

Alfonso Medina-Urrea

Affix Discovery based on Entropy and Economy Measurements

This paper briefly describes an entropy and economy-based word segmentation method. Also, results of its application to discover items belonging to affix subsystems of two unrelated American languages are presented; namely, a variant of Ralámuli or Tarahumara (Uto-Aztecan) and one of Chuj (Mayan). More importantly, an attempt is made to compare these experiments in order to evaluate this approach by means of precision and recall measurements.