

# Case and referential properties

February 12, 2009

## Abstract

In this paper we discuss a number of languages with a multidimensional Differential Object Marking (DOM) system. In such languages overt object marking is determined by more than one argument feature. We will show that such argument features can be related to case marking in different ways. On the one hand, they can trigger the occurrence of overt marking, on the other they can be the result of it. We will demonstrate that different languages may prioritize the different argument features in different ways. These cross-linguistic patterns call for a more flexible approach to DOM than hitherto developed. We develop a sign-based declarative model that does not rely on hierarchies. Hierarchies are argued not to be necessary for language-specific descriptions but only as a comparative concept in cross-linguistic comparison.

KEY WORDS: differential object marking, hierarchies, animacy, definiteness, specificity

## 1 Introduction

In recent years the phenomenon of Differential Object Marking (DOM) has received considerable attention in the literature (Bossong, 1985; Aissen, 2003; Jäger, 2003; de Swart, 2007; von Heusinger, 2008; Malchukov, 2008). In a language with a DOM-system direct objects are overtly (case) marked often depending on certain features of the argument. Most commonly these features are the animacy, definiteness, or specificity of the object argument or any combination thereof. For instance, in Spanish the presence of the object marker *a* depends on the animacy and definiteness of the object and interacts with its specificity.

Most of the discussion concerning DOM has focused on the rationale underlying it. A recurrent analysis is one in terms of ‘markedness

reversal’: what is unmarked for subjects is marked for objects and vice versa. Under this analysis those objects are most typically marked that resemble prototypical subjects to the largest extent: animate and definite objects. The naturalness or markedness of grammatical functions is often assessed in terms of ranking on a feature scale. That is, the relevant features animacy and definiteness are reinterpreted as a scale and features high on these scales (e.g. human, definite) are typical for subjects whereas those low on the scales (e.g. inanimate, indefinite) are typical for direct objects. With languages argued to choose different cut-off points on these scales, they are not only used in the overall explanation but also in the language-particular descriptions.

Our focus in this paper is not primarily to explain the rationale behind the DOM. Instead we will focus on the characterization of its manifestation in individual languages. This will be mainly concerned with languages exhibiting a multidimensional DOM system in which more than one argument feature is involved. Our main aim is to show that although the same features play a role in different languages they do so in different ways. A recurrent theme will be that in specific DOM patterns those features that are inherent to a noun or a noun phrase take priority over those features that are not. We will point out what the implications of this observation are for theoretical accounts of DOM and we will develop a formal proposal for the interaction of factors in DOM systems.

In our view three different levels of investigation should be distinguished: (i) the characterization of language-specific patterns, (ii) the formulation of cross-linguistic generalizations, and (iii) the explanation of these cross-linguistic generalizations. It seems that in most recent work on DOM these three levels have been conflated into one. Contra to what seems to be generally assumed, we will claim that hierarchies are not needed to characterize the DOM patterns found in the languages under discussion. In order to characterize those language-specific patterns we will make use of descriptive rules which themselves may be idiosyncratic in nature. We argue that hierarchies, if applicable at all, only enter at the higher level of language comparison, functioning as a type of *comparative concept* (Haspelmath, 2008a,b).

This paper is organized as follows: in section 2 we discuss some data from languages with a multidimensional DOM system, and show that inherent and non-inherent features relate differently to case – inherent properties trigger object marking, while non-inherent properties are often the results of object marking. In section 3 we discuss some previous accounts of DOM, and in section 4 we present a rule-based analysis of multidimensional DOM within the formalism of sign

grammars. Essentially, we postulate two rules for combining a direct object with a transitive verb, with optionality of object marking analysed as an overlap of the conditions under which rules apply. The formalism is flexible enough to accommodate the various types of split alternations, and accounts for fluid alternations by employing underspecification of features in the formulation of the rules. Finally, section 5 argues that the flexibility of the formalism is in fact necessary in order to characterize rare DOM patterns which provide exceptions to the cross-linguistic generalization about DOM, and that consequently cross-linguistic tendencies and generalizations should be understood as abstractions from the similar but not identical language-specific patterns. Section 6 concludes the paper.

## 2 Multidimensional DOM

In this section we will discuss data from languages with a multidimensional DOM system. Our goal is to show that when in a language more than one factor interacts with overt object marking these factors may stand in a different relation to the object marker. A clear distinction should be made between factors that *trigger* the occurrence of overt object marking and the ones that are the *result* of the occurrence of overt object marking. Triggers are properties that are either semantically or morphosyntactically intrinsic to an argument and are inert to change. For instance, every noun has a given animacy value which we cannot change by adding or removing overt case marking from the argument. In other words, animacy is a property which is semantically inherent to an argument. Moving beyond the level of the noun to that of the noun phrase or DP, DP-type or ‘syntactic definiteness’ may be a morphosyntactic argument feature that triggers case marking. Danon (2001, 2006) has shown that in modern Hebrew it is not semantic definiteness that triggers the occurrence of the object marker *’et*, but rather whether it is syntactically marked as being definite. In the domain of nouns this (roughly) means that only those objects containing the definite article *ha* can be preceded by the object marker. Syntactic definiteness or DP-type may then be seen as a morphosyntactically intrinsic property of an argument and by adding or removing case from an argument we do not change its DP-type.

In addition to argument features that function as a trigger for overt case marking, we also find features that are the result of the use of overt case marking. Such ‘result’ features are properties that are non-inherent to an argument and are subject to change. This means we are dealing with features that are either not semantically intrinsic features

of a noun or not morphosyntactically coded in the noun phrase. The clearest example of such a feature is the referentiality or specificity of an argument. It is well-known that in many languages the occurrence of overt case marking goes hand-in-hand with the specificity of an object. For instance, in Turkish by adding or removing case from an argument we can change its specificity (von Heusinger and Kornfilt, 2005).

This difference between triggers and results of overt object marking is closely related to the distinction between split and fluid case alternations as proposed by de Hoop and Malchukov (2007). The so-called triggers are always involved in a split case alternation in which overt case marking literally makes a split between categories of a certain dimension. For instance, in a given language objects that are animate may be obligatorily marked whereas those that are inanimate not. In this case absence of case marking on animate objects will result in ungrammaticality. The so-called result features, on the other hand, are always involved in a fluid case alternation in which the use of case applies within a category and has an effect on a dimension different from that category. For instance, accusative case may be used on inanimates in which case they are interpreted as specific or it may be absent in which case there is no claim with respect to their specificity. In a fluid case alternation the presence or absence of case does not correlate with (un)grammaticality but rather with a change in interpretation.

In the remainder of this section we want to demonstrate that in multidimensional DOM systems intrinsic properties ('triggers') take priority over extrinsic ones ('results') and as a result that split alternations take priority over fluid ones. This means that fluid alternations can only occur in those areas of the grammar where split alternations leave room for them. Below we will first show that split alternations indeed take priority over fluid ones. Then we will demonstrate that when different triggers are involved they may prioritize differently in different languages. Finally, we will show that fluid alternations may not only be dominated by properties of arguments but also by more general grammatical principles.

