Differential Object Marking and discourse prominence in Spanish

Klaus von Heusinger
Universität zu Köln
klaus.vonheusinger@uni-koeln.de

Tiago A. Duarte
Universität zu Köln
tiago.duarte@uni-koeln.de

Marco García García
Universität zu Köln
marco.garcia@uni-koeln.de

How to cite: von Heusinger, Klaus, Duarte, Tiago A. & Marco García García. 2024. Differential Object Marking and discourse prominence in Spanish. *Isogloss. Open Journal of Romance Linguistics* 10(1)/6, 1-37.
DOI: https://doi.org/10.5565/rev/isogloss.394

Abstract

Spanish, like many other Romance and non-Romance languages, shows Differential Object Marking (DOM), i.e., some direct objects are morphologically marked by the prepositional marker *a* 'to', while others remain unmarked. The literature has proposed different sentential parameters in order to capture this variation (Fábregas 2013, among others), including topicality (see Leonetti 2004, Iemmolo 2010, among others). In addition, Laca (1995: 82f.) has argued that DOM also depends on discourse

properties. She assumes that in Spanish the use of DOM with an indefinite direct object signals that more information about this object referent is to be expected in the upcoming discourse (see also Comrie 1981/1989). First empirical evidence for this hypothesis comes from DOM in Romanian (Chiriacescu & von Heusinger 2010). In this paper we explore the hypothesis that, in Spanish, human indefinite direct objects with DOM show more forward-looking potential than those without DOM. We present original results from two corpus studies and two paragraph continuation tasks. The corpus studies provide support for the discourse effect of DOM, while the paragraph continuation tasks do not, which might be due to the particular design of our experimental items. We evaluate the different parameters that contribute to the discourse prominence of a direct object with DOM and those that might mask such effects. We conclude that there is evidence that DOM contributes to discourse prominence, but that further studies are necessary.

Keywords: Differential Object Marking, DOM, Spanish, discourse prominence, corpus analysis, paragraph continuation task.

1. Introduction

Differential Object Marking (DOM) is traditionally characterized as a phenomenon in which overt morphological marking is applied only to a subset of direct objects, depending on various prominence features such as animacy and referentiality (Bossong 1982, 1998, Aissen 2003, among many others). In addition to these nominal features, further prominence features related to the verbal domain or to information structure have been shown to be relevant, including affectedness, agentivity and topicality (see Kabatek, Obrist & Wall 2021, Witzlack-Makarevich & Seržant 2018, among others). Based on these prominence features, a direct object can be characterized as more or less prominent with regard to its co-argument, i.e., the subject of a given sentence. In this paper, we extend the analysis of DOM from a sentence to a discourse perspective.

Comrie (1981:129; 1989: 136) observes that DOM also signals a certain forward-looking function: "The absence of the accusative suffix advises the hearer not to bother about identifying the referent, while presence of this suffix advises him that the referent of this noun phrase, though not yet determinable by the hearer, will be of relevance to the ensuing discourse." Nilsson (1985) provides first observations with respect to Turkish. The relevance of this discourse dimension for DOM in Spanish has previously been addressed by Laca (1995). According to Laca (1995: 83), instances of DOM in Spanish involving animate non-human direct objects are to be expected when the direct object is co-referential with a referent introduced in the preceding sentence. Furthermore, Laca emphasizes that referents repeatedly mentioned in the preceding discourse, serving as the central topic of a paragraph, are more likely to be associated with DOM than those not previously mentioned (see also Weissenrieder 1991: 152–154, Barraza 2003: 92, and the discussion in García García 2014: 47f.). Note that these discourse-based assumptions relate to the backward-looking discourse sensitivity of DOM.

In addition to this backward-looking perspective on the preceding discourse, Laca (1995: 82) also points out the connection of DOM to the subsequent discourse,

that is, to its forward-looking potential. She illustrates her assumption based on the optional case of DOM in (1), involving the animate (non-human), indefinite direct object *un tigre* 'a tiger'.

(1)   Juan mat-ó   Ø/a   un tigre.
       Juan kill-3SG.PST Ø/DOM   a tiger
       'Juan killed a tiger.'

Laca (1995: 82) states that *a*-marking of the indefinite direct object *un tigre* 'a tiger' signals that more information about the tiger in question is to be expected. More specifically, she hypothesizes that "the use of the preposition introduces an indefinite as a possible discourse topic" (Laca 1995: 82). First empirical evidence supporting this hypothesis has been provided for DOM in Romanian (see Chiriacescu & von Heusinger 2010 and Section 3.2 below).

To illustrate the aim of our investigation, compare the following two examples involving human indefinite direct objects:

(2)   La tecnología con naves robóticas se ha convertido en un factor importante de los planes de la NASA para enviar **un hombre** a la Luna en los próximos años. Uno de los objetivos de DART era poner a prueba la posibilidad de que naves robóticas lleven a cabo tareas que ahora realizan los astronautas.[1]
       'Robotic aircraft technology has become an important factor in NASA's plans to send a man to the Moon in the next few years. One of the aims of DART was to test the possibility that robotic aircraft can accomplish tasks that nowadays are performed by astronauts.'

(3)   El pasado 30 de diciembre, el Comité Electoral Nacional eligió **a una persona₁**. Ahí se cerró una fase y **la candidata designada₁** es la de todos, la del partido.[2]
       'On December 30, the National Electoral Committee elected DOM a person. At this point, a phase was closed, and the designated candidate is everyone's candidate, the party's one.'

In (2) the direct object without DOM *un hombre* 'a man' is not taken up in the subsequent sentence. In (3), however, the direct object with DOM *a una persona* 'DOM a person' is picked up by the coreferential definite noun phrase *la candidata designada* 'the designated candidate' in the subsequent sentence. These corpus examples illustrate our main hypothesis, i.e., that DOM not only reflects the prominence of the direct object with respect to the subject in the current sentence, but that it also signals discourse prominence with regard to the subsequent discourse. We follow recent research in referent management that reformulates Givón's (1983) notion of *topic continuity* in terms of accessibility (Ariel 1990, Arnold 2010), attentional focus (Gundel et al. 1993), salience (Chiarcos et al. 2011, Chiarcos 2011, Falk 2014), and

---

[1]   CORPES XXI, 6:
https://www.elmundo.es/elmundo/2006/05/16/ciencia/1147755253.html
[2]   CORPES XXI, 135:
https://www.losgenoveses.net/Personajes%20Populares/cascos/cascosvuelve.htm

prominence (Himmelmann & Primus 2015, von Heusinger & Schumacher 2019). Moreover, we will use the notion of discourse prominence, since this allows us to connect it to the well investigated local prominence relations of DOM at the sentence level. Discourse prominence can be operationalized by various parameters, including next mention bias (Kehler et al. 2008), topic shift (van Kampen 2007), referential persistence (see Chiriacescu & von Heusinger 2010; cf. "topic persistence" as put forward by Givón 1983, Fuchs & Schumacher 2020) and type of anaphoric expression (Ariel 1990, Gundel et al. 1993, Arnold 2010).

We base our investigation on Chiriacescu and von Heusinger's (2010) work on the discourse prominence of DOM in Romanian (see also von Heusinger & Chiriacescu 2011, Chiriacescu 2014, Tigău 2022, von Heusinger & Tigău 2024). They have shown that, in Romanian, direct objects with DOM are (i) more often used in the following discourse, (ii) more often topics and (iii) establish longer referential chains than direct objects without DOM. In the present paper, we focus on Spanish and formulate the hypothesis that human indefinite direct objects with DOM are not only more locally prominent, but also more discourse prominent than their unmarked counterparts. We first present results from two corpus studies which show that there is a significant effect for one parameter and marginal effects for the others of DOM on discourse prominence. Since it is very difficult to control for all the potentially relevant factors for DOM in corpus data, we further undertook two paragraph continuation tasks. In the first paragraph continuation task we provided a single-sentence context and asked for one continuation sentence, as in (4), where the underline is a place holder for the continuation elicited. In the second paragraph continuation task, we provided three-sentence contexts, as in (5), that resemble the corpus examples given in (2) and (3). We assumed that these larger contexts are more natural and allow better modelling of the interaction between two or more discourse referents. We therefore also manipulated the number of competitors for the direct object position, providing contexts with one appropriate human competitor for paragraph continuation, and contexts with more than one competitor, as in (5), where the direct object *pasajero* 'passenger' is in competition with *Roberto* and *uno de los oficiales del puente* 'one of the bridge officers'.

(4)   **Ana₁** oyó **(a) un bebé₂** en el piso de al lado. _____
        'Ana heard (DOM) a baby in the apartment next door.'

(5)   **Roberto₁,** capitán del crucero La gaviota, se encontraba en el puente de mando supervisando a su tripulación. Se acercaba una gran tormenta que podía poner en peligro la nave y ordenó **a uno de los oficiales del puente₃** que contactase con el servicio portuario. Después, decidió bajar a cubierta y al llegar, vio **(a) un pasajero₂**. _____
        'Roberto, captain of the cruise ship La Gaviota, was on the command bridge supervising his crew. A great storm was approaching that could endanger the ship and he ordered one of the bridge officers to contact the port service. Later, he decided to go down on deck and when he arrived, he saw (DOM) a passenger.'

For all of the studies presented, that is, the corpus searches and the paragraph continuation tasks, we annotated the subsequent sentence for the referents that were

anaphorically used, their grammatical function and their nominal form. We decided to focus on two of the different measurements for discourse prominence: (i) referential persistence (the number of anaphoric uptakes) and (ii) topic shift (measured in the shift from direct object to subject). Furthermore, we added a new measurement, (iii) the *discourse prominence advantage,* a relational measurement that is calculated as the relation between the forward-looking function of the direct object and the forward-looking function of the subject (see Section 3.3). Our predictions are that human indefinite direct objects with DOM show a higher rate of referential persistence and more topic shifts, and that they are stronger competitors with respect to the subject.

In Section 2, we provide a brief overview of the main parameters for local prominence that determine DOM in Spanish. In Section 3, we expand the view to discourse prominence and show that, in Romanian, DOM also depends on global parameters that establish a higher level of discourse prominence. In Section 3.3. we introduce and motivate *discourse prominence advantage* as a new relational measurement, which mirrors the assumption of prominence as a relational property between two or more units. In Section 4, we present the results from the two corpus studies, which support our claim with regard to Spanish. Section 5 is dedicated to the two paragraph continuation tasks. The general discussion in Section 6 compares the results from the corpus studies and from the paragraph continuation tasks and reviews these findings in the light of the Romanian data. Section 7 provides the conclusion.

2. Differential object marking in Spanish

Differential object marking (DOM) is defined as a split phenomenon in which only a subset of direct objects is morphologically marked, depending on certain semantic and/or pragmatic features of the direct object. These features include nominal parameters such as (i) animacy and (ii) referentiality, (iii) verbal parameters such as affectedness and (iv) information structural parameters such as topicality (see Comrie 1975, Bossong 1985, Aissen 2003, Iemmolo 2010, Witzlack-Makarevich & Seržant 2018, Kizilkaya 2024). In many Romance languages, DOM or *a*-marking depends on all four parameters (Gerards 2023, Mürmann 2023, García García & Caro Reina to appear). This holds particularly true for Spanish (Laca 1995, 2006, Torrego 1999, von Heusinger & Kaiser 2011, Leonetti 2004, Fábregas 2013, García García 2014, Romero Heredero 2022).

As for (i) animacy, only human direct objects can appear with DOM, as shown in (6), while inanimate direct objects must generally be realized without DOM, as in (7). However, there is a small class of verbs, such as verbs of substitution, that allow DOM for inanimate direct objects, as in (8) (see García García 2014, 2018 for an extensive discussion). In the remainder of this paper, we will not consider inanimate direct objects.

