

ALBANIAN-ENGLISH WORD REORDERING FOR STATISTICAL MACHINE TRANSLATION

Vehbi Neziri¹, Ali Caka², Ramadan Dervishi³

¹Faculty of Electrical and Computer Engineering, University of Prishtina
E.mail;vehbineziri@gmail.com

²Faculty of Philology, University of Prishtina
E.mail: ali.caka@gmail.com

³Kolegji Riinvest, Prishtinë, Republika E Kosovës
E.mail:dervishi@gmail.com

Abstract

Sequence or word order in the sentence, one of the important remedies of expression of syntactic relations, has a special role in the English language which is analytic language. In Albanian language, which by the structure belongs to the synthetic language group, syntactic relations within the sentence are expressed with inflection and word order in the sentence is free despite the English language in which the word order in the sentence is mostly fixed. Based on these facts and the need of translating from one language to another one and vice versa, and with all the information exchange and accelerated dissemination of it in all fields of knowledge, the need for use of additional software translation tools for faster, easier and at the same time accurate translation day by day is increasing. Today many software companies and specialized institutions are dealing with the issue of translation from one language into another, trying to do it so possibly more accurately. Thanks to the internet and search engines some of these tools are free; such is the Google translation service which currently provides service to 64 different world's languages, Indo-European and other. How it translates, and what are the requirements that must be met for a more accurate translation in Google and can be intervene in this process for a better result is the purpose of this paper. The paper aims to analyze the word order in the sentence and to suggest appropriate pattern that will give the best result of the translation.

Keywords: *Word order, statistical machine translation, Albanian language, English language, text-corpus, glossary*