

Article

# The Discourse Function of Differential Object Marking in Turkish

Klaus von Heusinger \*  and Haydar Batuhan Yıldız 

Department of German Language and Literature I, Linguistics, Faculty of Arts and Humanities,  
University of Cologne, 50923 Cologne, Germany; haydar.batuhan.yildiz@uni-koeln.de

\* Correspondence: klaus.vonheusinger@uni-koeln.de

## Abstract

Differential Object Marking (DOM) is a cross-linguistic phenomenon in which the overt marking of direct objects of certain transitive verbs exhibits distinct morpho-syntactic properties. In Turkish, DOM is realized by the accusative suffix *-(y)I* and is considered to be determined by parameters such as referentiality/specificity, affectedness, and topicality. In addition, Enç argues that discourse-linking, which is a backward-looking discourse function, is another relevant parameter. In this paper, we investigate whether DOM also serves a forward-looking discourse function, which has remained underexplored. Using corpus studies and offline experiments, we investigate the forward discourse function of DOM in Turkish by analyzing the frequency of anaphoric expressions referring to the direct object with vs. without DOM. Corpus data show that non-modified human indefinite direct objects with DOM are taken up significantly more often in the subsequent discourse than those without DOM. However, forced-choice and paragraph continuation tasks do not support these observations. We evaluate various parameters that might contribute to the discourse prominence of direct objects with DOM and those that might mask such effects. We conclude that there is some corpus evidence that DOM contributes to a forward-looking discourse function, though our experimental methods may be inadequate to capture it.

**Keywords:** Differential Object Marking (DOM); discourse prominence; referential persistence; topic shift; corpus linguistics; experimental linguistics



Academic Editor: Jaklin Kornfilt

Received: 13 May 2024

Revised: 24 June 2025

Accepted: 1 July 2025

Published: 18 July 2025

**Citation:** von Heusinger, K., & Yıldız, H. B. (2025). The Discourse Function of Differential Object Marking in Turkish. *Languages*, 10(7), 173. <https://doi.org/10.3390/languages10070173>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Turkish, like many other languages, shows Differential Object Marking (DOM), i.e., the phenomenon that overt morphological marking of the direct object is obligatory in some conditions, optional in others, and ungrammatical for further conditions (Comrie, 1975; Bossong, 1985; Aissen, 2003, among many others). DOM is argued to be determined by nominal parameters such as animacy, referentiality/specificity, verbal parameters such as telicity and affectedness, and information structure and discourse parameters such as topicality. These parameters are often described as scales, and together, they contribute to the prominence of the direct object. Based on the interaction of these prominence features, a direct object can be characterized as more or less prominent (Erguvanlı, 1984; Dede, 1986; Enç, 1991; Kornfilt, 1997; Aydemir, 2004; von Heusinger & Kornfilt, 2005). In this paper, we extend the analysis of DOM from a sentence to a discourse perspective, which we investigate with respect to the notion of topic and d(iscourse)-linking.

A topical argument is typically definite, but indefinites can also become topics if they have one of the following properties: either they are referentially 'strong', i.e., specific or

presuppositional, or they are familiar, i.e., d(iscourse)-linked or partitive (see İşsever, 2003 for an overview). Enç (1991) combines these two perspectives and argues that DOM in Turkish signals specificity, which is understood as partitivity, as illustrated by her seminal example given below: (1a) introduces a set of some children, and in (1b), the case-marked *iki kız-ı* ('two girl-ACC') refers to two of the children mentioned in the first sentence, while the unmarked *iki kız* ('two girl') in (1c) refers to some other set of two girls.

- (1) a. Oda-m-a                      birkaç                      çocuk      gir-di.  
       room-1.SG-DAT<sup>1</sup>            several                    child        enter-PST  
       'Several children entered my room.'
- b. İki              kız-ı              tanı-yor-du-m.  
       two              girl-ACC        know-PROG-PST-1.SG  
       'I knew two [of the] girls.'
- c. İki              kız              tanı-yor-du-m.  
       two              girl              know-PROG-PST-1.SG  
       'I knew two girls.' (Enç, 1991, p. 6)

In addition to this backward-looking perspective involving the preceding discourse, Comrie ([1981] 1989) and Nilsson (1985) argue that DOM also has an effect on the subsequent discourse. Nilsson (1985, p. 66) reports that the sentence in (2b) with DOM was 'judged incomplete' by her informants, which implies a certain expectation that the referent is used in the subsequent discourse.

- (2) a. Dün                      bir                      kitap                      al-dı-m.  
       yesterday              a                      book                      buy-PST-1.SG  
       'I bought a book yesterday.'
- b. Dün                      bir                      kitab-ı                      al-dı-m.  
       yesterday              a                      book-ACC                      buy-PST-1.SG  
       'I bought a book yesterday.'

This forward-looking discourse function can also be understood as the effect of the combination of the function of a topical argument and a 'strong' indefinite argument. First, topical arguments are typically also the subjects of subsequent sentences, following the principle of 'topic continuity' (Givón, 1983) or Centering Theory (Grosz et al., 1995). Second, a case-marked indefinite direct object can also be understood as a strong indefinite. These types typically have a cataphoric function, i.e., the speaker introduces a discourse referent and signals that they will provide more information on this referent (see Prince, 1981; Kamp, 2014).

Cross-linguistically, descriptive evidence for a forward-looking discourse function of DOM in Turkish also comes from other DOM languages, which leads to the assumption that we find similar patterns across DOM languages. For instance, Laca (1995, p. 82) discusses this function for Spanish and concludes that 'the use of the preposition [i.e., DOM, von Heusinger and Yıldız] introduces an indefinite object as a possible discourse topic'. Additional empirical evidence supporting this hypothesis has been provided for DOM in Romanian through a paragraph continuation task (Chiriacescu & von Heusinger, 2010; Tigău, 2022). Hence, we can put forward the following hypothesis:

- (3) Hypothesis on the forward-looking discourse function of DOM  
 Indefinite direct objects with overt accusative case in Turkish show more forward-looking discourse potential than those without overt accusative case.

On the basis of this hypothesis, we predict that an indefinite direct object will more often be taken up in the subsequent discourse than that without DOM—this corresponds to the ‘referential persistence’ measurement of Chiriacescu and von Heusinger (2010), which is a reformulation Givón’s (1983) ‘topic continuity’.

To our knowledge, only two previous studies have addressed the discourse function of DOM in Turkish. Özge and von Heusinger (2020) analyzed corpus attestations of around 200 indefinite direct objects with and without the accusative case and found no significant interaction between case marking and referential continuation. Similarly, Özge et al. (2016) conducted a paragraph continuation task with indefinite human direct objects and also found no significant discourse effect of DOM. While these studies offer valuable initial insights, we believe that both studies have some shortcomings and their scope is limited; thus, further empirical work is necessary to evaluate the forward-looking discourse function of DOM in Turkish. We contribute to this discussion with a series of complementary studies: a systematically annotated large-scale corpus analysis and two experimental tasks (a forced-choice task and a paragraph continuation task), employing both sentence-level variation and larger discourse. We believe that our study advances the broader understanding of DOM within and across languages.

In Section 2, we provide a brief overview of the main parameters that determine DOM in Turkish. In Section 3, we expand the view to discourse prominence and discuss various observations with respect to backward-looking and forward-looking discourse functions of DOM in Turkish. Importantly, we formulate different measurements to test the forward-looking discourse function of DOM. In Section 4, we present four empirical studies: two corpus studies, a forced-choice task, and a paragraph continuation task. The general discussion in Section 5 compares the results from the corpus studies with those of the forced-choice and paragraph continuation tasks, and evaluates these findings in the light of the function of the indefinite article in Turkish, arguing that it might have a more pronounced discourse function than DOM. Section 6 provides the conclusion.

## 2. Differential Object Marking in Turkish

Differential Object Marking (DOM) is a cross-linguistically well-documented phenomenon. DOM languages show a systematic variation in the morphological marking of their direct objects. Some direct objects are obligatorily marked, some are optionally marked, and for some, the morphological marker is ungrammatical. The conditions that control this variation can be described as an interaction of the syntactic function and the prominence of the direct object (Bossong, 1991, p. 158). The latter has been defined in terms of nominal prominence hierarchies (animacy, referentiality), an information structure hierarchy (topicality), and verbal hierarchies (affectedness, telicity, semantic roles) (Aissen, 2003; Bossong, 1985; Dalrymple & Nikolaeva, 2011; Malchukov & de Hoop, 2011; Witzlack-Makarevich & Seržant, 2018).

Turkish, one of the most-discussed DOM languages, is a head-final language with dominant SOV word order and rich case and verbal morphology. It has no definite articles but it has the indefinite article *bir* and demonstratives. DOM in Turkish depends primarily on the Referentiality Scale and to a lesser degree on animacy and affectedness. Topicality, as an information structure parameter, is also very important, particularly in combination with d(iscourse)-linking or partitivity, as it will be discussed in Section 3.3.1.

### 2.1. DOM and Referentiality

DOM in Turkish is primarily organized along the Referentiality Scale (Aissen, 2003). All personal and demonstrative pronouns and proper names in direct object positions must be marked. Demonstrative and definite noun phrases must also be marked, as in (4a-b),

although we find variation for noun phrases with the indefinite article *bir*. A noun phrase with the indefinite article and overt accusative (*bir N +ACC*) is typically interpreted as specific indefinite, as in (4c), while a noun phrase with the indefinite article without case (*bir N*) is interpreted as non-specific indefinite. A bare noun (N) is non-referential and considered as a (pseudo-)incorporated expression, as in (4e) (Erguvanli, 1984; Dede, 1986; Enç, 1991; Kornfilt, 1997; Kelepir, 2001; Aydemir, 2004; Öztürk, 2005; Ketrez, 2005; von Heusinger & Kornfilt, 2005):

- (4) a. Ali bu mektub-u yaz-dı. demonstrative  
A. this letter-ACC write-PST  
'Ali wrote this letter'
- b. Ali mektub-u yaz-dı. definite  
A. letter-ACC write-PST  
'Ali wrote the letter.'
- c. Ali bir mektub-u yaz-dı. (specific) indefinite  
A. a letter-ACC write-PST  
'Ali wrote a certain letter.'
- d. Ali bir mektup yaz-dı. (non-specific) indefinite  
A. a letter write-PST  
'Ali wrote a letter.'
- e. Ali mektup yaz-dı. bare  
A. letter write-PST  
'Ali did some letter-writing.'

The examples above illustrate the Turkish instantiation of the Referentiality Scale for DOM in Turkish in the version of Erguvanli (1984, p. 18), as shown in Table 1.

Table 1. Referentiality conditions for DOM in Turkish.

	Referential		Non-Referential
Definite	Indefinite		bare N
	Specific	Non-Specific	
N +ACC	<i>bir N +ACC</i>	<i>bir N (+ACC)</i>	

All definite direct objects in Turkish, i.e., personal pronouns, proper nouns, demonstratives, and definite NPs, obligatorily take the accusative case, while bare nouns never take case. Indefinite NPs can take case or not, i.e., they are optionally case-marked.<sup>2</sup> We will consider, first, the contrast between the caseless forms N vs. *bir N* and, second, the contrast between the indefinite forms *bir N* vs. *bir N +ACC*. To begin with, there is an ongoing discussion about the syntactic form and semantic representation of the two caseless forms (see Aydemir, 2004; Kornfilt, 2003; Kamali, 2015; Seidel, 2019, 2024; Driemel, 2023; Ketrez, 2023; see also Arslan-Kechriotis, 2009 for an overview and discussion). Ketrez (2023, pp. 228–229) summarizes some syntactic differences between these forms. Here we will focus on their discourse-pragmatic functions. The bare direct object is non-referential (Erguvanli, 1984) or non-definite (Dede, 1986), in the sense that it does not introduce a discourse referent that can be picked up by a personal pronoun, as in (5), while the *bir N* form introduces a discourse referent that can then be picked up, as in (6), taken from Aydemir (2004, p. 468) in the version of Ketrez (2023, p. 229).

