



**HAL**  
open science

## Discourse prominence and differential argument marking

Paul Compensis, Marco García García, Birgit Hellwig, Sonja Riesberg,  
Gertrud Schneider-Blum

► **To cite this version:**

Paul Compensis, Marco García García, Birgit Hellwig, Sonja Riesberg, Gertrud Schneider-Blum.  
Discourse prominence and differential argument marking. 2024. hal-04488084

**HAL Id: hal-04488084**

**<https://hal.science/hal-04488084>**

Preprint submitted on 4 Mar 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Discourse prominence and differential argument marking

Paul Compensis<sup>a</sup>, Marco García García<sup>b</sup>, Birgit Hellwig<sup>b</sup>, Sonja Riesberg<sup>b,c</sup>, & Gertrud Schneider-Blum<sup>b</sup>

<sup>a</sup>University of Erlangen-Nuremberg; <sup>b</sup>University of Cologne; <sup>c</sup>LACITO (CNRS)

**Abstract.** Special morphosyntactic encoding in language is typically motivated by the need to mark a deviation from a preferred and more frequently encountered pattern. Previous accounts often focused primarily on deviations concerning semantic features (e.g., animacy, definiteness, transitivity) or information structure (e.g., topicality), but many observations from the literature suggest that special encoding might also depend on discourse factors. In this contribution, we explore recent empirical work on one type of special encoding – differential argument marking – in a number of typologically diverse languages (Bulgarian, Romanian, Spanish, Tima, Yali) and illustrate the discourse-sensitivity of these strategies. In particular, we present evidence that discourse prominence, i.e., the dynamic ranking and keeping track of referents in discourse, interacts strongly with the phenomena under discussion. This contribution offers some new perspectives on differential argument marking and suggests that discourse prominence plays a critical role for special encoding in morphosyntax.

*Keywords.* Special encoding; differential argument marking; differential object marking; differential object indexing; discourse prominence

### 1. Introduction

#### 1.1 Coding asymmetries in language

Many languages exhibit special encoding with particular morphosyntactic markers, often in the form of differential marking.<sup>1</sup> In these phenomena, a linguistic element (an argument role, for instance) is marked with a specific morphosyntactic marker (e.g., a case flag or a person index) but — and this is the defining characteristic — only for a subset of elements of the same class (e.g., only animate patients instead of all members of this argument role). In terms of general function, these special encodings have in common that they highlight some deviation from a more frequently encountered and thus more commonly expected pattern (Comrie 1989; Haspelmath 2021a). More specifically, many argument markers only encode referents differentially that have some type of feature overlap with another argument role in the sentence (e.g., a patient bearing agentive features).

---

<sup>1</sup> In this study, we use *differential marking* as a cover term for different forms of differential argument marking.

Such coding asymmetries are probably best described in the literature for differential case marking. For example, in the most recent typological survey on optional case marking, Chappell & Verstraete (2019) identify a set of semantic and information-structural factors as recurring motivations across languages: Definiteness, animacy, degree of affectedness and/or topicality tend to be relevant for case-marked O arguments, while degree of agentivity and/or focal information status tend to be relevant for case-marked A arguments (see also Witzlack-Makarevich & Seržant 2018 and the literature cited therein).

These semantic-pragmatic features are usually described in terms of prominence hierarchies, as in (1), which have been put forward in different versions (see among others: Dixon 1979; Bossong 1998; Aissen 2003; Croft 2003; Iemmolo 2010).

(1) Prominence hierarchies

- a. Animacy: human > animate > inanimate
- b. Referentiality: pronoun > proper name > definite NP > indefinite specific NP > indefinite non-specific NP > bare noun
- c. Topicality: topical > non-topical

These scales provide a rough *descriptive* means to account for cross-linguistic recurrent tendencies concerning differential marking and other types of special encoding such as person-based ergative marking asymmetries, animacy-related locative marking or asymmetries in adnominal possession (see among others Bickel et al. 2014, Haspelmath 2013, 2021a). For instance, O arguments that are ranking higher on these scales are more likely to be overtly marked than those that rank lower – arguably due to the deviation from a more frequently found pattern (i.e., O arguments typically ranking lower on these scales).

This deviation has often been related to the level of semantic roles, that is, to proto-typical agents or patients and the features associated with these proto-roles in the sense of Dowty (1991) and Primus (1999) (see García García 2007, 2014; Primus 2012; García García et al. 2018; Mürmann 2023). Typically, arguments that exhibit non-role-prototypical features receive special encoding. This has been closely associated with markedness theory (Greenberg 1966; Croft 2003), according to which marked meanings align with marked forms. Accordingly, in this line of research differential marking has been explained in terms of markedness reversal (Aissen 2003 among others) or the distinguishing function, that is, the need to discriminate between different co-arguments (see Bossong, 1982, 1985, 1998; Comrie 1989, de Swart 2006 among others).

In a further prominent line of research, differential marking has instead been associated with the indexing or characterizing function of case (see Hopper & Thompson 1980; Kliffer 1995; Lazard 2003; Næss 2004, 2007). According to these accounts, differential marking does

not signal the alleged semantically marked status of an argument role (e.g., an animate, definite O), but a highly transitive event which typically involves an animate, definite and affected O argument (for a comprehensive discussion of the distinguishing function vs. the indexing or characterizing function, see Song 2001, Næss 2004, 2007; Mürmann 2023).

Irrespective of the exact mechanism, all these accounts have in common that they assume some form of ranking among linguistic elements of the same kind. This is in line with the general idea of linguistic prominence (in the sense of Himmelmann & Primus 2015), that is, linguistic elements are constantly ranked in relation to each other at several linguistic levels. In particular, it is assumed that linguistic elements are not prominent or central in absolute terms but only in relation to units of the same type and less prominent elements can exhibit features that are more common of more prominent elements – and vice versa. Prominent elements are the attentional centre (a-centre) at a given point of time in an utterance and some elements (especially A arguments) are more naturally occupying this a-centre. Shifting the a-centre to other elements is possible but requires “conventionalized specific formal means of indicating this shift” such as positional changes as in passives or changes in the morphosyntactic realization as in special encoding (Himmelmann & Primus 2015: 48-49).

Different linguistic dimensions, such as those stated in the hierarchies above, can contribute differently to the relative prominence status of a linguistic element in a given language. This is reflected in the observation that differential marking systems typically differ with respect to the sensitivity to these different scales. For instance, differential object marking (DOM) can be sensitive to only one of these prominence scales in some languages, as has been assumed, for example, for Turkish or Hebrew, or to two or more of these scales, as it is the case, for instance, in Romanian or Spanish (Bossong 1998, Aissen 2003, among others). Moreover, languages with DOM vary with respect to the relative cut-off point within the relevant scales (see section 3 for Romanian).

Whereas earlier accounts focussed primarily on animacy and definiteness, later accounts showed that differential marking is often also sensitive to different dimensions of information structuring (i.e., the organization of information at the sentence level in the sense of Chafe 1976) – especially to the structuring of *what the sentence is about*, i.e., topicality, but also focus. In these accounts, differential marking is either directly considered a topic-marking or focus-marking strategy or it is assumed that mismatches between argument structure and information structure (e.g., a topical object or a focal agent) condition the need to use special encoding. We will take up this issue in sections 2 to 4, where we discuss the information-structural accounts relevant to our case studies.

The information structure perspective made a fruitful contribution on differential marking research and uncovered an additional layer that impacts special encoding. Nevertheless, there are recurring issues that cannot be captured in these accounts. This is sometimes due to

conceptual and methodological issues such as the question what constitutes a topic or how to measure it but also due to contrasting findings (e.g., special encoding occurring with both topical and focal elements). Also, in some cases other factors are primarily responsible for the special encoding and only secondarily license topic marking, or there are interactions with other factors that are not yet sufficiently understood.

Following on from such information-structural accounts, some scholars argue that analyses should pay more attention to discourse factors in this respect, i.e., to the structuring of language beyond the sentence-level and particularly the management of discourse referents. For DOM and DOI, the role of discourse was emphasized by Schikowski & Lemmolo (2015) based on a sample of 178 languages. Besides, Chappell & Verstraete (2019: 16-17) report in their typological survey on case marking that, at least for some languages, discourse is the decisive level. This is acknowledged in some of the earlier accounts with an information-structural perspective (for optional O marking, see especially Lemmolo's 2011 account in terms of topic shift, and Dalrymple & Nikolaeva's 2011 concept of *secondary* topic; for optional A marking, see McGregor's 1992, 2006 analyses in terms of expectations within larger discourse). In addition, some authors refer explicitly to discourse topic (contrasted with sentence topic) in their accounts of explaining the function of a special encoding strategy (e.g., Leafgren 2002, for DOI in Bulgarian) or challenge previous topicality-based accounts by looking more closely at discourse factors (e.g., Compensis 2022, for DOI in Bulgarian; Riesberg 2018, for differential agent marking in Yali).

Our own contribution builds on the above accounts and proposes that the notion of discourse prominence (in the sense of von Heusinger & Schumacher 2019) can be fruitfully applied to research on special encoding in general, and differential marking in particular.

## **1.2 Our proposal**

As outlined in the last section, differential marking is typically determined by a combination of factors that impact the relative prominence of a given linguistic element and particularly a given argument role. Building on previous work on semantic and information-structural factors, we want to shift attention to discourse factors and argue that they play a larger role in this respect than sometimes assumed. In particular, we propose that the extra-sentential organization of language and the ranking (i.e., the management) of discourse referents needs to be taken into account for a fuller understanding of special encoding devices.

An important inspiration for this proposal are insights developed by McGregor in a series of papers (2006, 2010, 2013) on optionality in grammar, investigating cases where a linguistic form is optional in that its presence or absence does not change the grammatical structure. We are specifically interested in coding asymmetries in argument marking where an argument index or a case flag is optional in the above sense, and does not change the grammatical role

of the argument in question.<sup>2</sup> McGregor argues that the usage of such optional forms is inextricably linked to the notion of prominence and joint attention, whereby the “use of the item is associated with making it – or something nearby – prominent, drawing attention to that item, while non-use is associated with demoting it, and relegating it to the background” (McGregor 2013: 1148). The exact interpretation of use vs. non-use are language- and construction-specific, and it is the goal of this paper to develop these ideas further and to outline the important role that discourse prominence plays in this respect.

Throughout the paper, we use the broader term “differential marking” (in the sense of differential argument marking), and our case studies exemplify differential object indexing (DOI, Bulgarian; Section 2) and differential object marking (DOM, Romanian and Spanish; Section 3) as well as differential agent/subject marking (DAM/DSM, Tima and Yali; Section 4). Our prevailing interest is in the optionality of the marker, but our case studies show overlap with referent-based splits where some arguments are obligatorily (un-) marked depending on referential properties (animacy and definiteness; see especially Section 3) as well as overlap with construction-based splits where some word orders obligatorily trigger the (non-) use of a marker (see especially Section 4). We follow Chappell & Verstraete (2019) who consider optional (case) marking to be different from referent-based and construction-based splits, but who acknowledge functional and diachronic links between the systems. Such links are also evident in our case studies, and we have therefore opted for the term “differential marking”, but we are careful to distinguish between optional and split phenomena in our discussions.

To account for the role of discourse in the examination of differential marking, we apply the notion of discourse prominence as defined by von Heusinger & Schumacher (2019). Building on Himmelmann & Primus (2015), von Heusinger & Schumacher (2019: 119) argue that discourse prominence is “a relational property that singles out one element from a set of elements of equal type and structure” and that the “[p]rominence status [of a discourse referent] shifts in time (as discourse unfolds)”. Certain elements (i.e., discourse referents) are more commonly associated with a prominent role and this prominence status is typically stable over parts of a discourse but might occasionally shift to other referents available in common ground, or new referents might be added to the common ground and then referred to.

The prominence framework predicts that different discourse functions would typically be associated with particular structures or forms of marking. For the purpose of this study, we assume the following three broad discourse functions that potentially could be accompanied by differential marking:

- MAINTENANCE (i.e., continuation of a prominent referent)
- INTRODUCTION (of a discourse-new referent)

---

<sup>2</sup> McGregor (2013) lists a number of other candidate structures (complementizers, definiteness markers, person and number markers, NP ellipsis), but optional case marking continues to be the prime example, and the overall extent of optionality in grammar remains to be seen.

- SHIFT (to a less prominent referent)

Maintaining the most prominent element is typically associated with the shortest encoding or no marking at all (e.g., zero forms) whereas introducing a new referent typically requires longer and more complex marking (e.g., indefinite NPs) and affects discourse structuring as well as common ground management. In contrast, shifts to less prominent referents target elements that are already in the common ground but that received less attention in recent discourse. These shifts may either appear in the form of an elevation of a previously less prominent element or in the re-activation of a previously prominent referent whose status became obscure or lost – in contrast to INTRODUCTION, they require that the referent is already introduced.

Our case studies show that differential marking is largely associated with SHIFT and in some cases with INTRODUCTION, but not with MAINTENANCE. As such, differential marking can be associated both with backward- and forward-looking perspectives. Adopting a backward-looking perspective, we observe that differential marking is sensitive to the relative prominence of the preceding referents, picking out a less prominent or discourse-new referent (thereby excluding more prominent antecedents). By contrast, when adopting a forward-looking perspective, we observe that differential marking serves to establish a (less prominent or discourse-new) referent as the most prominent referent. While we can sometimes isolate the contributions of backward- vs. forward-looking perspectives to differential marking, it is usually not possible to make a clear-cut distinction. In natural discourse, it is analytically very difficult to decide to what extent differential marking is determined by the prominence structure of the preceding vs. by that of the upcoming discourse. See, however, the discussion of differential object marking in Romanian and Spanish (Section 3), where the experimental restriction to discourse-new referents allows for an isolation of the forward-looking potential.

In the following, we present analyses of differential marking phenomena in our convenience sample of typologically diverse languages to trace the role of discourse factors for differential marking. We have selected three different systems – differential object indexing, differential object marking, and differential agent/subject marking – in order to elaborate on our hypothesis that irrespective of the type or source of system, discourse is an important factor that needs to be taken into account in the investigation of differential marking in general. Section 2 deals with differential object indexing in Bulgarian showing how the consideration of wider context in qualitative and experimental research can help uncover discourse factors affecting differential marking. Section 3 extends this line of research to differential object marking in Romanian and Spanish. In Section 4, we turn to differential agent marking in Tima and Yali, also considering the interaction of this special encoding with word order. Finally, Section 5 summarizes and discusses our main findings.