(6)   Conozc-o *Ø/a   este   actor.
        know-1SG *Ø/DOM this   actor
        'I know this actor.'
(7)   Conozc-o Ø/*a   esta   película.
        know-1SG Ø/*DOM this   film
        'I know this film.'

(8)   En esta receta   la   leche   pued-e   sustituir   *el / al   huevo.
        in   this recipe   the   milk   can-3SG   substitute   *the/DOM.the egg
        'In this recipe the milk can replace the egg.'

With regard to (ii) referentiality, specific indefinite human direct objects and all direct objects that are higher on the Referentiality Scale (9) must be *a*-marked; see (11a) and (10), respectively. Even non-specific indefinites can optionally be *a*-marked. This is illustrated in (11b), where the subjunctive of the verb in the relative clause indicates that the head noun *un ayudante* 'an assistant' is non-specific. Determinerless noun phrases (bare nouns) such as *camarero* 'waiter' in (12) must not be *a*-marked.

(9)   Referentiality Scale:
        personal pronoun > proper noun > definite NP > specific indefinite NP
        > non-specific indefinite NP > bare noun

(10)   La   vi   *(a)   ella / *(a)   María / *(a) la   mujer /
        her see.PST[1SG] *(DOM) her / *(DOM)   María / *(DOM) the woman /
        (a) una mujer.
        (DOM) a woman
        'I saw her / María / the woman / a woman.'

(11)   a. Necesit-an *Ø/a   un ayudante que sab-e   inglés.
           need-3PL   *Ø/DOM   an assistent that know-3SG.PRS.IND English
           'They need an assistant who knows English.'
        b. Necesit-an Ø/a   un ayudante que sep-a   inglés.
           need-3PL   Ø/DOM   an assistent that know-3SG.PRS.SBJV English
           'They need an assistant who knows English.'

(12)   Necesit-an Ø/*a   camarero.
        need-3PL   Ø/*DOM   waiter
        'They need a waiter.'

Note that the contrast between a specific and a non-specific indefinite is best seen in sentences with intensional verbs such as in (11) and (12), where it has semantic impact: in the specific reading we assert the existence of the assistant, while we cannot do so in the non-specific reading. In transparent contexts, i.e. contexts without further operators, as in (10), the specificity expresses the speaker's intention to identify the referent. A specific reading clearly depends on the pragmatics of the sentence, and it is very difficult to detect for a reader. Other types of specific indefinites are partitives and topical indefinites (see the discussion of (14) below) and cataphoric or forward-looking indefinites (see Section 3.2; see von Heusinger 2019 for a comprehensive overview of different types of specificity).[3]

---

[3]   In a forced choice experiment, von Heusinger et al. (2024) presented indefinites with and without DOM in (i) intensional contexts, (ii) in sentences with a universal quantifier and (iii) in transparent contexts. Participants had to choose a continuation that clearly signaled either a specific or non-specific reading. The results showed that indefinites with DOM are significantly more often interpreted as specific in intensional contexts and contexts with a universal quantifier than in transparent contexts. In transparent contexts, DOM did not

As for (iii) verbal parameters, such as affectedness, it has been shown that they also influence DOM with human indefinite direct objects. While affected direct objects are preferentially realized with DOM, even if indefinite, as in (13a), non-affected direct objects only optionally appear with DOM, as in (13b) (Torrego 1999, Romero Heredero & García García 2023).[4]

(13) a. Golpe-aron/asesin-aron   [??]Ø/a   un turista.
          beat-3PL.PST/murder-3PL.PST   [??]Ø/DOM   a tourist
          'They beat/murdered a tourist.'
      b. Vi-eron/encontr-aron   Ø/a   un turista.
          see-3PL.PST/find-3PL.PST   Ø/DOM   a   tourist
          'They saw/found a tourist.'

As mentioned, (iv) topicality has been identified as a further relevant parameter for DOM in Spanish.[5] Similarly to other Romance languages, this is evidenced by the fact that human left-dislocated direct objects require *a*-marking, as in (14) (from Leonetti 2004: 86). It is, however, much harder to claim that *a*-marked direct objects that are not left-dislocated are also topical.

(14) *Ø/A   muchos estudiantes, ya   los   conoc-ía.
      *Ø/DOM many   students,   already them know-PST[1SG]
      'Many students I already knew.'

In general, topicality is an important parameter for DOM in Spanish, particularly with regard to its diachronic development (Melis 1995, 2021, Pensado 1995, von Heusinger & Kaiser 2005, Iemmolo 2011). Crucially, the relevance of this parameter is evidenced by constructions with left-dislocated direct objects. As has been argued, left-dislocated direct object pronouns seem to correlate with the rise of DOM in Spanish, supporting the view that DOM evolved from a topic-shift strategy (see Pensado 1995, Iemmolo 2010, Melis 2021). Moreover, it has been shown that in different Romance languages, including Spanish, left-dislocated constructions contributed to the diachronic expansion of DOM along the Referentiality Scale (Iemmolo 2010, García García & Caro to appear). In Spanish, this development led to a generalization of DOM with human definite direct objects in both left-dislocated and non-dislocated structures, suggesting a gradual loss of the link between DOM and

---

influence the interpretation of the indefinites. We therefore assume that in transparent contexts, DOM with regular indefinites is not (closely) related to specificity; however, for d(iscourse)-linked, topical or partitive indefinites, see the discussion of example (14) below.

[4]   Note that telicity, another verbal category, only influences DOM if the underlying verb belongs to the class selecting for affected direct objects (see Romero Heredero 2022). Put differently, the impact of telicity depends on affectedness. Note that there is a further verbal category that has been identified as favouring DOM in Spanish and other (Romance) languages independently of affectedness, namely agentivity (see García García, Primus & Himmelmann 2018 and Mürmann 2023 for extensive discussions).

[5]   Following the suggestion of a reviewer, we would like to point out that we also have DOM systems in Romance that take topicality as the main parameter, such as Balearic Catalan (see Escandell-Vidal 2009), where dislocated definite objects are marked, even if inanimate. For a recent overview of DOM in Catalan see Irimia & Pineda (2022).

topicality (von Heusinger & Kaiser 2005, Iemmolo 2010). Considering direct objects with DOM in canonical non-dislocated structures, scholars have proposed weaker notions that still maintain a link between DOM and topicality, such as *topic-worthiness* (Comrie 1989, Iemmolo 2010, 2011) or *secondary topic* (Nikolaeva 2001, Dalrymple & Nikolaeva 2011). All these authors assume that the secondary topic has a sentence-based function (e.g. as subject of a secondary predication). It is not quite clear how this notion can be properly integrated into a sentence-oriented information structure (Krifka 2007), but if we re-interpret the notion of secondary topic with respect to the discourse as an issue that should be resolved in the subsequent discourse, we come close to the forward-looking notion of "referential persistence" used by Chiriacescu & von Heusinger (2010). There is also a more pragmatic, discourse-oriented notion of specificity as a d(iscourse)-linked indefinite, i.e. an indefinite that is partly linked to items in the preceding discourse, such as presuppositional or partitive indefinites. Enç (1991) argues for Turkish that indefinites with DOM are all partitive, i.e. d-linked and thus specific. Ledgeway et al. (2019) show that in Italian dialects, DOM marking depends on indefinites that are presuppositional, i.e. assumed to be known by speaker and hearer. These are all types of specific indefinites that are discourse dependent or backward-looking. In the next section, we will instead discuss a forward-looking discourse function.

3. Differential object marking and discourse prominence

DOM marking depends on the prominence status of the direct object with respect to the subject. Both co-arguments compete with each other, with DOM signaling that a direct object is a better competitor for the subject than a direct object without DOM. The prominence status is generally described in terms of sentence-bound parameters such as animacy, referentiality and affectedness, and by the information structural parameter of topicality. We argue that we will gain a better understanding of DOM if we extend the parameters to a more global or discourse perspective and take discourse prominence as an additional parameter determining DOM.[6]

3.1. Local and discourse prominence

The concept of prominence as defined in recent work by Himmelmann & Primus (2015) for different grammatical domains and applied to discourse management by Jasinskaja et al. (2015) and von Heusinger & Schumacher (2019) allows us to combine the broad insights of the function of DOM in the sentence with its potential function in the discourse. Himmelmann & Primus (2015) have provided a clear definition of the intuitive concept of prominence in terms of three essential characteristics. First, prominence is a relational property that singles out one element from a set of elements of equal type and structure; second, prominence status shifts in time (as discourse unfolds); third, prominent elements are structural attractors, i.e., they serve as anchors

---

[6]   Laca (2006: 429-432) provides a quite different subclassification of DOM parameters: animacy and definiteness are local parameters; the properties of the verb, secondary predication, preverbal position and clitic doubling are global ones; and topicality and referential relations are textual ones.

for the larger structures they are constituents of, and they may license more operations than their competitors. With this definition, research on DOM was able to bring together and describe the interaction of sentential parameters for DOM (see Romero Heredero 2022, Kızılkaya 2024). The prominence concept was then applied to referential management in discourse according to the adapted definition for discourse prominence as a principle, that (i) singles out one discourse referent from among discourse referents of the same type; (ii) allows for dynamic changes in the discourse structure between discourse referents in an unfolding discourse; and (iii) assigns to the most prominent element a higher attraction for operations than to its competitors (see Jasinskaja et al. 2015, von Heusinger & Schumacher 2019). This concept of discourse prominence is a more recent instantiation of Givón's (1983) notion of *topic continuity*, which has often been reformulated in terms of accessibility (Ariel 1990, Arnold 2010), attentional focus (Gundel et al. 1993) or salience (Chiarcos et al. 2011, Chiarcos 2011, Falk 2014).

Furthermore, we think that the backward- and forward-looking functions of Centering Theory (Grosz et al. 1995) can also easily be integrated into this notion. The backward-looking function is the dependence of an expression on the discourse ranking of its antecedent, and the forward-looking function is the potential of an expression to structure the upcoming discourse. We focus on the forward-looking function, since we investigate indefinite direct objects, i.e., brand-new expressions, which cannot have a backward-looking function. We further operationalize the forward-looking function of the expressions under investigation into two measurable parameters (see Chiriacescu & von Heusinger 2010): referential persistence and topic shift. In addition, we introduce the new measurement of *discourse prominence advantage* (see Section 3.3).

3.2. Discourse prominence and DOM in Romanian

Chiriacescu & von Heusinger (2010) investigated whether DOM marking in Romanian makes the marked referent more discourse prominent. They conducted a paragraph continuation experiment with small contexts of four sentences, the pre-critical sentence containing a proper name as a subject, which was also then the null-subject of the critical sentences, together with an indefinite human direct object, in one condition with DOM and in the other condition without DOM, as in (14). Note that they contrasted human indefinite direct objects without DOM (*un copil* 'a child') with human indefinite direct objects realized with both DOM and clitic doubling (*(l-) a văzut (pe) un copil* 'CD DOM a child').

(14)   Stimulus item 1 (from Chiriacescu & von Heusinger 2010: 317)
        Aseară a fost extraordinar de cald. Pentru că nu mai rezista în casă, Graur s-a hotărât să se ducă în oraş. Pe drum (**l**-)a văzut (**pe) un copil** intrând într-un magazin.
        'It was extraordinarily warm outside yesterday evening. Because it was unbearable for him to stay home anymore, Graur decided to go downtown. On his way there he (CL-) saw (DOM) a child entering a store.'