- (5) \*Dün film<sub>i</sub> seyret-ti-m, on-u<sub>i</sub>/on-lar-ı sen de seyret-meli-sin.  
 yesterday film watch-PST-1.SG that-ACC/that-PL-ACC you too watch-NEC-2.SG  
 ‘I watched movies/did movie-watching yesterday, you should watch it/them too.’
- (6) Dün bir film<sub>i</sub> seyret-ti-m, on-u<sub>i</sub> sen de seyret-meli-sin.  
 yesterday a film watch-PAST-1.SG that-ACC you too watch-NEC-2.SG  
 ‘I watched a movie yesterday, you should watch it too.’

Bare nouns are interpreted as pseudo-incorporated objects, i.e., they are expressions that fill the thematic role of the verb but do not introduce a discourse referent that can be taken up in the following discourse (Öztürk, 2005; Kamali, 2015; Seidel, 2019, 2024; Driemel, 2023). Such an analysis predicts that such nouns are less accessible to following anaphoric expressions and that the anaphoric relation is established in a different way from regular anaphors. The anaphoric relation between a pronoun and a bare noun antecedent is different from the relation between a pronoun and a (discourse) referential expression (see Seidel, 2019, 2024 for an analysis). We can summarize that both caseless forms share certain properties, but they differ in important syntactic behavior. They show similar semantic behavior, as both always have a narrow scope (Öztürk, 2005), but their discourse function is different: *bir* N forms introduce discourse referents that are easily accessible to subsequent pronouns, while bare nouns do not introduce such referents. There are, less frequently, anaphoric pronouns, but these access the bare noun by other principles. Since bare nouns do not introduce discourse referents, we do not consider them in our further analysis, but see Section 4.1.

Second, the contrast between the two indefinite forms *bir* N vs. *bir* N +ACC is of central interest for our research. This contrast has been associated with discourse-linking (Nilsson, 1985; Enç, 1991; Zidani-Eroğlu, 1997), presuppositionality (Kennelly, 1997; Keleşir, 2001), individuation/particularization (Nilsson, 1985; Kılıçaslan, 2006), and specificity (Erguvanlı, 1984; Dede, 1986; von Heusinger, 2002; von Heusinger & Kornfilt, 2005), among other approaches. In the following, we present some semantic contrasts based on different types of specificity. We assume three different types of specificity: (i) referential specificity for indefinite referential readings in intensional (or opaque) contexts (Fodor & Sag, 1982), (ii) scopal specificity for indefinites with the widest scope with respect to extensional operators, and (iii) epistemic specificity for indefinites that introduce discourse referents that the speaker of the sentence can identify (see Fodor & Sag, 1982; von Heusinger, 2019 for a comprehensive overview of different types of specificity).

Note that Enç (1991) associates DOM in Turkish with partitivity or d-linking and defines specificity in terms of partitivity. Özge (2011) argues that DOM expresses an implicit domain restriction and is, therefore, orthogonal to the other types of contrasts mentioned above. von Heusinger and Bamyacı (2017a) tested for the three types of specificity (referential in intensional contexts, scopal in contexts with universal quantifiers, and epistemic in transparent contexts) by presenting sentences and providing two possible continuations: one that was congruent with a specific reading and one that was congruent with a non-specific reading. During the experiment, participants had to choose one of the continuations. The results show that DOM significantly triggers a specific reading in intensional contexts, as illustrated in (7), but not in contexts with the universal quantifier *her* (‘every’), as in (8), or in transparent contexts, as in (9).

(7) *test item for referential specificity*

Zeynep parti için bir elbise(-yi) ara-dı.  
 Zeynep party for a dress(-ACC) look.for-PST  
 ‘Zeynep looked for a dress for the party.’

- (a) *referential specific*: This was one of a kind dress made for her size and taste.  
 (b) *referential non-specific*: She tried many dresses but none of them were beautiful enough.

(8) *test item for scopal specificity*

Her oyuncu bir kostüm(-ü) dene-di.  
 Every player a costume(-ACC) try-PST  
 ‘Every actor tried a costume.’

- (a) *scopal specific*: It was hard to find an actor who can wear this costume.  
 (b) *scopal non-specific*: The actors had to try hard to find a fitting costume for their roles.

(9) *test item for epistemic specificity*

Mustafa bir sandalye(-yi) al-dı.  
 Mustafa a chair(-ACC) buy-PST  
 ‘Mustafa bought a chair.’

- (a) *epistemic specific*: This is a very similar one to the rocking chair I bought last month.  
 (b) *epistemic non-specific*: But I have not yet seen what type of chair this is.

The results for referential specificity support observations in the literature, but the results for scopal specificity contradict general assumptions in the literature (Keleşir, 2001; von Heusinger & Kornfilt, 2005). The data are more in line with Aygen (2007), who argues that *her* (‘every’) is strongly distributive and, therefore, forces an interpretation where we have different values for the direct object for each subject. von Heusinger and Bamyacı (2017b) report a follow-up experiment with the universal quantifier *bütün* ‘all’, which shows that scopal specificity significantly affects DOM. This means that DOM marks specific indefinites in contexts with other operators, but in transparent episodic contexts, DOM does not trigger specificity. In transparent contexts, a specific reading depends on the pragmatics of the sentence, which is very difficult to search for in a corpus and also difficult to control in production experiments. We, therefore, did not annotate for specificity in our corpus studies or experiments.

## 2.2. Further Parameters for DOM in Turkish

Cross-linguistically, animacy is an additional factor that controls DOM, and we have evidence that it is also effective in Turkish, even though most of the literature does not discuss it. Dede (1986, pp. 158–160) provides evidence that the animacy of the direct object interacts with its case marking in intensional contexts (her ‘non-factive verbs’). According to her, a human indefinite direct object without overt accusative marking in an intensional context with the verb *aramak*, ‘to look for’, can receive a specific or a non-specific interpretation, as in (10a), but a case-marked one is only licensed for the specific interpretation, as in (10b). Inanimate indefinite direct objects are only licensed without case and can be interpreted as specific or non-specific, as in (11).

- (10) a. Bir öğrenci arı-yor-um. Bul-a-mı-yor-um.  
 a student look.for-PROG-1.SG find-PSB-NEG-PROG-1.SG  
 'I am looking for a student. I can't find him./I can't find one.'
- b. Bir öğrenci-yi arı-yor-um. Bul-a-mı-yor-um.  
 a student-ACC look.for-PROG-1.SG find-PSB-NEG-PROG-1.SG  
 'I am looking for a student. I can't find him./\*I can't find one.'
- (11) a. Bir kitap arı-yor-um. Bul-a-mı-yor-um.  
 a book look.for-PROG-1.SG find-PSB-NEG-PROG-1.SG  
 'I am looking for a book. I can't find it./I can't find one.'
- b. \*Bir kitab-ı arı-yor-um. Bul-a-mı-yor-um.  
 a book-ACC look.for-PROG-1.SG find-PSB-NEG-PROG-1.SG  
 \*'I am looking for a book. I can't find it./I can't find one.'

While Dede's grammaticality judgements are controversial and the correct analysis of the anaphoric null pronoun in the second clause is not resolved, Krause and von Heusinger (2019) provide empirical evidence that animacy is a graded feature interacting with DOM in Turkish. According to the authors, referents that are high in animacy are more likely to trigger the overt accusative case on direct objects. They tested the acceptability of sentences with case-marked and unmarked indefinite direct object referents of the three main animacy levels: (i) human, (ii) animal, and (iii) inanimate. The results of their study showed that the acceptability of human direct objects with DOM is higher than that of unmarked human direct objects, while this was reversed for inanimates and balanced for non-human animate direct objects. Summarizing the discussion on animacy and DOM in Turkish, we can say that DOM is dispreferred on inanimate direct objects irrespective of their referential status. The fact that the accusative case can be dropped, despite semantic specificity, strengthens the need to investigate additional parameters that contribute to, and are determined by, DOM.

Besides the nominal parameters of referentiality and animacy, verbal parameters such as affectedness also play a role. Affectedness refers to the change an argument undergoes through the event. Moreover, affectedness is a linguistic notion that figures centrally in determining transitivity and defining direct objecthood (Dowty, 1991; Hopper & Thompson, 1980; Tenny, 1994). For Næss (2004), it is the central notion behind DOM. Kızılkaya et al. (2022) provide experimental evidence that affectedness interacts with DOM. Moreover, Kızılkaya (2024) tested sentences with verbs governing an affected direct object, as in (12a), and verbs selecting a non-affected direct object, as in (12b), where participants had to choose between two alternative object realizations, one with the accusative case and one without it. The results from the forced-choice task show that verbs that select affected objects, as *to topple* in (12a), trigger significantly more often case marking on their direct objects than verbs that do not select affected objects, as *to find* in (12b).

- (12) a. Ayşe bir lider(-i) yık-tı.  
 Ayşe a leader(-ACC) topple-PST  
 'Ayşe toppled a leader.'
- b. Selim bir sekreter(-i) bul-du.  
 Selim a secretary(-ACC) find-PST  
 'Selim found a secretary.'

Accusative marking in Turkish also depends on syntactic conditions such as word order and modification. Only direct objects in the immediate preverbal position can lack case marking (Erguvanlı, 1984; Dede, 1986; Kornfilt, 1997; von Heusinger & Kornfilt,

2005). In all other positions, they have to take case marking, as illustrated in (13a-b) (von Heusinger & Kornfilt, 2005, p. 11), irrespective of whether they are specific or non-specific. The direct object *çay-ı* ('tea-ACC') stands in a position that is not close to the verb; therefore, it must take the accusative case. An unmarked form in this position is ungrammatical (see also Göksel & Kerslake, 2005, p. 178).<sup>3</sup>

- (13) a. Bizim ev-de çay-ı her zaman Aytül yap-ar.  
 our house-LOC tea-ACC always Aytül make-AOR  
 'Aytül always makes the tea in our family.'
- b. \*Bizim ev-de çay her zaman Aytül yap-ar.  
 our house-LOC tea always Aytül make-AOR  
 Intended reading: 'Aytül always makes the tea in our family.'

A final issue to be discussed here is the modification of the direct object. There is a close relationship between modification and specificity (see Fodor & Sag, 1982). The more a noun phrase is modified, the more it tends to be specific. This relation is illustrated by the following example from Erguvanlı (1984, p. 27; her examples [75] and [76]), assuming that only specific expressions can be in the topic position.

- (14) a. \*Bir kitab-ı Murat aceleyle oku-yor.  
 a book-ACC Murat hurriedly read-PROG  
 (Intended reading: 'Murat is hurriedly reading a/some book.')
- b. Mavi kaplı bir kitab-ı Murat aceleyle oku-yor.  
 blue covered a book-ACC Murat hurriedly read-PROG  
 'Murat is hurriedly reading a (certain) blue-covered book.'

The case-marked noun *bir kitab-ı* ('a book-ACC') is non-specific and, therefore, not licensed in the topic position. However, the modified noun *mavi kaplı bir kitab-ı* ('a blue-covered book-ACC') is specific and, therefore, licensed in the topic position. This observation supports the close relationship between modification and specificity. However, there are also examples where this relationship does not hold. Göksel and Kerslake (2005, p. 326) show that the contrast between a specific and a non-specific reading in an intensional context is marked by the accusative case, even when the indefinite is modified by a long expression (for similar examples see Kornfilt, 1997, p. 262).

- (15) a. {Gürcistan folkloruyla ilgili bir kitap} arıyorum.  
 'I'm looking for {a book about Georgian folklore}.'
- b. {Gürcistan folkloruyla ilgili bir kitab-ı} arıyorum.  
 'I'm looking for {a certain book about Georgian folklore}.'

To the best of our knowledge, there are no empirical studies that investigate the interaction between the kind of modification and DOM. Whether modification positively or negatively correlates with case marking or whether there is no relation is an open question. We will come back to this issue in Section 4.

Summarizing, syntactic and semantic parameters determine whether the direct object is marked or not. Definite expressions are always marked, bare nouns are never marked, and indefinite noun phrases, i.e., noun phrases with the indefinite article *bir*, are marked, depending on further parameters; when used non-preverbally, they must be marked, independently of their referential status. Modification seems to correlate with case marking, even though this is not fully investigated. Variation in case-marking is typically attested when the direct object appears in the preverbal position. There, case marking depends on animacy, specificity, and affectedness. While this section has discussed sentence-bound

parameters, the next section is devoted to discourse-bound parameters, such as backward-looking discourse functions (e.g., topicality, partitivity) and forward-looking functions. We will focus on indefinite direct objects in the preverbal position.

