**Hokamp 2017 – Lexically Constrained Decoding for Sequence Generation using Grid Beam Search**

* Introduction of Grid Beam Search (GBS)🡪 basic idea: extend the beam search to allow the inclusion of pre-specified lexical constraints.
  + **Lexical constraints:** phrases or words, which **must be present** in the generating output sequence.
  + **Idea:** general way to incorporate additional knowledge into a models output without requiring any modification of the model parameters
  + **Field tests**  with ‘Neural interactive-predictive translation = user picks part of the hypothesis which is incorrect and provides the correct translation for that portion of the output 🡪 correction = constraint for the next coding cycle. Can be repeated as many times as somebody wants to.
* They can show, that it is more efficient and they can reach better results

**Post and Vilar, 2018 – Fast lexically constrained decoding with dynamic beam allocation for neural machine translation**

* Present an algorithm for lexically constrained decoding with a complexity of *O*(1) in the number of constraints 🡪 which is more efficient compared to the usual linear of exponential complexity with respect to the number of constraints.

**Neural Machine Translation of Rare Words with Subword Units**

**Rico Sennrich and Barry Haddow and Alexandra Birch**

* Addresses the problem of rare word translation within a neural network
* Propose the idea of subword (compounds of a word) e.g. “Abwasser|behandlungs|anlange” where subwords have more meaning as the whole word, especially consider the whole word will not be in the vocabulary.
* 🡺NMT models which operate on the level of subword units
  + Makes the process simpler
  + Better accuracy for translation of rare words compared to large-vocabulary models