They asked participants to continue small contexts such as those in (14) with five coherent sentences. They annotated these sentences according to (i) whether the

two referents of the last context sentences were anaphorically picked up, (ii) the corresponding sentence topics of each of the five continuation sentences, and (iii) the lexical form or nominal type (pronoun, full noun phrase, etc.) of the anaphoric expression.

They found that (i) there were more anaphoric continuations for direct objects with DOM than for unmarked direct objects and that (ii) direct objects with DOM became topics earlier and more often in the subsequent discourse than their unmarked counterparts. They did not find an effect of DOM on the lexical form of the anaphoric expression. The experiment is a very instructive pre-test, but the results could not be statistically tested since there were not enough data points.

However, this was addressed in a broader follow-up study involving 85 native speakers of Romanian providing 2,040 elicitated continuation sentences (von Heusinger & Tigău 2024). In contrast to Chiriacescu & von Heusinger (2010), in this study DOM was investigated independently of clitic doubling (CD) of the direct object. Paragraph continuations were elicited by using three different realizations of human indefinite direct objects: (i) with DOM and CD, (ii) with DOM, and (iii) with neither DOM nor CD. In a nutshell, findings from this study indicate that both DOM and CD contribute to the prominence of the direct object in discourse. Furthermore, the simultaneous presence of DOM and CD amplifies the discourse prominence of the direct object. Put differently, the introduction of an indefinite direct object marked with both DOM and CD generates an even higher expectation within the subsequent discourse compared to one marked solely with DOM. Tigau (2022) conducted a similar, but smaller study with fewer data points. She found only minor numerical effects of DOM and CD.

Note that, in contrast to Romanian, CD in Spanish is not possible with human indefinite direct objects, at least in most varieties of Spanish. As we will see, this difference between Romanian and Spanish might be responsible for the diverging results of DOM in terms of discourse prominence as found in these languages (see Section 6).

3.3. Hypothesis, discourse prominence measures and empirical data

Based on Laca's (1995) observations of the discourse function of DOM in Spanish and on the previous investigations of DOM in Romanian (Chiriacescu & von Heusinger 2010, von Heusinger & Tigău 2024), we hypothesize that direct objects with DOM in Spanish are more discourse prominent than their unmarked counterparts, as stated in (15).

(15)   Hypothesis
        Direct objects with DOM are more discourse prominent than their unmarked
        counterparts.

We will test this hypothesis on human indefinite direct objects, since in these cases DOM is generally optional. Recall that human definite direct objects require DOM and inanimate direct objects are only marked under very restricted conditions (see Section 2). We assume that discourse effects are most visible if DOM is optional with respect to the local prominence parameters mentioned above. As far as discourse prominence is concerned, the literature has used different measures: Givón (1983)

introduces the notion of topic continuity, which by itself is a complex of various parameters, such as the number of uptakes of an antecedent, the question of whether the antecedent becomes a topic, etc. Later work on discourse structure suggests that the explicitness, i.e., the lexical size of an anaphoric expression, reflects accessibility or the activation of the antecedent (Gundel et al. 1993, Ariel 1990). On the other hand, the antecedent itself might express a certain forward-looking function, such as the next mention bias, that is, the antecedent might condition its likelihood of being mentioned again in the subsequent sentence (Kehler et al. 2008, Arnold 2010). In the empirical studies presented in this paper, we will focus on two forward-looking functions: referential persistence and topic shift. Referential persistence is the potential of an expression to be taken up in the subsequent discourse. We can measure this by the frequency with which a referent is picked up in the next sentence. Investigating longer stretches of discourse beyond the next sentence following an antecedent, one can also assess referential persistence by the length of the anaphoric chain. The topic shift parameter, on the other hand, focuses on whether or not the critical referent continues to be part of the comment or shifts towards the topic in the subsequent sentence. Again, we can measure this parameter by the frequency with which a referent becomes the subject in the subsequent sentence, which can generally be identified with the topic of that sentence.

Besides referential persistence and topic shift we will introduce a new relational measure: the *discourse prominence advantage*. This measure is suggested by the relational view of DOM in the prominence account, as introduced in Section 3.2. Prominence is understood as the relation between at least two arguments, the subject and the direct object. If the direct object becomes more prominent, it will be a stronger competitor for the subject. We will describe this advantage with *the discourse prominence advantage*, a ratio between the forward-looking potential of the direct object and that of the subject. Importantly, we will base this relational parameter on continuations that either include the subject (Ref 1) or the direct object (Ref 2) of the critical sentence, but not both. Thus, we calculate the discourse prominence advantage as the ratio of single referent continuations with Ref 2 with respect to (divided by) single referent continuations with Ref 1 (see Section 4.1.3, Table 3). We think that including the competitors of the direct object in the measurement of discourse prominence is closer to the relational concept of discourse prominence, as we use it here (see von Heusinger & Schumacher 2019). Our three measurements for discourse prominence are summarized in (16).

(16)   Measures for discourse prominence
   (i)   **referential persistence**
         the number of anaphoric uptakes of a previously introduced referent
   (ii) **topic shift**
         whether the previously introduced referent becomes a topic
   (iii) **discourse prominence advantage**
         based on the number of single referent continuations with Ref 2 with respect to / divided by the number of single referent continuation with Ref 1

Following our main hypothesis (15), we predict that human indefinite direct objects with DOM have a higher referential persistence as well as a higher topic shift

potential than those without DOM. In addition, we expect that direct objects with DOM will also have a higher discourse prominence advantage than their unmarked counterparts. In order to test these predictions, we conducted two extensive corpus searches (Section 4) and two paragraph continuation experiments (Section 5).

4. Corpus studies

In order to investigate the discourse function of DOM we conducted two corpus studies. The first one is based on two corpus searches of the *Corpus del Español del Siglo XXI* (*CORPES XXI*) and Davies' *Corpus del Español*, yielding about 400 tokens (Section 4.1). However, this relatively high number of relevant tokens was only possible since we allowed for all kind of continuations following the direct object, including not only full sentences but also subordinated clauses. The second corpus study is basically a replication of the first one, though the selection of the data was confined to *CORPES XXI* and restricted to tokens where the continuation of the critical direct object always represented a full sentence. This corpus study was much more carefully controlled, but yielded only 112 relevant tokens for our analysis of the relationship between DOM and discourse prominence (Section 4.2). Comparing and discussing our findings we will show that despite the differences mentioned, both corpus studies show similar results that support our hypothesis on DOM and discourse prominence (Section 4.3).

4.1. Corpus Study 1

4.1.1. Study design

As Spanish does not provide a large annotated corpus, we searched "by hand" in two of the larger corpora, namely *CORPES XXI* and Davies' *Corpus del Español*. In both corpora, we looked up eight verbs and collected about 200 instances of indefinite human direct objects with DOM and 200 without DOM.[7] In terms of the verbs involved, we carefully confined our search to verbs such as *ver* 'to see' that were attested with direct objects both with and without DOM. Overall, we found more instances with DOM than without. However, we compensated for this asymmetric distribution by collecting the same number of tokens without DOM as with. Table 1 shows the numbers of tokens retrieved for both corpora according to the eight selected verbs.

---

[7]   More precisely, we queried both corpora for the following string of words: one of the selected verbs followed by a human indefinite direct object with or without DOM.

Table 1. Human indefinite direct objects with and without DOM selected by different verbs in Corpus Study 1.

	CORPES XXI			*Corpus del Español*			
Verbs	**DOM**	**no DOM**	**Total**	**Verbs**	**DOM**	**no DOM**	**Total**
-	-	-	-	*citar* 'to cite'	2	2	4
elegir 'to choose'	16	17	33	*elegir* 'to choose'	15	16	31
enviar 'to send'	15	13	28	*enviar* 'to send'	8	8	16
mandar 'to send'	3	2	5	*mandar* 'to send'	5	5	10
nombrar 'to nominate'	15	23	38	*nombrar* 'to nominate'	34	38	72
presentar 'to present'	8	10	18	*presentar* 'to present'	3	2	5
traer 'to bring'	9	13	22	*traer* 'to bring'	12	12	24
ver 'to see'	32	20	52	*ver* 'to see'	25	21	46
Total	**98**	**98**	**196**	**Total**	**104**	**104**	**208**

4.1.2. Annotation

Overall, Corpus Study 1 yielded 404 hits, 196 from *CORPES XXI* and 208 from *Corpus del Español*, for which we annotated the continuations following the human indefinite direct objects. However, as a first step, some of the tokens retrieved had to be excluded, since they contained unclear cases of continuations of the context sentence, involving direct speech, fragments and incoherent, inappropriate or ambiguous continuations. Thus, we based our analysis on 383 tokens. In a second step, we segmented the continuations to the human indefinite direct objects by annotating the first upcoming clause containing a finite verb, either pertaining to an independent matrix sentence or to a subordinate clause. Note that both of these continuations correspond to elementary discourse units (EDU), as they contain finite verbs that denote an event (Asher et al. 2012). However, we also included infinitival phrases as independent units functioning as continuations to the tokens in the critical sentences with and without DOM, respectively. In a third step, we concentrated on the anaphoric terms in the continuation sentences, annotating their grammatical role (subject, direct object, etc.) and their nominal type (null, clitic, personal pronoun, proper name, definite NP, etc.). See (17) for illustration.

(17)   Context   Ese político mediocre$_1$ que nombra **a un mediocre$_2$**
        sentences   'That mediocre politician$_1$ who appoints DOM a mediocre one$_2$'
        Continuation   para que Ø$_2$ no le$_1$ haga sombra […]
        clause   'so that he$_2$ does not overshadow him$_1$[...]'

    The context sentence includes the subject referent *ese politico mediocre* 'that mediocre politician', which contains a restrictive relative clause with the verb *nombrar* 'to appoint' governing the direct object *a un mediocre* 'a mediocre (one)'. We analyzed the subordinate clause starting with *para que* 'so that' as the relevant continuation clause. In this continuation clause, the subject referent of the context sentence (Ref 1) appears as a pronominal indirect object, while the direct object of the context sentence (Ref 2) is realized as a null-subject (Ø). Note that the indefinite direct

object could be understood as non-specific ('any mediocre person'); however, it still serves as an antecedent for the null-subject in the continuation clause. As discussed in Section 2, it is often difficult to judge whether the indefinite in a discourse is specific or non-specific. The indefinite in (18) refers to a particular individual whose identity is not known, while the indefinite *un ex marine* ('a former Marine') is clearly specific. The indefinite *un beneficiario* ('a beneficiary') in (19) is clearly non-specific, but can still be picked up with a demonstrative.

(18)   Context   Un testigo vio **a un hombre$_2$** en la casa
        sentences   'A witness saw a man$_2$ in the house'
        Continuation   y **su$_2$** descripción coincide con la de un ex marine muerto
        clause   hace 17 años
                        'his$_2$ description matches that of a former Marine who died 17
                        years ago.'

(19)   Context   El asegurado nombra **un beneficiario$_2$**.
        sentences   'The insured (person) appoints a beneficiary$_2$.'
        Continuation   **Este$_2$** debe estar informado de la existencia de la póliza, […]
        clause   'This one$_2$ must be informed of the existence of the policy,
                        [...]'