### 3. Differential Object Marking as a Discourse Function

We understand Differential Object Marking (DOM) as a result of a complex interaction of syntactic, semantic, and information structure parameters. This view is, however, restricted to the sentence structure, and in this section, we are extending it to discourse structure. We start by defining a notion of prominence that includes sentence and discourse prominence to describe different functions of DOM, in which we present different measures for testing discourse prominence. We move on to providing a brief cross-linguistic overview that deals with DOM as discourse functions in other languages such as Spanish and Romanian. This is followed by a subsection, in which we highlight that DOM in Turkish depends on the preceding context, documented in the marking of the topical function of direct objects or their partitive nature. We complement this backward-looking function of DOM with a forward-looking discourse function and suggest that DOM influences the referential structure of the subsequent discourse, on which we aim to elaborate using empirical data in the upcoming section.

#### 3.1. Measures of Discourse Prominence

We assume that the concept of prominence is able to provide a means to combine the sentential effects of DOM with the discourse effects and that DOM contributes to prominence relations in the sentence and prominence management in the discourse. In order to combine the concepts of local and global prominence, we apply the definition proposed by [von Heusinger and Schumacher \(2019\)](#) for discourse prominence, which is based on the general concept of prominence according to [Himmelmann and Primus \(2015\)](#). Discourse prominence is defined as a relation that (i) singles out one discourse referent among other 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 ([von Heusinger & Schumacher, 2019](#)). This concept of discourse prominence is similar to [Givón's \(1983\)](#) notion of *topic continuity*.

As far as discourse prominence is concerned, the literature has used different measures: [Givón \(1983\)](#) introduces the notion of topic continuity, which by itself covers a combination of various parameters such as the number of uptakes of an antecedent, the question of whether the antecedent becomes a topic, etc. Following [Chiriacescu and von Heusinger \(2010\)](#), we decompose Givón's notion of topic continuity into three measurable parameters. First, 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 it is picked up in the next sentence. Second, the topic shift parameter focuses on whether or not the critical referent continues to be part of the comment or shifts toward 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 the sentence. Third, the explicitness of the anaphoric expression, i.e., the size of an anaphoric expression, reflects the accessibility or activation of the antecedent (cf. [Ariel, 1990](#); [Gundel et al., 1993](#)).

- (16) Measures of 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) **Explicitness of the anaphoric expression:**  
The size of an anaphoric expression mirrors accessibility or activation of the antecedent.

While each of the measures described reflects one aspect of discourse prominence, we will primarily use referential persistence in our studies. First, this measure seems to be the most discourse- or thematic-oriented measure; the other two are more restricted by syntactic configuration. Second, even though we assume that the position of the topic in Turkish is sentence-initial, it can sometimes be difficult to identify a topic in a sentence, while finding an anaphoric expression is quite easy. Third, the explicitness of an anaphoric expression depends not only on the accessibility of the antecedent but also on the function of the anaphoric expression in its sentence, e.g., the pronoun rule in Centering Theory (Grosz et al., 1995; Song & Kaiser, 2023).

### 3.2. DOM in Other Languages as Discourse Functions

We present cross-linguistic evidence for a forward-looking discourse function of DOM. Laca (1995, p. 82) discusses this function for Spanish and concludes that the case marking introduces an indefinite object as a possible discourse topic.<sup>4</sup> Chiriacescu and von Heusinger (2010) report from a paragraph continuation experiment in Romanian with small contexts of four sentences. The pre-critical sentence contained a proper name as a subject, which is also the null subject of the critical sentences. Critical sentences also included an indefinite human direct object, in one condition with DOM and in the other without DOM, as in (17):

- (17) Stimulus item 1 (from Chiriacescu & von Heusinger, 2010, p. 317)
- Aseară a fost extraordinar de cald. Pentru că nu mai rezista în casă, Graur s-a hotărât să se ducă în oraș. Pe drum (I-) 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 these contexts with five coherent sentences, which were annotated for their anaphoric links, i.e., for the form and the function of the anaphoric expressions linked to the subject or the object of the preceding sentence. They found that (i) there were more anaphoric continuations for direct objects with DOM than for unmarked ones and that (ii) direct objects with DOM became topics in the subsequent discourse earlier and more often than their unmarked counterparts. Similarly, Tigău (2022) found that DOM and clitic doubling trigger a forward-looking discourse function in Romanian.

### 3.3. DOM as a Discourse Function in Turkish

There have been some early observations that DOM not only depends on the previous discourse but also affects the subsequent discourse. Comrie ([1981] 1989, pp. 129, 136) assigns a forward-looking function to DOM in Turkish:

*'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.'

Similar in spirit, Nilsson (1985, p. 66) elaborates on this observation and coins the notion of 'thematic prominence' for direct objects with DOM, as in (2), repeated as (18):

- (18) a. Dün bir kitap al-dı-m.  
yesterday a book buy-PST-1.SG  
'I bought a book yesterday.'
- b. Dün bir kitab-ı al-dı-m.  
yesterday a book-ACC buy-PST-1.SG  
'I bought a book yesterday.'

Nilsson (1985, p. 66) reports that (18b) was 'judged incomplete' by her informants; they would say that case marking makes the referent more prominent and signals that more is to be said about that referent (see also Nilsson, 1979, p. 125). The forward-looking discourse function can also be linked to the observation that topical indefinites must be strong or highly referential. According to Kamp (2014), we have at least three different types of indefinites: weak or incorporated indefinites, regular or existential indefinites, and strong or cataphoric indefinites. With a strong indefinite, the speaker signals to the hearer that more is to come from this referent. This function is realized by noun phrases with the indefinite article or unambiguously by the indefinite demonstrative in English (*And then I met this guy at the bar. He. . .*); see Prince (1981). Another aspect of topicality and discourse structure is that, according to Givón's (1983) concept of 'topic continuity', a topical element is the most likely argument to be reused in the next sentence. This is also one of the basic assumptions of Centering Theory (Grosz et al., 1995).

### 3.3.1. DOM, Topicality, D(iscourse) Linking, and Partitivity

Sentence-level parameters for DOM include referentiality and animacy, but cross-linguistically, the information-structural notion of topic is also a relevant parameter. The topic of a sentence is the argument or the element the sentence is about (Reinhart, 1981). This notion of an aboutness topic is complemented with the assumption that a topic is given, while the main contribution of its complement, the comment, contains new information. This account is also known as the familiarity theory of topic. Reinhart (1981), among others, assumes that referentiality is the central characteristic for topicality, while Kuno (1972) and Ward and Prince (1991) assume that familiarity is the central notion.

In Turkish, like in many other languages, topicality often clusters with definiteness and subjecthood, and the designated topic position is sentence-initial. However, Turkish licenses indefinite topics when they are either strong, i.e., specific or generic, or when they are discourse-linked, as in (19), taken from von Heusinger and Özge (2021).

- (19) (Dün gece onbeş kişi var-dı ya)  
(last night fifteen person there-PST PART)  
bir misafir sigara-sın-ı yanık unut-muş.  
a guest cigarette-3.SG-ACC burning forget-PERF  
'(There were fifteen people last night you know).  
A guest left his/her cigarette burning.'

The indefinite *bir misafir* ('a guest') is related to the set of 15 people and can act as topic (Erkü, 1983, p. 128). İşsever (2003, p. 1041) summarizes two accounts for topicality, as in (20).



set of verbs. They did not report any significant effect for the forward-looking function of DOM, possibly due to the high number of abstract objects in their dataset.

Another limitation was that the corpus they used in the study was the *Milliyet Corpus*, which had 21 million words and consisted of news articles of the daily newspaper *Milliyet* from the late 1990s. As the corpus focused on news articles, the direct objects throughout the study were often quite complex noun phrases with an abstract denotation (*this report, the position mentioned above, this panel, which will host. . .*). The referential properties of abstract objects (propositions, states, eventualities, properties, etc.) are not well understood. They typically score very low on the Referentiality and Animacy Scales. Since, to our knowledge, there are no studies investigating direct objects referring to such denotations, we will focus on concrete human and inanimate objects in our corpus study. Özge et al. (2016) investigated the referential persistence of human indefinite direct objects with or without DOM in a paragraph continuation task. In their stimuli, participants were given made-up news reports in the form of a discourse initiator, as in (22a), followed by the experimental item, as in (22b), designed to introduce discourse-new subjects and objects in ditransitive contexts. The direct objects were manipulated across  $\pm$ DOM.

- (22) a. Yeni markaların tanıtımına fırsat veren Nişantaşı moda günlerinin açılışı dün gece Safran Palas Oteli'nde yapıldı.  
'Nişantaşı fashion days, which serve for the introduction of new brands, started last night at Safran Palace Hotel.'
- b. Etkinliğin ilk defilesinde başarılı bir modacı izleyenlere **bir manken/bir manken-i** tanıttı.  
'In the first show of the night, a successful designer introduced a model/a model-ACC to the audience.'

Participants were asked to continue these two sentences with an additional continuation sentence. The continuation sentences were annotated with respect to the occurrence of the subject referent and the object referent. Across conditions, the subject was used in 8% of the continuations, the object in 41%, both in 21%, and neither in 30%. There was no significant effect of case marking. However, they were able to observe many more anaphoric links of the direct object than those of the subject. Therefore, they speculated, due to this particular design, that all direct objects were already highly salient and that case marking could not add to this high level of salience.

## 4. Empirical Studies

Given the theoretical background and observations outlined in previous sections, we have presented descriptive and theoretical arguments in favor of a forward-looking discourse function for DOM in Turkish and provided a cross-linguistic perspective on the contribution of DOM to discourse prominence. However, we lack empirical evidence for the discourse prominence of direct objects with DOM and their forward-looking discourse function. In this section, we present our empirical studies to test our assumptions and provide evidence for our theoretically and descriptively driven hypothesis. In doing so, although the literature employs different measures for discourse prominence, such as topic continuity, topic shift, and referential persistence (see Section 3.1), our focus will be on referential persistence, which is defined as the ability of an antecedent to be anaphorically retaken in the following discourse.

### 4.1. Pilot Study

To better understand the various properties of indefinite expressions with or without the accusative marking and gain an initial understanding of the observations as presented

in the research literature, we conducted a pilot corpus study using Sketch Engine (Kilgarriff et al., 2014), which provides access to morphologically annotated corpora in Turkish. We selected the trTenTen20 corpus (Çöltekin, 2013), which contains approximately 6 billion tokens, 5 billion words, and 370 million sentences. This corpus also provides metadata such as the source, the document number, and the token number for each attestation. We carried out phrase frequency analyses for the following types of indefinite expressions: (i) bare N; (ii) *bir* N  $\emptyset$ ; (iii) *bir* N +ACC; (iv) *bir* N +ACC (non-preverbal); (v) N +ACC (preverbal); see Ketrez (2023). We compared these five types in terms of their frequency of occurrence across the entire corpus; see Table 2 for the distribution of each type. The rationale is that we wanted to investigate the distribution of the kind of indefinite expressions across the referentiality/specificity scale (i.e., from non-referential and non-specific (bare N) to referential and specific (N +ACC)). Also, the reason why we were interested in the contrast between preverbal *bir* N +ACC and non-preverbal *bir* N +ACC is to test whether scrambled or topicalized direct objects are more prominent (i.e., more frequently picked up in the following discourse) than those placed preverbally.

