Dialect syntax through corpora: the subject doubling *se* construction

Ana Maria Martins
Universidade de Lisboa (FLUL/CLUL)

This talk will discuss novel syntactic data made available by the existence of the CORDIAL-SIN: *Syntax-oriented Corpus of Portuguese Dialects* (http://www.clul.ul.pt) and consider their empirical and theoretical relevance for Romance and general syntax. It will also illustrate how and why a corpus of spoken language can be an optimal tool to study syntactic variation.

European Portuguese (EP) dialects display a double subject construction where the impersonal clitic *se* shares with a strong pronoun, a null pronoun, or less commonly a full DP, the role of sentential subject. This construction, exemplified in (1)-(2) below, is fully ungrammatical in standard EP and has gone totally unobserved by philologists and linguists who dealt with dialect variation in European Portuguese. The corpus CORDIAL-SIN brought it into light.

(1) a. ***A gente*** *chama-**se** rãs a isto.* (CORDIAL-SIN. FLF)
pron.1PL calls-***SE*** *rãs* to that
'We call these ones *rãs* (frogs).'

b. *Chama-**se**-lhe* ***a gente*** *espigas.* (CORDIAL-SIN. AAL)
calls-***SE***-it.DAT **pron.1PL** *espigas*
'We call it *espigas* (spikes of cereal).'

(2) a. *Há várias qualidades que até ainda **nós** não **se***
there.is several species that even already **we** not ***SE***
*conhece**mos**.* (CORDIAL-SIN. ALV)
know-**1PL**
'There are so many species among fish that even we (fishermen) do not know all of them yet.'

b. *Nunca **se** vi**mos** este peixe.* (CORDIAL-SIN. CLC)
not ***SE*** saw-**1PL** this fish
'We never saw that kind of fish.'

The talk will describe in detail the subject doubling *se* construction, as it surfaces in the corpus, and discuss the relevance of the geolinguistic information offered by CORDIAL-SIN to make full sense of the particular features of the dialectal construction. It will finally be proposed that a nominal that lacks a person feature may also be devoid of Case. This is how the clitic *se* of the relevant EP dialects is like: it has number (plural), but no person or Case. Its deficient nature is what makes subject doubling a grammatical option in dialectal EP. In a few dialects, the expletive-like *se* drifted in the direction of a pure expletive:

(3) a. *Às vezes acontece; já **se** aconteceu.* (CORDIAL-SIN. PIC)
sometimes happens already ***SE*** happened.3SG
'Sometimes it happens; it has already happened.'

    b. Nunca ***se***   me   aconteceu   isso. (CORDIAL-SIN. CRV)

       never   ***SE***   me.DAT happened.3SG that

       'It never happened to me a thing like that.'

Under the proposed analysis, the subject doubling *se* construction of European Portuguese dialects brings new empirical evidence that supports the contentions (i) to (iii) below.

(i)   "If all other conditions are right, a given argument can be linked to, and coherently composed with, more than one expression". (Chung and Ladusaw 2004, on the *extra object* in Chamorro).

(ii)   The impersonal *se/si* of standard EP, Spanish and Italian is not deficient with respect to φ-features and Case (cf. Cinque 1988), but a full referential expression (D'Alessandro 2004). Therefore, subject doubling with impersonal *se/si* is not an option in standard EP, Spanish and Italian.

(iii) Overt pure expletives are not necessarily blocked when in competition with null pure expletives in a particular grammar (cf. Chomsky's 1981 *Avoid Pronoun Principle*).

References

Chung, Sandra & William A. Ladusaw. 2004. *Restriction and Saturation*. Cambridge, Mass.: MIT Press.

Cinque, Guglielmo. 1988. "On *si* constructions and the theory of *arb*". *Linguistic Inquiry* 19.521-582.

Chomsky, Noam. 1981. *Lectures on Government and Binding. The Pisa Lectures*. Dordrecht: Foris.

D'Alessandro, Roberta. 2004. *Impersonal si constructions. Agreement and interpretation*. Ph.D. dissertation. University of Stuttgart [Electronic Dissertation Series, Stuttgart University Library], http://elib.uni-stuttgart.de/opus/volltexte/2004/1630