It should be noted that in the discussion we will only focus on the indexing use of DOM and that we disregard any use as an actual disambiguation mechanism. It is well-known that in most languages in which animacy is involved in DOM it is not only in an indexing way but also in a disambiguation way (de Swart, 2007; Primus, 2009). Thus, although we want to show that multidimensional DOM systems are more complicated and varied than hitherto acknowledged, the reader should bear in mind that in reality the systems may be even more

complicated.

## 2.1 Split over Fluid: Accusative Case in Hindi and Kannada

In this section we will focus on the two-dimensional DOM systems of two South Asian languages, Hindi and Kannada. Starting with Hindi, in which direct objects can be marked with *ko*, the same marker that is used for indirect objects. In the present discussion we limit ourselves to the use of *ko* on direct objects that occur without a determiner. The differential use of *ko* on direct objects has received much attention in the literature (see Mohanan 1990; Butt 1993; Singh 1994; McGregor 1995; Aissen 2003; de Hoop and Narasimhan 2005; Kachru 2006; de Hoop and Narasimhan to appear, a.o.) and two factors can be distinguished that influence it. On the one hand, there is animacy as *ko* is obligatory for objects that are human, but not for objects that are animate or inanimate. On the other hand, the occurrence of *ko* is related to the definiteness or specificity of the direct object. Regarding the latter factor authors differ as to whether they take definiteness or specificity to be the primary factor. Mohanan (1990), for instance, seems to relate DOM in Hindi mainly to definiteness, with specificity playing a secondary role. Butt (1993), on the other hand, takes specificity to be the relevant notion but acknowledges that it interacts with definiteness. We will not make a principled choice for one or the other factor. Whether we call the interpretation given to a *ko*-marked direct object definite or specific, and that of an unmarked direct object indefinite or non-specific, does not affect our claim that animacy takes priority over definiteness/specificity in the use of *ko*.

Following Mohanan (1990), human objects have to be obligatorily marked with *ko*. When a human object is marked, it can be interpreted as definite or indefinite. When such an object occurs without *ko* this results in an ungrammatical sentence. This contrast is shown in (1) and (2) for the noun ‘child’:

HINDI (Indo-Aryan; Mohanan 1990:103)

(1) Ilaa-ne bacce-ko uthayaa.  
Ila-ERG child-KO lift.PF  
‘Ila lifted the/a child.’

(2) \*Ilaa-ne baccaa uthayaa.  
Ila-ERG child lift.PF

In the absence of a determiner, inanimate nouns, on the other hand, can either be marked with *ko* or be left unmarked. The use of *ko*

does have repercussions for the interpretation associated with the direct object. An unmarked inanimate can be interpreted as definite or indefinite, as is shown in (3):

HINDI (Indo-Aryan; Mohanan 1990:103)

- (3) Ilaa-ne haar uthaayaa.  
 Ila-ERG necklace lift.PF  
 ‘Ila lifted a/the necklace.’

Definiteness of an inanimate noun is expressed by using *ko*. This is shown for the noun ‘necklace’ in (4) below:

HINDI (Indo-Aryan; Mohanan 1990:104)

- (4) Ilaa-ne haar-ko uthayaa.  
 Ila-ERG necklace-KO lift.PF  
 ‘Ila lifted the necklace.’

The above examples show that both animacy and definiteness play a role in differential object marking in Hindi. Their roles are, nevertheless, clearly differentiated. Consider the table in (5):

(5)

	human	-human
ko	def/indef	def
∅	*	def/indef

From this table we can conclude two things: (i) the use of *ko* on direct objects is primarily triggered by the humanness of the direct object. Absence of this marker on human direct objects results in ungrammaticality indicating that this is a split case alternation; (ii) definiteness does not trigger the use of *ko* but rather is an effect of the use of this marker. If we were to claim that definiteness triggers case marking on Hindi direct objects we would have trouble explaining why indefinite human objects are marked with *ko* as well. Furthermore, it is left unexplained why in the absence of case marking both a definite and an indefinite reading are possible for non-human objects. If definiteness triggers case marking we would expect a definite reading always to co-occur with *ko*.

A similar situation can be found in Kannada, a Dravidian language with a differential object marking system very similar to that of Hindi (cf. Lidz 1999, 2006). As in Hindi the occurrence of accusative case on direct objects interacts with the animacy and specificity of the object. In Kannada, human and animate direct objects are obligatorily marked with accusative case. This is shown for human objects by the contrast in grammaticality between (6) and (7):

KANNADA (Dravidian; Lidz 2006:11)

- (6) \*Naanu sekretari huDuk-utt-idd-eene.  
 I.NOM secretary look.for-NPST-be-1SG  
 ‘I am looking for a secretary.’
- (7) Naanu sekretari-yannu huDuk-utt-idd-eene.  
 I.NOM secretary-ACC look.for-NPST-be-1SG  
 ‘I am looking for a secretary.’<sup>1</sup>

Inanimate objects, on the other hand, can occur with or without accusative case:

KANNADA (Dravidian; Lidz 2006:11)

- (8) Naanu pustaka huDuk-utt-idd-eene.  
 I.NOM book look.for-NPST-be-1SG  
 ‘I am looking for a book.’
- (9) Naanu pustaka-vannu huDuk-utt-idd-eene.  
 I.NOM book-ACC look.for-NPST-be-1SG  
 ‘I am looking for a book.’

As for the interpretation of the direct objects, Lidz notes that an animate direct object marked with accusative case can either be interpreted as non-specific or specific (*de dicto* or *de re* in the terminology used by Lidz). The same holds for inanimate objects without accusative case. Inanimate objects which occur with accusative case have to be interpreted as specific (*de re*). The pattern is summarized in the table in (10):

(10)		animate	inanimate
	ACC	<i>de dicto/de re</i>	<i>de re</i>
	∅	*	<i>de dicto/de re</i>

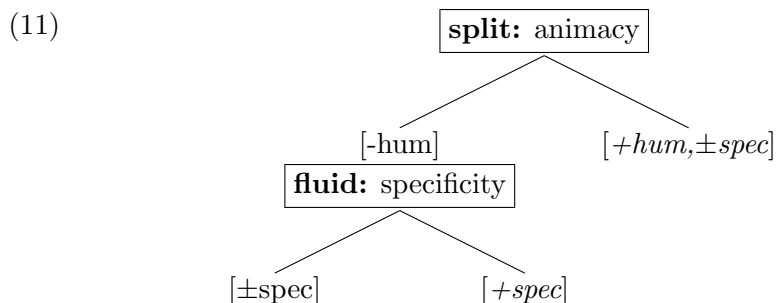
This pattern looks very similar to that of Hindi, as again we find that an analysis of the accusative case as a specificity marker breaks down in the domain of animate direct objects. It cannot be used as a specificity marker when it is required by the animacy of the direct object. In other words, animacy takes priority over definiteness/specificity. As a result, the correlation between accusative case and a strong (definite/specific) interpretation does exist, but not across-the-board. That is, it only holds in the domain of non-humans (Hindi) or inanimates (Kannada). In other words, the fluid alternation can only apply

---

<sup>1</sup>The occurrence of the glide *v* ([w]) or *y* ([j]) in the initial position of the accusative ending is determined by the preceding vowel.

there where the split alternation leaves room for it.