**Table 2.** The total number of hits of the construction type in the corpus and the hits per 1 M tokens.

TYPE	Number of Hits	Per 1 M Tokens
bare N (preverbal)	364,225	59.26
<i>bir</i> N $\emptyset$ (preverbal)	84,343	13.72
<i>bir</i> N +ACC (preverbal)	1799	0.29
<i>bir</i> N +ACC (non-preverbal)	5773	0.94
N +ACC (preverbal)	48,585	7.91

A first observation is that there is a stark asymmetrical distribution of bare N, *bir* N  $\emptyset$ , and *bir* N +ACC. There are four times more bare N direct objects than *bir* N  $\emptyset$  direct objects. And there are 40 times more *bir* N  $\emptyset$  direct objects than *bir* N +ACC direct objects. To investigate whether these types of indefinite expressions are taken up in the subsequent discourse, we randomly sampled 100 tokens for each type, which we then extracted separately into a spreadsheet, giving us a total of 500 tokens. Note that, at this stage of investigation, we did not control for or take into account parameters, such as modification and animacy, that might play a role in the discourse function of DOM.<sup>5</sup> For the annotation of the extracted tokens, we looked for anaphoric re-mentions of antecedents and annotated them for NEXT MENTION as *yes* or *no* (or *NA* for irrelevant observations, such as clausal elements like *sessiz durmamayı seçtik*, ‘we chose not to remain silent’). See Table 3 for the results; *NA* tokens were removed, and percentages were calculated per row.

**Table 3.** The distributional results of the anaphoric uptakes.

TYPE	NEXT MENTION
bare N (preverbal)	26% (23/89)
<i>bir</i> N $\emptyset$ (preverbal)	47% (43/92)
<i>bir</i> N +ACC (preverbal)	58% (51/88)
<i>bir</i> N +ACC (non-preverbal)	57% (52/92)
N +ACC (preverbal)	45% (33/73)

To report from the distributional results per row, we observe that bare N is the least frequent to be picked up in the following discourse, which is expected given their nature of non-referentiality (see Seidel, 2019, 2024). In contrast, *bir* N +ACC is the most powerful construction for the forward-looking function of DOM, regardless of its placement; that is, topicalization of direct objects does not change the frequency of referential uptakes.

More importantly, the contrast between *bir* N  $\emptyset$  and *bir* N +ACC shows that the overt accusative case influences the discourse prominence of the direct object in the expected direction. Furthermore, a similar contrast is observed between bare N and *bir* N  $\emptyset$ , where the realization of the indefinite article *bir* seems to also contribute to a forward-looking discourse function.

Based on the observed distributions, we tested the following pairs for statistical significance:<sup>6</sup> (i) bare N vs. N +ACC and (ii) bare N vs. *bir* N  $\emptyset$  are significant, while (iii) *bir* N  $\emptyset$  vs. *bir* N +ACC is not. We argue that not only the indefinite article *bir* but also the accusative case affects the frequency of anaphorical uptakes in discourse (although this effect did not yield a statistical significance, we believe that further research is needed), which seems to emerge only when the direct object is placed in its preverbal position since dislocated direct objects do not change this effect. Therefore, in what follows, we want to report from the more elaborate corpus study, which we employ to test our hypothesis contrasting *bir* N  $\emptyset$  with *bir* N +ACC. Taking up our hypothesis in (3), we can now formulate an extended hypothesis, as in (23), focusing on indefinite human direct objects.

- (23) Hypothesis on the forward-looking discourse function of DOM  
Indefinite human direct objects with overt accusative case in Turkish show more forward-looking discourse potential than those without overt accusative case.

To test this hypothesis, we conducted a corpus study, a forced-choice task, and a paragraph continuation task, which will be presented in the next sections in the same order.

#### 4.2. Corpus Study

We conducted a larger corpus study to carefully examine the nature of DOM and its effects on discourse structure. Our objective was to gather preliminary insights into the discourse function of DOM prior to analyzing the relationship between DOM and speakers' preferences in experimental settings. We opted for the TS Corpus, v.2 ([https://cqpweb.tscorpus.com/cqpweb/ts\\_corpus\\_ver\\_2](https://cqpweb.tscorpus.com/cqpweb/ts_corpus_ver_2), accessed on 7 February 2022; e.g., Sezer, 2017). This corpus offers a collection of texts taken from various internet sources, such as online newspapers, forums, and blog writings, and includes around 500 million tokens and 5 million types. In comparison to the previous corpus study testing the discourse effects of DOM (Özge & von Heusinger, 2020), the TS Corpus allowed us to obtain more hits and, therefore, compile a corpus with a more balanced number of data per construction type.

##### 4.2.1. Study Design

For this study, we employed a search query using CQP syntax (referring to tools that determine words or tags, effectively using regular expressions in order to look out for, among others, distinct word beginnings or endings of specific corpus entries). Our search string consisted of 3 blocks. The first block was used to indicate the indefinite article *bir* in Turkish. The second block included information on which word endings we do not want (represented by the exclamation mark after the word attribute). These pre-eliminated word endings included suffixes such as case markings and their allomorphs, including dative, locative, ablative, genitive, and instrumental, except for the accusative. Via this method, we were able to obtain as many words with an accusative or null ending as possible, as these two were the ones we were interested in on the basis of our hypothesis. The third block of the search query was allocated to the verbs using the lemma attribute in the corpus. With no emantic classification besides transitivity, we restricted our attention to the following six verbs: *gör* ('see'), *bekle* ('wait'), *seç* ('choose'), *getir* ('bring'), *gönder* ('send'), and *sun* ('present'), as they were attested for direct objects with or without the accusative

case. We wanted to investigate the variation across individual verbs based on the potential discourse-prominence effect of DOM.

We extracted 79,020 hits with 100 words preceding and following the target words and exported them to a spreadsheet, manually filtering the list of attestations based on nouns that occurred with both levels of DOM; we simply ignored those cases that did not. For instance, if the noun *balık* ('fish') always occurred unmarked in the list of hits in the corpus search, we disregarded *balık*. However, if the noun *adam* ('man') had only 5 attestations in the ACC-marked context, then we also extracted 5 attestations for this noun in the zero-marked context.

In this way, we created a balanced number of data points. The rationale was that we wanted to sample a dataset that included an equal number of observations for each type and noun. Some nouns (e.g., *kişi* ('person') and *kadın* ('woman')) had more hits than others. In this case, we sampled the maximum number of attestations that we could analyze for a balanced sample. For example, the noun *kadın* ('woman') had 24 attestations in ACC-marked contexts; accordingly, we also sampled 24 attestations for this noun in the unmarked context. In total, we determined a sampling limit of 500 randomly extracted attestations.

Before analyzing texts for coherence, we observed that the number of animal and collective nouns is relatively lower than that of human and inanimate nouns. Thus, we left animal and collective nouns out of consideration, not only because of the low number of attestations but also because the situation with collective nouns is rather complicated, as they refer to abstract collections of concrete entities. Consider the example below in (24), which shows an instance of an ACC-marked collective noun, *birlik* ('unit').

- (24) Ø Biz argümanlarımızı tamamen askeri kriterleri dikkate alarak, NATO'nun ihtiyaç duyacağı, NATO'nun vereceği bütün görevleri yapabilecek **bir birlik-i sunuyoruz** dedi.  
*'We are presenting our arguments by taking into account purely military criteria and presenting a unit -ACC that will be able to perform all the tasks that NATO will need and that NATO will assign.'*

A note must be made here regarding the manipulation of data sampling. We designed our sampling strategy to balance the comparison between ACC-marked and zero-marked nouns in this study. Instead of merely reflecting the raw distribution of noun types in the corpus, we aimed to control frequency imbalances that could skew results. By limiting the number of examples per noun, we aimed at reducing the influence of high-frequency nouns and/or lexical effects of specific nouns on our findings. Although the corpus includes various noun types, we focused on a specific subset relevant to our research question, i.e., concrete nouns. This does not mean abstract nouns are irrelevant to accusative marking (see Özge & von Heusinger, 2020); rather, it means that including them would have added extra semantic and discourse-related factors that are beyond the scope of this study. Our approach ensures that the patterns observed are not just a result of corpus biases but provide a clear and interpretable analysis of accusative marking in the chosen noun types. This approach further permitted an analysis of by-noun variation, which justified our sampling method, resulting in neutrality for both types of direct objects with or without DOM.

We analyzed texts in the subsequent discourse to check whether the continuations were coherent. (25) illustrates an example from the dataset, where the discourse following the string in bold forms a coherent continuation.

- (25) [...] Bir bay, arabaların geçmeyişinden yararlanıp yeşil ışığı beklemeden geçmek istedi. [Karşıdan hızla gelen **bir araba-yı**] **görüncü** yarı yoldan geri döndü. Yukarıdaki sözü anımsadım. [...].  
 [...] A gentleman, taking advantage of the fact that there were no cars passing wanted to cross without waiting for the green light. **When he saw [a car-ACC coming from the opposite direction at a fast speed]**, he turned around halfway. I remembered the above quote. [...].

After disregarding three animal nouns and eleven collective nouns, and checking the dataset to eliminate irrelevant and incoherent observations, such as mismatches between before- and after-contexts, the final dataset consisted of 276 data points with 138 per construction type: 170 tokens for human referents (85 for each type of direct object, i.e., *bir N +ACC* and *bir N Ø*) and 106 tokens for inanimate referents (53 for each type of direct object, i.e., *bir N +ACC* and *bir N Ø*) were selected for annotation.

#### 4.2.2. Annotation

We annotated the cleaned dataset for the independent variables of DOM, ANIMACY, and NOUN MODIFICATION and for NEXT MENTION as the dependent variable. DOM has two levels: +ACC vs. Ø. ANIMACY was considered with the levels human and inanimate, which had already been annotated during the data sampling. NOUN MODIFICATION is a predictor that we wanted to test in interaction with ANIMACY on the discourse effect of DOM. We annotated NOUN MODIFICATION because we argue that it is an independent variable that interacts with DOM, which we annotated as follows: no modification (e.g., *a car*), adjectival modification (e.g., *a crashed car*), or relative clause (e.g., *a car coming from the opposite direction at a fast speed*).

In line with our hypothesis, we predicted that the probability of ACC-marked direct objects being picked up in the following discourse is higher than that of unmarked direct objects. We also examined whether the syntactic modification of the direct object with the accusative marking interacts with the animacy of the direct object in predicting the probability of the antecedent in question being retaken in the subsequent discourse, i.e., NEXT MENTION.

- (26) Ø<sub>1</sub> koltuk değnekleriyle zorlukla yürümeye çabalayan 14 yaşında bir kız-<sub>12</sub> görünce, kendi<sub>1</sub> tekerlekli sandalye[sin]<sub>1</sub>-i ona<sub>2</sub> hediye etti.  
 When he<sub>1</sub> saw a 14-year-old girl-ACC<sub>2</sub> struggling to walk with crutches, he<sub>1</sub> gave her<sub>2</sub> his<sub>1</sub> own wheelchair as a gift.

(26) shows an instance of an ACC-marked human referent that is modified by a relative clause. The discourse following the occurrence of this referent shows that the referent is picked up. Example (27) provides an instance of a non-modified unmarked inanimate referent, which is used as an indefinite NP. The discourse following the occurrence of this referent has an anaphorical uptake of this referent by a definite NP in direct object position.

- (27) Eczaneci **bir paket** getirmiş. Adam **paket-i** açarak bir tanesini ambalajından çıkarmış.  
 The pharmacist brought a package. The man took one out of the wrapper by opening the package-ACC.

### 4.2.3. Results and Discussion

This section reports the results of the distributional and statistical analyses. Table 4 below shows the distribution of ANIMACY across case and non-case marking. We analyzed a total of 276 items, of which 67% (186 out of 276) re-mentioned the referent of the direct object. We also see that case-marked noun phrases are picked up more often than non-case-marked ones (71% vs. 64%). This contrast is more pronounced for human direct objects than for inanimate objects.

**Table 4.** The distributional results of the interaction between DOM and ANIMACY on NEXT MENTION.

	Human	Inanimate	Total
<i>bir</i> N +ACC	73% (62/85)	68% (36/53)	71% (98/138)
<i>bir</i> N ∅	64% (54/85)	64% (34/53)	64% (88/138)
total	68% (116/170)	66% (70/106)	67% (186/276)

The results above demonstrate a numerical contrast regarding the discourse function of DOM, with a 7% difference in favor of our prediction that *bir* N +ACC direct objects are more often picked up in the subsequent discourse than unmarked ones. If we compare these results with those presented above in Table 3, we observe a similar contrast (47% vs. 58%). Both differences, however, were not statistically significant. Across ANIMACY, the results show that the effect of human referents with DOM is the strongest. As a result, we can conclude that ACC-marked human direct objects are more often re-mentioned in the subsequent discourse than unmarked ones (73% vs. 64%). Although this contrast was statistically not significant, it provides initial support for our hypothesis.