4.1.3. Results

This section presents the results of Corpus Study 1 (*CORPES XXI* and *Corpus del Español*) according to our three measurements of discourse prominence. Table 2 shows the findings with regard to referential persistence, presenting the rate of anaphorical uptakes of the direct object referent (Ref 2), while, with regard to topic shift, it displays how often direct objects were realized as grammatical subjects in the continuation clauses. In terms of referential persistence, direct objects were anaphorically retaken in about 40% of the cases, across conditions. However, direct objects with DOM were picked up much more often than those without DOM (48% vs. 30%). With respect to topic shift, direct object referents (Ref 2) were anaphorically taken up as the subject in 25% of the continuation clauses. Again, there is a noticeable difference depending on DOM: direct objects with DOM show a higher frequency of topic shifts than those without DOM (32% vs. 18%). There is a significant effect of DOM on both referential persistence and topic shift.[8]

---

[8]   More precisely, we used a logistic regression analysis that revealed a significant effect of DOM on the referential persistence of Ref 2 (estimate = 0.7569, p < .001), as well as a significant effect of DOM on topic shift (estimate = 0.7481, p < .01). All analyses were performed using R (Team 2023).

Table 2. Referential persistence and topic shift of Ref 2. Percentage and absolute numbers of continuations of the direct object with and without DOM in Corpus Study 1.

Ref 2	Referential persistence	Topic shift (subject of continuation sentence)
DOM	48% (90/189)	32% (60/189)
no DOM	30% (58/194)	18% (35/194)
Total	**39% (148/383)**	**25% (95/383)**

In the following, we report on the results with regard to our newly introduced relational measurement of discourse prominence advantage. As discussed in Section 3.3, this parameter reflects the competition between the subject (Ref 1) and the direct object (Ref 2) of the context sentence with respect to its forward-looking function in discourse. In analyzing this competition, we exclusively considered single-referent continuations, that is, clauses and sentences that either contain Ref 1 or Ref 2 but not both. We assume that the discourse prominence advantage can be best instantiated in these single-referent continuations, since in configurations with only one referent we can most effectively measure the competition between Ref 2 and Ref 1. Based on the configurations of the continuation sentences, we calculated the relation between single-Ref 2 vs. single-Ref 1 continuation sentences with and without DOM. More specifically, we divided the absolute numbers of single-Ref 2 tokens by those of single-Ref 1 instances.

The results are provided in Table 3. As shown, indefinite direct objects (Ref 2) with DOM have a discourse prominence advantage of 1.1 over subjects (Ref 1), that is, they are more often picked up in single referent continuations than subjects (Ref 1). On the other hand, direct objects (Ref 2) without DOM exhibit a much lower discourse prominence advantage figure of 0.7, which means that they are taken up much less frequently than the corresponding subject (Ref 1).[9]

Table 3. Discourse prominence advantage in single referent continuations with and without DOM in Corpus Study 1.

	Ref 1	Ref 2	Discourse prominence advantage
DOM	44	49	1.1
no DOM	48	34	0.7

4.1.4. Discussion

As has been shown, Corpus Study 1, based on about 400 indefinite human direct objects, provides initial support for our hypothesis that DOM in Spanish signals discourse prominence. For our three selected measurements we found significant effects for referential persistence and topic shift and a numerical effect for the discourse prominence advantage, which confirmed our predictions: human indefinite direct objects with DOM exhibited a higher referential persistence and a higher topic shift potential than those without DOM. Moreover, direct objects with DOM also

---

[9]   We did not find an appropriate statistical method to test whether these relational values are significantly dependent on DOM or not.

clearly show a higher discourse prominence advantage over the subject than direct objects without DOM. Importantly, Corpus Study 1 provides a large number of tokens. We were able to achieve this large number as we allowed for all kinds of context clauses, including infinitives, and for different continuation clauses, including independent/full sentences but also subordinated and embedded infinitival clauses. However, this choice is not without problems, since topics (and topic shifts) are generally restricted to root clauses (see Emonds 2004, among others). Consequently, topic shifts typically occur between sentences but not between clauses inside a complex sentence.

4.2 Corpus Study 2

To verify our findings on the impact of DOM on discourse prominence we carried out a second corpus study based on a more carefully controlled dataset. Crucially, we only considered continuations representing full sentences, in order to provide a more appropriate structure for measuring topic shift.

4.2.1. Study design

Corpus Study 2 is based exclusively on *CORPES XXI.* Similar to Corpus Study 1, we restricted the search to transitive verbs attested with human indefinite direct objects, exhibiting instances with both DOM and no DOM. More specifically, we used six out of the eight verbs from Corpus Study 1 (see Table 4) and only considered verb forms in the 3rd person singular. Applying this strategy, we searched for the selected verbs, looked up tokens both with and without DOM and then manually selected those instances exhibiting both a human subject and a human direct object referent.

We further excluded (i) collective nouns; (ii) nominalizations; (iii) NPs referring to mythological entities; (iv) impersonal pronouns and constructions; (v) the indefinite pronoun *nadie* 'nobody'. Negation phrases were also ignored. Further tokens were discarded for the following reasons: (i) there were no continuation sentences following the context sentence; (ii) the continuation sentences had no finite verbs; (iii) we encountered combinations of direct and indirect speech; (iv) the continuation sentence used the 1st person (singular or plural); (v) the context sentence exhibited more than one direct object, as in coordination structures; and (vi) the sentence was ambiguous.

Table 4 gives an overview of the number of items retrieved in Corpus Study 2 and our subsequent manual selection of the data extracted. In total, the corpus search retrieved around 7,800 hits, of which we analyzed approximately 4,000 tokens. However, only a very small number of these hits, namely 112 tokens, turned out to meet the animacy and number restrictions, requiring a human subject and a human indefinite direct object both appearing in the 3rd person singular. Out of these 112 tokens, 83 exhibited direct objects with DOM, and 29 contained direct objects without DOM.

Table 4. Number of tokens with human subjects and human indefinite direct objects with and without DOM selected by different verbs in Corpus Study 2, based on *CORPES XXI*.

Verbs	Verbs + DOM				Verbs + no DOM			
	tokens retrieved	tokens analyzed	hum. obj.	hum. subj.	Tokens retrieved	tokens analyzed	hum. obj.	hum. subj.
elegir	41	41	25	13	398	398	15	6
enviar	82	82	26	13	819	720	8	3
mandar	22	22	12	5	370	370	4	4
presentar	104	104	18	5	3.597	720	0	0
traer	28	28	9	8	637	637	9	6
ver	393	100	54	39	1,432	720	13	10
Total	**670**	**377**	**144**	**83**	**7,253**	**3,565**	**49**	**29**

4.2.2. Annotation

For the annotation and selection of the continuation sentences, we basically followed the process established for Corpus Study 1 (see Section 4.1.2), though with one important exception: we restricted the continuation sentences to independent sentences, that is, we excluded all subordinated and embedded infinitival clauses. The following two examples from our corpus illustrate the high variation of indefinites. The indefinite with DOM in (20) is modified by a relative clause, but not taken up in the continuation sentence (note that this is not necessarily a specific indefinite, as the speaker might not have any further knowledge about that referent and is probably not able to identify him). The indefinite in (21) is also modified by a long relative clause, but not marked by DOM. However, it is picked up in the continuation clause by the personal pronoun.

(20) Context sentences En algún momento mi padre oyó un ruido de ramas a su espalda, se dio la vuelta y vio **a un miliciano₂ que le miraba.**
'At some point, my father heard a noise of branches behind him, he turned around and saw a militiaman₂ looking at him.`

Continuation clause Entonces se oyó un grito […]
'Then a scream was heard […]'

(21) Context sentences Judith pasó por el colegio como una niña más, sin destacar en nada, ni en lo malo ni en lo bueno, y a los dieciocho años eligió **un novio₂ que tampoco destacaba por nada en especial.**
'Judith went through school like any other girl, without standing out in anything, neither the bad nor the good, and at the age of eighteen she chose a boyfriend₂ who also did not stand out for anything in particular.'

Continuation clause Siempre se supo que, llegado el momento, se casaría con **él₂** […]
'It was always known that, when the time came, she would marry him₂ […]'

4.2.4. Results

Table 5 presents the results with regard to referential persistence and topic shift. As shown, direct objects with DOM are more often picked up in the continuation sentences than those without DOM (51% vs. 41%). Similarly, direct objects with DOM involved more cases of topic shifts than did direct objects without DOM (29% vs. 14%). Note that we measured topic shifts by counting those instances in which the direct object of the context sentence (Ref 2) is realized as grammatical subject of the continuation sentence.

Table 5. Referential persistence and topic shift of Ref 2 with and without DOM in Corpus Study 2.

Ref 2	Referential persistence	Topic shift (subject of continuation sentence)
DOM	51% (42/83)	29% (24/83)
no DOM	41% (12/29)	14% (4/29)
Total	**48% (54/112)**	**25% (28/112)**

In contrast to Corpus Study 1, the results concerning referential persistence and topic shift did not turn out to be statistically significant.[10] However, the numbers of both measurements indicate a noticeable numerical effect of DOM on discourse prominence that aligns with the findings from Corpus Study 1, where there was a significant difference for both parameters.

Table 6 shows the results for discourse prominence advantage, focusing on single-referent continuation sentences picking up either the subject (Ref 1) or the direct object (Ref 2) of the context sentence. Recall that this measurement informs us about the forward-looking potential of the direct object (Ref 2) with respect to that of the subject (Ref 1). Similar to Corpus Study 1, the findings from Corpus Study 2 reveal that human indefinite direct objects with DOM have a greater discourse prominence advantage than human indefinite direct objects without DOM (1.2 vs. 0.6).

Table 6. Discourse prominence advantage in single referent continuations with and without DOM in Corpus Study 2.

	Ref 1	Ref 2	Discourse prominence advantage
DOM	17	20	1.2
no DOM	8	5	0.6

Summing up, Corpus Study 2 demonstrates the numerical effects of DOM on discourse prominence. More specifically, the results show that human indefinite direct objects with DOM exhibit a greater discourse prominence than their unmarked counterparts based on all of the measurements employed, that is, referential

---

[10]   A logistic regression analysis indicated that DOM had no statistically significant effect on the referential persistence of Ref 2 (estimate = 0.3724, p > .05). Additionally, no statistically significant effect of DOM on the topic shift of Ref 2 was observed (estimate = 0.9331, p > .05).

persistence, topic shift and discourse prominence advantage. This provides additional support for the hypothesis outlined in (15).

4.2.5. Discussion

Corpus Study 2 yielded a much smaller dataset than Corpus Study 1, comprising only 112 relevant tokens out of the approximately 4,000 sentences analyzed. Moreover, the data showed a very unbalanced relation between instances of human indefinite direct objects with and without DOM (83 vs. 29). Nevertheless, we found noticeable numerical effects of DOM on discourse prominence with regard to all three of the measurements employed, referential persistence, topic shift and discourse prominence advantage. These results align with the findings from Corpus Study 1, even though the selection of the continuation sentences was much more constrained. While, in Corpus Study 1, we considered all kinds of (subordinate) clauses as continuations to the context sentence, in Corpus Study 2, we restricted relevant continuations to independent sentences, in order to more accurately measure topic shift. Still, both corpus studies show very comparable results.

4.3 Discussion of the corpus studies

We investigated the effect of DOM on the forward-looking function of human indefinite direct objects in two corpus studies. In Corpus Study 1, we collected about 200 human indefinite direct objects with DOM and about 200 without DOM, and in Corpus Study 2 we collected 83 instances with DOM and 29 without. The analysis of the first continuation clause (Corpus Study 1) or sentence (Corpus Study 2) showed a discernible effect in all three parameters (referential persistence, topic shift and discourse prominence advantage), which was significant for referential persistence and topic shift in Corpus Study 1. Note, however, that in Corpus Study 2 we had far fewer data points, which made it difficult to reach significance.

