In international and highly-multilingual environments, it often happens, that a talk, a document, or any other input, needs to be translated into a massive number of other languages. However, it is not always an option to have a distinct system for each possible language pair due to the fact that training and operating such kind of translation systems is computationally demanding.

Combining multiple target languages into one translation model usually causes a decrease in quality of output for each its translation direction. In this thesis, we experiment with combinations of target languages to see, if a specific grouping of them can lead to better results than just randomly selecting target languages.

We build upon a recent research on training a multilingual Transformer model without any change to its architecture: adding a target language tag to the source sentence.

We trained a large number of bilingual and multilingual Transformer models and evaluated them on multiple test sets from different domains. We found that in most of the cases grouping related target languages into one model caused a better performance compared to models with randomly selected languages. However, we also found that a domain of the test set, as well as domains of data sampled into the training set, usually have a more significant effect on improving or deterioration of multilingual model's translation quality compared to the bilingual one.