As already motivated in previous sections, we wanted to see whether the effect of ACC-marked human direct objects is in interaction with NOUN MODIFICATION. Tables 5 and 6 below illustrate the three-way interaction effect of DOM, ANIMACY, and NOUN MODIFICATION, where we first considered human referents across NOUN MODIFICATION, as in Table 5, and then inanimate referents, as in Table 6.

**Table 5.** The distributional results of the interaction between DOM and NOUN MODIFICATION on NEXT MENTION for human objects.

Human	No Mod	Adj Mod	Rel. Clause	Total
<i>bir</i> N +ACC	88% (14/16)	62% (16/26)	74% (32/43)	68% (62/85)
<i>bir</i> N ∅	53% (9/17)	66% (21/32)	67% (24/36)	64% (54/85)
total	70% (23/33)	64% (37/58)	71% (56/79)	68% (116/170)

**Table 6.** The distributional results of the interaction between DOM and NOUN MODIFICATION on NEXT MENTION for inanimate objects.

Inanimate	No Mod	Adj Mod	Rel. Clause	Total
<i>bir</i> N +ACC	63% (15/24)	67% (8/12)	76% (13/17)	73% (36/53)
<i>bir</i> N ∅	75% (15/20)	55% (11/20)	62% (8/13)	64% (34/53)
total	68% (30/44)	59% (19/32)	70% (21/30)	66% (70/106)

The results show that there is a remarkable difference between *bir* N +ACC and *bir* N ∅ with non-modified human referents on the forward-looking discourse function (cf. 88% vs. 53%), which turned out to be statistically significant.<sup>7</sup> The other distributional results across the three-way interaction of DOM, ANIMACY, and NOUN MODIFICATION did not yield any statistical significance. This means that human ACC-marked direct objects are

more likely to be re-mentioned in the following discourse when they are non-modified. In summary, the corpus study shows that non-modified human indefinite direct objects have a more forward-looking discourse function than those without DOM. However, this effect is only visible for unmodified human objects. Since the corpus study had limitations due to being unbalanced for animacy, modification, and the number of referents, further research is needed to explore the discourse function of DOM. Based on the results of our corpus study, which shows a significant discourse-prominence effect for non-modified human indefinite direct objects, we wanted to focus on direct objects with these attributes in the following experiments.

#### 4.3. Forced-Choice Experiment

Given the results of the corpus study, which showed that non-modified human indefinite ACC-marked direct objects are more discourse prominent, i.e., they license the forward-looking discourse function of DOM, we now report from a forced-choice experiment, focusing only on non-modified human indefinite direct objects.

##### 4.3.1. Study Design

In this questionnaire, we used a two-factorial experimental design with two conditioning factors: DOM and NUMBER OF REFERENTS. The former refers to the direct object (i.e., the critical item) being either marked by +ACC or  $\emptyset$  (zero); the latter refers to the number of competitors implemented within the text, controlled by certain considerations such as sentence structure, definiteness, and placement of the human referents. The rationale is, as observed by the corpus study, that the sentences we had analyzed included a variety of sentences with or without additional arguments; therefore, we wanted to control for the number of human referents in our paragraphs (i.e., two-competitor vs. three-competitor) to have a better understanding of whether the number of previously introduced human referents interacts with the effect of DOM on the choice of continuations, i.e., the choice of subject referents.

We prepared 12 critical items with short stories composed of around 40 words each and three sentences per paragraph to employ these parameters. We only used non-modified human singular referents as competitors and avoided proper names and collective nouns, possessive constructions, and kinship terms. The last sentence of each item was used in transitive contexts (i.e., a discourse topic plus a direct object that we controlled across DOM). We had four conditions manipulated across our factors (i.e., a  $2 \times 2$  design); that is to say, +ACC vs.  $\emptyset$  and two-competitor vs. three-competitor contexts (NUMBER OF REFERENTS). Regarding the selection of verbs for the transitive contexts, we opted for similar verbs, as in the corpus query, which selectively mark their direct objects, i.e., either marked or unmarked: *seç* ('to choose'), *görevlendir* ('to hire'), *fark et* ('to notice'), *ara* ('to look for'), *bekle* ('to wait'), and *gör* ('to see'), which were used in two different lexicalizations, giving us a total of 12 critical items.

The most important criteria we paid attention to for creating the experimental items are that the main topic was introduced as a definite NP in the subject position of the first sentence and then taken up by null subjects in the following sentence. The critical item, as the direct object, is always used as an indefinite non-modified NP, and the third referent is used in the second sentence of each paragraph as a definite NP. The lexicalization of the forced choice answers is designed to prevent ambiguity and increase naturalness. Note that the reason why we solely considered full NPs (e.g., *avukat* 'lawyer') lies in the fact that full NPs are the unmarked form of anaphoric uptakes. In contrast, null subjects strongly prefer subject antecedents, and overt pronouns are highly marked anaphoric expressions.

Examples (28) and (30) below show items from our stimuli, representing two- and three-competitor contexts in accordance with our experimental design and criteria. Examples (29) and (31) below demonstrate the different continuation sentences for the preceding examples, out of which the participants had to choose one. All continuations are congruent with the previous discourse.

(28) *Critical item with two-competitor context (Ref1 and Ref2)*

**Savcı<sub>1</sub>** adliyeye vardığında, sabahın erken saatleri olmasına rağmen kendisini çok yorgun hissetti ve günün hemen bitmesini umdu. Senelerdir kapanmamış bir dava dosyasıyla ve kafasında sorularla odasına doğru ilerledi. Adliye koridorunda yürürken **bir avukat<sub>2-1</sub>/bir avukat<sub>2</sub>** gördü.

*'When the prosecutor<sub>1</sub> arrived at the courthouse, although it was early in the morning, he felt exhausted and hoped the day would soon be over. He proceeded to his room with a case file that had not been closed for years and questions in his head. While he was walking in the courthouse corridor, he saw an attorney<sub>2-ACC</sub>/an attorney<sub>2-Ø</sub>.'*

(29) *Forced-choice items for (28)*

- a. **Savcı<sub>1</sub>** bir kahve otomatından kahve almak için cebinden bozuklukları çıkardı. *The prosecutor<sub>1</sub> took out coins from his pocket to buy coffee from a coffee vending machine.*
- b. **Avukat<sub>2</sub>** bir kahve otomatından kahve almak için cebinden bozuklukları çıkardı. *The attorney<sub>2</sub> took out coins from his pocket to buy coffee from a coffee vending machine.*

(30) *Critical item with three-competitor context (Ref1, Ref2, and Ref3)*

**Apartman yöneticisi<sub>1</sub>** Mayıs ayı aidatlarını toplamak için evden çıktığında alt kattan gelen bir ses duydu. Sakin bir apartmanda yaşadığı için bu sesi garipsedi, alt katta yaşayan **öğretmen<sub>3</sub>** ile merdivenlerde laflayarak aşağıya indi. Bina önüne vardığı anda **bir satıcı<sub>2-y1</sub>/bir satıcı<sub>2</sub>** gördü.

*'When the apartment manager<sub>1</sub> left the house to collect the May dues, he heard a noise from the ground floor. The voice was strange because he lived in a quiet apartment and he went downstairs while chatting on the stairs with the teacher<sub>3</sub> from downstairs. When he arrived in front of the building, he saw a seller<sub>2-ACC</sub>/a seller<sub>2-Ø</sub>.'*

(31) *Forced-choice items for (30)*

- a. **Apartman yöneticisi<sub>1</sub>** binanın önünde bir sigara içti. *The apartment manager<sub>1</sub> smoked a cigarette in front of the building.*
- b. **Satıcı<sub>2</sub>** binanın önünde bir sigara içti. *The salesman<sub>2</sub> smoked a cigarette in front of the building.*
- c. **Öğretmen<sub>3</sub>** binanın önünde bir sigara içti. *The teacher<sub>3</sub> smoked a cigarette in front of the building.*

For both filler and control items, we created similar sentences as the critical items. For filler items, we used subject-experiencer verbs *anımsa* ('recall'), *merak et* ('worry about'), and *hatırla* ('remember') and object-experiencer verbs *şaşırt* ('surprise'), *hüzünlendir* ('sadden'), and *etkile* ('impress'); for control items, we used source-goal verbs *ver* ('give'), *özetle* ('summarize'), and *gönder* ('send'); and goal-source verbs *kirala* ('rent'), *devral* ('take over'), and *al* ('take'), as illustrated in (32) and (34), respectively. The continuations of the filler items were both correct answers, as in (33), but those of the control items were created to judge whether the participants had given enough attention to the task, as in (35b), where the referent in italics is the only correct answer to the choice given.

- (32) *Filler item with two-competitor context*  
**Matematik profesörü<sub>1</sub>** bir pazartesi sabahı odasında çayını yudumlayarak haftalık ders programını hazırladı. O gün yarı dönem sınav sonuçlarını açıklaması gerektiği için fakülteye biraz erken gitti, tabii o sırada bina bomboştu. Bu titizlik ve çalışkanlığıyla **bir öğrenci<sub>2</sub>-yi** oldukça şaşırttı.  
*'The math professor<sub>1</sub> prepared his weekly syllabus while sipping tea in his room on a Monday morning. He went to the faculty a little early because he had to explain his half-term exam results that day; of course, the building was empty at the time. With this diligence and hard work, he surprised a student<sub>2</sub> -ACC quite a bit.'*
- (33) *Continuation item for (32)*  
 a. **Matematik profesörü<sub>1</sub>** üzerine o gün kalın bir palto giymişti.  
*The mathematics professor<sub>1</sub> was wearing a thick coat that day.*  
 b. **Öğrenci<sub>2</sub>** üzerine o gün kalın bir palto giymişti  
*The student<sub>2</sub> was wearing a thick coat that day.*
- (34) *Control item with three-competitor context*  
**Emlakçı<sub>1</sub>**, bürosunda maillere cevap verip internette yeni ev ilanlarını paylaştı. Öğleden sonra rahatsız olan **annesini<sub>3</sub>** arayıp bir ihtiyacı olup olmadığını sordu. Aynı gün büyük bir çalışma motivasyonu ile imara yeni açılan arazileri gezip **bir arsa<sub>2</sub>** devraldı.  
*'The realtor<sub>1</sub> responded to emails in his office and shared new house ads on the internet. In the afternoon, he called his mother<sub>3</sub>, who was unwell, to ask if she needed anything. On the same day, with great motivation to work, he toured the newly opened plots and took over a plot<sub>2</sub>.'*
- (35) *Continuation item for (34)*  
 a. **Emlakçı<sub>1</sub>** şehir merkezine biraz uzak olduğu için çok ucuza gelmişti.  
*The realtor<sub>1</sub> was very cheap because he was a little far from the city center.*  
 b. **Arsa<sub>2</sub>** şehir merkezine biraz uzak olduğu için çok ucuza gelmişti.  
*The plot<sub>2</sub> was very cheap because it was a little far from the city center.*  
 c. **Emlakçının annesi<sub>3</sub>** şehir merkezine biraz uzak olduğu için çok ucuza gelmişti.  
*The realtor's mother<sub>3</sub> was very cheap because she was a little far from the city center.*

All test items and continuation items can be accessed via the following OSF link:  
[https://osf.io/gfzp3/?view\\_only=39c2cc027d464b958902db139838ac8f](https://osf.io/gfzp3/?view_only=39c2cc027d464b958902db139838ac8f).

#### 4.3.2. Participants and Task

We created four lists, as we have two factors, each with two levels. Each unique condition was assigned to one of the lists, so that each item with its assigned condition appeared only once in each list. In the questionnaire, each list we created consisted of 24 items, with 12 critical, 6 filler, and 6 control items, which we implemented using Qualtrics (<https://www.qualtrics.com/>). The experiment was expected to take around 20 minutes. At the very beginning of the experiment, participants were instructed that they had only one choice among the given continuations and that they should mark the first option that came to their mind. We launched the experiment using Prolific (<https://www.prolific.com>).

A total of 60 monolingually raised Turkish speakers participated in the task. We paid each participant GBP 10.50 per hour after the task was completed. Two participants completed the task in less than 2 min; therefore, we removed their responses from the dataset. A total of 58 participants provided a total of 696 responses for the critical items.

