

ON THE SETTING OF SCALES IN THE DIACHRONY OF DOM
Monica Alexandrina Irimia and Anna Pineda

Background. Morphosyntactic splits of objects are common cross-linguistically (Comrie 1989, Bossong 1991, Aissen 2003, López 2012, a.o.). Under differential object marking (DOM), modern Spanish *definite animates* have an obligatory adposition, as opposed to inanimates.

- (1) He encontrado ***(a) la niña** / ***(a) el libro**.
have.1.SG found DOM the girl / DOM the book
'I have found the girl / the book.' (Ormazabal and Romero 2013, ex. 1a/b)

Functional-typological research connects such patterns to hierarchical generalizations known as *Scales* (Aissen 2003, a.o.), as in 2. The implication is that the higher (>) an object is on a scale, the stronger DOM trigger it will be. It has also been claimed that, besides implicational universals, *Scales* can also account for patterns of language change (von Heusinger & al. 2008, Leonetti 2008, a.o.). Thus, Romance DOM is typically seen to have started with objects higher on the scale (e.g. 1st/2nd pronouns, proper names, etc.) and then progressively extended to those lower down (Berretta 1989, Sornicola 1997, a.o.). One prediction is that there should not be stages where what is overtly marked are lower DPs, to the exclusion of the higher ones.

- (2) *Animacy/person*: 1/2 > 3 > proper name > human > animate > inanimate

Specificity/definiteness: pronoun > name > definite > specific indefinite > non-specific

Data, methodology, and aims. Our data come from the investigation of DOM in the earliest Catalan and Romanian texts, as compared to Spanish. Given that both modern Catalan (MC) and modern Romanian (MR) show DOM distinctions from modern Spanish (MS), a relevant question is whether diachrony gives the same picture. We address first some counterexamples to the *Scales*; a second goal is to provide a more comprehensive picture of Romance DOM both diachronically and synchronically, especially given that DOM in the earliest Catalan texts has not been extensively investigated. Formally, we examine two aspects which have not been extensively discussed for (diachronic) DOM: a) the composition of Romance 3rd person pronouns (restrictions to animacy/humanness or not, following Cardinaletti and Starke 1994); b) refined parameters of nominal licensing (Levin 2010, Nevins 2011, Baker 2012, 2013, 2015, a.o.). Our corpora are: for Catalan, the *Corpus informatizat del català antic*, comprising the first written texts up to the 16thc.; for Romanian, *Documente și însemnări românești (Romanian documents and notes, DÎ) – 16th c.* (see also Mardale 2015, and for other early texts von Heusinger and Onea Gáspár 2008 - H&OG, Avram and Zafiu 2017, a.o.); for Spanish, the data from Company (2002), von Heusinger and Kaiser (2005), Laca (2006).

Discussion. 1) Pronouns. H&OG's (2008) investigation of the earliest Bible Translations in *Romanian* has noted 97% of 3rd person pronouns have DOM (33 out of 34), as opposed to 50% for 1st/2nd. In DÎ (see also Mardale 2015), 3rd person pronouns are still DOM-ed more frequently than 1st/2nd. Early *Catalan* texts also show that it does not seem to be the case that DOM first consolidated with 1st/2nd person and then extended to 3rd person - in the knightly novel *Curial* (15th c.) some 3rd person pronouns have DOM, while 1st/2nd persons lack it (3 vs 4). These patterns require an explanation, as the *Scales* predict 1st/2nd person to be stronger DOM triggers.

- (3) a. vós havets honrat **a ell**. b. ell e un companyó seu combatrían **a ell**.
you.2PL have.2PL honoured DOM he he and a companion his fight.COND.3PL DOM he
(4) a. ¿què ha mogut **tu** e ton companyó a...? b. aquella senyora, qui mira **nosaltres**
what has moved you and your companion to that lady who look.3SG we

2) Pronouns and other DPs. The earliest Catalan texts also show that the *a*-marking did not necessarily consolidate first with strong pronouns and then appeared in DPs. This contrasts with old Spanish, where personal pronouns are claimed to have been systematically DOM-ed from the very beginning, with no exceptions (Company 2002: 207, Laca 2006: 426,). Old Catalan examples (5) and (6) all correspond to one and the same text, *Clams e crims (Court*

complaints), 13th c.; some strong pronouns do not show DOM (5), but some animate DPs do. In (6) we can see a proper name (which in Catalan takes the definite article) with DOM.