The hierarchical relation between split and fluid alternations in Hindi and Kannada can be schematically depicted as in (11):



## 2.2 Split over Split I: ‘a’-marking in Spanish

The distribution of the Spanish object marker *a* has been extensively discussed in the literature but still is not entirely understood (for discussion, see Brugé and Brugger 1996; Torrego 1998; Delbecque 2002; von Heusinger and Kaiser 2003; Leonetti 2004; Bleam 2005, among many others). We will show that the occurrence of *a* follows an intricate pattern in which different split alternations and a fluid case alternation interact. The factors underlying the case splits are animacy and syntactic definiteness, and the one underlying the fluid case alternation is specificity.

The primary split is between animate and inanimate noun phrases in that only the former can take the object marker. This contrast between animate and inanimate objects can be seen by comparing (12)-(13) to (14):

SPANISH (Romance; Brugé and Brugger 1996:3)

(12) Esta mañana he visto \*(a) Juan/la hermana de  
 this morning have.1SG seen A Juan/the sister of  
 María.  
 María  
 ‘This morning I saw Juan/María’s sister.’

(13) Esta mañana he visto \*(a) mi perro.  
 this morning have.1SG seen A my dog  
 ‘This morning I saw my dog.’

(14) Esta mañana he visto (\*a) la nueva iglesia.  
 this morning have.1SG seen A the new church  
 ‘This morning I saw the new church.’

Examples (12) and (13) show that human and animate objects have to be marked with the prepositional object marker. For the inanimate object in (14) it is prohibited.

The obligatoriness of the object marker with human and animate objects only holds if they are not preceded by an indefinite article. Hence, the second split case alternation in Spanish is determined by lexical definiteness. Indefinite human and animate objects can occur with or without the object marker:

SPANISH (Romance; Hopper and Thompson 1980:256)

- (15) Celia quiere mirar un bailarín.  
 Celia want.3SG watch.INF a ballet.dancer  
 ‘Celia wants to watch a (non-specific) ballet dancer.’
- (16) Celia quiere mirar a un bailarín.  
 Celia want.3SG watch.INF A a ballet.dancer  
 ‘Celia wants to watch (specific) a ballet dancer.’

The absence and presence of *a* correlate with a change in meaning. An indefinite object preceded by the object marker can be interpreted as specific, something which is not possible for an unmarked indefinite object.

In order to establish that lexical definiteness and specificity each play a separate role in Spanish DOM, consider the following examples:

SPANISH (Romance; Leonetti 2004:83, García García 2005:23)

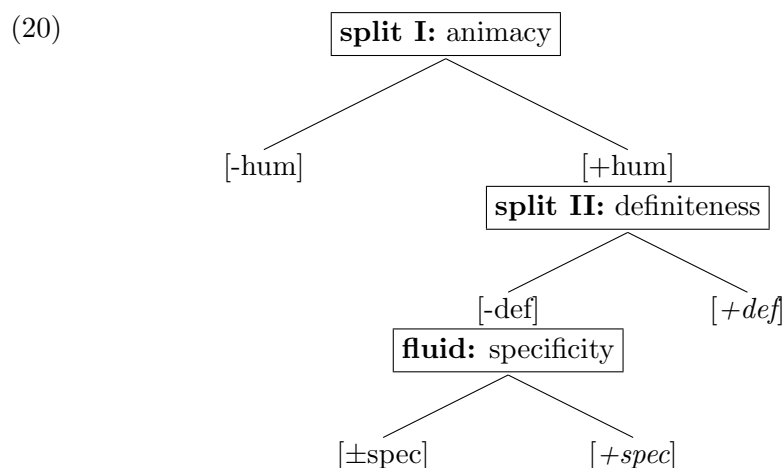
- (17) Está buscando a alguien.  
 be.3SG looking.for A someone  
 ‘(S)he is looking for someone.’
- (18) No está buscando a nadie.  
 not be.3SG looking.for A anyone  
 ‘(S)he is not looking for anyone.’
- (19) Besó a todo el mundo.  
 kissed.3SG A whole the world  
 ‘(S)he kissed everybody.’

The examples in (17)-(19) all involve lexically definite objects and have to be preceded by the object marker. Crucially, despite its presence none of the objects receive a specific interpretation. In fact, the examples all represent non-specific direct objects. This shows that lexical definiteness is a factor independent from specificity. Furthermore, it shows that lexical definiteness takes priority over specificity in determination of the occurrence of the object marker. Only when the

direct object is not lexically definite, can the object marker be used to indicate its specificity. Definiteness itself is in turn outweighed by animacy resulting in the following partial ordering of the relevance of the factors discussed so far: animacy > definiteness > specificity.

The correlation between specificity and the occurrence of *a* is in need of some further discussion. Leonetti (2004) argues that direct objects without *a* can only be interpreted as non-specific. Indefinite objects preceded by *a*, by contrast, can be interpreted as both specific and non-specific. If this is the case, the pattern in Spanish differs from the pattern found in Hindi and Kannada in which the absence of case goes with both a specific and a non-specific reading and the presence of case only with a specific reading.

Spanish thus presents a situation in which a split based on animacy takes priority over one based on DP-type. Moreover, specificity is involved in a fluid alternation which operates in the grammatical space left open by the split alternations. This is schematically depicted in (20):



Turning to Romanian, animacy is the primary factor triggering DOM, since we find the object marker on pronouns, names, definite and indefinite NPs, whereas with inanimate NPs we find the object marker as a rule only with pronouns:

ROMANIAN (Romance)

(21) Televiziunea m=a ales \*(pe) mine, nu eu  
 television ACC.1SG=has chosen PE me not I  
 \*(pe) ea.  
 PE 3SG.FEM

‘Television has chosen me, not I it.’

- (22) L=am                      văzut pe copil-ul  
ACC.MASC=have.1 seen PE child-DEF.M  
vecin-ului.  
neighbor-DEF.M.GEN  
‘I/we have seen the neighbour’s child.’
- (23) L=am                      văzut pe un prieten.  
ACC.MASC=have.1 seen PE a friend.MASC  
‘I/we have seen a friend.’
- (24) Am                      văzut (\*pe) calculatorul-ul  
ACC.MASC=have.1 seen PE computer-DEF.M  
vecin-ului.  
neighbor-DEF.M.GEN  
‘I/we have seen the neighbour’s computer.’
- (25) Am                      văzut (\*pe) un tractor.  
ACC.MASC=have.1 seen PE a tractor.MASC  
‘I/we have seen a tractor.’

So the notion of primary factor or trigger should not always be understood as implying that the distinction it makes is categorical (i.e. that objects with the triggering property can be marked while objects without it cannot be marked).