#### 4.3.3. Results and Discussion

According to our hypothesis, we expected to obtain more responses for ACC-marked direct objects than for unmarked direct objects, represented as the discourse effect of DOM. We also expected to observe whether this effect differs based on the number of referents (i.e., two-competitor and three-competitor contexts); that is to say, we tested for the interaction effect of DOM and CONTEXT on the RESPONSE.

The distributional results, however, do not support that DOM signals the discourse prominence of the direct object; in fact, the choice of Ref1 (subject referents) and Ref2 ( $\pm$ ACC-marked direct objects) is equally distributed. There are only very small numerical differences between the context with a subject and an object, on the one hand (see Table 7) and a subject, a direct object, and a further argument, on the other (see Table 8).

**Table 7.** The distributional results of the effect of DOM on NEXT MENTION for TWO-COMPETITOR CONTEXTS (subject = Ref1 and direct object = Ref2).

Two-Competitor Contexts		
	Ref1	Ref2
bir N +ACC	62% (108/174)	38% (66/174)
bir N $\emptyset$	59% (103/174)	41% (71/174)

**Table 8.** The distributional results of the effect of DOM on NEXT MENTION for THREE-COMPETITOR CONTEXTS (subject = Ref1, direct object = Ref2, and further argument = Ref3).

Three-Competitor Contexts			
	Ref1	Ref2	Ref3
bir N +ACC	52% (91/174)	37% (64/174)	11% (19/174)
bir N $\emptyset$	55% (95/174)	34% (59/174)	11% (20/174)

We observed that ACC-marked indefinite objects in two-competitor contexts are less often taken up than unmarked ones (Table 7), while those in three-competitor contexts are more often taken up than the unmarked ones. This numerical difference is in line with our prediction. As the distributional analysis summarized above cannot take random effects into account, we tested our hypothesis using a Bayesian mixed-effects model using brms (Version 2.21.0, e.g., Bürkner, 2017). The choice of such a model lies in the fact that it is more capable of dealing with convergence issues (e.g., Winter & Grice, 2021). Specifically, we fitted random slopes and random intercepts for by-participant, by-item, and by-verb variation as a full random effects structure, together with the interaction effect of DOM and CONTEXT as a function of RESPONSE. The model did not reveal any statistical significance. Nevertheless, we observed that the by-verb variation shows an interesting contrast between the verbs *seç*, ‘to choose’, and *görevlendir*, ‘to hire’, the former being in the positive direction for our hypothesis.

The forced-choice task showed, in two-competitor contexts, that about 40% of direct objects are taken up in the subsequent discourse, while this accounts to 60% for subjects. In three-competitor contexts, the additional argument was taken up in 11% of the continuation sentences. We think that this distribution mirrors the prominence relation between the arguments in the sentence. However, we could not find support for our hypothesis that direct objects with DOM show a much higher level of prominence. We speculate that this could have been caused by the nature of the task. We, therefore, designed a paragraph continuation task, which might reveal a more fine-grained interaction between the number of referents in the sentence and the marking with DOM.

#### 4.4. Paragraph Continuation Experiment

As the first experiment was conducted using a forced-choice task, we wanted to test our hypothesis further using a production experiment. We wanted the participants to continue the items in their own words rather than force them to choose from continuations that had already been prepared. All test items and statistical analysis, together with the dataset annotated for referential uptakes, can be accessed via the following OSF link: [https://osf.io/gfzp3/?view\\_only=39c2cc027d464b958902db139838ac8f](https://osf.io/gfzp3/?view_only=39c2cc027d464b958902db139838ac8f).

##### 4.4.1. Study Design

In this production experiment, we utilized the same design as the forced-choice task. This time, we had no control items but only filler items. We had four lists and 18 items per list, with 12 critical and 6 filler items.

##### 4.4.2. Participants and Task

The experiment was designed to have two tasks and was expected to take around one hour. A total of 80 participants provided a total of 960 responses. We created four Word documents for each list, displaying randomly ordered items. Our participants were monolingually-raised Turkish speakers and university students to whom we reached out through emails, phone calls, and video calls, inviting them to participate in the experiment. For the first task, we collected continuations from each participant, which were then analyzed for ambiguity in cases where participants referred to the competitors with overt and covert pronouns.

The second task was based on the ambiguous responses. Our priority for clarification was the subject of the continuation. If both the subject and the object were unclear, we only asked for the subject and retrieved the object from the answer. We asked additional filler questions to examine whether the participants were providing reliable answers, even if the continuation was clear. We let them know that there would not be a second task if there was no ambiguity. (36a) shows a continuation sentence for item (37) below, provided by one of the participants, in which the subject *pro* and the second argument used as a personal pronoun are ambiguous; (36b) shows one of our questions to disambiguate the covert pronoun.

- (36) a. Continuation sentence for (37) below  
 Ø Bir an onunla konuşma ihtiyacı hissetti.  
 ‘For a moment, *he/she* needed to talk to *her/him*.’
- b. Question: Who needed to talk?  
 Answer: The prosecutor.

##### 4.4.3. Annotation

After collecting data and disambiguating the cases as illustrated above, we annotated the continuations based on the following variables: NEXT-MENTION, FORM, and FUNCTION. FORM and FUNCTION, as defined in terms of the morphosyntactic form and the grammatical function of the anaphoric expression, are annotated with levels such as definite NP, demonstrative NP, personal pronoun, *pro*, and further forms for FORM, and subject, direct object, indirect object and experiencer datives, and further functions for FUNCTION. Note that, although these variables are annotated for each competitor, we wanted to restrict our attention to Ref2 only in this paper. (37) is one of our items showing a context with two competitors and a marked Ref2; (38) provides two continuation sentences for (37) that are given by different participants.

- (37) *Critical item with two-competitor context*  
**Moda tasarımcısı<sub>1</sub>** hazırladığı şahane elbiselerin en güzel şekilde sergilenmesi için kuliste savaş veriyordu. Tasarımını yaptığı son elbisenin kimin üstünde güzel duracağını planlamaya fırsatı bile olmadığından hızlı ve doğru bir karar almalıydı. Vakit kaybetmeden **bir manken<sub>2</sub>-i** seçti.  
*'The fashion designer<sub>1</sub> was fighting backstage to display the wonderful dresses she prepared in the most beautiful way. Since she didn't even have time to plan on whom the last dress she designed would look good, she had to make a quick and accurate decision. Without wasting time, she chose a model<sub>2</sub> -ACC.'*
- (38) Continuation items for (37)
- Ancak manken<sub>2</sub> elbiseyi giyerken yanlışlıkla yırttı.  
*'However, the model accidentally tore the dress while wearing it.'*
  - ø<sub>1</sub> İşlerin aksamasından nefret ederdi, çok telaşlıydı.  
*'She hated the disruption of work, she was very hectic.'*

In (38a), Ref2, *model*, is picked up in the sentence in the form of a definite NP and as the subject of the sentence; in (38b), Ref1 is picked up, realized as a null subject *pro*, which we disambiguated following the second task as illustrated above in (36).

#### 4.4.4. Results and Discussion

The results of the paragraph continuation task show that the interaction of DOM and CONTEXT has no visible effect on the re-mention of Ref2; there is no contrast between two-competitor and three-competitor contexts with regard to the discourse function of DOM. Although most responses consist of re-mentioned direct objects, i.e., 76% (725/960), the experiment does not support the hypothesis that DOM has a forward-looking discourse function, as evidenced by Table 9.

**Table 9.** The distributional results of the interaction between DOM and CONTEXT.

Two-Comp. Contexts	Ref2	Three-Comp. Contexts	Ref2
bir N +ACC	80% (191/240)	bir N +ACC	72% (172/240)
bir N ø	79% (189/240)	bir N ø	72% (173/240)

The results of the paragraph continuation task showed that about 80% of direct objects are taken up in the subsequent discourse if the direct object only competes with the subject and about 70% if it competes with the subject and another argument. Thus, the direct objects are reused similarly often as in the corpus study and more often than in the forced-choice task (which had about 40%). Note, however, that in the forced-choice task, we only measured subject continuation, while in the other tasks, we measured any continuation, following the definition of 'referential persistence'. Like in the forced-choice task, we do not observe a significant difference between the condition with DOM and the condition without it.

Independent of the type of context (i.e., two-competitor vs. three-competitor), we were interested in looking at the form and the function of the uptaken referent. As part of FORM, a variety of levels emerged, including definite NPs, demonstrative NPs, personal pronouns, and null arguments (which we annotated as *pro*). As part of FUNCTION, we looked at grammatical functions such as subject, direct object, and indirect object. The distributional results regarding the form of uptaken referents are as follows. *bir N +ACC* was taken up in 363 continuations with the following distributions: definite NPs (63%); demonstrative NPs (8%); personal pronouns (8%); *pro* (9%), and other forms (12%). *bir N ø* was taken up

in 362 continuations with the following distributions: definite NPs (66%); demonstrative NPs (6%); personal pronouns (7%); *pro* (8%); and other forms (13%). The distributional results regarding the function of uptaken referents are as follows. *bir* N +ACC was taken up with the following distributions: subject (56%); direct object (7%); dative object (20%); and further functions (17%). *bir* N  $\emptyset$  was taken up with the following distributions: subject (50%); direct object (10%); dative object (22%); and further functions (18%). While none of these contrasts are statistically significant, we do observe a marginal tendency for ACC-marked direct objects to be picked up as subjects in the following discourse, indicating a topic shift potential.

In summary, to test our hypothesis that direct objects with DOM show more forward-looking discourse effects than those without DOM, we undertook a corpus search, a forced-choice task, and a paragraph continuation task. For the latter two tasks, we provided two types of contexts of three sentences each. In one type of context, the direct object only competed with the subject, which was also the discourse topic. In the other context, there was one more competitor, which made the context more similar to the items from our corpus search. The analysis of the corpus search showed a significant effect of DOM with unmodified human direct objects but not with modified or inanimate ones. This led us to test only unmodified human indirect objects in the forced-choice and paragraph continuation tasks. However, there was no significant difference between the condition with DOM and the condition without it.

Our results are similar to the paragraph continuation task of [Özge et al. \(2016\)](#), which also did not support the hypothesis that DOM has a forward-looking discourse function. [Özge and von Heusinger's \(2020\)](#) corpus study did not provide any evidence for the hypothesis, but the material in that study was not fine-grained enough to test whether unmodified human indefinite direct objects with DOM show a higher forward-looking discourse function than those without DOM. In summary, we found a significant effect for unmodified human indefinite direct objects in the corpus study; however, this effect was not observed in other conditions or in other tasks.

## 5. General Discussion

Differential Object Marking (DOM) is a cross-linguistically widespread phenomenon that serves as a well-established and well-studied research topic. Research on DOM focuses on scales and parameters, such as the Animacy Scale, the Referentiality Scale, affectedness, and topicality, that determine the prominence relations between the subject and the direct object in a sentence. Only a few studies go beyond the sentence boundary and investigate discourse parameters. [Comrie \(\[1981\] 1989\)](#) and [Nilsson \(1985\)](#) argue that DOM in Turkish also has a forward-looking discourse function. [Laca \(1995\)](#) made a similar claim for Spanish, and [Chiriacescu and von Heusinger \(2010\)](#) provided the first empirical evidence, using a paragraph continuation task, that direct objects with DOM in Romanian show a higher level of referential persistence than their counterparts without DOM. In a more recent corpus study and a more extensive paragraph continuation task, [Tigău \(2022\)](#) corroborated these results. However, [Özge et al.'s \(2016\)](#) paragraph continuation task and [Özge and von Heusinger's \(2020\)](#) corpus study did not provide any support for a forward-looking function of DOM in Turkish. We, therefore, undertook a corpus study based on a much larger corpus, a forced-choice task, and a paragraph continuation task with a different design from [Özge et al.'s \(2016\)](#) design. But only in the corpus study did we find a significant effect for the forward-looking function of DOM and only for unmodified human indefinite objects. In all other conditions and in the other two tasks, we did not find any significant effect.