- (5) *darem a aquels [l]licència de peynorar vós*
 give.FUT.1PL to them permit to fine.INF you.2PL.HONORIFIC(=SG)
- (6) *com en Ca[ste]let, saig, volgués peynorar a n Ramon Sanç,*
 since the Castelet, executioner, want.SBJV.PST.3SG fine.INF DOM the Ramon Sanç,
lo dit Ramon dix a aquel que no-l peynoràs,
 the mentioned Ramon tell.PST.3SG to that.one that no him.ACC fine.SBJV.PST.3SG

Analysis. I. Lack of accusative morphology. H&GO (2008) tentatively attribute the obligatoriness of DOM (a means of signaling ACC) with 3rd person to NOM/ACC homomorphism (as opposed to 1st/2nd where there is no NOM/ACC homomorphism in MR). However, a closer examination of the Romanian data shows that in the 16th c., 3rd person pronouns could also exhibit distinct ACC morphology but still required DOM. (7) illustrates the form *sine* (he.ACC), as opposed to *elu* (he.NOM). Catalan 2nd and 3rd pronouns also show NOM/ACC homomorphism. (7) ... *ocină... ca să poată hrăni pre sine și oamenii lui.* (DÎ 1593)
 ... land... that SBJV can.SBJV feed.INF DOM he.ACC and men.the he.GEN

II. Discourse function. DOM has also been linked to discourse needs (as opposed to grammatical function), as foregrounding, familiarity, discourse-linking, or secondary topic (H&GO 2008, Darlymple & Nikolaeva 2011, Iemmolo 2010, Mardale 2015, Hill & Mardale 2017, a.o.). In old Catalan DOM this type of topicality seems to be playing a role “*the mentioned XX*” (8); however, we also find examples that are equally “topical” but do not show DOM (9):

- (8) *lo dit Castelet, saig, prè al dit Ramon Sans.*
 the mentioned Castelet, executioner, take.PST.3SG DOM.the mentioned Ramon Sans
- (9) *lo dit Castelet dix que peynorava lo dit Ramon.*
 the mentioned Castelet said that fine.IPFV.3SG the mentioned Ramon

Under topicalization, it is equally not easy to i) derive the pronominal split we noticed above; ii) more generally, test the secondary topic/foregrounding nature of DOM in diachronic corpora.

III. Distinct (Case) licensing strategies. We explore a connection of the patterns above to more than one object licensing strategy being active in early Romance: a) a mechanics (more similar to Latin) with a ϕ -probe licenser, which checks ACC (irrespective of animacy/specificity). This derives object participle agreement with no DOM, as in old Catalan (10, *Matrimoni* 14th c.):

- (10) a) *ell ... hac menjada la dita vianda* b) *n'Arnau ... hac aviada sa muller...*
 he had.3SG eaten.F.SG the said meat.F.SG the Arnau.M had.3SG sent.F.SG his wife

b) a novel strategy, probably the effect of enrichment of functional projections in the DP, where the licenser is endowed with a [PERSON] probe and sensitivity to animacy (and definiteness/specificity), resulting in DOM. Following Richards (2008), Adger & Harbour (2010), a.o., we take nominals containing *grammaticalized animacy* to have an added [PERSON] layer which requires *licensing* in narrow syntax (see also Zubizarreta & Pancheva 2017 for licensing of interpretable features). Another parameter is the featural composition of 3rd person pronouns (Cardinaletti & Starke 1994/1998) - whether they are restricted to animacy/humanness or not (cf. French *Il est beau* ‘He/it is nice’). In early Catalan texts, as the 3rd p. form started to get mapped to animacy only (cf. *Ell és bonic* ‘He/*It is nice’), DOM became more prominent with it. We interpret this as the result of *animacy starting to grammaticalize on pronouns*, entailing special licensing (as the other licensing strategy, not based on [PERSON], could not apply). With 1st/2nd persons animacy appears to have grammaticalized later, and the other strategy could still apply during the transition. In Spanish, on the other hand, a yet distinct nominal licensing strategy, namely (*pseudo*)*incorporation*, seems to be prominent with definite objects earlier than Catalan/Romanian (e.g. *non-specific* readings of definites are more easily available). This explains why Spanish specific pronouns tend to show DOM earlier and more robustly. As [PERSON] must escape (*pseudo*)*incorporation*, the only strategy left is DOM.