Moreover, we have to note that the annotation was difficult, in particular in Corpus Study 1, as we often encountered different kinds of embedded clauses following the critical direct object. In such cases, it is not entirely clear how to identify the next discourse unit to which the discourse prominence measures should be applied to. Besides this, both corpus studies contained some potentially relevant parameters that were annotated, but not controlled for, such as the nominal type of the subject in the context sentence or the number of tokens per verb. For example, in Corpus Study 1, we encountered quite different numbers for the selected verbs, ranging from 4 instances with *citar* 'to cite' to 106 for *nombrar* 'to nominate'. We consider these difficulties quite normal for working with corpus data, and we believe that the two corpus studies offer supporting evidence for the hypothesis regarding the impact of DOM on discourse prominence. Still, we decided to complement our corpus search with two paragraph continuation tasks.

5. Paragraph continuation task

The results of the corpus studies suggest that DOM in Spanish has a positive influence on the discourse prominence of the direct object. Although the two corpus studies yielded notably similar results, the collected data remains highly heterogeneous, in particular with regard to the structure of the context and the continuation text, the type of subject competitor in the context and the number of referents in the preceding context. To further examine our hypothesis (15) with better control for the relevant parameters, such as animacy, definiteness and additional referents, we conducted two paragraph continuation task experiments and one pilot study with native speakers of European Spanish. The first experiment included a very short context of one sentence, where participants were asked to continue with a single sentence (Section 5.1). In the second experiment, we provided a three-sentence context with either one competitor (subject) to the direct object or two or more competitor referents to the direct object. Again, participants were tasked with completing the context by generating a single sentence (Section 5.2). We conclude the section with an intermediate discussion comparing the two experiments (Section 5.3).

5.1 Experiment 1: single-sentence context

5.1.1. Study design

For Experiment 1, we employed a single-factor design with the independent variable comprising two levels: the presence or the absence of DOM in the critical item. More specifically, we selected 24 verbs[11] based on their balanced occurrence in *CORPES XXI* and created transitive sentences in two conditions (DOM and no DOM). The test items were distributed across two lists via a Latin Square design, such that each item only appeared once per list and all conditions were spread equally across the lists. Each list contained 12 filler items with similar structures but different verbs.[12]

5.1.2 Materials

In order to control for additional parameters influencing discourse prominence, we constructed the items using the following pattern: firstly, an overt human masculine or feminine subject, which could be either a proper name or a definite NP (Ref 1); secondly, one of the verbs mentioned above; thirdly, the occurrence or not of DOM; fourthly, a human masculine or feminine indefinite NP as the direct object (Ref 2); and lastly, an adjunct functioning either as a locative or a temporal PP. A sample item used

---

[11]   The verbs used in this experiment are: *asistir* 'to atend', *citar* 'to cite', *contemplar* 'to contemplate', *cuidar* 'to take care', *eligir* 'to choose/to elect', *elogiar* 'to praise', *entrevistar* 'to interview', *enviar* 'to send', *escuchar* 'to listen', *fotografiar* 'to photograph', *guiar* 'to guide', *investigar* 'to investigate', *invitar* 'to invite', *llevar* 'to take', *mandar* 'to send', *nombrar* 'to nominate', *oir* 'to hear', *ojear* 'to glance', *presentar* 'to introduce', *recoger* 'to pick up', *reconocer* 'to recognize', *rozar* 'to touch', *traer* 'to bring' and *ver* 'to see'.

[12]   Specifically: *animar* 'to encourage', *asustar* 'to scare', *bailar* 'to dance', *casarse* 'to marry', *competir* 'to compete', *encontrar* 'to meet', *hablar* 'to talk', *matar* 'to murder', *pasar* 'to spend', *sorprender* 'to surprise' and *tranquilizar* 'to calm down'.

in the experiment is provided in (18), where *Javier* is the subject, and *una cantante* 'a singer' is the human indefinite direct object.

(22)   Javier$_1$ escuchó a/Ø una cantante$_2$ en la plaza. _____
       'Javier$_1$ listened to DOM/no DOM a$_{FEM}$ singer$_2$ in the square.' _____

Filler items were created using the same structure, varying the verb and the definiteness of the direct object, but not displaying the condition tested in the critical items, i.e., an indefinite direct object. In (19), we provide a sample filler item.

(23)   La directora$_1$ habló con el conserje$_2$ antes de entrar en la escuela. _____
       'The director$_1$ talked to the doorman$_2$ before entering the school.' _____

5.1.3. Participants and task

We recruited 50 native speakers of Spanish to take part in this experiment. However, four participants were excluded from the annotation because they did not accomplish the task properly. Participants were recruited via Prolific and stem from different regions of Spain. Thirty of them hold university degrees while sixteen had high school certificates. The mean age is 26.9 years. Each participant received €5.50 as compensation for their participation. The experiment was launched using Qualtrics. Prior to starting the experiment, participants were provided with information about their rights and asked to provide their consent, as well as individual information. For the task, participants were instructed to create a continuation sentence for each (one-sentence) context provided; the experiment took an average of 19 minutes to complete. Given the simplicity of the task, no training was provided. Upon completion of the task, participants were given the opportunity to provide feedback related to the experiment.

5.1.4. Annotation

In total, we collected 1,104 data points. However, 92 continuation sentences needed to be discarded because they did not align with the specified parameters, such as ambiguous sentences (2), continuation sentences out of context (23), and sentence fragments (67), leaving us with 1,012 data points. We followed the same guidelines used in Corpus Study 2 for the annotation of the experiment, i.e., we focused on (i) the first anaphorical uptake of the referents, and (ii) their grammatical function (see Section 4.1.2). We considered the first sentence that participants produced as the continuation sentence.[13]
        (24) and (25) provide examples of continuation sentences to the context in (22) above, in which the direct object was preceded by DOM. In (24), Ref 2 is anaphorically taken up by means of the null-pronoun Ø. In (25), however, Ref 2 was not picked up in the continuation sentence, but Ref 1 was.

---

[13]   Even though the participants were instructed to create only one sentence for the continuation of the short story contexts, sometimes they provided more than just one.

(24)   Continuation of (22) with Ref 2:

        $\emptyset_2$ Estaba interpretando canciones de blues.

        'She$_2$ was performing blues songs.'

(25)   Continuation of (22) with Ref 1:

        $\emptyset_1$ Fue el único que aplaudió.

        'He$_1$ was the only one who applauded.'

5.1.5. Results

As illustrated in Table 7 below, participants have a much stronger tendency to continue with Ref 2 than what we found in the corpus studies. Across the DOM condition, participants continued in 76% (773/1,012) with the direct object, while in 54% (548/1,012) they shifted the direct object to the subject. Note that in the two corpus studies only 25% of direct objects became topics in the continuing text. In contrast to the corpus studies, in Experiment 1 DOM exhibited no impact on either referential persistence or topic shift.

Table 7. Referential persistence and topic shift of Ref 2: Percentage and absolute numbers of continuations of the direct object with and without DOM in single-sentence contexts with one-sentence continuation task (Experiment 1).

Ref 2	Referential persistence	Topic shift (subject of continuation sentence)
DOM	77% (392/507)	54% (273/507)
no DOM	75% (381/505)	54% (275/505)
Total	**76% (773/1,012)**	**54% (548/1,012)**

In Table 8, we also see the strong bias towards Ref 2 in the comparison between the subject and the direct object in single referent continuations mirroring the forward-looking function. There are more than twice as many continuations of the direct object as of the subject. Here, we do observe a numerical effect of DOM. Subjects in sentences exhibiting a direct object with DOM have less continuations (60) than those in sentences featuring a direct object without DOM (76), while the continuations of the direct object are similar, independently of DOM. Nonetheless, we can still see that the discourse prominence advantage in sentences with DOM is higher than in sentences without DOM (2.8 vs. 2.2).

Table 8. Discourse prominence advantage in single referent continuations with and without DOM in single-sentence contexts with one-sentence continuation task (Experiment 1).

	Ref 1	Ref 2	Discourse prominence advantage
DOM	60	167	2.8
no DOM	76	168	2.2

5.1.6. Discussion

We conducted a paragraph continuation task with 24 different transitive verbs that allow for DOM variations. The subject of these sentences, i.e., the competitor to the direct object, was either a proper name or a definite description. The inspection of the data showed a strong bias towards a continuation of the direct object referent, which might be close to a ceiling effect. We did not find any effect of DOM on referential persistence or topic shift, but we found a minor effect on our relational parameter of discourse prominence advantage. More specifically, test items with DOM elicited fewer continuations of the subject (Ref 1) than test items without DOM. The results of this paragraph continuation task differ from those of the corpus studies. We speculate that this might be due to the very short context and we therefore conducted a second paragraph continuation task with a three-sentence context, assuming that this presents a more natural environment to test the discourse effect of DOM.

5.2 Experiment 2: three-sentence context - one continuation sentence

5.2.1. Study design

The second paragraph continuation task experiment employed a 2x2 factorial design, crossing the factors DOM vs. no DOM and the presence or absence of (an) additional referent(s) in the context (AddRef vs. NoAddRef). We used a smaller set of verbs, namely the six verbs utilized in Corpus Study 2: *enviar* 'to send', *elegir* 'to choose', *mandar* 'to send', *presentar* 'to introduce', *traer* 'to bring' and *ver* 'to see'.[14] For each verb we constructed two contexts with three sentences each: one context had only one competitor to the direct object, namely the subject and discourse topic (NoAddRef), while the other context had additional referents (AddRef). The 12 contextual sentences were presented in two conditions (DOM vs. no DOM), and they were allocated across two lists using a Latin Square design. This ensured that each item appeared only once per list, and the conditions were evenly distributed across the lists. No filler items were included. To assess discourse prominence, we based our analysis on the measurements used previously (referential persistence, topic shift and discourse prominence advantage).

5.2.2. Materials

Each item consisted of two parts, (i) a short narrative context containing at least three sentences, where the presence or absence of additional referents was manipulated; and (ii) the critical sentence where a human indefinite direct object is introduced and manipulated with respect to DOM. A sample item providing a context with no additional referents (NoAddRef) is given in (26).

(26)   Context with NoAddRef:

Aquella mañana hacía muy buen tiempo y **María₁** estaba tomando café sentada en el balcón de **su₁** apartamento. Era muy temprano, pero **le₁** encantaba esa

---

[14]   While some of the verbs can be used as ditransitive predicates, we constructed the items using them solely as transitive predicates, without including the indirect object.

hora porque la calle estaba aún vacía y eso **le₁** permitía comenzar el día relajada. Justo cuando **Ø₁** estaba terminando de beber**se₁** el café, **Ø₁** oyó un ruido que provenía de la esquina, **se₁** giró y vio **(a/Ø) un hombre₂**.

'That morning the weather was very good and Maria₁ was drinking coffee sitting on the balcony of her₁ apartment. It was very early, but she₁ loved that time because the street was still empty and that allowed her₁ to start the day relaxed. Just as she₁ was finishing drinking the coffee, she₁ heard a noise coming from the corner, she₁ turned around and saw (DOM/no DOM) a man₂.'

Context sentences such as (26) contain two referents, realized as subject and direct object. The subject in (26) is introduced using the proper name *María*. The indefinite direct object *un hombre* 'a man' was presented in one condition with DOM and in the other condition without DOM. The test items always have a singular subject, which could be either a definite NP (such as *el director* 'the director', *la presentadora* 'the host') or a proper name (*Maria*). Both masculine and feminine characters were used as subjects. Regarding the direct object, it was consistently a human indefinite NP such as *un niño* 'a boy' or *una profesora* 'a teacher'.

