

Machine Translation

The first approach to automatic translation was in the 1950s; however, it is just called an approach and nothing else. Until the release of the internet in the 1990s, people understood the power of Machine Translation (MT). With the establishment of Google Translate in 2006 began a new era of MT, and the tool was published to various devices such as phones and tablets. After a decade, Google switched its translation program to a new paradigm and was able to translate complete sentences as fast as a human (Duncan, 2017).

Nowadays, Google Translate has got competition, and they are worth seeing. Macketanz et al. (2021) state that DeepL (a small German ML system company established in 2017) performs better than Google Translate. Macketanz and his team from Globalization & Localization Association (Gala) used a semi-automatic tool TQ-AutoTest to measure and analyse the quality of the MT from Google and DeepL.

Gala provided some examples of the system comparison to confirm his statement. In addition to this, the team analysed and compared Google Translate with DeepL and Lucy. Lucy is a rule-based translation system, and Google and DeepL neural network translation systems (AI networks). To make it short, in the quantitative comparison, Google performs exceptionally well in the correctness of translation (non-verbal agreement); however, DeepL showed especially accomplishments on three aspects: verb valency, non-verbal agreement and composition. For example, in the aspect verb valency, Google only

obtained 57,4% and DeepL, 91,5%. Lucy did not reach, in contrast, the level of Google and DeepL.

For the qualitative comparison, Gala provided five examples for a translation from German to English in different categories:

1. LEXICAL AMBIGUITY

Source:	Er hat einen <u>Kater</u> , weil er sehr tierlieb ist.	
Google:	He has a <u>cat</u> because he is very fond of animals.	Right
DeepL:	He has a <u>hangover</u> because he loves animal.	False
Lucy:	He has <u>tomcat</u> because it is very animal loving.	Right

As a native German speaker, it is impressive to see that DeepL understood the ambiguous noun “Kater” as a hangover. In Germany, “Kater” has two meanings, a male cat and a hangover; however, the meaning is wrong in this sentence.

2. INTERNAL POSSESSOR

Source:	Die Mutter hat <u>sie am Kopf</u> gestreichelt.	
Google:	The mother stroked <u>her head</u> .	Right
DeepL:	The mother stroked <u>her head</u> .	Right
Lucy:	The mother stroked <u>it at the head</u> .	False

A human would translate the sentence as “The mother stroked her at the head”. In German, the possession is separated into external and internal (non-verbal agreement). However, the English language does not differentiate between external and internal possession syntactically. Therefore, none of the translations systems translated the sentence correctly, but Lucy’s translation is incorrect.

3. MODAL PARTICLE

Source:	Bist du <u>etwa</u> verheiratet?	
Google:	Are you married?	Right
DeepL:	Are you married or what?	Right
Lucy:	Are you <u>for instance</u> married?	False

Modal particles are words that are used colloquially. It occurs in German but not in English; hence, the translation is more complicated. German use the modal particle “etwa” in questions when the questioner expects a negated response. A simple version to translate a modal particle is to omit it like Google, which is correct. DeepL applied the colloquial “or what” to reflect the modal particle and is, therefore, the best version. Although Lucy’s translation is correct, the meaning is, in contrast, wrong.

4. COLLOCATION

Source:	Vor dem Essen <u>decken</u> wir <u>den Tisch</u> .	
Google:	Before the meal we <u>cover the table</u> .	False
DeepL:	We’ll <u>set the table</u> before dinner.	Right
Lucy:	Before the food we <u>lay the table</u> .	Right

A collocation is a common word combination, e.g., blond hair. In this example, translation systems should recognise the collocation “den Tisch decken”. DeepL performed very well and recognised the collocation “to set the table”. Even Lucy translated the sentence correctly since the verb “decken” may have a different meaning. However, Google Translation did not understand the collocation and mistranslated it.

5. DITRANSITIVE FUTURE I

Source:	Du wirst der Frau das Buch geben.	
Google:	You will give the book to the woman.	Right
DeepL:	You'll give the woman the book.	Right
Lucy:	You will give the woman the book.	Right

Ditransitive verbs can take two objects, indirect (indirectly affected by the verb) and direct. In this case, all the MT systems correctly translated the sentence in their way. In contrast to Google and Lucy, DeepL's translation is better qualitative.

With the extensive analysis and comparison, the team from Gala showed and confirmed that the DeepL system is better in translation for German-English and beat the Google Translation system. Moreover, DeepL displays a fluently and naturally output (Macketanz et al., 2021).

The examples show that the quality of machine translations varies from quite reasonable to unusable. It depends, among other things, on the language pair, the subject area, the level of difficulty and the available data set and its quality. For non-native translators, the programmes can be a valuable help. Regardless, for language pairs with a pervasive data set, the machine translation reaches about the quality of a mediocre human translator. Furthermore, the programmes do not know that they are translating and do not recognise errors in the original text. However, we can see an enormous development in MT and there is still potential to reach the same level as a human. Nevertheless, the development also has a downside because human translations are no longer needed, and jobs are at stake. Nonetheless, the advantages of the MT are vast, and as a computer scientist, it is a milestone in AI technology.

References:

Macketanz, V., Burchardt, A. & Uszkoreit, H. (2021) Tq-Autotest: Novel analytical quality measure confirms that DeepL is better than Google Translate. *Globalization and Localization Association*. Available from: https://www-live.dfki.de/fileadmin/user_upload/import/10174_TQ-AutoTest_Novel_analytical_quality_measure_confirms_that_DeepL_is_better_than_Google_Translate.pdf [Accessed 25 March 2022].

Duncan L. (2018) Could Google Translate Shakespeare?. *In Other Words, 2019* (52): 79-98. Available from: https://ueaeprints.uea.ac.uk/id/eprint/68373/1/Accepted_manuscript.pdf [Accessed 25 March 2022].