With indefinite human NPs, Romanian displays a fluid alternation comparable to Hindi or Kannada. When the NPs is object marked, it is interpreted as specific, and when it is not object marked it can be interpreted either as specific or as non-specific.

ROMANIAN (Romance; Dobrovie-Sorin 1994)

- (26) Caut                      o                      secretară.  
look-for.1.SG a.F:DEF secretary.F  
‘I’m looking for a (any or specific) secretary.’
- (27) Caut                      pe o                      secretară.  
look-for.1.SG PE a.F:DEF secretary.F  
‘I’m looking for a specific secretary.’

### 2.3 Split over Split II: DOM in Mongolian

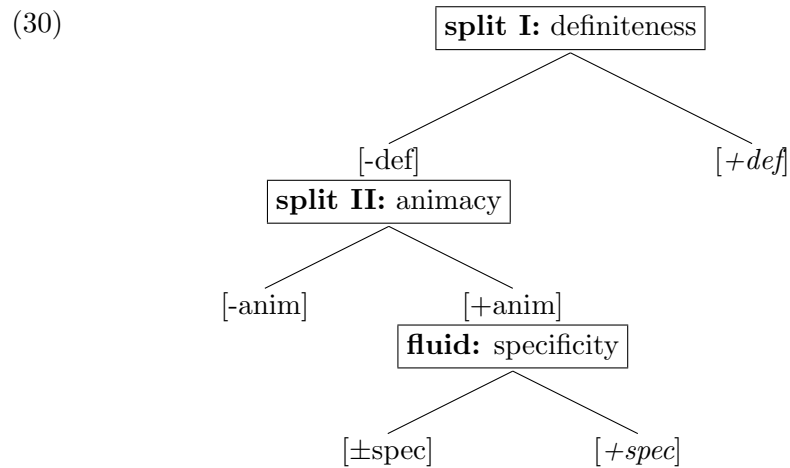
Mongolian presents a multidimensional DOM system in which we find the same type of split and fluid alternations as in Spanish (Guntsetseg, 2009). However, the splits are ordered in a different way than in

Spanish. In this language all definite objects are obligatorily marked with accusative case irrespective of their animacy. This means that the first split is based on definiteness. With indefinite noun phrases the use of the accusative case marker becomes dependent on animacy as it is only found with animate objects. This represents the second split. The object marker is however not obligatory with animate indefinite objects. Instead it is involved in a fluid alternation: when accusative case is used the animate indefinite has to be interpreted as specific, cf. (28), in absence of overt case marking it can be either specific or non-specific, cf. (29).

(28) Bold neg ohin-ig unssen.  
 Bold a girl-ACC kissed  
 ‘Bold kissed a certain girl.’ [specific reading]

(29) Bold neg ohin unssen.  
 Bold a girl kissed  
 ‘Bold kissed a girl.’ [specific or non-specific reading]

This means that the fluid alternation attested in Mongolian follows the general pattern we have seen several times by now. The full range of split and fluid alternations in this language can be schematically depicted as in (30).



## 2.4 Grammatical Principles over Fluid: Accusative Case in Turkish

In Turkish accusative case on a direct object corresponds with a specific reading. The contrast between marked and unmarked direct objects is demonstrated in (31):

TURKISH (Turkic; von Heusinger and Kornfilt 2005:8)

- (31) (Ben) bir kitap oku-du-m.  
I a book read-PST-1SG  
'I read a book.'
- (32) (Ben) bir kitab-ı oku-du-m.  
I a book-ACC read-PST-1SG  
'I read a certain book.'

Von Heusinger and Kornfilt (2005) show, however, that accusative case is only a reliable indicator of specificity when the direct object immediately precedes the verb. In any other position, the use of accusative is obligatory and is compatible with a non-specific reading of the object. This is demonstrated in the following example in which the object 'tea' receives a non-specific (generic) reading:

TURKISH (Turkic; von Heusinger and Kornfilt 2005:11)

- (33) Bizim ev-de çay-ı her.zamen aytül yap-ar.  
our house-LOC tea-ACC always Aytül make-AOR  
'Aytül always makes the tea in our family.'

A similar thing can be observed with the marking of embedded subjects. When they directly precede the verb and are unmarked they receive a non-specific reading, cf. (34), but when they are marked with genitive case they have to be interpreted as specific, cf. (35):

TURKISH (Turkic; von Heusinger and Kornfilt 2005:15)

- (34) /Yol-dan bir araba geç-tiğ-in/-ı gör-dü-m.  
road-ABL a car pass-NMZ-3SG-ACC see-PST-1SG  
'I saw that a car [non-specific] went by on the road.'
- (35) /Yol-dan bir araba-nın geç-tiğ-in/-ı gör-dü-m.  
road-ABL a car-GEN pass-NMZ-3SG-ACC see-PST-1SG  
'I saw that a car [specific] went by on the road.'

Like direct objects, when the embedded subject is moved away from the preverbal position it has to be marked with genitive case and can receive either a specific or non-specific reading. This is illustrated in

(36):

TURKISH (Turkic; von Heusinger and Kornfilt 2005:16)

- (36) /bir araba\*(-nın) yol-dan geç-tiğ-in/-1 gör-dü-m.  
a car-GEN road-ABL pass-NMZ-3SG-ACC see-PST-1SG  
'I saw that a car [non-specific or specific] went by on the road.'

Yet another environment in which the correlation between specificity and overt case marking breaks down involves partitives. When partitive direct objects occur with a lexical head they can surface with or without accusative case, resulting in the familiar difference in interpretation. This can be seen by comparing (37) and (38):

TURKISH (Turkic; von Heusinger and Kornfilt 2005:32)

- (37) Ali büro-ya çocuk-lar-dan iki kız al-acak.  
Ali office-DAT child-PL-ABL two girl take-FUT  
'Ali will hire, for the office, two (non-specific) girls of the children.'
- (38) Ali büro-ya çocuk-lar-dan iki kız-ı al-acak.  
Ali office-DAT child-PL-ABL two girl-ACC take-FUT  
'Ali will hire, for the office, two (specific) girls of the children.'

When the lexical head of the partitive, *kız* in (37) and (38), is missing it has to be replaced by an agreement marker, *sin* in (39). This agreement marker, however, comes with the morphological requirement that it has to be followed by accusative case (in transitive contexts). Due to this formal requirement, accusative case can no longer be used to indicate specificity and as a result partitive objects without a lexical head can be interpreted as both specific and non-specific. This is illustrated in (39):

TURKISH (Turkic; von Heusinger and Kornfilt 2005:34)

- (39) Kitap-lar-dan iki-sin\*(-ı) al, geri-sin-ı  
book-PL-ABL two-AGR(3)-ACC buy remainder-AGR(3)-ACC  
kutu-da bırak.  
box-LOC leave  
'Take (any) two of the books and leave the remainder [of the books] in the box.'