## 2. Discourse prominence and differential object indexing in Bulgarian

A case in point that nicely illustrates how discourse prominence interacts with special encoding is differential object indexing (DOI), that is traditionally also known as *object reduplication*, *clitic doubling* (Jaeggli 1982, 1986), *optional object agreement* (e.g., Nikolaeva 1999) or *differential object agreement* (Lazard 2003). DOI is a form of object marking that makes use of person indexes in a cross-indexing relationship in the sense of Haspelmath (2013, 2019). We therefore use the term *differential object indexing* – synonymously to lemmolo’s (2011) *differential object indexation* – to describe this phenomenon. In the following, we restrict our discussion on how discourse prominence interacts with differential object indexing in Bulgarian, mainly based on recent experimental work (Compensis 2022), but we refer the reader also to other recent work on discourse factors of DOI in other languages (for an overview, see Just 2022; and Just & Čéplö 2022 for the discussion of discourse in their work on DOI in Ruuli and Maltese).

The example in (2) illustrates DOI in Bulgarian. In the first clause, *orehi* (‘walnuts’) and *kopǎr* (‘dill’) are unmarked objects whereas in the second clause the object argument ‘dill’ is represented by both, a full noun and a short pronoun (i.e., person index) in that sentence. This form of marking (or cross-indexing) is optional for both direct and indirect objects in Bulgarian and can be used to cross-refer to full nouns or long pronouns. In each case, the person index thereby agrees with the object argument in number and gender (and case when cross-indexing long pronouns). Note that Bulgarian (and Macedonian, unlike other Slavic languages) has lost case marking (except for pronominal forms and some remains among definite masculine nouns in formal language). In the following, we confine the discussion to DOI with full nouns in direct object function.

(2) Context: On the preparation of *tarator* (a cold soup) (item 486920412 in the BG-WEB2012 corpus)

<i>ako</i>	<i>imate</i>	<i>oreh-i</i>	<i>i</i>	<i>kopǎr</i> ,	<i>ne</i>	<i>vseki</i>	<b><i>go</i></b>
when	have.PRS.2PL	walnut-PL	and	dill	NEG	everybody	3SG.M.ACC
<i>običa</i>	<b><i>kopǎr-a</i></b> ,	<i>no</i>	<i>e</i>	<i>prekrasen</i> ,	<i>tarator-ǎt</i>	...	
love.PRS.3SG	dill-ART.SG.M	but	be.PRS.3SG	great	tarator-ART.SG.M		

‘when you use walnuts and dill, but not everybody likes the dill, but it [the dill] is great, the tarator ...’

DOI in Bulgarian is typically restricted to definite (or at least specific) referents and occurs more frequently (but not exclusively) with O arguments in non-canonical object-initial word order. Both aspects were previously assumed to be the conditioning factor underlying DOI in Bulgarian, for instance in the definiteness/specificity hypothesis (e.g., Cyxun 1962; Minčeva 1969; Jaeger and Gerassimova 2002) or in the word order/case disambiguation hypothesis (Georgieva 1974; Popov 1983). This last perspective was based on the impression that DOI is always obligatory with objects in non-canonical and particularly sentence-initial word order.



Interestingly, the potential obligatoriness of DOI with marked word orders resembles the situation of differential agent marking in Tima and to some extent Yali (see below). However, neither approach can predict the (optional) use of DOI because not all definite O arguments are cross-indexed with DOI and there are also cases where sentence-initial objects remain unmarked (see Compensis 2022). Hence, definiteness is only a necessary but not a sufficient condition for DOI in Bulgarian. Interestingly, there is no restriction in terms of animacy of the O argument.

To overcome the limitations of the previous accounts, a prominent line of research claimed that DOI in Bulgarian is a topic-marker – highlighting that the object of the sentence is the topic of the sentence. The underlying assumption is that due to the less prototypical association of objecthood and topicality, these objects require additional marking (see among others, the accounts by Rudin, 1990; Guentchéva, 1994, 2008; Leafgren, 1997, 2002). Nevertheless, at closer examination, there is evidence that speaks against such a direct link between topicality and DOI in Bulgarian. For instance, there is clear evidence that in special cases (contrastive and non-contrastive) focal elements but also newly introduced entities can be cross-indexed in Bulgarian (see examples in Leafgren, 1997; Ovcharova, 2018).

Also, in many cases it rather seems as if DOI is used to single out one particular discourse referent from a set of similar elements. For instance, in (2) the two ingredients of *tarator* are both ranked almost equally in terms of discourse prominence and then *dill* is briefly singled out, a comment is made on it but then the story continues with reference to the more prominent element *tarator*. Thus, rather than pointing to a topicalization of the cross-indexed direct object, the use of DOI seems to be motivated by the above-mentioned discourse function of SHIFT (Section 1.2), that is, by the necessity to mark the singling out and the (albeit brief) shift to a less prominent referent that is not the topic.

Interestingly, some of the earlier topicality-based accounts on Bulgarian also expanded their account beyond a classical (sentence-initial) topicality perspective and considered discourse factors to some extent. For instance, Leafgren (1997) stated that Bulgarian DOI is used when “the [sentence-level] topicality of an object seems less obvious in terms of the discourse structure”. In other words, he assumes that local deviations of the sentence topic from the broader discourse topic are licensing the use of DOI and suggests an explanation taking into account larger discourse.

More recently, Ovcharova (2018) claimed that DOI is used to (re-)introduce inactive referents in Bulgarian (see also Belloro, 2007, 2015; García-Miguel, 2015; Melis, 2018, for similar ideas on DOI in Spanish). Similar to Leafgren (1997, 2002), Ovcharova’s (2018) account also depends on a consideration of factors concerning common ground updating and discourse structuring. This can be illustrated in examples where a referent is re-activated that

was previously highly prominent but lost its status (again a discourse process in line with the SHIFT function described in section 1.2).

In (3), *Jesusa* is introduced and taken up several times (also by S/A indexes, indicated in the translated version of the previous discourse with *i*). At that point of time, it is clearly the most prominent referent in that piece of discourse. Then, several other referents turn up (indicated by different colours) and potentially obscure the discourse status of what was before the most prominent referent. *Jesusa* is then re-introduced and DOI is applied in this context.

- (3) Context (translated): *In my family we've always had a speedy metabolism. My sister Jesusa, may God rest her soul, was capable of eating a six-egg omelette with blood sausage in the middle of the afternoon and then tucking in like a Cossack at night. Poor thing. She was, just like me, you know? Same face and same classic figure, rather on the lean side. A doctor from Caceres once told my mother that the Romero de Torres family was the missing link between man and the hammerhead, for ninety percent of our organism is cartilage, mainly concentrated in the nose and the outer ear. ...* (Ruiz Zafón 2021: chapter 26)

*Hesusa* često *ja* bǎrkaha s *men* v seloto ...  
 Jesusa often 3SG.F.ACC confuse.PST.3PL with me in village ...  
 'Jesusa was often mistaken for me in the village ...'

Elaborating on Leafgren (2002) and Ovcharova (2018) and by looking at wider parts of the surrounding discourse, Compensis (2022) described DOI in terms of discourse prominence and defined it as “a type of differential marking of a P referent by means of a person index in cases when there is a certain level of unpredictability with respect to the (re)establishment or elevation of the discourse prominence status of this referent” (p. 113).

In this study, it was pointed out that this shift concerning the discourse prominence status can show up in different situations, for example in

- the (deictic) selection of a referent that is directly accessible to the interlocutors due to proximity
- the anaphoric singling out of one out of two (more or less) equally prominent referents from the common ground or previous discourse
- the reactivation of a referent that was active before but whose status was lost due to the interference with other referents that turned up in discourse.

DOI can be found with all of these situations and thereby fulfils a backward-looking function since it excludes more prominent or equally prominent but competing referents and selects a referent at a lower prominence level (see also the experimental data below). At the same time, it fulfils aspects of a forward-looking function by shifting to another referent – potentially allowing for additional operations on that referent. For instance, the cross-indexed object may also be topicalized (e.g., by word order). Example (3) illustrates this point: The proper name *Jesusa* is cross-indexed and marked as topic at the same time through its sentence-initial position. This is predicted by the prominence framework, since more prominent elements (or

elements raised in prominence) have more linguistic operations at their disposal (Himmelmann & Primus 2015; von Heusinger & Schumacher 2019).

The qualitative analysis of discourse contexts helped us uncover an association of DOI with discourse functions. In our research, we have substantiated these associations by means of a number of experiments that presented test subjects with scenarios of two (and more) referents whose relative ranking toward each other was manipulated (in terms of question type, number of mentions, and other linguistic means). In this way, Compensis (2022) was able to rule out the (sentence) topic marking function for Bulgarian DOI.

Three acceptability judgment studies either used questions probing the (binarily defined) topic or focus of a given sentence or manipulated the givenness in a brief introduction sentence in scenarios with two referents. None of these studies showed any difference in the rating (on a 7-point Likert scale) with respect to DOI – providing empirical evidence that DOI is not associated more closely with a pure topic-marking function. These three studies were aiming at the sentence-level structuring of the referents but no wider discourse was tested. To overcome that limitation and to test further the discourse prominence hypothesis for DOI in Bulgarian, a fourth web-based experimental study was conducted – inspired by scenarios used in Vogelzang et al. (2020).

In this web-based experiment, larger contexts established the ranking of three referents and this was combined with comprehension questions, acceptability judgment and reaction time measurement. Example stimuli are presented in (4). In each trial, a discourse topic was established by repeated mention in the discourse (e.g. *Peter* in the example stimuli below) and a second referent (*a woman*) introduced right before the critical sentence. In the target sentence, the most prominent element was referred to as direct object with DOI (condition b) or without DOI (condition a). This was contrasted with target sentences either presenting the second (less prominent) referent as object (condition c) or a discourse-new (not prominent but potentially inferable) referent (condition d), both with DOI. The main hypothesis here was that if DOI serves as a marker of referential shifting to a less prominent referent (i.e., SHIFT), the cross-indexing of the referent in condition c should be preferred over the cross-indexing of the most prominent element in condition b (i.e., MAINTENANCE) and the cross-indexing of a discourse-new referent (i.e., INTRODUCTION) in condition d (condition a served as a neutral comparison). In the example, agreement relations between both objects and subjects are indicated by the same colour as the respective referent. DOI is additionally highlighted in bold face.

(4) Example stimuli for the four conditions (object index and co-nominal in bold):

Context (translated): *The next story is about Peter. Yesterday, Peter was at a party and Ø talked to a beautiful woman for a long time. Suddenly, ...*

**a. Most prominent referent without DOI**

*žena-ta*                      *go*                      *celuna-la*  
woman-ART.SG.F    3.SG.M.ACC    kiss-PTCP.SG.F  
'the woman kissed him.'

**b. Most prominent referent with DOI**

*žena-ta*                      *go*                      *celuna-la*                      *Petăr.*  
woman-ART.SG.F    3.SG.M.ACC    kiss-PTCP.SG.F              Peter  
'the women kissed Peter.'

**c. Less prominent referent with DOI**

*Petăr ja*                      *celuna-l*                      *žena-ta.*  
Peter 3.SG.F.ACC    kiss-PTCP.SG.M              woman-ART.SG.F  
'Peter kissed the woman.'

**d. Not prominent (but inferable) referent with DOI**

*Petăr ja*                      *celuna-l*                      *domakinja-ta.*  
Peter 3.SG.F.ACC    kiss-PTCP.SG.M              host-ART.SG.F  
'Peter kissed the (female) host.'

The suggested hypothesis on DOI was confirmed and differences between these conditions were found for all three behavioural measures (i.e., accuracy, acceptability and reaction time). Among the three conditions with DOI, condition c (less prominent referent) received more accurate responses to the comprehension questions ( $M = 91.00\%$ ,  $SD = 28.67$ ) than condition b (most prominent referent,  $M = 89.67\%$ ,  $SD = 30.49$ ) and condition d (not prominent referent,  $M = 85.67\%$ ,  $SD = 35.10$ ). This difference was highly significant (main effect of condition on accuracy:  $\chi^2(3) = 18.10$ ,  $p < .001$ ) and group-wise comparisons between the single conditions were highly significant as well (for details, see Compensis 2022: 201-202).

Cross-indexing the less prominent referent also rated highest in terms of acceptability ( $M = -0.01$ ,  $SD = 0.82$ )<sup>3</sup> in contrast to the non-prominent referent that received the lowest rating ( $M = -0.45$ ,  $SD = 0.93$ ) and in comparison to the most prominent referent ( $M = -0.17$ ,  $SD = 0.80$ ). The main effect of condition on acceptability was also significant ( $\chi^2(3) = 18.10$ ,  $p < .001$ ) but after group-wise comparison only the contrast between condition c (less prominent) and condition d (not prominent) and between condition a (most prominent) and d (not prominent) remained significant (for details, see Compensis 2022: 201-202).

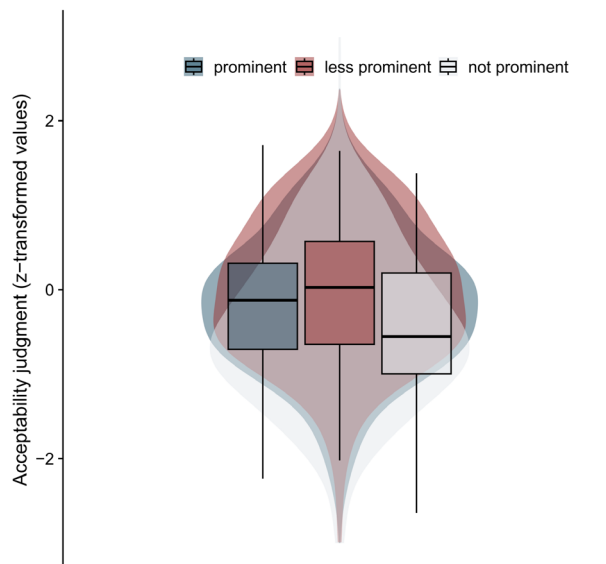
Overall, the distribution of the values suggests a systematically higher rating of DOI with less prominent referents. To illustrate that briefly, the (z-transformed) values from the acceptability judgment task are visualized in Figure 1 in form of a boxplot and a violin plot (capturing the relative density distribution of each of the three conditions that used differential object indexing). Again, cross-indexing less prominent referents (condition c, in red) yielded

---

<sup>3</sup> Acceptability ratings were assessed on a 5-point Likert scale and later z-transformed (following the suggestion by Schütze & Sprouse 2018). Values were also adjusted based on accuracy so that only ratings following accurately comprehended trials were considered in the subsequent analysis to exclude items that were potentially misunderstood. Nevertheless, effects were comparable and also significant in control calculations with the raw and non-adjusted data (see Compensis 2022: 201).

the strongest behavioural response and received the highest rating in comparison to indexing the most prominent (condition a, in light blue) or the non-prominent (condition d, in white-grey) referent.

