

The annotation of causative constructions in a dialectal corpus

SANDRA PEREIRA
CLUL

CORDIAL-SIN is a corpus designed to study aspects of European Portuguese (EP) dialectal syntax. It uses automated tools (POS tagger and a syntactic parser) and, in this sense, it is a specialized resource that allows (i) a fast access to systematically organized and structured data and its geographical distribution and (ii) an automatic searching of precise morphological and syntactic information. In this poster, it will be shown how causative constructions are annotated and how to find them.

In standard EP, Gonçalves (1999) describes three types of causative constructions: (i) construction of inflected infinitive (1); (ii) construction of Exceptional Case Marking (2) and (iii) Faire-INF (3).

- (1) A mãe fez os filhos comerem a sopa.
The mother made the_{MASC.PL} kids_{NOM} to-eat_{3PL} the soup
- (2) A mãe fez os filhos comer a sopa.
The mother made the_{MASC.PL} kids_{ACC} to-eat the soup
- (3) A mãe fez comer a sopa aos filhos.
The mother made to-eat the soup to-the kids

‘The mother made the kids eat the soup.’

What is interesting about these constructions is the fact that the causee can assume different forms and positions: while in (1) the causee is a nominative NP which is the subject of the infinitival clause, in (2) it is an accusative NP and the subject of the infinitival clause; in (3) the causee follows the verbal complex and can be a PP (or a dative pronoun).

Due both to the POS tags and to a rich syntactic annotation system (cf. Carrilho 2010; Magro 2010), causative constructions are codified in a way that makes it easy to retrieve them. The annotation of this particular subject inside the infinitival domain and its co-indexation with the matrix domain allows automatic searches on these constructions, making possible to identify and describe non-standard syntactic variants, associated with geographic information (Pereira 2012).