

### **On Differential Object Marking (DOM) in French: a corpus-based analysis**

Among Romance languages, only Spanish and Romanian are known to have systematic DOM (with contexts of obligatory, optional or excluded marking). Other languages or dialects (such as Portuguese, Galician, Sardinian, Corsican, Engadinese) also display DOM, but less systematically. Unlike these varieties, French is considered to be ‘DOM-free’, which may explain why there is no specific study of this phenomenon in French – it is merely mentioned in passing by a few authors (Hills 1920, Rohlf 1971, Roegiest 1979, Lois 1982, Charolles & Lamiroy 2011, Iemmolo 2011). Constructions such as (1) are generally considered as mistakes due to the influence of neighboring Romance languages, like Spanish (i.a. Blanche-Benveniste 1997):

(1) *Voyez l'hypocrisie: on **nous** empêche à **nous** de produire des maïs OGM, par contre les commerciaux sont autorisés à rendre le maïs OGM.*

See the hypocrisy: they keep [DOM] us from producing GMO corn, but sales people are authorized to sell GMO corn (French National Radio “France Info”, Feb. 22<sup>nd</sup>, 2012, a speaker from “Haute Garonne”, in the South of France)

Our goal is (i) to check whether French DOM is indeed a phenomenon resulting from (Romance) language contact, and (ii) to examine the contexts that can trigger its occurrence. In order to do this, we propose to go from what we already know for the other Romance languages, in which DOM has been extensively studied, including in a diachronic perspective (i.a. Bossong 1985, von Heusinger & Kaiser 2005, 2007, 2009, Iemmolo 2010, Stark 2011). These studies have shown that there are several competing factors at play (Laca 2002, 2006, Aissen 2003, de Swart 2008). Some of these factors are object-internal: its categorial status (Pronoun, Proper N, DP, etc.), animacy, definiteness and specificity. Other factors are (object-)external, and linked more specifically to the context: topicalization, modification, the lexical nature/type of the verb, coordination.

One of the problems we encountered in the study of this phenomenon is the fact that it is extremely rare, with, besides, a clear regional bias. Thus, although our corpus is composed of 7 different databases<sup>1</sup> (amounting to more than 13 *billion* words), we have gathered little more than a hundred clear occurrences of DOM. One first conclusion is that these occurrences do not seem characteristic of the South-West of France, as is sometimes claimed, but also found in other French varieties:

(2) a. *j'aidais à **mon**, à **papa** pendant que mon frère était parti* (PFC database) (Normandy)

I helped [DOM] my dad while my brother was gone

b. *j'suis même zoophile, la preuve j'**te** nique à **toi*** (Bielefeld) (Clapi database) (Alsace)

I'm even a zoophile, look, I fuck [DOM] you

(3) *comme tout bon enfant, on fait chier à **son** baby-sitter.* (PFC) (Switzerland)

like any regular child you piss off [DOM] your baby-sitter.

(4) *Et vous? Qu'est-ce qui **vous** empêche à **vous** et à **vos chums cyclistes** d'entretenir votre bout de piste?* (Internet, frTenTen) (Québec)

How about you? What is it that keeps [DOM] you and [DOM] your bike-rider chums from keeping your road-side clean?

(5) *il faut aller voir à **mes** parents* (PFC) (French-speaking Africa)

I've got to go see [DOM] my parents

Our claim is that – without excluding a possible influence of Spanish for the South-West French dialects – French DOM is not only the result of language contact (since it occurs in other dialects that are not in contact with DOM-languages (2) - (5)), but a recent language-

---

<sup>1</sup> Including six different databases of Spoken Modern French (CFPP2000, Clapi, Debaisieux, PFC, CFPQ and Valibel) and a large database of ‘Web’ Written Modern French (frTenTen (SketchEngine)).

internal evolution phenomenon. Crucially, its contexts of appearance are similar – but not identical – to those of other Romance languages.

More specifically, our study brings to light the following facts:

**A)** unlike Spanish and Romanian, DOM in French is always an optional phenomenon (whatever the category it affects) and appears essentially for pragmatic reasons (such as contrast, emphasis, etc.). This is why it is generally associated to topicality and to (right- or left-)dislocation (see Lois 1982, Pensado 1984, Iemmolo 2010, 2011):

(6) a. *à moi, il ne m'attrapera pas* (Iemmolo 2011: 257)

[DOM] me, he won't catch me. (lit. "[DOM] me, he NEG me catch.FUT.1SG NEG")

b. *Il le va blesser, à cet enfant* (Rohlf's 1971: 68, apud Immolo *id.*)

He is going to injure him, [DOM] this child.

**B)** the object affected by DOM is (almost) always clitic-doubled (73 occurrences out of 90, cf. (4), (6), (7)) and quite frequently associated with a modifier such as *-même* "itself", *aussi* "also" (13 occurrences out of 90, cf. (7)):

(7) *Oui, je le déteste à lui aussi, mais je me répugne en même temps.*

Yes, I hate [DOM] him too, but at the same time I find myself repulsive.

**C)** unlike Spanish and Romanian where the inherent properties of the DO can, in some cases, be not only necessary but also sufficient, in French these properties are never sufficient, and must always be reinforced by O-external properties. As we saw in A) and B) above, factors such as topicality and modification are particularly relevant for French, as are also the verb's semantics: only V's with a highly-affected object (Hopper & Thompson 1980, von Heusinger & Kaiser 2005) such as *kill, hurt, adore* seem capable of triggering DOM in French (8):

(8) *Non, il ne faut pas les tuer car sinon ça nous tuer[a] à nous les hommes.* (FrTenTen)

No, we should not kill them, otherwise this will kill [DOM] us human beings.

**D)** as for which objects can be marked, we propose the following hierarchy:

**Pronouns** referring to humans > **Specific Definite DPs** referring to roles/figures which are unique in the "speaker sphere" (i.e. parents, brother/s, boss, God, priest) > **Specific Definite DPs** referring to humans, with a strong D (i.e. a possessive or a demonstrative) > **Proper Ns** referring to humans

One significant difference with respect to other Romance languages is the extreme rarity of examples with proper nouns referring to humans (only 5 occurrences out of 90):

(9) *Jésus aimait à Jean le pêcheur de Galilée* (FrTenTen)

Jesus loved [DOM] John, the fisherman from Galilea

**E)** finally, a diachronic consideration is in order: the contexts of appearance of DOM in French are similar to what is reconstructed for its initial phases in other Romance languages (cf. D) above). In other words, we are dealing with a not yet grammaticalized phenomenon, which emerges from the highest positions in Aissen's 2003 hierarchy (i.e. Pronouns and DPs with a human unique referent). It seems to extend (sporadically) to lower positions, starting with strong specific human definite DPs and human Proper Nouns (cf. von Heusinger & Onea 2008 for a similar hypothesis on the development of DOM in Romanian). Besides, we should add that the development of DOM in French is also influenced by the verb's syntax, since many of the occurrences we found include verbs with double object constructions, i.e. an IO in the dative (10a) and a DO in the accusative (10b) (see also Pensado 1984, Laca 2006 for Spanish; Müller 1979, Sornicola 1997, Parry 2003 for late Latin):

(10) a. *il a bien satisfait à ses obligations fiscales et sociales* (IO, Dative)

he has satisfied fully his fiscal and social obligations

b. *peut-être qu'il n'envie pas ton style de vie, mais se satisfait à lui-même* (DO, Accusative)

maybe he doesn't envy your lifestyle, but (lit.) satisfies [DOM] himself (i.e. with his own)