Thus, Turkish provides a good illustration of how formal requirements of the grammar, i.e., word order and agreement, can overrule the otherwise robust correlation between overt case marking and specificity. These formal requirements can be seen as split case alternations, e.g., a split between preverbal and non-preverbal position or between pres-

ence and absence of agreement morphology. In this way, the Turkish system provides additional evidence for the fact that split alternations take priority over fluid ones.

We have seen in this section that in the case of fluid alternations the relation between overt case and the interpretation of the object may be different from one language to the other. In Turkish, overt case indicates specificity and absence of case non-specificity. In Spanish, by contrast, overt marking of the object seems to result in an underspecified interpretation. In most languages (Hindi, Kannada, Romanian, Mongolian), however, overt case indicates specificity whereas absence of case results in underspecification of the specificity of the object.

## 2.5 Procedural and declarative perspectives on fluid case alternations

From a comprehension perspective, the relevant generalizations involving the distribution of the object marker on indefinite human NPs in Romanian can be formulated as follows:

- if a sign  $x$  is an indefinite human ACC-marked NP then  $x$  is specific, or
- if a sign  $x$  is an indefinite human NP which is NOM then  $x$  is either specific or non-specific

From a production perspective, the relevant generalizations are that:

- if sign  $x$  is an indefinite human specific NP then  $x$  is NOM or ACC, or
- if sign  $x$  is an indefinite human non-specific NP then  $x$  is NOM

Note the following two asymmetries in this fluid case alternation. First, if an indefinite human NP is ACC then it is specific, but it does not hold that a specific indefinite NP must be ACC – it may also be NOM. Secondly, if an indefinite human NP is non-specific then it is NOM, but the reverse does not hold. It is not the case that indefinite human NOM-marked NPs are non-specific. Despite these asymmetries it is possible to formulate the relevant descriptive generalization (GEN) from a declarative perspective (i.e. by abstracting from the production and comprehension perspectives):

(GEN) The combination of a indefinite human direct object NP with a transitive verb is grammatical, provided that:

- the NP is NOM-marked, or
- NP is ACC-marked and specific

Note that this descriptive generalization is a disjunction and that it captures the possibility of interpreting a NOM-marked indefinite human NP as specific or non-specific simply by leaving the information about specificity underspecified.

### 3 Previous Accounts of DOM

In the previous section we have demonstrated that argument features can be related to overt object marking in different ways. Such features can be triggers that make the occurrence of overt case marking obligatory. Alternatively, certain argument features such as specificity can be the result of the occurrence of overt case marking. We have seen that the triggers are always involved in a split case alternations, whereas the results partake in fluid alternations. Different languages may prioritize splits in a different way but they always let them dominate fluid alternations. The latter take up the grammatical space left over by split alternations. Thus, multidimensional DOM systems although perhaps looking similar to one another on the surface present a range of variation when one closely examines the interaction of different argument features. It is this variation that has to be captured in a formal approach to DOM.

In the literature on DOM we can distinguish two kinds of approaches. On the one hand, there are researchers who have mainly focused on the shifts in interpretation associated with the occurrence of overt object marking. In the terminology used here, they have concentrated on the fluid alternations in DOM systems. Indeed, many authors have proposed a systematic correlation between case and semantic interpretation (cf. Enç 1991; de Hoop 1992; Butt 1993; Ramchand 1997; Bleam 2005; Danon 2006). This correlation always seems to fall out in the following way that overt/accusative case corresponds with a strong interpretation, i.e., a definite, specific, *de re*, or presuppositional interpretation, and absence of case with a weak interpretation, i.e., an indefinite, non-specific, *de dicto*, or non-presuppositional interpretation. Although such a strong correlation may be observed in certain parts of the grammar, it certainly does not hold across the board. We have seen that the association between case and a strong interpretation can be counteracted by the association of this case with an inherent argument feature (a split case alternation). This means that these approaches although right in pointing out the existence of a correlation between case and certain interpretations are too limited as they can only account for part of the data observed in (multidimensional) DOM systems.

Alternative accounts try to characterize the complete language-

specific patterns of DOM by means of reference to a hierarchy (Bossong, 1985; Aissen, 2003). The formally best worked-out proposal is that of Aissen (2003) who starts from the following simplex hierarchies:

(40) Human > animate > inanimate

(41) Pronoun > proper name > definite NP > indefinite specific NP > indefinite non-specific NP

In order to describe multidimensional DOM systems she crosses the two hierarchies which results in a large matrix ranking a set of composite properties. Although certain property combinations are not ranked universally with respect to one another they have to assume a certain ranking in the language-specific characterization of DOM patterns. We believe that such a hierarchy-based approach is not viable for a number of reasons. First, the crossing of the two hierarchies in order to describe multidimensional DOM systems seems to assume that the two features play an equal role in such systems. However, as we have argued in the previous section, languages seem to give priority to one feature over the other and they may differ in their prioritizing. Thus, instead of fixing the ranking between universally unordered sets of composite properties, languages seem to apply one dimension after the other.

Secondly, this kind of approach seems to presuppose that all features involved in DOM are involved in the same way. However, as we have shown in the previous section some features may trigger the occurrence of overt case marking, whereas others are the result of the occurrence of overt case marking. The inability of hierarchy-based approaches to account for this variation becomes clearest in the case of fluid alternations where they often have to resign to analyses in terms of optionality. This is simply due to the fact that they cannot acknowledge that instead there is a change in the relation between the argument feature and the overt case marking.

On a more fundamental level, in our opinion hierarchies are not needed to state the language-specific patterns of DOM for the languages under discussion. These languages are fundamentally different from those that do make reference to such a hierarchy. For instance, in the Papuan language Awtuw, overt object marking is dependent on the ranking of the object with respect to the subject on an animacy hierarchy such that object marking only occurs when the object is at equal or higher rank than the subject de Swart (2007); Malchukov (2008). This makes this type of language similar to the inverse type where verbal agreement is dependent on the ranking of the object with respect to the subject on a hierarchy. By contrast, in order to deter-

mine object marking in the languages under discussion in this paper no reference has to be made to the subject. This seems to obviate the need for a hierarchy in the characterization of these languages. In the next section we will show how a formal model can be formulated without reference to hierarchies.

## 4 A rule-based analysis of multidimensional DOM

The aim of this section is to provide an analysis of multidimensional DOM in a rule-based sign grammar formalism, in which the notion *sign of language L* is defined in terms of deriving a composite sign by applying a rule to some component signs. The analysis captures both the split as well as the fluid case alternations patterns observed in DOM. Importantly, in this analysis we characterize language-specific multidimensional DOM patterns without making reference to hierarchies.