We will now discuss possible reasons for the different evidence. On the one hand, the literature reports strong intuitions about the discourse effect of DOM; on the other hand, we have only scant empirical evidence. One could say that the hypothesis is unfounded and that DOM is a phenomenon that is limited to the sentence-level variation. This does not mean that DOM might still be dependent on discourse but just that it does not determine the further development of the discourse. As an alternative to this negative position, we would like to discuss three possible explanations for this discrepancy: (i) the design of the experiments, (ii) the difference between specificity and discourse prominence, and (iii) the complex noun paradigm in Turkish and its distribution in the corpus, which suggests that it is not case marking but the use of the indefinite article *bir* that makes an essential contribution to discourse prominence.

For the design of the forced-choice and paragraph continuation tasks, we created smaller texts that consisted of at least one further argument, the subject, in addition to the direct object. The subject was the discourse topic in all stories and also the sentence topic in the sentence with the critical item as the direct object. The direct object, with or without DOM, was preverbal and was, therefore, always the focus of the sentence concerned. Participants, however, very often continued the story with the direct object referent, sometimes much more than with the subject referent. In the paragraph continuation task, the object was continued in 80% of cases (whereas in Özge et al. (2016), it was only 40%), but the subject was continued in only 9% of cases. Similar strong biases for the direct object were found in comparable studies for Romanian (Tigău, 2022) and Spanish (von Heusinger et al., 2024). We think that this strong prominence status of the direct object also has to do with the fact that direct objects are typically inanimate. Since we only conducted the experiments with human direct objects, this could have led to a disproportionate prominence shift, with a certain ceiling effect. That is, this strong bias may have masked a more fine-grained forward-looking function.

A second explanation for the discrepancy could be that DOM co-determines the referential properties of direct objects but not their discourse properties. In Turkish, direct objects with DOM are specific; thus, they show a referential reading in intensional contexts and the widest possible scope with regard to extensional operators such as quantifiers or negation (Erguvanli, 1984; Kelepir, 2001; von Heusinger & Kornfilt, 2005; Ketrez, 2023). According to Enç (1991), specificity is associated with d-linking, i.e., with a backward-looking discourse function. Additionally, there is a certain assumption that strong or specific indefinites also show a strong or prominent discourse behavior. This is shown for indefinite demonstratives (Prince, 1981), and it is in general assumed for strong indefinites, even in English (see Kamp, 2014). However, we do not see evidence in our experiments. This might be connected to the next issue.

Turkish indefinites do not only alternate between a case-marked and an unmarked form of the direct object; they also alternate in the use of the indefinite article with a noun vs. a bare noun in the preverbal position. Turkish has four options to present an argument in direct object position: a bare noun, a noun with the indefinite article *bir*, a noun with the indefinite article and morphological case, and a noun without an indefinite article and morphological case, as discussed in Section 4.1. In our study, we compared nouns with indefinite articles with and without case marking. Other approaches compare bare nouns with nouns with the indefinite article; see Ketrez (2023). The syntactic discussion of the contrast concerns the status of the noun as an independent argument or as a (pseudo-)incorporated noun (Erguvanli, 1984; Kornfilt, 2003; Öztürk, 2005; Ketrez, 2023). The semantic discussion concerns the scopal properties, and the discourse pragmatic discussion concerns the accessibility of the referent introduced by the bare noun vs. the indefinite noun (Erguvanli, 1984; Aydemir, 2004; Arslan-Kechriotis, 2009; Seidel, 2019, 2024).

Summarizing, we found preliminary evidence that the use of the indefinite article *bir* with direct objects in Turkish provides a higher level of discourse prominence and a larger forward-looking function. This effect might be associated with the first step in the grammaticalization path from the numeral to the indefinite article, which is described as a pragmatically-salient entity (Givón, 1981, p. 36). While this effect is robust, we cannot determine whether it also accounts for the pattern for unmodified human indefinite direct objects, as observed in our corpus study, or whether its strength might be masking a potential forward-looking discourse function of DOM.

## 6. Conclusions

Differential Object Marking (DOM) is a research topic at the interface of the lexicon, semantics, syntax, and information structure. DOM systems exhibit cross-linguistic patterns, which reflect the interplay between nominal parameters such as animacy, verbal parameters such as affectedness, and information structure and discourse parameters such as referentiality and topicality. In our contribution, we have investigated DOM from a discourse perspective in order to better understand how different prominence scales determine DOM. We assumed that the notion of prominence allows us to integrate both sentential and discourse parameters. Focusing on the forward-looking function, we hypothesized that indefinite direct objects with DOM in Turkish are more discourse prominent than their unmarked counterparts. To test this hypothesis, we conducted a corpus study, a forced-choice task, and a paragraph continuation task. However, only the corpus study supported our hypothesis for unmodified human indefinite direct objects. The other two tasks did not provide any evidence for our hypothesis. We have, therefore, speculated that either the hypothesis is wrong or that other interfering parameters might mask the predicted effect. As there are different kinds of evidence for our hypothesis in Turkish, which was also tested cross-linguistically, we think that it is fruitful to retain the hypothesis and investigate other confounding parameters. We think that the paragraph continuation task might not be optimal for testing the forward-looking function of indefinite direct objects, given that they are introduced in the discourse-prominent focus position and that human direct objects also contribute to a high prominence level. Furthermore, we think that online experimental tasks, such as self-paced reading tasks and eye-tracking, might be promising methods. As observed by our pilot corpus study looking into the discourse distribution of different morphosyntactic forms of direct objects, the use of the indefinite article *bir* has a strong discourse effect, potentially masking the effect of DOM. We believe that a larger corpus or even utilizing large language models will offer more interesting data, enabling us to explore additional parameters such as frequency and register, and that our research is a first step toward a better understanding of DOM, connecting local-sentential prominence scales with global-discourse ones. Further cross-linguistic research is needed, employing different methods.

**Author Contributions:** Conceptualization, K.v.H. and H.B.Y.; methodology, K.v.H. and H.B.Y.; software, H.B.Y.; validation K.v.H. and H.B.Y.; formal analysis, H.B.Y.; investigation, K.v.H. and H.B.Y.; resources, H.B.Y.; data curation, H.B.Y.; writing—original draft preparation, K.v.H. and H.B.Y.; writing—review and editing, K.v.H. and H.B.Y.; supervision, K.v.H.; project administration, K.v.H.; funding acquisition, K.v.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project-ID 281511265—SFB “Prominence in Language” in Project B04 “Interaction of nominal and verbal features for Differential Object Marking” at the University of Cologne, Department of German Language and Literature I, Linguistics.

**Institutional Review Board Statement:** The studies involving humans were approved by Ethics Board of the German Linguistic Society (#2016-09E2\_200213). The studies were conducted in accordance with the local legislation and institutional requirements.

**Informed Consent Statement:** Informed consent was obtained from all participants involved in the experimental studies.

**Data Availability Statement:** All test items used in our experiments, annotated dataset for referential uptakes, and statistical analysis can be accessed via the following OSF link, together with relevant material for the pilot corpus study: [https://osf.io/gfzp3/?view\\_only=39c2cc027d464b958902db139838ac8f](https://osf.io/gfzp3/?view_only=39c2cc027d464b958902db139838ac8f).

**Acknowledgments:** We would like to thank the audiences at the 10th International Contrastive Linguistics Conference (Mannheim, 18–21 July 2023) and at the 21st International Conference on Turkish Linguistics (Mainz, 2–4 August 2023) for helpful discussions and feedback. We appreciate the four reviewers for their very constructive comments to improve this paper, and in particular, we would like to thank Jaklin Kornfilt for her support and always helpful comments. Furthermore, we would like to thank our student assistant, Sümeyye Uluer, for helping us with the preparation of the experimental studies and with the annotation of corpus and experimental data.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Notes

<sup>1</sup> List of abbreviations: ACC: ‘accusative’; AOR: ‘aorist’; DAT: ‘dative’; DOM: ‘Differential Object Marking’; LOC: ‘locative’; NEC: ‘necessitative’; NEG: ‘negation’; PERF: ‘perfective’; PROG: ‘progressive’; PSB: ‘possibilitative’; PST: ‘past tense’; 1.SG: ‘first-person singular’; 3.SG: ‘third-person singular’; Ø: ‘zero morpheme’.

<sup>2</sup> One reviewer suggested that the term ‘optional’ is not appropriate, as we are assuming that there are parameters that can determine case assignment. However, we think, as long as we do not have clear evidence for such parameters, that the term seems appropriate. For example, the function of DOM in transparent episodic sentences without further operators is often assumed to express specificity; however, experiments have not been able to support this assumption; see Section 2.1, ex. (9).

<sup>3</sup> One reviewer points out that there is the possibility that even an unmarked direct object can take the topic position under a particular discourse condition, as in (i) from Ketrez (2023, p. 216; adapted from Uygun, 2006):

(i) Kahve Ali de iste-miş-ti.  
 coffee Ali also want-PERF-PST  
 ‘Ali too wanted coffee.’

<sup>4</sup> Laca (1995, p. 82) illustrates this with (i) where the indefinite direct object *un tigre* (‘a tiger’) with the DOM marker *a* signals that more information about the tiger in question is to be expected, in comparison to a direct object without DOM.

(i) Juan mat-ó (a) un tigre.  
 Juan kill-3.SG.PST (DOM) a tiger.  
 ‘Juan killed a tiger.’

<sup>5</sup> Modification and animacy are parameters that we want to test for the probability of direct objects to be taken up in the following discourse. For this reason, in our corpus study presented below, we annotated these parameters; see Section 4.2.2.

<sup>6</sup> To test for statistical significance, we opted for Fisher’s exact test, which is preferred for categorical data with a small sample size (e.g., Agresti, 2002). Using R (RStudio Team, 2020), we created three different  $2 \times 2$  matrices for the pairs we were interested in: (i) bare N vs. N +ACC; (ii) *bir* N Ø and *bir* N +ACC; (iii) bare N and *bir* N Ø. The distribution between bare N and N +ACC yielded a statistical significance ( $p = 0.01273$ ). The distribution between *bir* N Ø and *bir* N +ACC did not yield a statistical significance ( $p = 0.139$ ). The distribution between bare N and *bir* N Ø yielded a statistical significance ( $p = 0.0053$ ).

<sup>7</sup> To justify our observations in more elaborate statistical terms and to draw reliable conclusions, we conducted a statistical analysis using R (RStudio Team, 2020). We employed a logistic regression model using GLM (e.g., McCullagh & Nelder, 2019) with a three-way interaction fitted as a function of NEXT MENTION, predicting the probability of the level re-mention. This effect yielded a statistical significance (Estimate = 3.08501,  $p = 0.0345$ ).