The critical sentences, that is, the last sentence within the three-sentence context, consisted of a transitive verb (one of the six selected verbs) and its two arguments, one of which was the subject that was previously introduced in the context and the other a human indefinite direct object. Importantly, the direct object is always a new referent to the scene, which is not mentioned in the context until its introduction in the last sentence, such as *un hombre* 'a man' in (26). In order to avoid any co-reference between the referents, we created contexts in which the direct object referent had no pre-existing relationship with the subject referent, such as in the case of *amigo* 'friend'*, colega* 'colleague', etc.

A sample item providing a context with an additional referent (AddRef) is given in (27). In this example, we have the subject *Enrique*, the direct object *un vigilante* 'a guard' and one additional referent, *una enfermera* 'a nurse'. Other examples had even more additional referents.

(27)   Context with AddRef:

Enrique₁, jefe de seguridad del hospital psiquiátrico, **se₁** encontraba en la centralita junto a la puerta como cada noche. En torno a las dos de la madrugada, **Ø₁** oyó un fuerte golpe que provenía del piso de arriba. De repente, **Ø₁** escuchó el grito de **una enfermera₃** y, para comprobar qué estaba ocurriendo, **Ø₁** inmediatamente mandó **(a/Ø) un vigilante₂**.

'Enrique₁, head of security at the psychiatric hospital, was at the control room by the door, as he₁ was every night. At around 2 a.m., he₁ heard a loud knock coming from the upper floor. He₁ immediately heard the scream of **a nurse₃** and, in order to check what was happening upstairs, he₁ immediately sent (DOM/no DOM) a guard₂.'

5.2.3. Participants and task

We recruited 80 participants via Prolific, but three had to be excluded as they did not complete the task properly. Consequently, we annotated the answers of the remaining 77 participants. All of them were native speakers of European Spanish stemming from various regions of Spain, holding either a high school certificate (19) or a university degree (58). The group comprised 44 men, 31 women and 2 participants identifying as non-binary. The participants' average age was 30 years, with ages ranging from 18 to 59 years. They received €5.50 as compensation for their participation. After the participants had filled out a form with personal data and confirmed that they agreed to participate in the experiment, they were asked to read the short narrative contexts and create a single continuation sentence for each context. They were free to use their imaginations. This process was repeated 24 times per participant, as we had 24 items and no fillers per list. The questionnaires were created using Qualtrics.

5.2.4. Annotation

In total, we collected 1,848 responses, of which 163 needed to be discarded due to non-conformity with our predefined parameters. This included instances of ambiguous sentences (16), combinations of direct and indirect speech (47), sentence fragments (94) and sentences out of context (6). After filtering these responses, (see also Section 5.1.4), we retained a total of 1,685 responses. We present in (28) a context with two competitors for the direct object and two continuations each for the version with DOM and the one without, respectively.

(28) Context sentence:
Aquella noche, **la presentadora** estaba bastante nerviosa porque **uno de los concursantes** se había ido y tenía que ser reemplazado. Y, aunque el resto de participantes no parecía estar de acuerdo**, la presentadora** bajó del escenario y eligió (**a) un espectador$_2$**.
'That night, the presenter was quite nervous because one of the contestants had left and had to be replaced. And, although the rest of the participants didn't seem to agree, the presenter came down from the stage and chose a spectator$_2$.'

Continuation 1 with context with DOM (with Ref 2 being retaken):
El espectador$_2$ jugo muy bien el concurso, dejando en ridiculo a los otros participantes.[15]
'The spectator$_2$ played the contest very well, making a fool of the other participants.'

Continuation 2 with context with DOM (with Ref 2 not being retaken):
Es algo muy inusual
'It is something very unusual.'

---

[15] The continuation is kept exactly as the participant wrote it, including the typos.

Continuation 3 with context without DOM (with Ref 2 being retaken):
El espectador₂ montó una escena ya que no quería participar
'**The spectator₂** made a scene because he did not want to participate'

Continuation 4 with context without DOM (with Ref 2 not being retaken):
El concurso comenzó en medio de un clima bastante tenso.
'The contest began in a rather tense atmosphere.'

5.2.5. Results

Table 9 below summarizes the results regarding the referential persistence of Ref 2, split by NoAddRef vs. AddRef. Similar to Experiment 1, there was a very strong bias towards continuing with Ref 2, i.e., the direct object. This is the case with 73% (1,235/1,685) of the continuations across all test items. Moreover, we see a minimal difference in the referential persistence of Ref 2 in context sentences with DOM compared to those without DOM (75% vs. 72%). Interestingly, this difference is primarily due to contexts with more than two referents (AddRef). Here, we observe a more pronounced contrast: with context sentences showing DOM, the direct object (Ref 2) is taken up more frequently than with those lacking DOM (74 % vs. 68%).

Table 9. Referential persistence of Ref 2 regarding DOM vs. no DOM and NoAddRef vs. AddRef in three-sentence contexts with one-sentence continuation task (Experiment 2).

Ref 2	NoAddRef	AddRef	All
DOM	76% (320/423)	74% (317/430)	75% (637/853)
no DOM	76% (317/417)	68% (281/415)	72% (598/832)
Total	76% (637/840)	71% (598/845)	73% (1,235/1,685)

With regard to our second measurement of discourse prominence, that is, topic shift, no noticeable effect was observed, regardless of whether the context involved additional referents or not. However, it is noteworthy that we observed an overall high topic shift rate of 60% (1,012/1,685), contrasting with the findings of our corpus studies, yet resembling the 54% (548/1,012) rate of topic shifts documented in Experiment 1 (see Table 7).

Table 10 provides a summary of the findings regarding discourse prominence advantage. Similar to referential persistence, we note a strong bias towards the direct object and a slight preference for direct objects without DOM over those with DOM. In contrast to the results of Experiment 1, we observe a four to five times higher frequency of single-referent continuations with direct objects compared to subjects. Furthermore, there is a minor preference for direct objects without DOM over those with DOM.

Table 10. Discourse prominence advantage in single referent continuations with and without DOM in three-sentence contexts with one-sentence continuation task (Experiment 2).

	Ref 1	Ref 2	Discourse prominence advantage
DOM	74	326	4.4
no DOM	66	304	4.6

5.2.6. Discussion

With Experiment 2 we carried out a second paragraph continuation task, this time involving a three-sentence context entailing either one competitor (subject) or more competitors to the direct object referent. This specification was motivated by the differing results between Experiment 1 and the corpus studies. We assumed that more context and more referents – as in the corpus – might provide a more natural environment in which to test the discourse prominence behaviour of DOM. The results for referential persistence and topic shift in Experiment 2, however, were not very different from the paragraph continuation task of Experiment 1. They showed a very strong bias towards the continuation of the direct object, and very few continuations of the subject. We speculate that this might create a ceiling effect masking the discourse effect of DOM. A minor advantage of DOM was observed in the AddRef condition but not in the NoAddRef condition (see Table 9). This suggests that our contexts with more competitors resemble natural narratives, as reflected in the corpus data. No further contrasts were evident in Experiment 2, although it was intriguing to note that we observed four to five times more continuations with the direct object than with the subject (see Table 10).

5.3 Discussion of the paragraph continuation tasks

We conducted two paragraph continuation tasks in order to control for more conditions in the context sentence. In both experiments, we employed a balanced list of verbs, comprising 24 verbs in the first experiment and 6 verbs in the second. The subjects were consistently definite noun phrases or proper names. In Experiment 1, the transitive context sentence in SVO order concluded with a temporal or local adverb, whereas, in Experiment 2, it concluded with the direct object. In Experiment 1, a single sentence served as the context, whereas, in Experiment 2, three sentences were used as context. Additionally, we manipulated the number of competitors to the direct object in Experiment 2. The results for the two experiments are very similar. They both showed a much higher referential persistence (about 75%) and topic shift (50–60%) than in the corpus studies (39–48% and 25%) for each measurement, but no clear effect of DOM. In Experiment 2, we see a noticeable effect of the additional competitors to the direct object. In this condition the referential persistence is more pronounced than in the condition with only one competitor. No further effects are observed. There is a substantial difference in the discourse prominence advantage between the two experiments, ranging from 2.5 in Experiment 1 to 4.5 in Experiment 2, with unclear effects of DOM. This variation may also be influenced by the recency of the direct object. In Experiment 2, the context sentences concluded with the direct object, leading to a higher topic shift rate than in Experiment 1, which employed an adverbial clause.

Most of our corpus examples also featured diverse linguistic material at the end of the context sentence.

To examine whether the task of generating only one continuation sentence contributed to the elevated number of continuations with Ref 2, we conducted a pilot experiment using the three-sentence contexts from Experiment 2 as test items. However, in this pilot, participants were instructed to produce short narratives comprising five continuing sentences. The objective was to determine whether the task of generating a more extensive narrative continuation would yield a different relationship between the continuation of Ref 2 as compared to Ref 1. In this pilot study we employed the same setup as in Experiment 2, namely: (i) the same 2x2 factorial design, i.e., DOM and additional referents as independent variables, (ii) the same verbs, and (iii) the same parameters for the annotation. However, there are two main differences. First, the task assigned to the participants differed. Instead of generating a single continuation sentence, they were instructed to create a short narrative comprising a minimum of five sentences. Second, the number of critical items was reduced to 12, which were distributed across two lists. Therefore, each participant (26 native Spanish speakers) provided a short narrative for 6 critical items, producing a total of 156 short narratives, of which we had to discard 53 based on the parameters listed above (direct speech, fragments, etc.). The remaining 103 short narratives were only annotated for each occurrence of Ref 2. We then counted whether or not Ref 2 appeared in the first four sentences (as not all short narratives contained five sentences), treating this as an extended referential persistence measure. No discernible effect of DOM on the continuations was observed.

Moreover, we examined the potential influence of the variable additional referents (NoAddRef vs. AddRef), as this turned out to be relevant in Experiment 2. Table 11 presents the results concerning the interaction between DOM and additional referents. In instances where there are additional referents in the context, we observed an increase in the uptake of Ref 2, aligning with our findings in Experiment 2. Specifically, direct objects with DOM were mentioned again in 93% of cases, while those without DOM were taken up in 78%.

Table 11. Referential persistence of Ref 2 with or without DOM and NoAddRef vs. AddRef in three-sentence contexts within the first four continuation sentences (Pilot Experiment).

Ref 2	NoAddRef	AddRef	All
DOM	85% (22/26)	93% (25/27)	89% (47/53)
no DOM	96% (26/27)	78% (18/23)	88% (44/50)
Total	**91% (48/53)**	**86% (43/50)**	**88% (91/103)**

To sum up, the paragraph continuation experiments did not provide support for our hypothesis. We think that this might be partly due to a ceiling effect, as participants showed a very high bias towards continuations with the direct object, an inclination not mirrored in our corpus data. The length of the context provided (one sentence vs. three sentences) did not yield any discernible difference. The pilot experiment suggests that writing a short narrative strengthens the bias towards the direct object, but not the effect of DOM. Experiment 2 and the pilot study featured an additional manipulation between one and more than one competitor to the subject. This manipulation appears to create an environment where the effect of DOM becomes evident, whereas in the

one-competitor context, such an effect is not observable. We think that such multi-referential environments are more representative of natural discourse, as we also find them in our corpus data. They suggest that the prominence structure of a discourse is more complex than a simple competition between two referents.