Figure 1: Acceptability judgments of DOI in Bulgarian depending on the prominence of the direct object referent



These results support the idea that DOI is sensitive to fine-grained prominence differences and serves as a prominence-lending cue for referents that initially have a medium-level discourse prominence rank based on the previous context. Of course, this does not completely exclude the possibility of DOI occurring with discourse-new or the most discourse-prominent element in special cases (see Leafgren 2022: 149, Compensis 2022: 121-122) but demonstrates that DOI in Bulgarian is more commonly associated with a SHIFT function in discourse. It is less clear, however, whether DOI only leads to a short-lived elevation in prominence or a more sustainable change in the discourse representation. Future research should focus more closely on the subsequent discourse to evaluate that question.

The discussed web-based experiment provided evidence that DOI in Bulgarian is directly concerned with reference tracking in discourse and particularly the status of mid-level discourse prominence. In a way, this interaction with discourse structure is not surprising given the fact that DOI involves the use of a person index – since person marking is associated with speech roles or highly accessible third person referents (Haspelmath 2013), and the core function of person marking is “primarily a means of keeping track of referents in the discourse” (Siewierska & Bakker 2012: 293). It is thus to be expected that this core function should also turn out to be an integral part in the function and usage of differential object indexing. However, as plausible as this hypothesis may be, it requires further substantiation, both for Bulgarian (to clarify open questions such as the ones raised above) and for other languages more generally

(to establish the extent to which discourse prominence and DOI are linked in the languages of the world), also with respect to interactions of DOI with different types of object arguments (e.g., direct vs. indirect objects), other types of object marking (e.g. case flagging), word order and also prosody.

### **3. Discourse prominence and differential object marking in Romanian and Spanish**

This section deals with differential object marking (DOM), more precisely with the relation between DOM and discourse structure in Romanian and Spanish. We will start by providing a general overview on DOM and the prominence features that are said to determine this type of special encoding of verbal arguments – with a focus on the relation of DOM to information and discourse structure (Section 3.1). On the basis of Chiriacescu & von Heusinger's (2010) account on Romanian, and by applying the notion of discourse prominence as outlined above, we will argue that DOM does not only hinge on the given prominence status of the direct object, but that it also contributes to the discourse structure by signalling discourse prominence of the marked direct object referent with respect to the subsequent discourse (Section 3.2).

#### **3.1 DOM, topicality and discourse prominence**

DOM is a cross-linguistically wide-spread type of special encoding that is traditionally conceived of as a split-phenomenon: Whereas some direct objects require morphological marking others must be left unmarked, or are only optionally marked depending on certain semantic-pragmatic features, such as animacy, referentiality and topicality. These features are usually associated with the prominence hierarchies mentioned in section 1 (see among others: Dixon 1979; Bossong 1998; Aissen 2003; Croft 2003; Iemmolo 2010; García García & Caro Reina accepted). As was said before, languages with DOM can be sensitive to only one or a combination of these prominence scales and vary with respect to the cut-off point within the relevant scales.

In Romanian – and similarly also in Spanish – the cut-off point is between human definite and indefinite direct objects. Thus, human direct objects that are definite are obligatorily coded with the differential object marker *pe*, as in (5a), and the same holds for all human direct objects ranging higher on the referentiality scale.<sup>4</sup> For human indefinite direct objects *pe*-marking is optional, as in (5b), while human direct objects realized as bare nouns do not allow *pe*-marking, as in (5c).

---

<sup>4</sup> Note that in Romanian human definite direct objects are realized differently depending on whether the direct object is modified or not. While modified human direct objects require a definite article, DOM and clitic doubling (*L-am văzut pe băiatul bolnav* 'I saw the sick boy') unmodified ones can be realized by two alternative constructions: without definite article and presence of clitic doubling and DOM, as in (5a), or with definite article and lack of both clitic doubling and DOM (*Am văzut băiatul* 'I saw the boy') (see Onea & Hole 2017: 361, 371–373 and Hill & Mardale 2019: 20 for details).

### (5) DOM in Romanian

- a. *L-am văzut pe băiat.*  
Him-have1SG seen DOM boy  
'I saw the boy.'
- b. *(L-)am văzut (pe) o băiat.*  
Him-have1SG seen DOM a boy  
'I saw a boy.'
- c. *Am văzut \*pe băieți.*  
have1SG seen DOM boys  
'I saw boys.'

Note that Romanian exhibits not only DOM but also DOI, that is, clitic doubling of the direct object, which is optional for human indefinite direct objects (5b) and required for human definite direct objects (5a) and all other direct objects ranging higher on the referentiality scale. We will come back to the interplay of DOM and DOI in Section 3.2.

In addition to the nominal factors animacy and referentiality, as well as some verbal factors such as agentivity and affectedness (Witzlack-Makarevich & Seržant 2018, Romero Heredero 2022), the information structural category of topicality has been identified as a further crucial feature for DOM, in particular with regard to the diachronic development of this construction in Romance languages. More specifically, it has been shown that in languages such as Spanish or Sicilian topics marked via left-dislocation structures correlated with the rise of DOM (Iemmolo 2010, among others). Besides, topicality served as a facilitating category for a diachronic expansion of DOM along the referentiality scale (among others Laca 2006, Melis 2021, García García & Caro Reina accepted). Most Romance languages including Spanish, Romanian and Sicilian experienced a gradual loss of the link between DOM and topicality and a strengthening of animacy and referentiality in their diachronic development (among others Iemmolo 2010, Onea & Mardale 2020, Hill & Mardale 2021).<sup>5</sup> Still, it has been argued that even in languages such as Spanish topicality continues to be a driving force for DOM. One argument in favour of this view is that while DOM with human indefinite direct objects is optional in postverbal position it is required in a left-dislocation structure, that is in a construction that generally serves to mark the dislocated constituent as the sentence topic. This is illustrated by the Spanish example in (6), where DOM is realized with the prepositional marker *a* 'to'.

### (6) DOM with left-dislocated human indefinite direct objects in Spanish (Leonetti 2004: 86)

- a. *Ya conoc-ía Ø/a muchos estudiantes.*  
already know-1SG.PST Ø/DOM many students

<sup>5</sup> An interesting though rather rare exception within in the Romance languages is Balearic Catalan, where to present topicality and DOM continues to be very closely related. In this language, DOM is narrowed to topical objects only (among others Rohlfis 1971, Escandell-Vidal 2007, 2009, Iemmolo 2010).

- 'I already knew many students.'
- b. \* $\emptyset$ /A *muchos estudiantes*      ya      los      *conoc-ía*.  
 \* $\emptyset$ /DOM many students      already them      know-1SG.PST  
 'Many students, I already knew them.'

The view that in languages such as Spanish topicality continues to be crucial for DOM is not without problems. While it might be less controversial to assume that the left-dislocated direct object in (6b) is a topic it is much more difficult to maintain this interpretation with respect to the direct object in postverbal position in (6a), regardless of whether it appears with or without DOM. Thus, similar to DOI in Bulgarian, it is not fully convincing to account for DOM relying on the information structural notion of topicality, at least as far as languages such as Modern Spanish and Modern Romanian are concerned. Adhering to the view that topicality maintains a close relation to DOM some related weaker notions have been proposed, such as topic-worthiness (Comrie 1989: 197-199, lemmolo 2010: 248f), secondary topics (Nikolaeva 2001, Dalrymple & Nikolaeva 2011), embedded topics (Onea & Mardale 2020), discourse topics (Laca 1995) and discourse prominence (Chiriacescu & von Heusinger 2010).

For example, Laca (1995: 83) argues that optional or unexpected cases of DOM in Spanish such as those with animate non-human direct objects are quite regular, if the direct object in question is coreferential with a referent introduced in the preceding sentence (see also Isenberg 1968: 108ff).<sup>6</sup> Laca further points out that referents that are repeatedly mentioned within a previous stretch of discourse and do thus function as the main topic of a paragraph are much more likely to appear with DOM than those that have not been mentioned before (see also Weissenrieder 1991: 152–54, Barraza 2003: 92 and the discussion in García García 2014: 47f.). Note that these discourse-based assumptions relate to the backward-looking sensitivity of DOM pointing to the discourse function of a SHIFT to a less prominent referent.

In addition to the link of DOM to the preceding discourse (backward-looking perspective), there are also indications that DOM may influence the subsequent discourse (forward-looking potential) fulfilling the discourse function of INTRODUCTION. Based on the optional case of DOM in (7) Laca (1995: 82) states that *a*-marking signals that more information about the tiger in question is to be expected. More specifically, she hypothesizes that “the use of the preposition introduces an indefinite as a possible discourse topic” (Laca 1995: 82).

---

<sup>6</sup> Similar observations have also been made for Turkish, where the use of DOM with indefinite direct objects seems to force a coreferential link to a previously mentioned referent causing a partitive interpretation of the indefinite direct object (Enç 1991: 6f.). Note that although Enç (1991: 7) is primarily concerned with the semantic notion of specificity and its impact on DOM, she explicitly acknowledges the relevance of discourse here.



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

Summing up, different observations within and beyond an information structural perspective relying on topicality point to a close relationship between DOM and discourse prominence. However, there are only very few empirical studies on this subject, both with regard to the backward-looking and the forward-looking function. Following the notion of discourse prominence from Chiriacescu & von Heusinger (2010), which focuses on the forward-looking perspective, the next section will examine the relationship between discourse prominence and DOM in Romanian and Spanish in more detail.

### 3.2 DOM and discourse prominence

#### 3.2.1 Romanian

Chiriacescu & von Heusinger (2010) start from the observation that animacy and referential factors such as specificity are not sufficient to account for DOM in Romanian, in particular with regard to human indefinite direct objects in transparent contexts, that is, in sentences without further operators as in (5b). Note that in these cases DOM in Romanian is generally optional. Moreover, a non-specific interpretation of human indefinite direct objects is compatible with both a direct object with and without *pe*-marking. Focusing on these optional cases of DOM with human indefinite direct objects Chiriacescu & von Heusinger (2010) provide empirical evidence for the hypothesis that *pe*-marking signals discourse prominence (see also Chiriacescu 2014).

In their study discourse prominence is measured by three different parameters: (i) *referential persistence*, (ii) *topic shift potential* and (iii) *type of anaphoric referring expressions*. In the following, we focus on the first two of these parameters. Both of them are based on Givón's (1983) graded notion of *topic continuity*. Importantly, both referential persistence and topic shift potential relate to the forward-looking perspective, that is, on the question of how linguistic elements structure the upcoming discourse. More specifically, referential persistence stands for the number of anaphoric uptakes of a referent after its first introduction into the discourse. Here, the idea is that a prominent referent is cataphorically persistent showing a relatively high number of anaphoric uptakes. The parameter topic shift potential refers to "the potential of a non-topic to become a topic in the subsequent discourse" (Chiriacescu & von Heusinger 2010: 307).

Note that while referential persistence does not relate properly to any of the main discourse functions mentioned in section 1.2, topic shift potential is associated with INTRODUCTION, more precisely, with the introduction of a new referent that has a relatively

high potential to be reused as a topic. In this sense, topic shift potential resembles the notions of “secondary topic” (Givón 1981), “pre-topic” (Endriss & Gärtner 2005) as well as Laca’s (1995) above-mentioned idea concerning the introduction of “a possible discourse topic”.

Chiriacescu & von Heusinger (2010) analyzed the impact of DOM on discourse prominence measured by referential persistence and topic shift potential on the basis of an online story continuation experiment. Participants were confronted with a small context that contained three different stimuli sentences displaying a transitive clause with a human indefinite direct object. Importantly, in contrast to the web-based experiment on DOI in Bulgarian (see Section 2) the direct object stimulus was not a previously mentioned definite, but always a discourse-new indefinite referent. Note also that, although *pe*-marking of indefinite direct objects may appear without clitic doubling (5b), the test items were either presented with both *pe*-marking and clitic doubling or without *pe*-marking and clitic doubling, as in their test item 2 given in (8).

(8) Story continuation test item – DOM and discourse prominence (cf. Chiriacescu & von Heusinger 2010: 318)

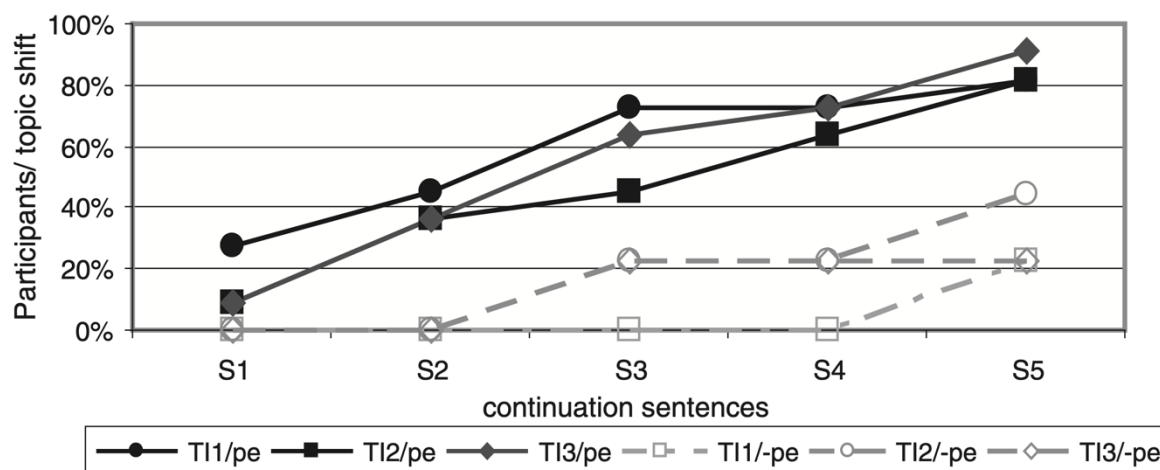
*Anul trecut când am fost la mare am cunoscut un salvamar. Era tot timpul activ. La sfârșitul sejurului meu, a salvat (-o pe) o fată de la înec.*

‘Last year when I was at the seaside, I met a lifeguard. He was very active all day long. At the end of my stay there, he saved (**pe-**) a girl from drowning.’

