

# Multilingual Translator

**A training method for multilingual neural machine translation systems that is efficiently extendable to new languages and data modalities.**

## The Challenge

Training neural machine translation systems is challenging both in terms of data and computational resources. These factors become more critical in the multilingual setting, where several languages may be input or generated by the system. The most widespread strategy consists of training a single sequence-to-sequence system shared between all languages. This architecture allows knowledge transfer but it forces a dependency between languages. This dependency limits the ability of the system to efficiently extend to new languages or modalities. The whole system has to be retrained using data for all languages which implies variations in the overall performance.

## The Technology

This technology based on the sequence-to-sequence architecture for neural machine translation defines how to efficiently train a multilingual system extendable to new languages and modalities while allowing knowledge transfer. The process consists of two main steps. Firstly, **joint training** of languages specific encoder and decoders to a common language representation without parameter sharing. Secondly, **incremental training** of new languages and modalities to the system, including an additional module to mitigate the differences between speech and text representations.

## Innovative aspects and advantages

Multilingual neural machine translation system that is able to:

- Converge to a common language representation without sharing parameters between languages.
- Extend to new languages by a fraction of the cost of previous methods.
- Extend to spoken language even to language pairs without specific training (i.e zero-shot translation).

## Current stage of development (Development stage or status)

The technology has been explored and verified by a proof of concept system with positive results on text and speech translation. Partners to further develop the system and/or to establish commercial agreements along with technical cooperation are sought.

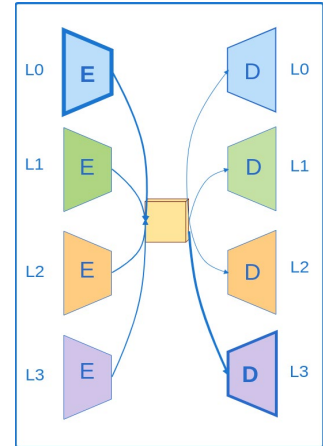
## Applications and Target Market

This method could be applied to any machine translation solution available to increase the efficiency of their systems to extend to new languages or input modalities.

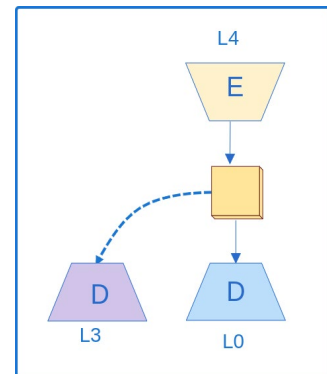
## Reference Number

MKT2021/0177\_I

## Joint training of several languages



**Incremental training of new languages. Allowing zero-shot translation to all languages in the system**



## Commercial Opportunity

Licensing opportunity with technical cooperation

## Patent Status

Priority application  
US17229,657

## Contact

Mr. Gerard Margalef  
Licensing Manager  
T. + 34 93 413 76 27  
gerard.margalef @upc.edu

## See more technologies at

[www.upc.edu/patents](http://www.upc.edu/patents)  
UPC—BarcelonaTech