The following rule-based analysis of multidimensional DOM will be couched in a variant of sign grammars as defined in Kracht (2003, 181). The basic notions of this grammar formalism are the notion of *sign* as a form-category-meaning triple, and the notion of *rule*, which specifies (i) an operation on the formal entities of the component signs, (ii) an operation on the categories of the component signs, and (iii) an operation on the meaning of the component signs. Given a set BS of basic signs and a set of such rules G, a sign  $s_n$  is part of a language L iff (i) it is a basic sign, or (ii) it can be derived relative to BS and G. A sign  $s_n$  can be derived relative to BS and G iff (i)  $s_n$  is in BS, or (ii) there are signs  $s_0, \dots, s_{n-1}$  and an  $n$ -ary rule R such that  $R(s_0, \dots, s_{n-1}) = s_n$ , and the signs  $s_0, \dots, s_{n-1}$  can be derived relative to BS and G (in finitely many steps). Categories are feature structures, and are represented by means of attribute-value matrices. For example, the attribute-value matrix  $[SEM [ANIM : human]]$  represents the category of signs whose semantic value involves (a mental representation of) a human individual.

To account for the two possible combinations of direct objects with a transitive verb, we can now begin with the formulation of the two rules. What they both have in common, is that they assign the patient-like role of the verbal sign to the semantic value of the direct object sign. To account for this we postulate the same semantic operation for both rules, namely:

$$O(ARG, PRED(x, y)) = PRED(x, ARG)$$

For simplicity, we assume that the semantic value of the transitive verb sign is the Curried function  $\lambda y.\lambda x.PRED(x, y)$  and that the semantic operation is functional application (FA):

$$FA(ARG, \lambda y.\lambda x.PRED(x, y)) = [\lambda y.\lambda x.PRED(x, y)](ARG)$$

We also assume that the two rules share the syntactic operation, so that the formal result of combining the direct object sign with the verbal sign is the concatenation of the direct object exponent to the right of the verb exponent.

$$SYN(x, y) = y\_x$$

where  $x$  is the formal exponent of the direct object sign, and  $y$  is the formal exponent of the verb sign.

The categories are combined by means of an operation which results in the same category as the transitive verb sign, except that the value for the SYN attribute is *itv*.

$$CAT([ SYN: pro \sqcup name \sqcup def \sqcup indef ], [ SYN: tv ]) = [ SYN: itv ]$$

Putting these two operations together we get the common core of both rules, namely:

$$\begin{aligned} R_{tr}(\langle e_1, c_1, m_1 \rangle, \langle e_2, c_2, m_2 \rangle) &= \\ &= \langle SYN(e_1, e_2), CAT(c_1, c_2), FA(m_1, m_2) \rangle = \\ &= \langle e_2\_e_1, [ SYN: itv ], m_2(m_1) \rangle \end{aligned}$$

The range of application of the particular rules is determined by specifying the categories of the component signs. So the rules implementing for example the multidimensional DOM pattern in Romanian need to specify (among other things) the following category information:

- the first rule combines a direct object with a transitive verb, provided that the indefinite human NP is NOM-marked
- the second rule combines a direct object with a transitive verb, provided that the indefinite human NP is ACC-marked and specific

So the information that needs to be encoded in the direct object category of the first rule is that the sign is indefinite, human and marked as nominative. To implement these restrictions we need to distinguish two formal features, namely SYN and CASE, where SYN can have (among others) the values *pro, name, def, indef* for nominal signs, and *tv, itv* for verbal signs, and CASE has (among others) the values *nom* and *acc*. In addition, we need to distinguish two semantic features, namely HUM with values  $+, -$  and SPEC with values  $+, -$ .

The first rule can then be formulated as follows:

$$\begin{aligned}
& R_{tr_1}(\langle e_1, c_1, m_1 \rangle, \langle e_2, c_2, m_2 \rangle) = \\
& = \langle SYN(e_1, e_2), CAT(c_1, c_2), FA(m_1, m_2) \rangle = \\
& = \langle e_2 - e_1, [ SYN: \quad itv ], m_2(m_1) \rangle,
\end{aligned}$$

where  $c_1 = \begin{bmatrix} SYN: & indef \\ CASE: & nom \\ HUM: & + \end{bmatrix}$  and  $c_2 = [ SYN: \quad tv ]$

Note that since the feature structure does not specify any value for specificity, it represents the set of all human indefinite NOM-marked signs, irrespective of their value for specificity. Therefore this rule applies to both specific and non-specific human indefinite NOM-marked signs, and thus captures the generalization that NOM-marked indefinite NPs are not restricted to non-specific interpretations.

The second rule requires both ACC-marking and specificity:

$$\begin{aligned}
& R_{tr_2}(\langle e_1, c_1, m_1 \rangle, \langle e_2, c_2, m_2 \rangle) = \\
& = \langle SYN(e_1, e_2), CAT(c_1, c_2), FA(m_1, m_2) \rangle = \\
& = \langle e_2 - e_1, [ SYN: \quad itv ], m_2(m_1) \rangle,
\end{aligned}$$

where  $c_1 = \begin{bmatrix} SYN: & indef \\ CASE: & acc \\ HUM: & + \\ SPEC: & + \end{bmatrix}$  and  $c_2 = [ SYN: \quad tv ]$

The application of the second rule has to be extended to pronouns, names and definite NPs denoting humans, in order to capture the fact that these NPs can also be object marked. The fact that they must be marked follows from the fact that they can be combined with the verb only by this rule. The extension can be achieved by adding a disjunct to the category of the direct object sign.

$$\begin{aligned}
& R_{tr_2}(\langle e_1, c_1, m_1 \rangle, \langle e_2, c_2, m_2 \rangle) = \\
& = \langle SYN(e_1, e_2), CAT(c_1, c_2), FA(m_1, m_2) \rangle = \\
& = \langle e_2 - e_1, [ SYN: \quad itv ], m_2(m_1) \rangle,
\end{aligned}$$

where  $c_1 = \begin{bmatrix} SYN: & indef \\ CASE: & acc \\ HUM: & + \\ SPEC: & + \end{bmatrix} \sqcup \begin{bmatrix} SYN: & pro \sqcup name \sqcup def \\ CASE: & acc \\ HUM: & + \end{bmatrix}$

and  $c_2 = [ SYN: \quad tv ]$

However, since pronouns for inanimate entities are also marked, it is necessary to split the second disjunct into two disjuncts, one accounting for personal pronouns irrespective of the animacy of the entity they stand for, and the other accounting for names and definite NPs

standing for humans.

$$\begin{aligned}
& R_{tr_2}(\langle e_1, c_1, m_1 \rangle, \langle e_2, c_2, m_2 \rangle) = \\
& = \langle SYN(e_1, e_2), CAT(c_1, c_2), FA(m_1, m_2) \rangle = \\
& = \langle e_2 \_ e_1, [ SYN: \text{ itv } ], m_2(m_1) \rangle, \\
& \text{where } c_1 \left[ \begin{array}{l} SYN: \text{ indef} \\ CASE: \text{ acc} \\ HUM: + \\ SPEC: + \end{array} \right] \sqcup \left[ \begin{array}{l} SYN: \text{ pro} \\ CASE: \text{ acc} \end{array} \right] \sqcup \left[ \begin{array}{l} SYN: \text{ name} \sqcup \text{ def} \\ CASE: \text{ acc} \\ HUM: + \end{array} \right] \\
& \text{and } c_2 = [ SYN: \text{ tv } ]
\end{aligned}$$