23 native speakers of Romanian were asked to read the stimulus sentences and add five natural continuation sentences. With regard to referential persistence the results show that the persistence of the *pe*-marked referents clearly exceeds that of the non *pe*-marked ones. Considering, for example, the test item in (8), there were notably more anaphoric uptakes for the *pe*-marked indefinite direct object *o fată* ‘a girl’ compared to its non *pe*-marked realization. On average, *pe*-marked direct objects were taken up in the subsequent discourse twice as often than their non *pe*-marked counterparts (Chiriacescu & von Heusinger 2010: 321).

As for topic shift potential, only direct objects with *pe* but not those without *pe*-marking exhibited a clear preference to become grammatical subjects and topics in the subsequent discourse. Interestingly, this preference usually becomes evident two or three sentences after the *pe*-marked direct object is introduced into the discourse and it increases considerably along the further continuation sentences. This is illustrated in Figure 2, where in all three test items (TI) the direct objects with DOM (*pe*) clearly show a greater topic shift potential than the ones without DOM (–*pe*).

Figure 2: Topic shifts associated with direct objects in Romanian across five continuation sentences (S1 to S5) in the DOM (*pe*) and the no DOM (*-pe*) condition (Chiriacescu & von Heusinger 2010: 323).



On the basis of the mentioned parameters of discourse prominence, that is, referential persistence and topic shift potential, the authors conclude that “the relevant discourse contribution of *pe* is to signal to the hearer that subsequent information about that referent will follow” (Chiriacescu & von Heusinger 2010: 298).

The findings from Chiriacescu & von Heusinger (2010) have recently been confirmed by an empirically broader follow-up study involving 85 native speakers of Romanian (Tigău & von Heusinger 2020). Contrary to Chiriacescu & von Heusinger (2010), in this study DOM was investigated independently of DOI. Paragraph continuations were elicited by using three different realizations of human indefinite direct objects: (i) with DOM and DOI, (ii) with DOM, (iii) neither with DOM nor DOI. In a nutshell, the results of this unpublished study suggest that both DOM and DOI contribute to the discourse prominence of the direct object. Moreover, the co-occurrence of DOM and DOI reinforces the discourse prominence of the direct object in question. Put differently, an indefinite direct object realized with both DOM and DOI creates an even greater expectation within the subsequent discourse than one marked with DOM only.

So far, we have gathered evidence on the influence of discourse prominence with respect to DOI in Bulgarian, as well as to DOI and DOM in Romanian, where both special encodings contribute to discourse prominence. The following section deals with DOM in Spanish. Since at least in European Spanish DOI is much more restricted than in Romanian the relation between DOM and discourse prominence can naturally be analyzed in absence of DOI.

### 3.2.2 Spanish

Recall that for Spanish already Laca (1995) claimed that DOM has an influence on the structuring of the subsequent discourse (see Section 3.1). More specifically, she hypothesized

that the use of DOM introduces an indefinite direct object as a possible discourse topic (Laca 1995: 82). Surprisingly, this hypothesis has only been followed up by a single study (von Heusinger, Duarte & García García, in prep.). In this study, discourse prominence was measured by similar parameters as in the experimental investigation on Romanian from Chiriacescu & von Heusinger (2010). Here, we concentrate on the two more closely related parameters: (i) referential persistence, that is, the question whether the direct object referent is mentioned again in the following sentence after being introduced into the discourse, and (ii) grammatical function, that serves to check whether the direct object is realized as a subject and thus as a sentence topic in the following clause. Importantly, grammatical function closely relates to Chiriacescu & von Heusinger's (2010) parameter of topic shift potential and the second of the above-mentioned broad discourse functions, that is, INTRODUCTION of a discourse-new referent (see Section 1.2).

Von Heusinger, Duarte & García García's (in prep.) study comprises both a corpus analysis and a paragraph continuation experiment focusing on DOM with human indefinite direct objects in Modern European Spanish. Here, we confine ourselves to the discussion of the experimental part of the investigation, which resembles Chiriacescu & von Heusinger's (2010) on Romanian. The experiment on Spanish is based on 48 test items. Each of the test items provided a two-sentence context in addition to the following critical sentence containing a human definite subject and a human indefinite direct object, as in (9).

- (9) Paragraph continuation test item – DOM and discourse prominence (cf. von Heusinger, Duarte & García García, in prep.)

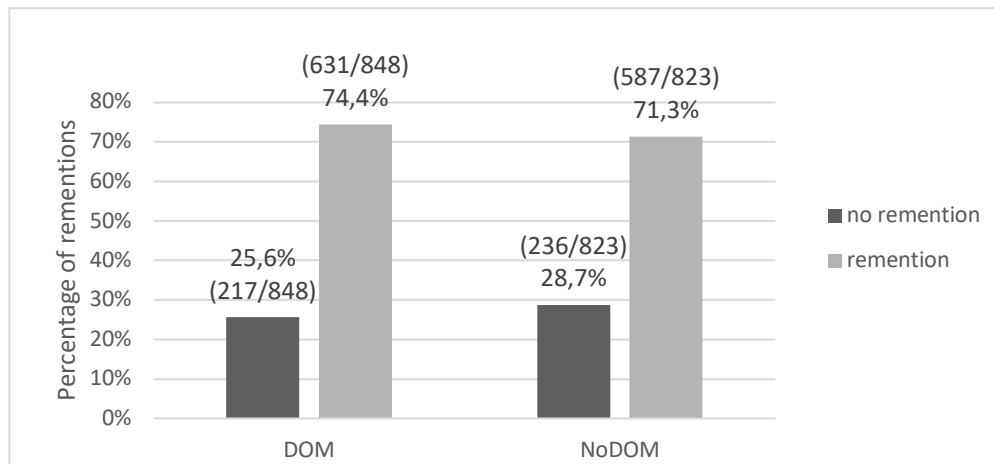
*Aquella mañana hacía muy buen tiempo y María estaba tomando café sentada en el balcón de su apartamento. Era muy temprano, pero le encantaba esa hora porque la calle estaba aún vacía y eso le permitía comenzar el día relajada. Justo cuando estaba terminando de beberse el café, oyó un ruido que provenía de la esquina, **se giró y vio (a/∅) un hombre.***

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

Contrary to the experiment on Romanian, participants were asked to add only one natural continuation sentence per test item. For the experiment on Spanish 80 native speakers of European Spanish were recruited providing around 1,700 continuation sentences. These continuation sentences were manually annotated and analyzed with regard to the discourse

prominence parameters (i) referential persistence and (ii) grammatical role as well as some further parameters. Figure 3 summarizes the results with respect to referential persistence.

Figure 3: Experimental results on referential persistence of human indefinite direct objects in Spanish depending on DOM and no DOM (von Heusinger, Duarte & García García in prep.)



As shown in Figure 3 the experimental results on Spanish reveal the same trend as the Romanian data reported in the previous section with regard to referential persistence: direct objects with DOM were anaphorically taken up in the continuation sentences more often than their unmarked counterparts (74,4 % vs. 71,3%). However, in contrast to Chiriacescu & von Heusinger's (2010) data on Romanian, where direct objects were taken up twice as often than those without DOM, for Spanish we only observe a numerical effect that did not turn out to be statistically significant.

As far as the second parameter of discourse prominence is concerned, namely grammatical function, in the experiment on Spanish no effect of DOM could be observed. Direct objects with DOM were reused as often as grammatical subjects in the subsequent sentence as it was the case for direct objects without DOM. Thus, unlike in Chiriacescu & von Heusinger's (2010) data on Romanian, where direct objects with DOM exhibited a clear preference to become sentence topics in the subsequent discourse (see Figure 2), the experimental data on Spanish did not show any effect of DOM on topic shift potential.

Taken together with the rather small numerical effect observed for referential persistence (Figure 3), this might lead to the conclusion that contrary to Laca's (1995) and our own assumptions DOM in Spanish does only have a marginal influence on the structuring of the upcoming discourse. We are currently exploring whether the missing effect of DOM on the discourse prominence parameter of topic shift potential might be due to the experimental design. Whereas in the Spanish experiment participants had to provide only one continuation sentence, participants in the Romanian experiment were asked to add five continuation sentences. This may make a crucial difference. Recall that in the Romanian data the frequency

of a direct object with DOM to become a topic increased considerably along the five annotated continuation sentences, whereby the preference to become a topic did not become evident until the second or third sentence after the direct object was introduced into the discourse (see Figure 2). A closer replication of the Romanian experiment asking for at least three rather than just one continuation sentence might exhibit an effect of DOM on topic shift potential in Spanish as well.

Summing up, we have discussed three related paragraph continuation studies focusing on the forward-looking discourse potential of DOM in Romanian and Spanish, where discourse prominence has been measured by several related parameters including referential persistence and topic shift potential. As shown, the discourse contribution of DOM differs with regard to these prominence parameters and the languages in question. While in Romanian DOM has a clear effect on discourse prominence, that is additionally reinforced by DOI, the data on Spanish is less supportive since it only exhibits a small numerical effect of DOM regarding referential persistence. Clearly, further research considering larger stretches of discourse is needed in order to show whether and to what extent DOM in Spanish contributes to the discourse prominence of a direct object referent.

In view of the positive findings for DOI in Bulgarian (see Section 2) and for DOM and DOI in Romanian, it seems promising to extend the investigation on discourse prominence to other languages and varieties. In the case of Spanish, it would be particularly interesting to consider varieties such as those spoken in Buenos Aires and Lima, where DOI has a similar distribution as in Romanian (see among others Sánchez & Zdrojewski 2013, Fischer et al. 2019). Comparing these varieties with European Spanish should reveal whether discourse prominence is indicated by DOM, DOI, or an interplay of these special encoding mechanisms.

#### **4. Discourse prominence and differential agent marking in Tima (Niger-Congo) and Yali (Trans-New Guinea)**

This section turns from differential object marking to differential agent marking, presenting case studies on Tima (Niger-Congo; Sudan) (Section 4.1) and Yali (Trans-New Guinea; Indonesia) (Section 4.2). In their typological survey, Chappell & Verstraete (2019) show that the optional marking of A arguments is cross-linguistically either associated with semantic factors (“with A arguments that are especially potent or agentive”, p. 5) or with information structure (specifically, with “[f]ocus and/or unexpectedness”, p. 17). Of particular relevance to our paper is their observation that information-structural factors, and especially the notion of unexpectedness, can go beyond the sentence level and may relate to the discourse level (pp. 15-16). This is very much the case for both Tima and Yali, where the key to understanding case marking clearly lies beyond the sentence level, as we will show in the following. While semantic factors play a role in both languages, they seem to be epiphenomenal in the sense

that discourse-prominent participants tend to be human, definite or agentive. Similarly, while information-structural accounts capture important insights (in that case-marked A arguments show an overall affinity to focal information), case-marking and topic/focus structures are independent of each other.

While there are some differences in detail, case-marked structures in both languages are implicated in the broad discourse function of SHIFT, i.e. they single out a less prominent referent from the preceding discourse and lend prominence to it. This SHIFT function plays out in two distinct ways. First, case marking allows speakers to elaborate on a non-agent referent in O function. In such cases, the non-agent argument outranks the agent argument on some – but not all – prominence scales (usually it is given and hence discourse-prominent), and speakers resort to case marking (on the A argument) in order to make the non-agent argument the attentional centre. This shift in attentional centre does not entail a shift in grammatical role (the referent continues in O function), the effect is temporary and local (in that it applies to one sentence only) and – importantly – correlates with a marked word order (in particular, a word order that places O before A). Second, speakers use case marking at (narrative) boundaries to establish an agent as the most prominent referent that is expected to persist in subject function throughout the upcoming discourse segment. Here, case marking serves to either elevate a less prominent referent (e.g. the preceding non-agent argument) or to re-activate a prominent referent (that has not been mentioned for a while), but usually does not introduce a discourse-new referent. In either case, case marking has a global effect (in that the shift in prominence status is expected to last throughout the episode) and does not (necessarily) correlate with a marked word order.

Before we substantiate the above summary for Tima (Section 4.1) and Yali (Section 4.2), we want to draw attention to one aspect that was mentioned above in passing: the role of word order. In both our case studies, differential marking is not entirely optional, as it interacts with word order: obligatorily so in Tima, and partially in Yali. Such construction-based splits are sometimes difficult to clearly distinguish from optional case marking, and Chappell & Verstraete (2019: 20) explicitly name word order as a candidate structure that could “be regarded as standing in between construction-based splits and optional marking.” Specifically, factors that condition optional case marking are also known to condition word order variation (e.g. agentivity features or topicalization processes), and there are known diachronic links between the systems (Chappell & Verstraete 2019: 21-26). It is thus probably no coincidence that studies of optional ergativity also tend to include discussions of potential interactions between case marking and word order (Meakins & O’Shannessy 2010; Suter 2010; *inter alia*). With respect to split ergativity, by contrast, discussions center on TAM or polarity values, not word order. From an analytical perspective, the close connection between case marking and word order in both our case studies makes it difficult to always clearly separate the contribution

of each structure to the overall interpretation. Through the comparison between Tima and Yali, we hypothesize that the first function (to accord local prominence to a non-agent argument) is primarily triggered by word order, while the second function (to establish a prominent agent) is the hallmark of the case marker. It remains an important desideratum for the future to extend this line of research to languages that do not exhibit such a close connection between word order and case marking.