## References

- Agresti, A. (2002). *Categorical data analysis*. John Wiley and Sons. [CrossRef]
- Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory*, 21(3), 435–483. [CrossRef]
- Aksan, Y. (1995). *Functional universals and some aspects of sentence topic in Turkish* [Ph.D. thesis, Hacettepe University].
- Ariel, M. (1990). *Accessing noun-phrase antecedents*. Routledge. [CrossRef]
- Arslan-Kechriotis, Z. C. (2009). *Determiner phrase and case in Turkish: A minimalist account*. Verlag Dr. Müller.
- Aydemir, Y. (2004). Are Turkish preverbal bare nouns syntactic arguments? *Linguistic Inquiry*, 35(3), 465–474.
- Aygen, G. (2007). Specificity and subject-object position. Scope interaction in Turkish. *Dil ve Edebiyat Dergisi/Journal of Linguistics and Literature*, 4(2), 11–43.
- Bossong, G. (1985). *Empirische Universalienforschung: Differentielle Objektmarkierung in den Neuiranischen Sprachen*. Narr. [CrossRef]
- Bossong, G. (1991). Differential object marking in Romance and beyond. In D. A. Kibbee, & D. Wanner (Eds.), *New analyses in Romance linguistics* (pp. 143–170). Benjamins.
- Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1), 1–28.
- Chiriacescu, S., & von Heusinger, K. (2010). Discourse prominence and pe-marking in Romanian. *International Review of Pragmatics*, 2(2), 298–332. [CrossRef]
- Comrie, B. (1975). Definite and animate direct objects: A natural class. *Linguistica Silesiana*, 3, 13–21. [CrossRef]
- Comrie, B. (1989). *Language universals and linguistic typology: Syntax and morphology* (2nd ed.). University of Chicago Press. First published 1981.
- Çöltekin, Ç. (2013). *TRmorph: A morphological analyzer for Turkish*. Available online: <http://coltekin.net/cagri/trmorph/trmorph-manual.pdf> (accessed on 16 October 2023).
- Dalrymple, M., & Nikolaeva, I. (2011). *Objects and information structure*. Cambridge University Press. [CrossRef]
- Dede, M. (1986). Definiteness and referentiality in Turkish verbal sentences. In D. I. Slobin, & K. Zimmer (Eds.), *Studies in Turkish linguistics* (pp. 147–163). Benjamins. [CrossRef]
- Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, 67, 547–619. [CrossRef]
- Driemel, I. (2023). *Pseudo-noun incorporation and differential object marking*. Oxford University Press.
- Enç, M. (1991). The semantics of specificity. *Linguistic Inquiry*, 22, 1–25. Available online: <https://www.jstor.org/stable/4178706> (accessed on 28 November 2018).
- Erguvanlı, E. E. (1984). *The function of word order in Turkish grammar*. University of California Press.
- Erkü, F. (1983). *Discourse pragmatics and word order in Turkish* [Ph.D. thesis, University of Minnesota].
- Fodor, J., & Sag, I. (1982). Referential and quantificational indefinites. *Linguistics and Philosophy*, 5, 355–398. Available online: <https://link.springer.com/content/pdf/10.1007/BF00351459.pdf> (accessed on 28 November 2018).
- Givón, T. (1981). On the development of the numeral “one” as an indefinite marker. *Folia Linguistica Historica*, 11(1), 35–53. [CrossRef]
- Givón, T. (Ed.). (1983). Topic continuity in discourse: An introduction. In *Topic continuity in discourse: A quantitative cross-language study* (pp. 1–42). Benjamins. [CrossRef]
- Göksel, A., & Kerslake, C. (2005). *Turkish: A comprehensive grammar*. Routledge.
- Grosz, B., Joshi, A., & Weinstein, S. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21, 203–226. Available online: <https://aclanthology.org/J95-2003.pdf> (accessed on 16 October 2023).
- Gundel, J., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69(2), 274–307. [CrossRef]
- Himmelman, N. P., & Primus, B. (2015, December 12–14). Prominence beyond prosody—A first approximation. In A. De Dominicis (Ed.), *pS-prominenceS: Prominences in linguistics. Proceedings of the international conference, Viterbo, Italy* (pp. 38–58). DISUCOM Press. Available online: [https://ifl.phil-fak.uni-koeln.de/sites/linguistik/Personen/ASW/Himmelman/Publikationen/2011-2015/himmelman\\_primus\\_2015\\_on\\_prominence\\_pS-prominenceS.pdf](https://ifl.phil-fak.uni-koeln.de/sites/linguistik/Personen/ASW/Himmelman/Publikationen/2011-2015/himmelman_primus_2015_on_prominence_pS-prominenceS.pdf) (accessed on 28 November 2018).
- Hopper, P. J., & Thompson, S. A. (1980). Transitivity in grammar and discourse. *Language*, 56(2), 251–299.
- İşsever, S. (2003). Information structure in Turkish: The word order–prosody interface. *Lingua*, 113(11), 1025–1053. [CrossRef]
- İşsever, S. (2007). Towards a unified account to clause-initial scrambling in Turkish: A feature analysis. *Turkic Languages*, 11(1), 93–123.
- Kamali, B. (2015). Caseless direct objects in Turkish revisited. In A. Meinunger (Ed.), *Byproducts and side effects: Nebenprodukte und Nebeneffekte* (pp. 107–123). ZAS Papers in Linguistics 58. ZAS.
- Kamp, H. (2014). *Dividing the province of indefinite noun phrase uses into three parts* [Manuscript, University of Stuttgart].
- Keleşir, M. (2001). *Topics in Turkish syntax: Clausal structure and scope* [Ph.D. thesis, Massachusetts Institute of Technology].
- Kennelly, S. D. (1997, August 12–14). The presentational focus position of nonspecific objects in Turkish. In K. Imer, & N. Engin Uzun (Eds.), *Proceedings of the 8th international conference on Turkish linguistics, Oxford, UK*. Ankara Üniversitesi Basımevi.
- Ketrez, F. N. (2005). *Children’s scope of indefinite objects* [Ph.D. thesis, University of Southern California].
- Ketrez, F. N. (2023). Are Turkish non-case-marked objects with and without *bir* interpreted and acquired differently? *Languages*, 8, 229. [CrossRef]

- Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography*, 1, 7–36.
- Kılıçaslan, Y. (1994). *Information packaging in Turkish* [Ph.D. thesis, University of Edinburgh].
- Kılıçaslan, Y. (2006). A situation-theoretic approach to case marking semantics in Turkish. *Lingua*, 116, 112–144.
- Kızılkaya, S. (2024). *Affectedness at the morphosyntax-semantics interface. Evidence from Differential Object Marking*. De Gruyter. [CrossRef]
- Kızılkaya, S., Forsythe, Z. L., & von Heusinger, K. (2022). Affectedness and Differential Object Marking in Turkish and Uzbek. *Linguistics*, 60(6), 1907–1941. [CrossRef]
- Kornfilt, J. (1997). *Turkish. A descriptive grammar*. Routledge.
- Kornfilt, J. (2003). Scrambling, subscrambling, and case in Turkish. In S. Karimi (Ed.), *Word order and scrambling* (pp. 125–155). Blackwell.
- Krause, E., & von Heusinger, K. (2019). Gradient effects of animacy on Differential Object Marking in Turkish. *Open Linguistics*, 5(1), 171–190. [CrossRef]
- Kuno, S. (1972). Functional sentence perspective: A case study from Japanese and English. *Linguistic Inquiry*, 3(3), 269–320.
- Laca, B. (1995). Sobre el uso del acusativo preposicional en español. In C. Pensado (Ed.), *El complemento directo preposicional* (pp. 61–91). Visor.
- Malchukov, A. L., & de Hoop, H. (2011). Tense, aspect, and mood based differential case marking. *Lingua*, 121, 35–47.
- McCullagh, P., & Nelder, J. A. (2019). *Generalized linear models*. Routledge. [CrossRef]
- Nakipoğlu, M. (2009). The semantics of the Turkish accusative marked definites and the relation between prosodic structure and information structure. *Lingua*, 119, 1253–1280. [CrossRef]
- Næss, A. (2004). What markedness marks: The markedness problem with direct objects. *Lingua*, 114, 1186–1212.
- Nilsson, B. (1979). Definiteness and reference in relation to the Turkish accusative. In *Orientalia suecana* (vols. XXVII–XXVIII, pp. 118–131). Uppsala.
- Nilsson, B. (1985). *Case marking semantics in Turkish* [Ph.D. thesis, University of Stockholm].
- Özge, U. (2011). Turkish indefinites and accusative marking. In A. Simpson (Ed.), *Proceedings of the 7th workshop on Altaic formal linguistics (WAFL 7)* (pp. 253–268). MIT Working Papers in Linguistics 62. MIT Press.
- Özge, U., Özge, D., & von Heusinger, K. (2016). Strong indefinites in Turkish, referential persistence, and salience structure. In A. Holler, & K. Suckow (Eds.), *Empirical perspectives on anaphora resolution* (pp. 169–192). De Gruyter. [CrossRef]
- Özge, U., & von Heusinger, K. (2020). Case marking and forward and backward discourse function. In D. Zeyrek, & U. Özge (Eds.), *Discourse meaning: The view from Turkish* (pp. 81–102). Mouton de Gruyter. [CrossRef]
- Öztürk, B. (2005). *Case, referentiality and phrase structure*. Benjamins. [CrossRef]
- Prince, E. F. (1981). On the inferencing of indefinite-this NPs. In B. Webber, A. K. Joshi, & I. Sag (Eds.), *Elements of discourse understanding* (pp. 231–250). Cambridge University Press.
- Reinhart, T. (1981). Pragmatics and linguistics: An analysis of sentence topics. *Philosophica*, 27, 53–94. [CrossRef]
- RStudio Team. (2020). *RStudio: Integrated development environment for R*. RStudio, PBC. Available online: <http://www.rstudio.com> (accessed on 16 October 2023).
- Seidel, E. (2019, August 11–13). Anaphoric potential of pseudo-incorporated nouns in Turkish. In H. Sofu, D. Abik, M. Y. Özezen, C. Can, & A. Kilimci (Eds.), *Proceedings of the 18th international conference on Turkish linguistics (ICTL 18), Adana, Turkey* (pp. 259–266). Harrassowitz.
- Seidel, E. (2024). *Anaphoric potential of bare nouns and event structure in Turkish*. University of Cologne. Available online: <https://kups.uni-koeln.de/72720/> (accessed on 2 May 2024).
- Sezer, T. (2017). TS corpus project: An online Turkish dictionary and TS DIY corpus. *European Journal of Language and Literature*, 9(1), 18–24. [CrossRef]
- Song, J., & Kaiser, E. (2023). Effects of referential structure on pronoun interpretation. *Language, Cognition and Neuroscience*, 39(1), 98–117. [CrossRef]
- Tenny, C. L. (1994). *Aspectual roles and the syntax-semantics interface*. Kluwer Academic Publishers.
- Tigău, A. (2022). An experimental study on the discourse properties of Romanian direct objects. *Studii de Lingvistică*, 12(2), 85–107. Available online: <https://www.proquest.com/scholarly-journals/experimental-study-on-discourse-properties/docview/2781297729/se-2?accountid=10218> (accessed on 13 May 2024).
- Uygun, D. (2006, August 16–20). *Scrambling bare nominal objects in Turkish*. International Conference on Turkish linguistics (ICTL 13), Uppsala, Sweden.
- von Heusinger, K. (2002). Specificity and definiteness in sentence and discourse structure. *Journal of Semantics*, 19(3), 245–274. [CrossRef]
- von Heusinger, K. (2019). Indefiniteness and specificity. In B. Abbott, & J. Gundel (Eds.), *Oxford handbook of reference* (pp. 146–167). Oxford University Press. [CrossRef]
- von Heusinger, K., & Bamyacı, E. (2017a). Specificity effects of Turkish Differential Object Marking. In L. Zidani-Eroğlu, M. Ciscel, & E. Koulidobrova (Eds.), *Proceedings of the 12th workshop on Altaic formal linguistics (WAFL12)* (pp. 309–319). MIT Press.

- von Heusinger, K., & Bamyacı, E. (2017b, June 26). *Specificity effects of Turkish Differential Object Marking in monolingual and bilingual speakers*. Workshop on (In)definiteness and (Non)specificity: Meaning and Usage, Berlin, Germany.
- von Heusinger, K., Duarte, T., & García, M. G. (2024). Differential Object Marking and discourse prominence in Spanish. *Isogloss*, 10(1), 1–37. [[CrossRef](#)]
- von Heusinger, K., & Kornfilt, J. (2005). The case of the direct object in Turkish: Semantics, syntax and morphology. *Turkic Languages*, 9, 3–44.
- von Heusinger, K., & Özge, U. (2021). Inferable and partitive indefinites in topic position. In A. Holler, K. Suckow, & I. de la Fuente (Eds.), *Information structuring in discourse* (pp. 112–140). Brill. [[CrossRef](#)]
- von Heusinger, K., & Schumacher, P. B. (2019). Discourse prominence: Definition and application. *Journal of Pragmatics*, 154, 117–127. [[CrossRef](#)]
- Ward, G. L., & Prince, E. F. (1991). On the topicalization of indefinite NPs. *Journal of Pragmatics*, 16, 167–177.
- Winter, B., & Grice, M. (2021). Independence and generalizability in linguistics. *Linguistics*, 59(5), 1251–1277. [[CrossRef](#)]
- Witzlack-Makarevich, A., & Seržant, I. A. (2018). Differential argument marking: Patterns of variation. In I. A. Seržant, A. Witzlack-Makarevich, & K. Mann (Eds.), *Diachrony of differential argument marking* (pp. 1–40). Language Science Press. [[CrossRef](#)]
- Zidani-Eroğlu, L. (1997). *Indefinite noun phrases in Turkish* [Ph.D. thesis, University of Wisconsin-Madison].

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.