

Part-of-speech (POS) tagging is one of the most basic and crucial tasks in Natural Language Processing (NLP). Supervised POS taggers perform well on many resource-rich languages i.e. English, French, Portuguese etc, where manually annotated data is available. However, it is impossible to use a supervised approach for the vast number of resource-poor languages. In this thesis, we apply a multilingual unsupervised method for building taggers for resource-poor languages base additionally on parallel data (Universal Tagger), that is, we use parallel data as the bridge to transfer tag information from resource-rich to resource-poor languages. On average, our tagger performs on par with the state of the art on the same test set of eight languages. However, we use less data and a less sophisticated method which also results in significant difference in speed. In an effort to further improve performance, we investigate the choice of source language. We found that English is rarely the best source language. We successfully built a model that can predict the best source language only based on monolingual data. However, even better predictions can be made if we additionally use parallel data. Finally, we show that, if multiple source languages are available, it is possible to get further improvement by incorporating them.