervorhebung)

Anhang: Der ‚Priming'-Effekt der maschinellen Übersetzung

„[A]nother challenge is the fact that **the widespread use of online MT tools in the language classroom can influence the language learning process**. For instance, thanks to a **'syntactic priming study'** with Brazilian Portuguese learners of English as a second language, Resende & Way (2021) have shown that **syntactic constructions provided by Portuguese-to-English MT output has an influence on learners' production** through the reuse of syntactic constructions. The authors conclude that **MT has a 'robust long-lasting priming effect'** (Resende & Way, 2021: 82).
All of these challenges suggest that **students might place excessive trust in what the machine has to offer**, and might not be aware of the limits of the technology. This kind of attitude is nothing new and is not related to the development of MT tools: students have often been found to tend to 'overtrust' translation aids like general bilingual dictionaries […], translation memories […], or electronic corpora […].“ (Loock/Léchauguette 2019:207, meine Hervorhebung)

Anhang: Sozialer *Bias* in der maschinellen Übersetzung





(vgl. Hackenbuchner 2022:40)

Zur weiteren Lektüre

Machine translation for language learners

Alice Carré[a], Dorothy Kenny[b], Caroline Rossi[a], Pilar Sánchez-Gijón[c] & Olga Torres-Hostench[c]

[a]Université Grenoble-Alpes [b]Dublin City University [c]Universitat Autònoma de Barcelona

Machine Translation (MT) has been controversial in second and foreign language learning, but the strategic integration of MT might be beneficial to language learning in certain contexts. In this chapter we discuss the conditions in which MT can be useful in language learning, set out digital alternatives to MT, and provide examples of how MT can support language learners.

1 Introduction

Machine translation (MT) has been controversial in second and foreign language learning,[1] with some commentators arguing that it can encourage plagiarism, promote errors or deflect learners from what they should be doing. In some cases, however, MT has been found to help students complete certain tasks, and there appears to be merit in considering MT as just one among many digital resources that contemporary language learners can use. The successful integration of MT into language learning requires us to understand, even at a basic level: how the technology works, how we can judge the quality of its outputs, how those outputs can be improved through intervention either before or after the fact of translation (through pre-editing or post-editing), and what the ethical issues in using

[1]Note that we use the generic terms *language learning* and *language learner* in this chapter to cover instances of foreign language learning and second and subsequent language learning. If a student's first language is their L1, then the language learning to which we refer corresponds to their learning of an L2, L3 or L*n*.

(https://zenodo.org/record/6760024/files/342-Kenny-2022-10.pdf?download=1)

Technology Arts Sciences
TH Köln

Literatur

- Bowker, Lynne (2021): „Promoting Linguistic Diversity and Inclusion: Incorporating Machine Translation Literacy into Information Literacy Instruction for Undergraduate Students“, in: *The International Journal of Information, Diversity, & Inclusion* 5(3), 127-151. https://doi.org/10.33137/ijidi.v5i3.36159

- Brisset, Annie; Colón Rodríguez, Raúl (2021): „World Translation Flows. Preferred Languages and Subjects“, in: Bielsa, Esperança; Kapsaskis, Dionysios (Hrsg.): *The Routledge Handbook of Translation and Globalization*. London/New York: Routledge, 230-250.

- Carré, Alice; Kenny, Dorothy; Rossi, Caroline; Sánchez-Gijón, Pilar; Torres-Hostench, Olga (2022): „Machine Translation for Language Learners“, in: Kenny, Dorothy /Hrsg.): *Machine Translation for Everyone. Empowering Users in the Age of Artificial Intelligence*. Berlin: Language Science Press. https://doi.org/10.5281/zenodo.6653406

- Delorme Benites, Alice; Lehr, Caroline (2021): „Neural Machine Translation and Language Teaching – Possible Implications for the CEFR“, in: *Bulletin suisse de linguistique appliquée* 114, 47-66. https://www.researchgate.net/publication/359619476_Neural_machine_translation_and_language_teaching_-_possible_implications_for_the_CEFR

- EU-Parlament (2017): *Language Equality in the Digital Age. Towards a Human Language Project*. Science and Technology Options Assessment (STOA) Study. European Parliamentary Research Service. https://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_STU(2017)598621

Technology
Arts Sciences
TH Köln

Literatur

- Hackenbuchner, Janiça (2022): „Gender Bias in Machine Translation – A Comparison of Systems Trained with Biased and Unbiased Datasets“. Masterarbeit am Institut für Translation und Mehrsprachige Kommunikation der TH Köln.

- Krüger, Ralph (2021): „Die Transformer-Architektur für Systeme zur neuronalen maschinellen Übersetzung – eine popularisierende Darstellung“, in: *trans-kom* 14(2), 278-324. https://www.trans-kom.eu/bd14nr02/trans-kom_14_02_05_Krueger_NMUe.20211202.pdf

- Loock, Rudy; Léchauguette, Sophie (2019): „Machine Translation Literacy and Undergraduate Students in Applied Languages: Report on an Exploratory Study“, in: *Revista Tradumàtica* 19, 204-225. https://doi.org/10.5565/rev/tradumatica.281

- Manning, Christopher; Cho, Kyunghyun; Luong, Thang (2016): *ACL 2016 Tutorial on Neural Machine Translation*. Stanford Natural Language Processing Group. https://sites.google.com/site/acl16nmt/home

- Vashee, Kirti (2020): „American Machine Translation Association (AMTA2020) Conference Highlights.“ *eMpTy Pages*. http://kv-emptypages.blogspot.com/2020/12/american-machine-translation.html

Technology
Arts Sciences
TH Köln