#### 4.1 Tima

Tima is a Niger-Congo language that is spoken by about 7,000 people in the Nuba Mountains in Sudan. It belongs to the Katloid group of languages, where it occupies a unique position in that its grammatical structure differs sometimes considerably from that of the other two languages in this group, Katla and Julut. This also includes the phenomenon under discussion, which is only attested in Tima. Outside of the Katloid group, however, the differential marking of agents (and subjects) is an areal phenomenon that is present in numerous languages of Northeast Africa, where it probably originated within the Southern branch of Eastern Sudanic (of the Nilo-Saharan phylum) and spread to languages from other families (see Casaretto et al. 2020; Creissels et al. 2009: 87-91; Dimmendaal 2014; König 2008; Mitchell et al. submitted). Depending on the language, differential marking is either restricted to transitive structures like in Tima (differential agent marking, usually termed ‘split ergativity’) or includes intransitive structures as well (differential subject marking, termed ‘marked nominative’). Typically, agents or subjects can either occur in preverbal position and then remain unmarked for case, or else occur in postverbal position, and then receive ergative or nominative case marking. This correlation is illustrated by means of the contrastive Tima examples below. In Tima, speakers can choose to present any two-participant event<sup>7</sup> in either a non-ergative or an ergative structure. In a non-ergative structure, the A argument is preverbal and unmarked for case (as in 10a). In an ergative structure, the A argument is postverbal and obligatorily marked by means of the ergative proclitic *N=* (as in 10b). If it is realized as a pronoun, a special set of ergative pronouns is used. Note that the ergative argument needs to be overtly expressed, i.e. it cannot be reduced to zero, while all other arguments (A/S, O, BEN, INS) allow for zero realization.

---

<sup>7</sup> Two-participant events in Tima cover two types of structures, and differential agent marking is available for both: structures consisting of an A and an O argument (as in the example above); and structures consisting of an A argument, a verb derived with benefactive or instrumental morphology, and an oblique BEN or INS argument. Such oblique BEN and INS arguments differ from O arguments in several respects, e.g. they occur in different positions, and they need overt flagging by means of proclitics. Nevertheless, they are arguments that are required by the verb (not adjuncts), and both ergative and non-ergative structures are therefore possible.



(10) Contrasting non-ergative (10a) and ergative (10b) structures in Tima

- a. *káwùh á-hàmbìr-i tíyà*  
 stone 3PRF-trip.up-TR Tiya  
 A V O  
 ‘the stone tripped Tiya up’
- b. *tíyá á-hàmbìr-i ñ=káwùh*  
 Tiya 3PRF-trip.up-TR ERG=stone  
 O V AERG  
 ‘the stone tripped Tiya up’

This phenomenon is so widespread across the region that König (2008: 240) coined it the “‘no case before the verb’ rule”. Given the close connection between case marking and word order, it is not surprising that the literature tends to evoke information-structural factors as conditioning factors. For the verb-initial languages across the region, it is often assumed that preverbal non-case-marked S/A arguments originated diachronically in structures with left-dislocated participants, whereby the dislocated constituents lost their (ergative or nominative) case markings and were subsequently reanalyzed as constituents of the main clause (see e.g. König 2008). That is, the differential marking of S/A arguments (and their correlation with word order) possibly arose as a by-product of information-structural processes.<sup>8</sup> And synchronically, there are accounts that information-structural factors continue to be of relevance. In three Nilotic languages – Lopit (Moodie & Billington 2020: 271), Maasai (Arusa variety) (Levergood 1987: 43) and Tugen (Jerono 2011: 144) – the preverbal non-case-marked position is analyzed in terms of focus. More commonly, though, this position seems to be reserved for unfocused constituents and/or topics, cf. pertinent discussions for Akie (König et al. 2015: 92), Anywa (Reh 1996: 368-269), Cherang’any (Mietzner 2016: 108), Datooga (Gisamjanga variety) (Kießling 2007: 171), Jur-Luwo (Storch 2014: 201-202), Nandi (Creider & Creider 1989: 125), Northern Burun (Andersen 2015: 512), Pári (Andersen 1988: 293-294), Samburu (Camus variety) (Heine 1980: 104), Shilluk (Miller & Gilley 2001: 36), or Turkana (Dimmendaal 1986: 136). Similarly in Tima, the initial position is associated with topicality rather than focus (Becker & Schneider-Blum 2020; Schneider-Blum 2023). Furthermore, in our earlier research on Tima,

<sup>8</sup> Much of this literature centers on Nilotic languages, which tend to be verb-initial, i.e. analysts strive to explain the preverbal non-case-marked agent/subject position. See also Longacre (1995) and Payne (1995) on information-structural accounts for preverbal S/A arguments in verb-initial languages more generally. Note, though, that the verb-initial syntax is in itself an innovation, as the relevant Eastern Sudanic languages not only developed a system of differential agent/subject marking correlating with word order, but also lost their inherited verb-final word order (in favour of verb-medial or verb-initial orders) (Dimmendaal 2017). Furthermore, unlike Nilotic languages, the Katloid languages are verb-medial with basic AVO order. That is, it is the postverbal case-marked position that needs an explanation (see Casaretto et al. 2020 for the presumed diachronic scenarios). Given the many uncertainties in diachronic developments in this region, the present contribution adopts an exclusively synchronic perspective.

we assumed that postverbal ergative case marking correlated with focus (reported in König 2008: 285).

Overall, patterns in this region thus look to be in line with the cross-linguistic observation that case-marked A arguments tend to be associated with focus, not with topicality (see especially McGregor 2010: 1619-1620; Schultze-Berndt 2017: 1102). However, there are accounts that express an uneasiness with such explanations, suggesting to go beyond the sentence level to larger discourse structure. Most notably, Creider & Creider (1983: 12) argue for the Southern Nilotic language Nandi that there is an interaction between topic shift, continuity, and listener expectation. They propose that preverbal non-case-marked S/A arguments occur in contexts of “marked topicalization”, i.e. when an S/A referent unexpectedly continues, or unexpectedly does not continue, in the S/A role of the subsequent clause. For other languages, there are reports that either preverbal unmarked S/A arguments (in Anywa; Reh 1996: 310) or postverbal case-marked S/A arguments (in Tugen; Jerono 2011: 137) serve to introduce new discourse participants.

Such remarks notwithstanding, in-depth investigations into the interaction of case marking, word order and discourse structure are still few and far between (but see Longacre 1990 for case studies of word order in narratives). To date, the conditioning factors are probably best understood for Tima, and it remains to be seen if and to what extent the findings can generalize to other languages. For the moment, we can minimally state that they are not incompatible with what we know from other languages. On-going corpus research furthermore suggests that central insights also hold for the Southern Nilotic language Datooga (Lorenzen, in prep.; Mitchell & Lorenzen 2023a, 2023b; see also Kießling 2007). While this section deals exclusively with Tima, our expectation is that findings will transfer at least in part to other languages across the region.

Stimuli-based research has given us a good idea of the semantic, referential and discourse properties that favour the use of one structure over another. In a series of publications (Becker & Schneider-Blum 2020; Casaretto et al. 2020; Dolscheid et al., submitted; Schneider-Blum 2022, 2023; Schneider-Blum & Hellwig 2018; Schneider-Blum et al. 2022), it was shown that ergative structures tend to be used in contexts where there is a prominence conflict between the two arguments of a verb. In such contexts, the A argument is inherently prominent by virtue of the grammatical relations hierarchy (with subjects being more prominent than non-subjects) and the semantic role hierarchy (with agents being more prominent than non-agents). At the same time, it is less prominent than the non-A argument on at least one other prominence hierarchy, e.g., on the animacy hierarchy (in particular, when a non-human agent acts upon a human patient), the referentiality hierarchy (especially when the agent is not identifiable as an individual), and/or the givenness hierarchy (when the agent is new, while the patient is given). That is, case marking in Tima is not determined by the

properties of the agent (or patient) argument alone, but by the configuration of the two arguments (see also Silverstein’s 1976 notion of “global ergativity”, Witzlack-Makarevich et al.’s 2016 “co-argument-conditioned split”, or Haspelmath’s 2021b “scenario splits”). Overall, its distribution exhibits the hallmarks of special encoding as outlined at the beginning of this contribution: all things being equal, we would expect prominence cues from different dimensions to cluster on the A argument. And any deviation from this expectation favours ergative structures, thereby reflecting the increased prominence of the non-A argument vis-à-vis the A argument.

While such prominence conflicts robustly favour ergative structures in our controlled data set, they do not obligatorily trigger them, and speakers can always choose to use non-ergative structures instead. We hypothesized that discourse factors play an important role in this respect, and we therefore conducted a corpus analysis of case marking in natural discourse, with a focus on narrative monologues.<sup>9</sup> This analysis shows that ergative structures play a central role for structuring discourse, as they shift attention away from a prominent referent to a less prominent referent. As such, case marking occurs in two contexts, both of which are illustrated with the help of the corpus extract in (11) below (with ergative A arguments highlighted in boldface).

(11) Discourse uses of the Tima ergative

- i. *[ihiná]<sub>A</sub> †kólól [y-áná=wèéŋ]<sub>O</sub>*  
 PRON3PL directed PL-COW=DEM.REF  
 ‘they<sub>Katla</sub> directed those cows’
- ii. *[∅]<sub>INS</sub> túñèl-àà=†án [ŋ=ihiná]<sub>ERG</sub> l-úmúrik*  
 returned-INS=LOC3P ERG=PRON3PL LOC-Tima  
 ‘**they**<sub>Katla</sub> returned with them<sub>cows</sub> to Tima’
- iii. *ádíyàá túñèl-àà=†án [ŋ=ihiná]<sub>ERG</sub> [y-†áná]<sub>INS</sub> l-úmúr†ík=i*  
 when returned-INS=LOC3P ERG=PRON3PL PL-COW LOC-Tima=SEL  
 ‘when **they**<sub>Katla</sub> returned with the cows to Tima’
- iv. *màk dàh=†ì [ŋ=†rbá]<sub>ERG</sub> [∅]<sub>BEN</sub> nλ-diyáŋ=†è*  
 then said=APPL ERG=Irba IMP:PL-come=REP  
 ‘then **Irba** said to them<sub>Katla</sub>’

In the first context (illustrated in line 11ii), ergative structures temporarily shift the attentional centre towards a non-agent referent in order to elaborate on it. In this example, the O referent from line (11i) *yánáwèéŋ* ‘those cows’ is continued as the INS referent in the ergative structure

<sup>9</sup> Future research will have to show to what extent the findings are valid for non-narrative monologues and conversations.

of line (11ii). This referent is of intermediate prominence status: it is given (i.e. it is discourse-prominent), but it is not the most prominent referent, nor does it continue as the most prominent referent beyond the sentence. Interestingly, the ergative agent tends to be given, too. It may be new, but most commonly it simply continues the preceding agent – in this case, *ihìná* ‘they’ referring to the Katla, the most prominent referent throughout the entire episode. The ergative structure thus tends to maintain both referents in their respective grammatical roles, but conveys a subtle shift in their prominence status, which allows the narrator to elaborate on the non-agent referent. This elaboration typically takes the form of repeating or paraphrasing the preceding proposition, plus additional adjunct phrases and/or subordinate clauses. Here, non-ergative ‘they directed those cows’ in (11i) is paraphrased as ergative ‘they returned with them’ in (11ii), and is elaborated upon by means of a prepositional phrase detailing the destination of the cows (‘to Tima’). This first use of the ergative resembles most closely our stimuli-based findings and helps us place them into a discourse perspective: the increased prominence status of the non-agent argument is not a sufficient condition for the use of the ergative structure – but it increases the likelihood that the narrator would want to shift attention to it and elaborate on it, and thus resorts to the ergative. Following this elaboration, the narrator then either terminates all referential chains and concludes the episode (more common; exemplified in example 11), or resorts to a non-ergative structure and continues both referents within their respective grammatical roles (less common).

In the second context, ergative structures occur after a narrative boundary where they (re-) introduce an agentive referent in subject function at the start of a new referential chain, establishing it as the most prominent referent. Lines (11iii) and (11iv) above initiate such a new episode. They exhibit some of the formal cues that tend to cluster at major boundaries: tail/head linkage, where the adverbial clause in line (11iii) recapitulates preceding events before the main clause in (11iv) advances the story-line; as well as the presence of overt sequential markers such as *ádíyâá* ‘when, after’ (line 11iii) and *mâk* ‘then’ (line 11iv). The main clause in line (11iv) uses ergative marking to establish the agent, in this case, a shift from given ‘Katla’ to new ‘Irba’. Most commonly, the ergative agent is a previously prominent referent that has not been mentioned for a while, and is now re-established as prominent (as in example 11). It is also possible, but less common, for discourse-new referents to occur in this context. And, interestingly, there are even cases where a referent from before the transition point is continued – yet, speakers present this continuing referent as if it were new (and use maximal referring expressions), which suggests that there is a prominence re-set at boundaries. Once the referent is (re-) introduced, it is expected to be maintained as non-ergative subject throughout the upcoming episode. The most surprising finding in this respect is the central role of boundaries, which did not show up in our controlled data set: a shift to new agent referents

is not enough to trigger ergative structures in themselves – narrators use them in boundary contexts, i.e. when establishing referential chains at narrative boundaries.

In both contexts, ergative marking thus serves to shift attention to a referent that is currently not prominent and to establish it either locally or globally as the attentional centre. A surprising finding is that Tima uses the same structure for lending prominence to both non-agent referents (in the first context) and agent referents (in the second context). As we will see below, Yali also uses case-marked structures in both contexts, but different word orders: a marked word order (of O before A) (mostly) in the first context, and an unmarked order (of A before O) in the second context. For Yali it is thus possible to separate the contributions of case marking vs. word order – whereas in Tima, case marking and word order are inextricably linked.

## 4.2 Yali

Yali is a Trans-New Guinea language of West Papua (Indonesia). It belongs to the Ngalik branch of the Dani family, and is spoken by approximately 30,000 speakers in the mountainous area east to the Baliem valley. Differential argument marking, and particularly differential subject marking, is a widespread feature in New Guinea, not only among the Dani languages, but also within the Trans-New Guinea (TNG) linkage in general (Foley 2000; Fedden 2020). Many languages exhibit both optional ergative marking, and split ergativity conditioned by word order. In previous studies, the use of the ergative has mostly been attributed either to semantic factors like animacy (e.g., Scott 1986 for Fore) and agentivity (e.g., Anderson & Wade 1988 for Fologa; Christensen 2010 for Yongkom), or, less frequently so, to information-structural factors. Suter (2010), e.g., argues that the use of the ergative in Kâte correlates with high rhematicity, and that its primary function is to mark new, non-topical elements.

Yali, like many other TNG languages, exhibits AOV word order, in which the subject can either be unmarked for case, as in (12)a, or differentially case-marked by the ergative clitic =*en*, as in (12)b below. In marked (OAV) word order, just like for Tima postverbal subjects, ergative marking is obligatory (cf. (13)).