6. General discussion

Differential Object Marking (DOM) is a typologically well-established research topic. Previous studies have focused on the prominence scales and parameters that contribute to a higher prominence of the direct object with respect to the subject, including the Animacy Scale, the Referentiality Scale, affectedness and topicality. These parameters interact to determine the prominence status of the direct object in a particular sentence. Only a few studies go beyond the sentence boundary and investigate discourse parameters. Comrie (1981/1989), Nilsson (1985) and Laca (1995) extends the sentence perspective to the broader view that DOM may also depend on and interact with the discourse. Laca (1995) observes that, in Spanish, DOM with animate non-human direct objects is to be expected when the direct object is co-referential with a referent introduced in the preceding sentence. Furthermore, a discourse topic is more easily associated with DOM than a less discourse prominent referent. Laca (1995) also points out that indefinite direct objects with DOM have a higher forward-looking potential than those without DOM. In other words, she suggests that DOM signals a preference for continuations of the direct object referent in the subsequent discourse.

The first empirical evidence for this discourse function of DOM comes from Romanian (see Chiriacescu & von Heusinger 2010 and Section 3.2). Through a paragraph continuation task, this study illustrated that indefinite direct objects with DOM show a higher referential persistence and a higher level of topic shift than their counterparts without DOM. These findings were subsequently corroborated through a more extensive paragraph continuation task and a corpus study (von Heusinger & Tiğau 2024). Based on the theoretical assumptions of Comrie (1981/1989), Nilsson (1985) and Laca (1995) and the first empirical evidence from Romanian, we formulated the hypothesis that human indefinite direct objects with DOM exhibit greater discourse prominence than those without DOM. We employed three measures for discourse prominence: referential persistence and topic shift, which are adaptations of Givón's (1983) concept of topic continuity, and a new relational measure, introduced by us, termed the discourse prominence advantage. This measure evaluates the forward-looking potential of the direct object relative to the subject, enabling us to model the relational nature of prominence as outlined by von Heusinger & Schumacher (2019).

We conducted two corpus studies. Although the approaches to searching the corpus and some contextual settings differed between the two studies, the results were very similar. Across all three measurements, we observed a numerical effect of DOM, and in Corpus Study 1, we detected a significant impact of DOM on referential persistence and topic shift.

Furthermore, we carried out two paragraph continuation task experiments to ensure a more homogeneous context and a balanced frequency of the verbs employed. Experiment 1 used a single-sentence context, whereas Experiment 2 utilized a three-sentence context, enabling us to manipulate the number of competitors to the direct

object. This approach aimed to create a more natural context with multiple competing referents, similar to our corpus examples. In both paragraph continuation experiments, we observed a high bias towards the continuation of human indefinite direct object referents regardless of DOM. This finding contrasts sharply with the results from our corpus studies. Across conditions, the referential persistence rates in the corpus studies were 39% (Corpus Study 1) and 48% (Corpus Study 2), whereas in the paragraph continuation task experiments, they were both approximately 75% (see Tables 2, 5, 7 and 9, respectively). Similarly, the topic shift rates of human indefinite direct objects were 25% in both corpus studies, whereas in Experiments 1 and 2, they were much higher, namely 54% and 60%, respectively. We assume that this large difference, particularly regarding the topic shift parameter, may be attributed to a recency effect. In the corpus studies, the continuation sentences varied in distance from the direct object, depending on the specific examples. In Experiment 1, an adverbial phrase followed the direct object, whereas in Experiment 2, the direct object was the last word in the context sentence, directly adjacent to the continuation sentence. Consequently, recency may account for the pronounced bias towards picking up the direct object referent in subsequent sentences. However, this assumption requires further investigation and support through additional research.

In terms of DOM, Experiment 1 demonstrated a certain effect on the discourse prominence of the direct objects tested, although not as pronounced as the effect observed in the corpus studies. Experiment 2 presents a less clear picture: while DOM shows a small positive effect on discourse prominence for referential persistence, no noticeable effect was observed with respect to topic shift. As for discourse prominence advantage, a slight negative effect of DOM was detected. Note that in Experiment 2, we additionally manipulated the number of competitors to the direct object. With more competitors, the referential persistence of the direct object decreases slightly, specifically, from 76% without additional competitors to 71% with additional competitors. However, in the DOM condition there are clearly more anaphorical uptakes than in the no DOM condition (74% vs. 68%). This observation suggests that DOM may reinforce the prominence status of the referent in a setting with multiple competitors, rather than in a scenario with only one competitor.

We must conclude that our two corpus studies provide evidence supporting our hypothesis that DOM contributes to discourse prominence, while the two paragraph continuation tasks show either a very marginal effect or no effect. We believe that the disparity in results may also be attributed to the differing types of empirical data: corpus data provide insights into referential management in authentic narratives, often within intricate structures. On the other hand, paragraph continuation tasks are specialized tasks involving the elicitation of one or more sentences given a specific input. While they certainly reflect certain preferences for antecedents and the form of anaphoric expression, they may be less suitable for examining more subtle referential relations.

Interestingly, our experimental findings on DOM and discourse prominence in Spanish contrast with those on DOM in Romanian (Chiriacescu & von Heusinger 2010), but are quite similar to some recent results in Turkish (see von Heusinger & Yıldız submitted). In a similar experiment to the one conducted for Spanish, Chiriacescu & von Heusinger (2010) provide clear evidence, despite being based on a small dataset, that, in Romanian, DOM signals the discourse prominence of the direct object. However, their manipulation of DOM also involved clitic doubling. While

direct objects with DOM may appear without clitic doubling in Romanian, the stimulus items were presented with both DOM and clitic doubling, or without DOM and clitic doubling; thus, DOM was not tested independently of clitic doubling. This was addressed in a more recent study by von Heusinger & Tigău (2024), which presented material from a corpus study and an extensive paragraph continuation task with 24 items and 87 participants. Three types of direct object realizations were distinguished: without any marking, with DOM and with both DOM and clitic doubling (CD). The corpus study revealed a numerical advantage for CD+DOM compared to an unmarked direct object. The paragraph continuation task showed a similar effect for referential persistence and topic shift, but a strong effect on the discourse prominence advantage (1.0 for direct objects without DOM, 1.8 for those with DOM, and 1.9 for those showing both DOM and CD).[16] This advantage is much higher than in our Spanish data. It seems that the observed effect of DOM on discourse prominence in Chiriacescu & von Heusinger's (2010) study is attributable to DOM, or rather, to an interaction of DOM and clitic doubling, as demonstrated by the study of von Heusinger & Tigău (2024). As for Standard Spanish, clitic doubling with human (in)definite direct objects is ungrammatical. However, it seems to be acceptable in other Spanish varieties, such as those spoken in Buenos Aires and Lima (Sánchez & Zdrojewski 2013). Future research must show, whether in these varieties, DOM together with clitic doubling has a comparable impact on discourse prominence to that found in Romanian.

7. Conclusion

Differential Object Marking (DOM) is a linguistic phenomenon at the interface of the lexicon, semantics, syntax and information structure, where factors such as animacy and affectedness interact with referentiality and topicality. DOM systems exhibit intricate crosslinguistic patterns, reflecting the interplay of these and other parameters. In our contribution, we have argued that examining DOM from a discourse perspective enhances our understanding of its competitive nature. Specifically, we suggested that, within a sentence, the direct object competes with the subject for prominence in a more abstract prominence structure that integrates both sentential and discourse parameters. This perspective enables us to incorporate discourse parameters related to both backward-looking functions (see Laca 1995, Ledgeway et al. 2019 for Italian dialects, Enç 1991 for Turkish DOM) and forward-looking functions. Focusing on the forward-looking function, we have formulated the hypothesis that direct objects with DOM in Spanish are more discourse prominent than their unmarked counterparts. To examine this hypothesis, we conducted two corpus studies, as well as two paragraph continuation task experiments. While the corpus studies support our hypothesis, the paragraph continuation tasks only showed marginal effects. Interpreting the specific differences between the empirical data, we have speculated that paragraph continuation tasks may not capture subtle narrative principles such as the forward-looking function as properly as corpus studies. We consider our research a first step towards a unified theory of DOM as a device for marking a prominent sentential

---

[16]   Tigău (2022) presents results from a similar, but much smaller study with only 12 items and 84 participants. Based on numerical differences, Tigău concludes that CD, rather than DOM, is the marker that makes the direct object more prominent.

argument, as well as a prominent discourse referent. Clearly, further research is needed, incorporating diverse languages and methodologies. This inquiry may extend beyond the scope of DOM. Recent typological studies on related phenomena, such as clitic doubling in Bulgarian and differential agent marking in Tima (Niger-Congo) and Yali (Trans-New Guinea), pursue a similar line of investigation (see Compensis et al. submitted and references cited therein). They suggest that discourse prominence generally plays a critical role in the morphosyntactic encoding of verbal arguments.

Acknowledgments

References

Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory* 21(3), 435–483. https://doi.org/10.1023/A:1024109008573.

Ariel, Mira. 1990. *Accessing noun-phrase antecedents*. London: Routledge. https://doi.org/10.4324/9781315857473.

Arnold, Jennifer E. 2010. How speakers refer: The role of accessibility. *Language and Linguistics Compass* 4(4), 187–203. https://doi.org/10.1111/j.1749-818x.2010.00193.x.

Asher, Nicholas, Vladimir Popescu, Philippe Muller, Stergos Afantenos, Anais Cadilhac, Farah Benamara, Laure Vieu & Pascal Denis. 2012. Manual for the analysis of settlers data. Manuscript.

Barraza, Georgina. 2003. *Evolución del objeto directo inanimado ein espanol*. Mexico, D.F.: Universidad Nacional Autónoma de México.

Bossong, Georg. 1982. Der präpositionale Akkusativ im Sardischen. In Otto Winkelmann (ed.), *Festschrift Johannes Hubschmid zum 65. Geburtstag. Beitrag zur allgemeinen, indogermanischen und romanischen Sprachwissenschaft*, 579–599. Bern: Francke

Bossong, Georg. 1985. *Empirische Universalienforschung: Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr. https://doi.org/10.5281/zenodo.4697660.

Bossong, Georg. 1998. Le marquage différentiel de l'objet dans les langues d'Europe. In Jack Feuillet (ed.), *Actance et valence*, 193–258. Berlin: De Gruyter.

Chiarcos, Christian. 2011. Evaluating salience metrics for the context-adequate realization of discourse referents. In *Proceedings of the 13th European Workshop on Natural Language Generation (ENLG), 32– 43.)*, 32–43. Association for Computational Linguistics. https://dl.acm.org/doi/pdf/10.5555/2187681.2187688.

Chiarcos, Christian, Berry Claus & Michael Grabski (eds.). 2011. *Salience: Multidisciplinary perspectives on its function in discourse*. Berlin: De Gruyter. https://doi.org/10.1515/9783110241020.

Chiriacescu, Sofiana. 2014. *The discourse structuring potential of indefinite noun phrases. Special markers in Romanian, German and English*. Stuttgart: Universität Stuttgart dissertation. http://dx.doi.org/10.18419/opus-5395.

Chiriacescu, Sofiana & Klaus von Heusinger. 2010. Discourse prominence and *pe*-marking in Romanian. *International Review of Pragmatics* 2(2), 298–332. https://doi.org/10.1163/187731010X528377.

Compensis, Paul, Marco García García, Birgit Hellwig, Sonja Riesberg & Gertrude Schneider-Blum. Submitted. Discourse prominence and special encoding in morphosyntax.

Comrie, Bernard. 1979. Definite and animate direct objects: A natural class. *Linguistica Silesiana* 3, 13–21. https://doi.org/10.5281/zenodo.3862697.

Comrie, Bernard. 1981/1989. *Language universals and linguistic typology: Syntax and morphology*. Chicago: University of Chicago Press.

Dalrymple, Mary & Irina Nikolaeva. 2011. *Objects and information structure*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511993473.

