After learning the basic principles of building parallel corpora, the student will focus on the Czech-English parallel corpus Czeng. The main goal of the work is to improve quality of the Czeng part created from Czech/English movie and series subtitles. Above all, it is necessary to design and implement methods for detecting wrongly aligned (or otherwise problematic) subtitle files or their parts. Impact of the cleaning methods on the corpus quality will be evaluated quantitatively.