- (12) a. [**he itno**]<sub>A</sub> [*wam ibma=no*]<sub>O</sub> *oho nibag fug*  
 woman DET wam DIST=GIV also eat:3SG.REM.PST NEG  
 ‘the woman didn’t eat that pig, either’ [man\_and\_pig 129]
- b. [**mene ari=en**]<sub>Aerg</sub> [*hiyap eke*]<sub>O</sub> *na-ruk*  
 dog MED=ERG woman one eat-PROG  
 ‘the dog is biting a woman’ [quis\_task12\_Yorina 012]

- (13) [not                      ari]o [wamen]<sub>Aerg</sub> narision  
n-ot                      ari      wam=en      na-tisi=on  
1SG.GEN-brother    MED      pig=ERG      eat-3SG.PST=AM  
ari    fahet    wilahi      o      ibagmare...  
ari    fahet    wilat-ihī      o      ibag=mu=te  
MED    for      stay-1SG.PRS DMV    say:3SG.REM.PST=DS=TOP  
‘because my brother was eaten by a pig, I am staying here, she said...’ [man\_and\_pig  
184]

In a previous study, Riesberg (2018) investigated the factors that facilitate the use of the ergative in clauses with unmarked word order, and demonstrated that optional ergative marking in Yali shares many of the features that have been described for differential argument marking in other Papuan languages. That is, the optional ergative usually requires a dynamic event with a prototypical agent. However, in certain contexts this preference can be ‘overridden’, e.g., to mark contrast, or to evoke a dynamic/volitional reading of otherwise unergative or stative predicates. Yet, just like in Tima, none of these features obligatorily trigger ergative marking. In natural discourse, Yali speakers always have the choice between a non-ergative and an ergative construction, and, as pointed out in Riesberg (2018: 38ff) and Schneider-Blum et al. (2022: 214f), it is likely that discourse-related factors also play a role in the use (or non-use) of ergative structures. For the current study, we therefore conducted a pilot corpus study of traditional narratives to investigate the discourse contexts in which ergative constructions preferably occur. We were able to identify three different discourse functions for the two Yali ergative constructions, two of which – prominence lending of the non-agent argument, and demarcating larger discourse units – are very similar to the ones identified for Tima. Both of them involve a SHIFT to a referent that was formerly less prominent. The third function is neither directly tied to the discourse participants themselves nor to the structuring of larger units, but rather pertains to the highlighting of events that are crucial for the development and outcome of the narrative. We will discuss these three functions in turn.

The first function is exclusive to the marked-word-order ergative construction. Like in Tima, it serves to temporarily shift the attentional centre to the non-agent argument, without, however, promoting it to become the (most) prominent participant in the following discourse. Both the agent and the non-agent in these contexts are given, and often both participants will be continued in their respective grammatical functions in the following discourse. Interestingly, this is, however, not obligatorily so the case for the agent. Example (13) cited above nicely illustrates this. At the specific point in the narrative when (13) is uttered, both the brother and the pig are given, but both have not been mentioned for quite a while. The pig continues not

to play a role in the following discourse, but the brother, or, more precisely, his bones, become central for following chain of events – even if only as a passive undergoer.

The second function, the demarcation of larger discourse units, is the primary function of the optional ergative in unmarked word order. As already pointed out by Schneider-Blum et al. (2022: 214), it is closely related to what Heeschen (1998: 308ff) calls “forms of summarizing”. Heeschen’s basic observation is that in Eipo (TNG), as in many other Papuan languages, most clauses in narratives will exhibit one full NP at most, and more often only involve pronominal or zero referents (see also de Vries 2006, and Himmelmann & Riesberg, to appear). Only in specific points in the narrative, namely in “concluding statement at the end of an episode” (1998: 311), or in “the summarizing headline of the story the narrator is going to tell” (1998: 309), will all relevant participants be expressed by full referential expressions. In Yali, these summarizing forms frequently involve a differentially marked subject. They thus occur at narrative boundaries, but unlike in Tima, where the ergative construction signals the beginning of a new episode, in Yali they can both introduce or conclude a unit. The following two examples illustrate the introducing, forward-looking perspective in (14), and the concluding, backward-looking perspective in (15).

- (14) i. **Tenus hahon itnoente** ti eke ehek ari  
 Tenus hag=on itno=**en**=te ti eke ehek ari  
 Tenus like=AM DET=**ERG**=TOP song one speak:3SG.IM.PST MED
- ii. ti ehekteg itare  
 ti ehek-teg ita=te  
 song speak:3SG.IM.PST DIST=TOP
- iii. hinare itnore we ti ibmano <urus->  
 hin-are itno=te we ti ibma=no  
 2PL.GEN-friend DEM=TOP ?? song DIST.LOC=GIV
- iv. surukmen warukmen surukmen warukmen  
 su-tuk-men wa-tuk-men su-tuk-men wa-tuk-men  
 do-PROG-CONJ bring-PROG-CONJ do-PROG-CONJ bring-PROG-CONJ  
 surukmen warukmente  
 su-tuk-men wa-tuk-men=te  
 do-PROG-CONJ bring-PROG-CONJ=TOP
- v. o itno hup atuk ahama itare  
 o itno hup at-tuk aha=mu ita=te  
 day DET night become-PROG become:3SG.IM.PST=SIT DIST=TOP

‘(i) **the one like Tenus** sang a song (ii) after he sang a song (iii) he, the song (iv) he did that for a very long time (v) (until) it became night’ [suit\_13 208-214]

- (15) i. *hiyap itnore sa hahon uruk*  
 hiyap itno=te sa hag=on u-tuk  
 woman DET=TOP who like=AM speak-PROG
- ii. *we yohore pinggi pinggi warefesul itno*  
 we yoho=te pinggi pinggi wat-e-fe-ul itno  
 ?? today=TOP dead dead kill-SF-COMPL-1PL.IM.FUT DET
- iii. *ari eluk weregma, eluk atisiregte*  
 ari eluk wereg=ma eluk at-tisi-teg=te  
 MED alive 3SG:be=SIT alive become-3SG.PST-SS.PRIOR=TOP
- iv. *ari roho uruk ari ibag ari*  
 ari tu-oho u-tuk ari ibag ari  
 MED do-ADV speak-PROG MED speak:3.REM.PST MED
- v. ***ir hiyap itnoen***  
 it hiyap itno=en  
 3PL woman DET=ERG

'(i) the women (said) who is talking like that? (ii) we will kill him today  
 (iii) he is alive, he became alive, (iv) he is talking like that, they said (v) **the women**'  
 [suit\_2 586-592]

In rare cases, optional ergative marking also occurs within an episode and without any change in the prominence relations of the participants. In this third function, it seems that the case-marked structure (usually in unmarked AOV order) highlights an event that is important for the further development of the storyline. Consider the following example:

(16) Preceding context: *The Wamena people and the Yali people came in groups.*

***hinare llabahun itnoen hilalug...***  
 hin-are llabahun itno=en hila  
 2SG.GEN-friend llabahun DET=ERG take:SB-SS.SEQ

'llabahun had taken him away' [suit\_15 115]

Following context: *So when they came back to Warikma, they were angry.*

*(They asked) why did you invite us?!*

The case-marked structure in (16) is uttered within a passage about the Wamena and the Yali people. Neither does it introduce a new event chain with llabahun as the main actor, nor does it conclude a preceding one. It also does not shift the attentional centre to the victim (llabahun's brother), who was about to be cooked by the Wamena and Yali people. However, the fact that llabahun has taken his brother constitutes a turning point in the story. It leads to the cruel (but for the audience happy) end of the other two main protagonists of the story, Tanghe and Temhe



(see example below), who had kidnapped Ilabahun’s brother and invited the Yali and Wamena people to come and eat him.

Now, while it generally seems to be the case that the two functions of episode demarcating and event highlighting are predominantly achieved by unmarked word order ergative structures in Yali, marked word order ergative constructions can potentially also occur in these contexts. This is, for example, the case at the very end of narratives, when the narrator concludes the story by letting the audience know what happened to the main protagonists. (17) is an example of such a concluding remark (it follows closely to the utterance in (16) and it is the direct consequence of the fact that Ilabahun freed his brother).

(17) *Tanghe itno Yali fil=en isag sa-lug eke*  
 Tanghe DET **east direction=ERG** cook SB:do-SS.SEQ and  
*mel Temhe itno Hubula filen watsalug...*  
 and Temhe DET **west direction=ERG** kill-SS.SEQ

‘Tanghe was cooked by the people from the East, and  
 Temhe was killed by the people from the West, and...’ [suit\_15 123-128]

Thus, lending prominence to an (at that point) non-prominent non-agent, on the one hand, and boundary marking or event highlighting, on the other hand, do not exclude each other.

Two issues are important to note at this point. First, and as mentioned at the beginning of this section, the event-highlighting function of differential case marking is not linked to reference-related prominence, but rather to temporal prominence structures (as established in Becker & Egetenmeyer 2018). Second, we can raise the question of how the unit-demarcating function interacts with reference-related discourse prominence. While in most cases in the sub-corpus used for this pilot study, unit-demarcating involves a shift to a less prominent referent, a prominent referent in Yali can actually also be continued after an episode break. As mentioned in Section 4.1, similar observations hold for Tima, too, where a continuing prominent referent is re-introduced with a maximal referring expression after a boundary. In certain ways this is not surprising, as it is well known that episode boundaries interfere with the “default” mechanisms of referent management. For example, full noun phrases are frequently used when a referent is re-mentioned at the beginning of a new paragraph or episode, even if its antecedent was mentioned in the immediate previous discourse and is thus highly topical and would otherwise be referred to by a pronominal or zero expression (see, amongst others, Tomlin 1987; Vonk et al. 1992; Same 2022). It therefore has been claimed that boundaries constitute “barriers” for anaphora, which basically ‘deactivate’ all preceding information (e.g., Hofmann 1989: 241). In a similar way Riesberg et al. (submitted) claim that at episode boundaries in Totoli a deactivation or re-set of prominence values takes place. It is highly

interesting that ergative structures seem to play a central role in (re-) establishing a prominent referent in this context.

## **5. Conclusion**

In this contribution, we pursued the hypothesis that special encoding strategies in morphosyntax – in particular differential argument marking – are sensitive to discourse prominence, specifically, to the ranking of referents in discourse, taking into account the prominence status of the referents as well as any effects on subsequent discourse structuring. We presented case studies of differential object indexing (DOI), differential object marking (DOM) and differential agent/subject marking (DAM/ DSM) in five typologically diverse languages (Bulgarian, Romanian, Spanish, Tima, Yali). In each language, previous accounts had highlighted the importance of semantic and information-structural factors, but also provided (more or less explicit) indications that discourse factors might be of relevance. Explanations in terms of discourse factors have received much less attention, though, and they were often proposed in the form of hypotheses that might be able to account for otherwise unexplained patterns, but that require further empirical testing. Building on these previous accounts, each case study presented recent and on-going experimental and/or corpus-based research that targeted the discourse functions of differential marking in these languages. Given the exploratory character of our own research, as well as the dearth of relevant previous studies, we are aware that we have only scratched the surface, often raising more questions than we can answer. But we believe that the results do suggest an important place for discourse prominence in our understanding of differential marking, irrespective of the type or source of the marker.

Clearly, it is not surprising that the connection of discourse to morphosyntax is more transparent or more direct for some markers than for others. This is to be expected in the case of person markers (including pronominal markers) that bear a more direct link to reference management in discourse (Siewierska, 2004; Siewierska & Bakker, 2012), and we believe that this insight transfers to the pronominal elements in the case of DOI. In section 2, we presented recent evidence that DOI in Bulgarian is sensitive to fine-grained differences in discourse prominence and serves as a prominence-lending cue for referents that initially have a medium-level discourse prominence rank based on the previous context. It is less clear however, whether this shifting causes only a short-lived elevation in prominence or a more sustainable change in the discourse representation. Hence, future research on the forward-looking potential of DOI in Bulgarian and other languages is needed. Bulgarian might be a case where that link is still highly available since it is still fully transparent that DOI consists of a pronominal element – highlighting the direct link to discourse.

For Romanian, there is initial evidence that DOI – in interaction with DOM – has such a forward-looking function, establishing a referent that becomes prominent in subsequent discourse. As was discussed in section 3.2, the study on Romanian revealed that direct object referents with DOM were more frequently picked up in subsequent discourse than those without DOM – an effect that was reinforced by DOI – but it was not possible to replicate that finding in a follow-up study with a comparable design on Spanish.

This could suggest that differential object marking (with case flags) is less affected by discourse prominence than differential object indexing, but further research on Spanish (and other languages) is clearly needed, including paying attention to interactions between DOM and DOI in Spanish. In this context, it has to be kept in mind that the design of the empirical studies differed across the three languages, i.e., differences may arise because the studies capture different aspects of discourse prominence: the experimental work on Bulgarian focused primarily on the backward-looking potential, while the paragraph continuation studies on Spanish and Romanian both focused on the forward-looking potential, but differing in the number of continuing sentences (one in Spanish vs. five in Romanian).

Finally, evidence from DAM systems presented in section 4 suggested that ergative marking in Tima and Yali is implicated in the discourse function of shifting attention to a referent. In both languages, case marking is used in two contexts. First, to elaborate on a non-agent referent in O function (for a short period of time only), thus making the non-agent argument the attentional centre. A second function was associated with narrative boundaries. There, speakers use case marking to establish an agent as the most prominent referent for a wider stretch of discourse, or – less commonly – to confirm a participant as the most prominent referent of the preceding episode. Comparable to DOI in Bulgarian, this marking tends to either elevate a less prominent referent or re-activate a prominent referent.

In both languages, case marking interacts with word order, i.e., it is not possible to clearly separate the contributions of case marking and word order. However, the comparison between Tima and Yali allows us to hypothesize that the marked word order (with O occurring before A) is a central aspect of the first function, but not of the second. We would expect that this generalization would hold for a larger sample of languages. More generally, it is noteworthy that interactions of differential marking and word order seem to be frequent not only in DAM systems, but also in DOI and DOM. As was mentioned in section 2, DOI in Bulgarian appears more frequently (but, despite earlier claims, not obligatorily) with (non-canonical) sentence-initial objects. This observation is particularly relevant since word order is often interacting with information structuring, and an important aspect of future studies would be to clearly distinguish between the effects of word order and differential marking.

