

## The relation between syntactic annotation and information structure annotation

In this paper, I present a general problematic issue with multi-layered corpus analysis, namely how our analysis of ambiguous syntactic structures will influence the results we get from an IS analysis. I exemplify with Old Portuguese causative constructions and similar sentences with verbs of permission and perception.

In sentences like a), *os alunos* may be analysed either as the direct or indirect object of the main clause or as the subject of an accusative with infinitive. In b), the infinitive of the subordinate clause is inflected according to person (3<sup>rd</sup> pl) and *os alunos* is clearly the subject of the infinitive, while the inflection in c) is ambiguous since the 3<sup>rd</sup> singular shares formal likeness with the non-inflected form.

- a) *mandei* (a)os *alunos* *fazer* o *bolo*  
ask 1<sup>st</sup> sg (to) the pupils make INF the cake  
I asked the pupils to make the cake
- b) *mandei* os *alunos* *fazerem* o *bolo*.  
ask 1<sup>st</sup> sg the pupils make INF3<sup>rd</sup>pl. the cake  
I asked the pupils to make the cake
- c) *mandei* o *aluno* *fazer* o *bolo*  
ask 1<sup>st</sup> sg the pupil make INF3<sup>rd</sup>sg./INF the cake  
I asked the pupil to make the cake

Our choice of analysis for ambiguous syntactic structures like the ones in question will have an impact on the search results for information structure. Should we annotate *os alunos/o aluno* as the object of a main clause or the subject of a subordinate clause? Our choice will affect the universal correlation between old information and subjects and new information and objects. A simple solution would be to classify these constructions separately, as ambiguous, and thus remove possible noise from the statistics. This would increase the retrievability of these particular constructions, but it would make it harder to retrieve data for instance on subordinate or main clauses in general, which is relevant to our study on word order and information structure.