Emonds, Joseph. 2004. Unspecified categories as the key to root constructions. In David Adger, Cécile de Cat & Georges Tsoulas (eds.), *Peripheries: Syntactic edges and their effects,* 75–120. Dordrecht: Kluwer.

Enç, Mürvet. 1991. The semantics of specificity. *Linguistic Inquiry* 22, 1–25. http://www.jstor.org/stable/4178706.

Escandell-Vidal, Victoria. 2009. Differential object marking and topicality: The case of Balearic Catalan. *Studies in Language* 33(4), 832–885. https://doi.org/10.1075/sl.33.4.02esc.

Fábregas, Antonio. 2013. Differential Object Marking in Spanish: State of the art. *Borealis: An International Journal of Hispanic Linguistics* 2(2), 1–80. https://doi.org/10.7557/1.2.2.2603.
Falk, Simone. 2014. On the notion of salience in spoken discourse - prominence cues shaping discourse structure and comprehension. *TIPA. Travaux interdisciplinaires sur la parole et le langage* 30, 1–19. https://doi.org/10.4000/tipa.1303.

Fuchs, Melanie & Petra B. Schumacher. 2020. Referential shift potential of demonstrative pronouns – evidence from text continuation. In Åshild Næss, Anna Margetts & Yvonne Treis (eds.), *Demonstratives in discourse*, 185–213. Berlin: Language Science Press. https://doi.org/10.5281/zenodo.4055826.

García García, Marco. 2014. *Differentielle Objektmarkierung bei unbelebten Objekten im Spanischen*. Berlin: De Gruyter. https://doi.org/10.1515/9783110290974.

García García, Marco. 2018. Nominal and verbal parameters in the diachrony of differential object marking in Spanish. In Ilja A. Seržant & Alena Witzlack-Makarevich (eds.), *Diachrony of differential argument marking*, 209–242. Berlin: Language Science Presss. https://doi.org/10.5281/zenodo.1228257.

García García, Marco, Beatrice Primus & Nikolaus P. Himmelmann. 2018. Shifting from animacy to agentivity. *Theoretical Linguistics* 44 (1–2), 25–39. https://doi.org/10.1515/tl-2018-0002

García García, Marco & Javier Caro Reina. To appear. Differential object marking. In Matthias Heinz, Marc-Olivier Hinzelin & Lorenzo Filipponio (eds.), *Manual of classification and typology of the Romance languages*. Berlin: De Gruyter.

Gerards, David. 2023. Differential Object Marking in the Romance languages. In Michele Loporcaro (ed.), *Oxford Encyclopedia of Romance Linguistics*. Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.648.

Givón, Talmy. 1983. Topic continuity in discourse: An introduction. In Talmy Givón (ed.), *Topic continuity in discourse: A quantitative cross-language study*, 1–42. Amsterdam: Benjamins. https://doi.org/10.1075/tsl.3.01giv.

Grosz, Barbara, Aravind Joshi & Scott Weinstein. 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics* 21, 203–226.

Gundel, Jeanette, Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69(2), 274–307. https://doi.org/10.2307/416535.

von Heusinger, Klaus. 2019. Indefiniteness and specificity. In Barbara Abbott & Jeanette Gundel (eds.), *Oxford Handbook of Reference*, 146–167. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199687305.013.9.

von Heusinger, Klaus & Sofiana Chiriacescu. 2011. *Pe*-marked definite NPs in Romanian and discourse prominence. In Natascha Pomino & Elisabeth Stark (eds.), *Proceedings of the V Nereus International Workshop "Mismatches in Romance"* (Arbeitspapier 125), 33–54. Konstanz: Fachbereich Sprachwissenschaft, Universität Konstanz. http://nbn-resolving.de/urn:nbn:de:bsz:352-159309.

von Heusinger, Klaus, Marco García García & Georg A. Kaiser. 2024. Differential Object Marking and (non-)specificity in Spanish – Some empirical evidence for Leonetti's Generalization. In Silvia Gumiel, Cristina Sánchez & Ignacio Bosque (eds.), *Homenaje póstumo a Manuel Leonetti*, Alcala: Universidad de Alcalá. Preprint: https://semanticsarchive.net/Archive/GJlYTlmM/

von Heusinger, Klaus & Georg A. Kaiser. 2005. The evolution of differential object marking in Spanish. In Georg A. Kaiser, Klaus von Heusinger & Elisabeth Stark (eds.), *Proceedings of the Workshop "Specificity and the Evolution / Emergence of Nominal Determina-tion in Romance"* (Arbeitspapier 119), 33–69. Konstanz: Fachbereich Sprachwissenschaft, Universität Konstanz. https://d-nb.info/1079847138/34.

von Heusinger, Klaus & Georg A. Kaiser. 2011. Affectedness and Differential Object Marking in Spanish. *Morphology* 21(3–4), 593–617. https://dx.doi.org/10.1007/s11525-010-9177-y.

von Heusinger, Klaus & Petra B. Schumacher. 2019. Discourse prominence: Definition and application. *Journal of Pragmatics* 154, 117–127. https://doi.org/10.1016/j.pragma.2019.07.025.

von Heusinger, Klaus & Alina Tigău. 2024. Clitic Doubling and Differential Object Marking in Romanian as Marker for Discourse Prominence. Ms., University of Cologne.

von Heusinger, Klaus & Haydar Batuhan Yıldız. Submitted. The discourse function of Differential Object Marking in Turkish.

Himmelmann, Nikolaus P. & Beatrice Primus. 2015. Prominence beyond prosody – A first approximation. In Amedeo De Dominicis (ed.), *pS-prominenceS: Prominences in Linguistics. Proceedings of the International Conference*, 38–58. University of Tuscia, Viterbo: DISUCOM Press. https://kups.ub.uni-koeln.de/24935/

Iemmolo, Giorgio. 2010. Topicality and differential object marking. Evidence from Romance and beyond. *Studies in Language* 34(2), 239–272. https://doi.org/10.1075/sl.34.2.01iem.

Iemmolo, Giorgio. 2011 *Differential object marking*. Pavia: University of Pavia dissertation. https://zenodo.org/records/4388380.

Irimia, Monica Alexandrina & Anna Pineda. 2022. Differential object marking in Catalan. Descriptive and theoretical aspects. *Linguistic Variation* 22(2), 325–385. https://doi.org/10.1075/lv.20009.iri.

Jasinskaja, Katja, Sofiana Chiriacescu, Marta Donazzan, Klaus von Heusinger & Stefan Hinterwimmer. 2015. Prominence in discourse. In Amedeo De Dominicis (ed.), *pS-prominenceS: Prominences in Linguistics. Proceedings of the International Conference*, 134–153. Viterbo: DISUCOM Press. https://kups.ub.uni-koeln.de/9893/

Kabatek, Johannes, Philipp Obrist & Albert Wall (eds). 2021. *Differential Object Marking in Romance – The third wave*. Berlin: De Gruyter. https://doi.org/10.1515/9783110716207.

van Kampen, Jacqueline. 2007. Anaphoric topic-shift devices. Paper presented at the 'Workshop on anaphoric uses of demonstrative expressions', Siegen.

Kehler, Andrew, Laura Kertz, Hannah Rohde & Jeffrey L. Elman. 2008. Coherence and coreference revisited. *Journal of Semantics* 25(1), 1–44. https://doi.org/10.1093/jos/ffm018.

Kızılkaya, Semra. 2024. *Affectedness at the morphosyntax-semantics interface. Evidence from Differential Object Marking.* Berlin: De Gruyter.

Krifka, Manfred. 2007. Basic notions of information structure. In Caroline Féry, Gisbert Fanselow & Manfred Krifka (eds.), *Interdisciplinary Studies on Information Structure 6*, 13–55. https://www.jstor.org/stable/26190905.

Laca, Brenda. 1995. Sobre el uso del acusativo preposicional en español. In Carmen Pensado (ed.), *El complemento directo preposicional*, 61–91. Madrid: Visor.

Laca, Brenda. 2006. El objeto directo. La marcación preposicional. In Concepción Company Company (ed.), *Sintaxis historica del español. Primera parte: La frase verbal*, vol. 1, 423–475. Mexico: Fondo de cultura económica – Universidad Nacional Autónoma de México.

Ledgeway, Adam, Norma Schifano & Giuseppina Silvestri. 2019. Differential Object Marking and the properties of D in the dialects of the extreme south of Italy. *Glossa: a journal of general linguistics* 4(1), 51, 1–25. https://doi.org/10.5334/gjgl.569.

Leonetti, Manuel. 2004. Specificity and Differential Object Marking in Spanish. *Catalan Journal of Linguistics* 3, 75–114. https://doi.org/10.5565/rev/catjl.106.

Melis, Chantal. 1995. El objeto directo personal El Cantar de Mio Cid. Estudio sintáctico-pragmático. In Carmen Pensado (ed.), *El complemento directo preposicional*, 133–163. Madrid: Visor.

Melis, Chantal. 2021. From topic marking to definite object marking: Focusing on the beginnings of Spanish DOM. In Johannes Kabatek, Philipp Obrist & Albert Wall (eds.), *Differential Object Marking in Romance – The third wave*, 39–64. Berlin: De Gruyter. https://doi.org/10.1515/9783110716207-003.

Mürmann, Sophie. 2023. *Differential Object Marking and role semantics in Romance*. Köln: Universität zu Köln dissertation. https://kups.ub.uni-koeln.de/70360/.

Nikolaeva, Irina. 2001. Secondary topic as a relation in information structure. *Linguistics* 39, 1–49. https://doi.org/10.1515/ling.2001.006.

Nilsson, Birgit. 1985. *Case marking semantics in Turkish*. Stockholm: University of Stockholm dissertation.

Pensado, Carmen. 1995. El complemento directo preposicional. Estado de la cuestión y bibliografía comentada. In Carmen Pensado (ed.), *El complemento directo preposicional*, 11–60. Madrid: Visor.

Pensado, Carmen. 1995. La creación del complemento directo preposicional y la flexión de los pronombres personales en las lenguas románicas. In Carmen Pensado (ed.), *El complemento directo preposicional*, 179–233. Madrid: Visor.

R Core Team. 2023. _R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Romero Heredero, Diego. 2022. *Marcado diferencial de objeto y semántica verbal en español*. Berlin: De Gruyter. https://doi.org/10.1515/9783110751734.

Romero Heredero, Diego & Marco García García. 2023. Differential object marking in Spanish: The effect of affectedness. *Caplletra* 74(1), 259–285. https://doi.org/10.7203/caplletra.74.26043.

Sánchez, Liliana & Pablo Zdrojewski. 2013. Restricciones semánticas y pragmáticas al doblado de clíticos en el español de Buenos Aires y de Lima. *Lingüística* 29(2), 271–320.

Tigău, Alina. 2022. An experimental study on the discourse properties of Romanian direct objects. *Studii de lingvistică* 12(2), 85–107. https://studiidelingvistica.uoradea.ro/docs/12-2-2022/pdf-uri/Tigau.pdf.

Torrego, Esther. 1999. El complemento directo preposicional. In Ignacio Bosque & Violeta Demonte (eds.), *Gramática descriptiva de la lengua española. vol. 2*, 1779–1805. Madrid: Espasa Calpe.

Weissenrieder, Maureen. 1991. A functional approach to the accusative A. *Hispania* 74(1), 146–156.

Witzlack-Makarevich, Alena & Ilja A. Seržant. 2018. Differential argument marking: Patterns of variation. In Ilja A. Seržant, Alena Witzlack-Makarevich & Kelsey Mann (eds.), *Diachrony of differential argument marking*, 1–40. Berlin: Language Science Press. https://doi.org/10.5281/zenodo.1219168.