In this contribution, we have followed up on intriguing remarks and proposals in the literature on the five languages that indicate a more prominent role for discourse factors in

differential argument marking than previously assumed. While we cannot provide conclusive answers, and while many questions have to remain open for now, we nevertheless believe that we have accumulated enough evidence to substantiate the hypothesis that differential argument marking supports reference management in discourse. Specifically, we hope to have shown that the concept of discourse prominence provides a promising framework for any future research on differential argument marking.

### **Acknowledgments**

The research for this paper has been funded by the German Research Foundation (DFG) as part of the SFB 1252 *Prominence in Language* (Project-ID 281511265) at the University of Cologne.

### **Abbreviations**

1 – first person; 2 – second person; 3 – third person; ACC – accusative; ADV – adverbial; APPL – applicative; ART – article; BEN – benefactive; COMPL – completive; CONJ – conjunction; DEM – demonstrative; DET – determiner; DIST – distal; DOM – differential object marker; DS – different subject; ERG – ergative; F – feminine; FUT – future; GEN – genitive; GIV – given; IM – immediate; IMP – imperative; INS – instrumental; LOC – locative; M – masculine; MED – medial; NEG – negation; P – person; PL – plural; PRS = present tense; PST – past tense; PRF – perfect; PRIOR – prior (to the event denoted in the previous clause); PRON – pronoun; PTCP – participle; PROG – progressive; REF – referential; REM – remote; REP – reportative; SEL – selective; SEQ – sequential; SG – singular; SIT – situational; SS – same subject; TOP – topic; TR – transitive

### **References**

- Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory* 21(3). 435–483.
- Andersen, Torben. 1988. Ergativity in Pāri, a Nilotic OVS Language. *Lingua* 75. 289–324.
- Andersen, Torben. 2015. Syntacticized topics in Kurmuk: A ternary voice-like system in Nilotic. *Studies in Language* 39(3). 508–554.
- Anderson, Neil & Martha Wade. 1988. Ergativity and control in Folopa. *Language and Linguistics in Melanesia* 19. 1–16.
- Barraza, Georgina. 2003. *Evolución del objeto directo inanimado en español*. Ciudad de México: Universidad Nacional Autónoma de México PhD dissertation.
- Becker, Laura & Gertrud Schneider-Blum. 2020. Morphological marking of contrast in Tima. *Glossa: A Journal of General Linguistics* 5(1). 125.
- Becker, Martin & Jakob Egetenmeyer. 2018. A prominence-based account of temporal discourse structure. *Lingua* 214. 28–58.

- Belloro, Valeria A. 2007. *Spanish Clitic Doubling: A Study of Syntax-Pragmatics*. Albany: State University of New York PhD dissertation.
- Belloro, Valeria A. 2015. *To the Right of the Verb: An Investigation of Clitic Doubling and Right Dislocation in Three Spanish Dialects*. Newcastle upon Tyne: Cambridge Scholars.
- Bickel, Balthasar, Alena Witzlack-Makarevich & Taras Zakharko (2014). Typological evidence against universal effects of referential scales on case alignment. In: Ina Bornkessel-Schlesewsky, Andrej Malchukov & Marc Richards. *Scales and Hierarchies: a cross-disciplinary perspective on referential hierarchies*, 7–44. Berlin & Boston: De Gruyter.
- Bossong, Georg. 1982. Historische Sprachwissenschaft und Empirische Universalienforschung. *Romanistisches Jahrbuch* 33(1): 17–51.
- Bossong, Georg. 1985. *Empirische Universalienforschung: Differentielle Objektmarkierung in Den Neuiranischen Sprachen*. Tübingen: Narr.
- Bossong, Georg. 1998. Le Marquage Différentiel de l'Objet Dans Les Langues d'Europe. In Jack Feuillet (ed.). *Eurotyp 2, Actance et Valence Dans Les Langues de l'Europe*, 193–258. Berlin & Boston: De Gruyter.
- Casaretto, Antje, Gerrit J. Dimmendaal, Birgit Hellwig, Uta Reinöhl & Gertrud Schneider-Blum. 2020. Roots of ergativity in Africa (and beyond). *Studies in African Linguistics* 49(1). 111–140.
- Chafe, Wallace. L. 1976. Givenness, contrastiveness, definiteness, subjects, topics and point of view. In Charles. N. Li (ed.), *Subject and topic*, 27–55. New York: Academic Press.
- Chappell, Hilary & Jean-Christophe Verstraete. 2019. Optional and alternating case marking: Typology and diachrony. *Language and Linguistics Compass* 13(3). e12311 <https://doi.org/10.1111/lnc3.12311>
- Chiriacescu, Sofiana. 2014. *The discourse structuring potential of indefinite noun phrases: special markers in English, German and Romanian*. Stuttgart: University of Stuttgart. <http://dx.doi.org/10.18419/opus-5395>
- Chiriacescu, Sofiana & Klaus von Heusinger. 2010. Discourse prominence and pe-marking in Romanian. *International Review of Pragmatics*, 2(2), 298–332.
- Christensen, Steve. 2010. Yongkom discourse: Ergativity and topic. In Joan Hooley (ed.), *Papers on six languages of Papua New Guinea*, 1–39. Canberra: Pacific Linguistics.
- Comrie, Bernard. 1989. *Language universals and linguistic typology: Syntax and morphology*. 2nd edn. Chicago: University of Chicago Press.
- Compensis, Paul. 2022. *Differential object indexing in Bulgarian: The role of discourse prominence and predictability*. Cologne: University of Cologne PhD dissertation.
- Creider, Chet A. & Jane T. Creider. 1983. Topic – comment relations in a verb-initial language. *Journal of African Languages and Linguistics* 5(1). 1–15.
- Creider, Chet A. & Jane T. Creider. 1989. *A grammar of Nandi*. Hamburg: Helmut Buske

- Creissels, Denis, Gerrit J. Dimmendaal, Zygmunt Frajzyngier & Christa König. 2009. Africa as a morphosyntactic area. In Bernd Heine & Derek Nurse (eds.), *A linguistic geography of Africa*, 86–150. Cambridge: Cambridge University Press.
- Croft, William. 2003. *Typology and Universals*. 2nd edn. Cambridge: Cambridge University Press.
- Cyxun, Gennadij. 1962. Mestoimennata enklitika i slovoredät v bälgarskoto izrečenie [Clitic pronouns and word order in Bulgarian spoken language]. *Bälgarski ezik* 12(4), 283–291.
- Dalrymple, Mary & Irina Nikolaeva. 2011. *Objects and information structure*. Cambridge: Cambridge University Press.
- Dimmendaal, Gerrit J. 1986. Prominence hierarchies and Turkana Syntax. In Gerrit J. Dimmendaal (ed.), *Publications in African languages and linguistics*, 127–148. Berlin & Boston: De Gruyter.
- Dimmendaal, Gerrit J. 2014. Marked nominative systems in Eastern Sudanic and their historical origin. *Afrikanistik Online* 11(3).
- Dimmendaal, Gerrit J. 2017. Areal contact in Nilo-Saharan. In Raymond Hickey (ed.), *Cambridge handbook of areal linguistics*, 446–470. Cambridge: Cambridge University Press.
- Dixon, Robert M. W. 1979. Ergativity. *Language* 55. 59-138.
- Dolscheid, Sarah, Maria Bardají i Farré & Gertrud Schneider-Blum. Submitted. Key challenges for agent prominence – and reasons for why it may still play a (proto-)role.
- Dowty, David. 1991. Thematic Proto-Roles and Argument Selection. *Language*, 67(3). 547–619.
- Enç, Mürvet. 1991. The Semantics of Specificity. *Linguistic Inquiry* 22. 1–26.
- Endriss, Cornelia & Hans-Martin Gärtner. 2005. Relativische Verb-Zweit Sätze und Definitheit. In: Franz-Josef. d’Avis (wd.): Proceedings of the Symposium Deutsche Syntax: Empirie und Theorie, 195-220. Göteborg: Acta Universitatis Gothoburgensis.
- Escandell-Vidal, Victoria. 2007. Topics from Ibiza: Differential object marking and clitic dislocation. Definiteness, specificity and animacy in Ibero-Romance languages. In In Georg A. Kaiser & Manuel Leonetti (eds.). *Proceedings of the Workshop "Definiteness, Specificity and Animacy in Ibero-Romance Languages*, 23–43. Konstanz: University of Konstanz.
- Escandell-Vidal, Victoria. 2009. Differential object marking and topicality. *Studies in Language*, 33(4). 832–884.
- Fedden, Sebastian. 2020. Morphology of Trans New Guinea languages. *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press.
- Fischer, Susann, Mario Navarro, & Jorge Vega Vilanova. 2019. The clitic doubling parameter: Development and distribution of a cyclic change. In Miriam Bouzouita, Anne Breitbarth,

- Lieven Danckaert & Elisabeth Witzzenhausen (eds.), *Cycles in Language Change*, 52-70. Oxford: Oxford University Press.
- Foley, William A. 2000. The languages of new guinea. *Annual review of anthropology*, 29(1). 357–404.
- García García, Marco 2007. Differential Object Marking with inanimate objects. In Georg A. Kaiser & Manuel Leonetti (eds.), *Proceedings of the Workshop "Definiteness, Specificity and Animacy in Ibero-Romance Languages"*, 63-84. Konstanz: University of Konstanz.
- García García, Marco. 2014. *Differentielle Objektmarkierung bei unbelebten Objekten im Spanischen*. Berlin & Boston: De Gruyter.
- García García, Marco & Javier Caro Reina. accepted. Differential Object Marking. In Matthias Heinz, Marc-Olivier Hinzelin & Lorenzo Filipponio (eds.), *Manual of Classification and Typology of the Romance Languages* (Manuals of Romance Linguistics). De Gruyter.
- García García, Marco, Beatrice Primus and Nikolaus P. Himmelmann. 2018. Shifting from animacy to agentivity. *Theoretical Linguistics* 44 (1-2). 25-39.
- García-Miguel, José M. 2015. Variable coding and object alignment in Spanish: A corpus-based approach. *Folia Linguistica* 49(1). 205-256.
- Georgieva, Elena. 1974. *Slovo red na prostoto izrečenie v bălgarskija knižoven ezik* [Word order of the simple sentence in Bulgarian standard language]. Sofia: Bulgarian Academy of Science.
- Givón, Talmy. 1981. On the development of the numeral 'one' as an indefinite marker. *Folia Linguistica Historica* 2. 35-53
- Givón, Talmy. 1983. Topic continuity in discourse: An introduction. In Talmy Givón (ed.), *Topic Continuity in Discourse*, 1-42. Amsterdam & Philadelphia: John Benjamins.
- Greenberg, Joseph. 1966. *Universals of Language*. Cambridge (MA): MIT Press.
- Guentchéva, Zlatka. 1994. *Thématistion de l'objet en bulgare*. Bern: Peter Lang.
- Guentchéva, Zlatka. 2008. Object clitic doubling constructions and topicality in Bulgarian. In Dalina Kallulli and Liliane Tasmowski (eds.), *Clitic doubling in the Balkan languages*, 203–226). Amsterdam & Philadelphia: John Benjamins.
- Haspelmath, Martin. 2013. Argument indexing: a conceptual framework for the syntactic status of bound person forms. In Dik Bakker & Martin Haspelmath (eds.), *Languages across boundaries*, 197–226. Berlin & Boston: De Gruyter.
- Haspelmath, Martin. 2019. Indexing and flagging, and head and dependent marking. *Te Reo*, 62(1). 93–115.
- Haspelmath, Martin. 2021a. Explaining grammatical coding asymmetries: Form–frequency correspondences and predictability. *Journal of Linguistics*, 57(3). 605–633.
- Haspelmath, Martin. 2021b. Role-reference associations and the explanation of argument coding splits. *Linguistics* 59(1). 123–174.

- Heine, Bernd. 1980. Camus, a Maa dialect. In Bernd Heine (ed.), *The non-Bantu languages of Kenya*, 99–140. Berlin: Dietrich Reimer.
- Heeschen, Volker. 1998. *An ethnographic grammar of the Eipo language spoken in the central mountains of Irian Jaya (West New Guinea), Indonesia*. Berlin: Dietrich Reimer Verlag.
- von Heusinger, Klaus & Petra B. Schumacher. 2019. Discourse prominence: Definition and application. *Journal of Pragmatics*, 154. 117–127.
- von Heusinger, Klaus, Tiago Augusto Duarte & Marco García García. In prep. Differential object marking and discourse prominence in Spanish.
- Hill, Virginia & Mardale, Alexandru. 2019. Patterns for differential object marking in the history of Romanian. *Journal of Historical Syntax* 3(5). 1–47.
- Hill, Virginia & Alexandru Mardale. 2021. *The diachrony of Differential Object Marking in Romanian*. Oxford: Oxford University Press.
- Himmelman, Nikolaus P. & Beatrice Primus. 2015. In Amedeo de Dominicis (ed.), *pS-prominenceS: Prominence in Linguistics*, 38-58. Viterbo: Disucom Press.
- Himmelman, Nikolaus P. & Sonja Riesberg. to appear. Discourse patterns and emerging grammar. In Nicholas Evans and Sebastian Fedden (eds.), *The Oxford Guide to the Papuan languages*. Oxford: Oxford University Press.
- Hofmann, Thomas R. 1989. Paragraphs, & anaphora. *Journal of Pragmatics* 13.2, 239–250.
- Hopper, Paul. J., & Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56(2). 251–299.
- lemmolo, Giorgio. 2010. Topicality and differential object marking: Evidence from Romance and beyond. *Studies in Language*, 34(2). 239–272.
- lemmolo, Giorgio. 2011. *Towards a typological study of differential object marking and differential object indexing*. Pavia: Università degli Studi di Pavia, PhD dissertation.
- Isenberg, Horst. 1968. *Das direkte Objekt im Spanischen*. Berlin: Akademie-Verlag.
- Jaeger, T. Florian & Veronica A. Gerassimova. 2002. Bulgarian word order and the role of the direct object clitic in LFG. In Miriam Butt & Tracy. H. King (eds.), *Proceedings of the LFG 02 Conference*, 197–219. Stanford: CSLI Publications.
- Jaeggli, Osvaldo. 1982. *Topics in Romance syntax*. Dordrecht: Foris.
- Jaeggli, Osvaldo. 1986. *The Syntax of Pronominal Clitics*. Leiden: Brill.
- Jerono, Prisca. 2011. *Tugen Word Order: A minimalist perspective*. Nairobi: University of Nairobi PhD dissertation.
- Just, Erika. 2022. *A functional approach to differential indexing: combining perspectives from typology and corpus linguistics*. Amsterdam: LOT Publications.
- Just, Erika & Slavomír Čéplö. 2022. Differential object indexing in Maltese – a corpus based pilot study. In Turek Przemysław & Julia Nintemann (eds.) *Maltese: Contemporary Changes and Historical Innovations*, 105-132. Berlin & Boston: De Gruyter



