

## **Linguistic Resources for Natural Language Processing in Pakistan**

Dr. Sarmad Hussain

There are more than 60 languages spoken in Pakistan. Such diversity in languages makes language computing a necessary field of research in Pakistan. There has been recent spark of interest in this area and as researchers and programs are focusing more on language computing research, the first impediment they face is the lack of linguistic resources to do this work. A cohesive approach to develop these resources is required to effectively propagate this research interest in natural language processing in Pakistan. This paper focuses on identifying the variety of linguistic resources required for Pakistani languages. The paper also introduces some of the details of these resources, including lexica, corpora and grammars, and the challenges in developing them. The paper will finally list some of the current initiatives in this area.

---