Analogously, it is necessary to extend the application of the first rule also to cases in which names and definite NPs refer to animals or inanimate entities, resulting in:

$$\begin{aligned}
& R_{tr_1}(\langle e_1, c_1, m_1 \rangle, \langle e_2, c_2, m_2 \rangle) = \\
& = \langle SYN(e_1, e_2), CAT(c_1, c_2), FA(m_1, m_2) \rangle = \\
& = \langle e_2 \_ e_1, [ SYN: \text{ itv } ], m_2(m_1) \rangle, \\
& \text{where } c_1 = \left[ \begin{array}{l} SYN: \text{ indef} \\ CASE: \text{ nom} \\ HUM: + \end{array} \right] \sqcup \left[ \begin{array}{l} SYN: \text{ name} \sqcup \text{ def} \\ CASE: \text{ nom} \\ HUM: - \end{array} \right] \text{ and} \\
& c_2 = [ SYN: \text{ tv } ]
\end{aligned}$$

These two rules account for the following patten of NOM/ACC case alternation in Romanian differential object marking:

	pro	name	def	indef	
				+spec	-spec
human	+	+	+	±	-
non-human	+	-	-	-	-

Note, that in addition to allowing both NOM- and ACC-marking for human specific indefinite NPs (since both rules can apply for this type of direct object), the two rules capture (i) that an ACC-marked indefinite can only be interpreted as specific (only the second rule can apply) and (ii) that a NOM-marked indefinite can be both specific and non-specific (due to the underspecification of the feature SPEC). It should, however, be clear that the above figure is a simplification of the facts in Romanian. For example, we ignored the fact that non-pronominal NPs referring to animals or human definite NPs can be object marked. Nevertheless, we think that the formalism is flexible enough to allow for the characterisation of these facts, too.

In conclusion, then, the generalization (GEN) about the multidimensional DOM pattern in Romanian is captured as follows.

(GEN) The combination of a indefinite human direct object NP with a

transitive verb is grammatical, provided:

- the NP is NOM-marked, or
- NP is ACC-marked and specific

First, by choosing a rule-based formalism, we can implement the two parts of this generalization by means of distinct categorial requirements on two different rules. Given that either one or the other rule has to apply, we capture the disjunction in this generalization. Secondly, the optionality of the ACC on human indefinite NPs is analyzed by means of overlapping conditions under which the two rules apply. And thirdly, the interpretation of an ACC-marked human indefinite NP as specific also follows, because (i) the only rule that can be used to combine this NP with a transitive verb sign is  $R_{tr_2}$ , and this rule imposes a specific interpretation on the direct object which could otherwise be left underspecified for the specificity value. This third point may hold the key to understanding the different behavior of inherent and non-inherent properties associated with DOM. The relevant difference is that non-inherent semantic properties can be left underspecified, whereas inherent semantic features are usually fully specified. What this means is that only with non-inherent properties is there a possibility of specifying an otherwise underspecified property. This may explain why only non-inherent properties partake in fluid case alternations.

## 5 Language description, abstraction and language comparison

The first point to note about the formalism within which this analysis has been expressed is that it does not prevent the formulation of DOM patterns which go against the cross-linguistic DOM generalization. For example, it is possible to formulate a pattern in which names, definite and indefinite NPs are ACC-marked whereas pronouns are morphologically unmarked. The following two rules characterize precisely this pattern:

$$\begin{aligned} R_{tr_3}(\langle e_1, c_1, m_1 \rangle, \langle e_2, c_2, m_2 \rangle) &= \\ &= \langle SYN(e_1, e_2), CAT(c_1, c_2), FA(m_1, m_2) \rangle = \\ &= \langle e_2-e_1, [ SYN: \quad itv ], m_2(m_1) \rangle, \end{aligned}$$

$$\text{where } c_1 = \begin{bmatrix} SYN: & pro \\ CASE: & nom \end{bmatrix} \text{ and}$$

$$c_2 = [ SYN: \quad tv ]$$

$$R_{tr_4}(\langle e_1, c_1, m_1 \rangle, \langle e_2, c_2, m_2 \rangle) =$$

$$\begin{aligned}
&= \langle SYN(e_1, e_2), CAT(c_1, c_2), FA(m_1, m_2) \rangle = \\
&= \langle e_2 \_ e_1, [ SYN: \quad itv ], m_2(m_1) \rangle, \\
\text{where } c_1 &= \left[ \begin{array}{l} SYN: \quad \text{name} \sqcup \text{def} \sqcup \text{indef} \\ \text{CASE:} \quad \text{acc} \end{array} \right] \text{ and} \\
c_2 &= [ SYN: \quad tv ]
\end{aligned}$$

In the face of so much flexibility it is of course justified to ask whether the framework is not too liberal. Should the formalism not better be restricted so that such patterns cannot be characterized? The simple answer to such questions is that this flexibility is necessary, because this pattern of DOM is actually attested in Nganasan (Samoyedic) as discussed by Filimonova (2005).

But then which linguistic patterns of DOM should the grammar formalism exclude? Should it exclude any pattern at all? In view of such a rare but attested pattern, it may be wiser not to exclude any logically possible DOM pattern by restricting the framework. This leads to a more general question: which linguistic patterns should the grammar formalism exclude? In line with Hawkins (2004) and many others, we assume that some missing patterns should be allowed in principle by the grammatical framework, but ruled out by factors which are external to the grammatical framework itself. Patterns which are not universal do not require explanation in terms of restrictions on the grammatical framework. Generally, DOM patterns follow a certain tendency. However, exceptions are attested as discussed in Bickel and Witzlack-Makarevich (2008). Therefore they do not require explanation in terms of restrictions on the grammatical framework.

In our view, cross-linguistic patterns of DOM are abstractions from language-specific patterns, which may or may not be universal. If, for example, the languages of a particular language family (or set of language families) show theoretically interesting similarities in their language-specific pattern of e.g. DOM, this justifies analyzing these similar but not identical patterns as instantiations of one abstract pattern. Note that this does neither require nor imply that the abstract pattern has to be instantiated by all languages with DOM. This cross-linguistic abstract pattern need not be universal, it only needs to be of sufficient theoretical interest. Viewing things this, it is not surprising that that different hierarchies have been proposed for different classes of languages to capture different cross-linguistic generalizations of a cluster of related patterns. So, comparative concepts in the sense of Haspelmath do not need to be universal – they only need to be sufficiently abstract so that theoretically relevant generalizations can be formulated.

Finally, should hierarchies be used not just to formulate but also to explain cross-linguistic generalizations? If these hierarchies represent innate constraints on grammatical knowledge, and there were no exceptions to the cross-linguistic pattern, then appeal to hierarchies may explain the cross-linguistic pattern. However, since cross-linguistic DOM patterns allow for exceptions, they cannot be explained solely by internal factors (innateness), but should be explained by an interaction between internal and external factors.

