

Differential Object Marking and discourse prominence in Spanish: a corpus-based study

Tiago Augusto Duarte
(University of Cologne)

In Spanish, differential object marking (DOM) is characterized by the morphological marker *a* ‘to’ (García García 2018), which may precede human-indefinite direct objects. The literature has proposed several possible hypotheses to explain this phenomenon (see Fábregas 2013). One of them claims that DOM is triggered by topicality (Leonetti 2004). Following this line of research, we hypothesize that the presence of DOM increases the discourse prominence of the direct object (DO). Hence, it increases the probability that an *a*-marked DO will be re-used in the following discourse (cf. Chiriacescu & von Heusinger 2010). Moreover, we assume that in a transitive construction with a human subject and a human-indefinite direct object, as in (1), DOs with DOM trigger more rementions than DO without DOM.

- (1) Juan vio (*a*) un muchacho.
‘Juan saw (DOM) a boy.’

In order to test this hypothesis, we conducted a corpus-based search on the *Corpus del Español del Siglo XXI*. The study focuses on written Peninsular Spanish and analyzes sentences, such as (1), constructed using the verbs *ver* ‘to see’, *elegir* ‘to choose’, *enviar* ‘to send’, *mandar* ‘to send’, *traer* ‘to bring’, and *presentar* ‘to present’ with the feature [\pm DOM]. In total, we analyzed 3942 sentences. However, after manually isolating all the hits with both a human subject and a human DO, we were left with 112 sentences. Then, we looked at the subsequent sentences and annotated all the cases in which referent 1 and/or referent 2, i.e., the respective subject and direct object of the preceding sentence, were rementioned.

The results of the corpus search regarding the remention of Ref1 and Ref2 show (through a numerical difference) that Ref2 (48,2%) is more rementioned than Ref1 (42,2%) when it is preceded by DOM. Moreover, when we compare the number of rementions of Ref2 according to [\pm DOM], Ref2 is more rementioned in constructions with DOM (48,2%) than without DOM (37,9%), which seems to corroborate our hypothesis.

Besides presenting the corpus-based study and all the decisions that were taken, the talk will address the problems/difficulties of using corpora to investigate discourse and the challenges of using a corpus that is not annotated for animacy.

Acknowledgment:

The research for this paper has been funded by the German Research Foundation (DFG) as part of the SFB 1252 “Prominence in Language” in the project B04 “Interaction of nominal and verbal features for Differential Object Marking” at the University of Cologne.

Corpus:

REAL ACADEMIA ESPAÑOLA: Banco de datos (CORPES XXI) [en línea]. *Corpus del Español del Siglo XXI (CORPES)*. <<http://www.rae.es>> [last access: 06.01.2022]

References:

- Chiriacescu, S. & von Stechow, K. (2010). Discourse Prominence and Pe-marking in Romanian. *International Review of Pragmatics* 2. 298–332.
- Fábregas, A. (2013). Differential Object Marking in Spanish: State of the art. *Borealis: An International Journal of Hispanic Linguistics* 2(2). 1–80.
- García García, M. (2018). Nominal and verbal parameters in the diachrony of differential object marking in Spanish. In I. A. Seržant & A. Witzlack-Makarevich (Eds.), *Diachrony of differential argument marking*, 209–242. Berlin: Language Science Press.
- Leonetti, M. (2004). Specificity and differential object marking in Spanish. *Catalan Journal of Linguistics* 3. 75–114.