- König, Christa. 2008. *Case in Africa*. Oxford: Oxford University Press.
- König, Christa, Bernd Heine & Karsten Legère. 2015. *The Akie language of Tanzania: A sketch of discourse grammar*. Tokyo: Research Institute for Languages, Cultures of Asia & Africa.
- Kießling, Roland. 2007. The “marked nominative” in Datooga. *Journal of African Languages and Linguistics* 28(2). 149–191.
- Kliffner, Michael. 1995. El a personal, la kinesis y la individuación. In Carmen Pensado (ed.), *El complemento directo preposicional*, 93–111. Madrid: Visor Libros.
- Laca, Brenda. 1995. Sobre el uso del acusativo preposicional en español. In Carmen Pensado (ed.), *El complemento directo preposicional*, 61–91. Madrid: Visor.
- Laca, Brenda. 2006. El objeto directo. La marcación preposicional. In Concepción Company Company (ed.), *Sintaxis histórica de la lengua española. Primera parte: La frase verbal*, 423–475. México, D.F.: Universidad Nacional de México.
- Lazard, Gilbert. 2003. What is an object in a cross-linguistic perspective. In Giuliana Fiorentino (ed.), *Romance Objects. Transitivity in Romance languages*, 1-16. Berlin & Boston: De Gruyter.
- Leafgren, John. R. 1997. Bulgarian clitic doubling: Overt topicality. *Journal of Slavic Linguistics*, 5(1). 117–143.
- Leafgren, John. R. 2002. *Degrees of explicitness: Information structure and the packaging of Bulgarian subjects and objects*. Amsterdam & Philadelphia: John Benjamins.
- Leonetti, Manuel. 2004. Specificity and differential object marking in Spanish. *Catalan Journal of Linguistics* 3. 75–114.
- Levergood, Barbara Jo. 1987. *Topics in Arusa phonology and morphology (Tanzania)*. Ann Arbor: University of Michigan PhD dissertation.
- Longacre, Robert E. 1990. *Storyline concerns and word order typology in East and West Africa*. Los Angeles: The James S. Coleman African Studies Center and the Department of Linguistics, University of California.
- Longacre, Robert E. 1995. Left shifts in strongly VSO languages. In Pamela Downing & Michael Noonan (eds.), *Word order in discourse*, 331–354. Amsterdam & Philadelphia: Benjamins.
- Lorenzen, Mandy. In prep. *Word order variation in Datooga*. University of Cologne: PhD Thesis.
- McGregor, William B. 1992. The semantics of ergative marking in Gooniyandi. *Linguistics* 30. 275–318.
- McGregor, William B. 2006. Focal and optional ergative marking in Warrwa (Kimberley, Western Australia). *Lingua* 116(4). 393–423.
- McGregor, William B. 2010. Optional ergative case marking systems in a typological-semiotic perspective. *Lingua* 120(7). 1610–1636.

- McGregor, William B. 2013. Optionality in grammar and language use. *Linguistics* 51. 1147–1204.
- Meakins, Felicity & Carmel O’Shannessy. 2010. Ordering arguments about: Word order and discourse motivations in the development and use of the ergative marker in two Australian mixed languages. *Lingua* 120(7). 1693–1713.
- Melis, Chantal. 2018. Spanish indexing DOM, topicality, and the case hierarchy. In Ilija. A. Seržant & Alena Witzlack-Makarevich (eds.), *Diachrony of differential argument marking*, 97–128. Berlin: LangSci Press.
- Melis, Chantal. 2021. From topic marking to definite object marking: Focusing on the beginnings of Spanish DOM. In Johannes Kabatek, Philipp Obrist & Albert Wall (eds.), *Differential object marking in Romance*, 39–64. Berlin & Boston: De Gruyter.
- Mietzner, Angelika. 2016. *Cherang’any: A Kalenjin Language of Kenya*. Cologne: Rüdiger Köppe.
- Miller, Cynthia L. & Leoma G. Gilley. 2001. Evidence for ergativity in Shilluk. *Journal of African Languages and Linguistics* 22(1). 33–68.
- Minčeva, Angelina. 1969. Opit za interpretacija na modela na udvoenite dopälnenija v bälgarskija ezik [An attempt for an interpretation model of doubled complements in the Bulgarian language] In Ljubomir Andrejčin (ed.), *Isvestija na instituta za bälgarski ezik* 17. Sofia: Bulgarian Academy of Sciences.
- Mitchell, Alice & Mandy Lorenzen. 2023a. *Constituent order variation in Datooga narrative and conversation*. Talk presented at the 25th Afrikanist:innentag 2023, May 4–6, Leipzig.
- Mitchell, Alice & Mandy Lorenzen. 2023b. *The dynamics of coding human arguments in Datooga: Integrating functional-typological and socio-interactive approaches*. Talk presented at the 18th International Pragmatics Conference, July 9–14, Brussels.
- Mitchell, Alice, Mandy Lorenzen, Gertrud Schneider-Blum & Birgit Hellwig. Submitted. Split case marking and constituent order in Nilotic languages.
- Moodie, Jonathan & Rosey Billington. 2020. *A grammar of Lopit: an eastern Nilotic language of South Sudan*. Leiden & Boston: Brill.
- Mürmann, Sophie. 2023. *Differential Object Marking and Role Semantics in Romance*. Cologne: University of Cologne PhD dissertation.
- Næss, Åshild. 2004. What markedness marks: the markedness problem with direct objects. *Lingua*, 114(9-10). 1186–1212.
- Næss, Åshild. 2007. *Prototypical transitivity*. Amsterdam & Philadelphia: John Benjamins.
- Nikolaeva, Irina. 1999. Object agreement, grammatical relations, and information structure. *Studies in Language*, 23(2). 331–376.
- Nikolaeva, Irina. 2001. Secondary topic as a relation in information structure. *Linguistics*, 39(1). 1–49.

- Onea, Edgar, and Daniel Hole. 2017. Differential object marking of human definite direct objects in Romanian. *Revue Roumaine de Linguistique* 62(4). 359–377.
- Onea, Edgar & Alexandru Mardale. 2020. From Topic to Object: Grammaticalization of Differential Object Marking in Romanian. *Canadian Journal of Linguistics/Revue Canadienne De Linguistique* 65 (3). 350–92.
- Ovcharova, Bilyana. 2018. Clitic doubling of the proposed direct object in Bulgarian. *Slavia Meridionalis*, 18.
- Payne, Doris. 1995. Verb initial languages and information order. In Pamela Downing & Michael Noonan (eds.), *Word order in discourse*, 449–485. Amsterdam & Philadelphia: Benjamins.
- Popov, Konstantin. 1983 *Gramatika na sǎvremenija bǎlgarski ezik: Vol. III Sintaksis* [Grammar of the contemporary Bulgarian language: Vol. III Syntax]. Sofia: Bulgarian Academy of Sciences.
- Primus, Beatrice. 1999. *Cases and thematic roles: Ergative, accusative and active*. Berlin & Boston: De Gruyter.
- Primus, Beatrice. 2012. Animacy, generalized semantic roles, and differential object marking. In Monique Lamers & Peter de Swart (eds.) *Case, word order, and prominence. Interacting Cues in Language Production and Comprehension*, 65-90. Dordrecht: Springer.
- Reh, Mechthild. 1996. *Anywa language: Description and internal reconstructions*. Cologne: Rüdiger Köppe.
- Riesberg, Sonja. 2018. Optional ergative, agentivity and discourse prominence – Evidence from Yali (Trans-New Guinea). *Linguistic Typology*, 22(1). 17–50.
- Riesberg, Sonja, T. Mark Ellison, Maria Bardají & Nikolaus P. Himmelmann. submitted. Word order and the interaction of prominence hierarchies A case study from Totoli, a western Austronesian symmetrical voice language.
- Rohlf, Gerhard. 1971. Autour de l'accusatif prépositionnel dans les langues romanes. *Revue de Linguistique Romane* 35. 312–334.
- Romero Heredero, Diego. 2022. *Marcado diferencial de objeto y semántica verbal en español*. Berlin & Boston: De Gruyter.
- Rudin, Catherine. 1990. Topic and focus in Bulgarian. *Acta Linguistica Hungarica*, 40(3/4), 429–447.
- Ruiz Zafón, Carlos. 2021. *Sjankata na vjatǎra [The shadow of the wind]*. Sofia: Iztok-Zapad.
- Same, Fahime. 2022. *Referring expression generation in context: Combining linguistic and computational evidence*. Cologne: University of Cologne PhD dissertation.
- Sánchez, Liliana, & Pablo Zdrojewski. 2013. Restricciones semánticas y pragmáticas al doblado de clíticos en el español de Buenos Aires y de Lima. *Lingüística*, 29(2), 271-320.

- Schneider-Blum, Gertrud. 2022. On noncausal/causal alternations in Tima (Nuba Mountains, Sudan). *Linguistique et Langues Africaines* 8(2). 1–34.
- Schneider-Blum, Gertrud. 2023. The impact of attentional centering on ergative marking in Tima. *STUF: Language Typology and Universals* 76(1). 87–112.
- Schneider-Blum, Gertrud & Birgit Hellwig. 2018. Reference tracking in Tima and its interplay with split ergative marking. *Studies in Language* 42(4). 970–993.
- Schneider-Blum, Gertrud, Sonja Riesberg, Birgit Hellwig & Nikolaus P. Himmelmann. 2022. On resolving prominence conflicts – exceptional case marking in Tima and Yali. In Chiara Gianollo, Łukasz Jędrzejowski & Sofiana I. Lindemann (eds.), *Paths through meaning and form*, 211–215. Cologne: USB Monographs.
- Schikowski, Robert. & Giorgio Iemmolo. 2015. *Commonalities and differences between differential object marking and indexing*. Available at: <https://www.comparativelinguistics.uzh.ch/schikowski/publications/2015-domvs-doi.pdf>
- Schultze-Berndt, Eva. 2017. Interaction of Ergativity and Information Structure in Jaminjung (Australia). In Jessica Coon, Diane Massam & Lisa Demena Travis (eds.), *The Oxford handbook of ergativity*, 1089–1113. Oxford: Oxford University Press.
- Schütze, Carson. T. & Jon Sprouse. 2018. Judgment data. In Robert. J. Podesva & Devyani. Sharma (eds.), *Research Methods in Linguistics*, 27-50. Cambridge (UK): Cambridge University Press.
- Scott, Graham. 1986. On ergativity in Fore and other Papuan languages. *Papers in New Guinea Linguistics* 24. 167–175. Canberra: Pacific Linguistics.
- Siewierska, Anna. 2004. *Person*. Cambridge: Cambridge University Press.
- Siewierska, Anna, and Dik Bakker. 2012. Case and alternative strategies: Word order and agreement marking. In Andrej. L. Malchukov and Andrew Spencer (eds.), *The Oxford handbook of case*, 290-303. Oxford: Oxford University Press.
- Silverstein, Michael. 1976. Hierarchy of features and ergativity. In R.M.W. Dixon (ed.), *Grammatical categories in Australian languages*, 112–171. Canberra: Australian National University.
- Song, Jae Jung. 2001. *Linguistic typology: Morphology and syntax*. Harlow: Longman.
- Storch, Anne. 2014. *A grammar of Luwo: An anthropological approach*. Amsterdam & Philadelphia: John Benjamins.
- Suter, Edgar. 2010. The optional ergative in Kâte. In John Bowden, Nikolaus P. Himmelmann, and Malcolm Ross (eds.), *A journey through Austronesian and Papuan linguistic and cultural space. Papers in the honour of Andrew Pawley*, 423–437. Canberra: Pacific Linguistics.

- de Swart, Peter. 2006. Case markedness. In Leonid Kulikov, Andrej L. Malchukov & Peter de Swart (eds.), *Case, valency and transitivity*, 249–267. Amsterdam & Philadelphia: John Benjamins.
- Tigău, Alina & Klaus von Heusinger. 2020. Clitic doubling, DOM and discourse prominence in Romanian. Paper presented at the 10<sup>th</sup> international workshop on Referential Expressions in Discourse *Prominent Inferences*, University of Cologne.
- Tomlin, Russell S. 1987. Linguistic reflections of cognitive events. In Russell S. Tomlin (ed.), *Coherence and grounding in discourse*, 455–479. Amsterdam/Philadelphia: John Benjamins.
- Vogelzang, Margreet, Francesca, Foppolo, Maria T. Guasti, Hedderik van Rijn and Petra Hendriks. 2020. Reasoning about alternative forms is costly: The processing of null and overt pronouns in Italian using pupillary responses. *Discourse Processes*, 57(2), 158–183.
- Vonk, Wietske, Letticia G.M.M. Hustinx, and Wim H.G. Simons. 1992. The use of referential expressions in structuring discourse. *Language and Cognitive Processes* 7(3). 301–333.
- de Vries, Lourens. 2006. Areal pragmatics of New Guinea: Thematization, distribution and recapitulative linkage in Papuan narratives. *Journal of Pragmatics* 38, 811–828.
- Weissenrieder, Maureen. 1991. A functional approach to the accusative a. *Hispania* 74(1). 146–156.
- Witzlack-Makarevich, Alena, Taras Zakharko, Lennart Bierkandt, Fernando Zúñiga & Balthasar Bickel. 2016. Decomposing hierarchical alignment: Co-arguments as conditions on alignment and the limits of referential hierarchies as explanations in verb agreement. *Linguistics* 54(3). 531–561.
- Witzlack-Makarevich, Alena & Ilja Seržant. 2018. Differential argument marking: Patterns of variation. In Ilja A. Seržant & Alena Witzlack-Makarevich (eds.), *Diachrony of differential argument marking*, 1–40. Berlin: LangSci Press.