LABURPENA / ABSTRACT

Euskararako murriztapen-gramatika: lehen urratsak

Txosten honen helburua bikoitza da: batetik, “Constraint Grammar” (murriztapen-gramatika) izeneko formalismoa deskribatu; bestetik, euskararen sintaxiaren tratamenduan erabiltzeko egin duguna azaldu, aurkitu ditugun arazoak-eta komentatuz.

Euskararen tratamendu sintaktiko automatikoari ekiteko, eredu hau hartu dugu oinarritzat. Formalismo honen arabera, tratamendu sintaktikoan lau pauso nagusi bereizten dira: analisi morfologikoa, desanbiguazio morfologikoa, esaldien arteko mugen bereizkuntza, eta azaleko funtzio sintaktikoen asignazioa. Analisi morfologikoan oinarritua izateak bereziki aproposa egiten du euskararen kasurako. Parserraren ezaugarrien artean aipatzekoa da benetako testuak analizatzeko helburua izatea. Formalismoa egoki bilatu dugu euskararen tratamendu sintaktiko orokorrean erabiltzeko. Beste formalismo batzuek ezartzen duten zurruntasunetik urrun, eta arazoak arazo, perpaus barneko elementuen ordena librea-edo duten hizkuntzentzat egokia dela uste dugu.

Aurkitutako arazoak aipatzen hasita, batik bat, analisi sintaktikoaren emaitzaren "azaltasuna" eta horrek ekar ditzakeen ondorioak, eta hitz barruko elipsiaren arazoa komentatzen dira txostenean.

A first approach to the Constraint Grammar of Basque

The objective of this report is twofold: first, the description of the Constraint Grammar formalism (CG) and second, the presentation of the work done on the treatment of Basque syntax, commenting the problems we have found.

We have taken the CG model for the automatic treatment of Basque. According to this formalism, there are four major steps in the syntactic treatment of texts: morphological analysis, morphological disambiguation, determination of clause boundaries, and the assignment of syntactic functions. The fact that it is based on morphological analysis makes this formalism adequate for Basque. Among the features of the parser is the objective of analyzing real texts. We have found the formalism convenient for the automatic treatment of Basque. Far from the rigidity imposed by other formalisms, and despite some problems, we think that it is satisfactory for languages with some degree of free order of sentence elements.

As we have mentioned, some problems are discussed in this report. Firstly, the "flatness" of the result of syntactic analysis, with its consequences, and the problem of intraword ellipsis.