## 6 Conclusion

After discussing a number of multidimensional differential object marking patterns, we observed that the various argument properties are related to case marking in different ways. After discussing previous accounts of DOM, which typically characterize the language-specific patterns by reference to cross-linguistic animacy or referentiality hierarchies, we propose a rule-based analysis (couched in a sign grammar formalism) which characterizes DOM patterns without reference to such hierarchies. Importantly, the analysis also captures the different ways in which the argument properties relate to case, given that rules can impose e.g. a specific interpretation on an NP which would otherwise be left underspecified for the specificity value. Since only non-inherent properties can be left underspecified, we predict that only such properties can partake in fluid alternations. Finally, we show that the flexibility of the formalism is in fact necessary in order to characterize rare DOM patterns which provide exceptions to the cross-linguistic generalization about DOM.

## References

- Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory* 21(3), 435–483.
- Bickel, B. and A. Witzlack-Makarevich (2008). Referential scales and case alignment: reviewing the typological evidence. Volume 86 of *Linguistische Arbeitsberichte*, University of Leipzig.
- Bleam, T. (2005). The role of semantic type in differential object marking. *Belgian Journal of Linguistics* 19, 3–27.
- Bossong, G. (1985). *Differentielle Objektmarkierung in den Neuiranischen Sprachen*. Tübingen: Narr.

- Brugé, L. and G. Brugger (1996). On the accusative *a* in Spanish. *Probus* 8, 1–51.
- Butt, M. (1993). Object specificity and agreement in Hindi/Urdu. In K. Beats, G. Cooke, D. Kathman, S. Kita, K. McCullough, and D. Testen (Eds.), *Papers from the 29th Regional Meeting of the Chicago Linguistic Society. Volume 1: The Main Session*, pp. 89–104.
- Danon, G. (2001). Syntactic definiteness in the grammar of Modern Hebrew. *Linguistics* 39(6), 1071–1116.
- Danon, G. (2006). Caseless nominals and the projection of DP. *Natural Language and Linguistic Theory* 24(4), 977–1008.
- Delbecque, N. (2002). A construction grammar approach to transitivity in Spanish. In K. Davidse and B. Lamiroy (Eds.), *The Nominative/Accusative and Their Counterparts. Case and Grammatical Relations across Languages*, pp. 81–130. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Dobrovie-Sorin, C. (1994). *The Syntax of Romanian: Comparative Studies in Romance*. Berlin: Mouton De Gruyter.
- Enç, M. (1991). The semantics of specificity. *Linguistic Inquiry* 22(1), 1–25.
- Filimonova, E. (2005). The noun phrase hierarchy and relational marking: Problems and counterevidence. *Linguistic Typology* 9(1), 77–113.
- García García, M. (2005). Differential object marking and informativeness. In K. von Stechow, G. A. Kaiser, and E. Stark (Eds.), *Proceedings of the Workshop ‘Specificity and the evolution/emergence of nominal determination systems in Romance’*, pp. 17–32. Konstanz: Universität Konstanz.
- Guntsetseg, D. (2009). Differential Object Marking in Khalkha Mongolian. In R. Shibagaki and R. Vermeulen (Eds.), *Proceedings of the 5th Workshop on Formal Altaic Linguistics (WAFL5)*, Volume 58. MIT Working Papers in Linguistics.
- Haspelmath, M. (2008a). Comparative concepts and descriptive categories in cross-linguistic studies. Manuscript.
- Haspelmath, M. (2008b). Descriptive scales versus comparative scales. Volume 86 of *Linguistische Arbeitsberichte*, University of Leipzig.

- Hawkins, J. A. (2004). *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.
- von Heusinger, K. (2008). Verbal semantics and the diachronic development of differential object marking in Spanish. *Probus* 20, 1–33.
- von Heusinger, K. and G. Kaiser (2003). Animacy, specificity, and definiteness in Spanish. In K. von Heusinger and G. Kaiser (Eds.), *Proceedings of the Workshop “Semantic and Syntactic Aspects of Specificity in Romance Languages”*, pp. 41–65. Universität Konstanz: Arbeitspapier 113.
- von Heusinger, K. and J. Kornfilt (2005). The case of the direct object in Turkish: Semantics, syntax and morphology. *Turkic Languages* 9, 3–44.
- de Hoop, H. (1992). *Case Configuration and Noun Phrase Interpretation*. Ph. D. thesis, Groningen University. Groningen Dissertations in Linguistics 4.
- de Hoop, H. and A. Malchukov (2007). On fluid differential case marking: A bidirectional OT approach. *Lingua* 117(9), 1636–1656.
- de Hoop, H. and B. Narasimhan (2005). Differential case-marking in Hindi. In M. Amberber and H. de Hoop (Eds.), *Competition and Variation in Natural Languages: The Case for Case*, pp. 321–345. Amsterdam: Elsevier.
- de Hoop, H. and B. Narasimhan (to appear). Ergative case-marking in Hindi. In H. de Hoop and P. de Swart (Eds.), *Differential Subject Marking*. Dordrecht: Springer.
- Hopper, P. J. and S. A. Thompson (1980). Transitivity in grammar and discourse. *Language* 56(2), 251–299.
- Jäger, G. (2003). Learning constraint sub-hierarchies: The bidirectional gradual learning algorithm. In R. Blutner and H. Zeevat (Eds.), *Pragmatics in OT*, pp. 251–287. Houndmills: Palgrave MacMillan.
- Kachru, Y. (2006). *Hindi*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Kracht, M. (2003). *Mathematics of language*. Studies in Generative Grammar No. 63. Berlin: Mouton de Gruyter.

- Leonetti, M. (2004). Specificity and object marking: The case of Spanish *a*. *Catalan Journal of Linguistics* 3, 75–114.
- Lidz, J. (1999). The morphosemantics of object case in Kannada. In *Proceedings of West Coast Conference on Formal Linguistics* 18. Cambridge, MA: Cascadilla Press.
- Lidz, J. (2006). The grammar of accusative case in Kannada. *Language* 82(1), 10–32.
- Malchukov, A. (2008). Animacy and asymmetries in differential case marking. *Lingua* 118(2), 203–221.
- McGregor, R. (1995). *Outline of Hindi Grammar. Third edition*. Oxford: Oxford University Press.
- Mohanan, T. (1990). *Arguments in Hindi*. Ph. D. thesis, Stanford University.
- Primus, B. (2009). Animacy, generalized semantic roles, and case variation. *Lingua*. to appear.
- Ramchand, G. (1997). *Aspect and Predication*. Oxford: Oxford University Press.
- Singh, M. (1994). Thematic role, word order, and definiteness. In M. Butt, T. H. King, and G. Ramchand (Eds.), *Theoretical Perspectives on Word Order in South Asian Languages*, pp. 217–236. Stanford: CSLI Publications.
- de Swart, P. (2007). *Cross-linguistic Variation in Object Marking*. Ph. D. thesis, Radboud University Nijmegen. LOT Publications.
- Torrego, E. (1998). *The Dependencies of Objects*. Cambridge, MA: The MIT Press.