

Differential Object Marking: Iconicity vs. Economy*

Judith Aissen
UCSC

MAY, 2000

DRAFT

aissen@cats.ucsc.edu

* This research was supported by the National Science Foundation, Grant #SBR-9818177. Valuable research assistance was provided by Laura Bozzolo, Rodrigo Gutierrez Bravo, and Oscar Velasco. Earlier versions of this paper were presented at the Stanford/CSLI Workshop 'Is Syntax Different?' (December 1998), the Joint Stanford/UCSC Workshop on Optimal Typology (October 1999), and the UCSD Conference on Explanation in Linguistics (December 1999). Comments from the participants have been very helpful. I especially thank Joan Bresnan for her comments on this work, Donka Farkas, for discussion of issues related to definiteness and to Romanian, Vida Samiian for help with the Persian examples, and my father, Michael Aissen, for discussion of some of the structures which figure in this paper.

1 Differential Case Marking

It is common for languages with overt case marking of direct objects to mark some objects, but not others, depending on semantic and pragmatic features of the object. Following Bossong (1985), I call this phenomenon DIFFERENTIAL OBJECT MARKING (DOM). DOM takes many forms, including the three listed in (1):

- (1) a. Sinhalese, in which case marking is optional, but only animate-referring objects may be case marked (Gair 1970);
- b. Hebrew, in which object case marking is obligatory, but is limited to definite objects (Givón 1978);
- c. Romanian, in which object case marking is obligatory for some objects, optional for others, and excluded for a third set. Those for which it is obligatory are animate-referring personal pronouns and proper nouns (Farkas 1978).

Although these three examples may seem quite disparate, DOM is in fact a highly principled phenomenon. Bossong (1985, VIII) observed [my translation]:

- (2) The structural uniformity of this phenomenon in at least 300 (presently known) languages around the earth is so obvious that one wonders why linguistics has up to now dealt so little with this topic.

What we understand about DOM is due to theoretical work in functional and typological syntax, and to a large body of descriptive work on individual languages.¹ The general understanding of DOM which has emerged can be characterized as in (3):

- (3) The higher in prominence a direct object, the more likely it is to be overtly case marked.

The dimensions along which prominence is assessed include, and are perhaps limited to, animacy and definiteness.²

- (4) a. Animacy scale: Human > Animate > Inanimate
- b. Definiteness scale: Personal pronoun > Proper noun > Definite NP > Indefinite specific NP > Non-specific NP

The relation between (3) and (4) is this: if in some language a direct object at some rank can be case marked, then higher-ranked direct objects in that language can be case marked, but not necessarily lower ranked ones. Returning to our initial examples: in Sinhalese, only objects high in prominence on the animacy hierarchy are case marked; in

¹ Especially important among the former are (Silverstein 1976; Comrie 1979; Comrie 1980; Silverstein 1981; Lazard 1982; Lazard 1984; Bossong 1985; Comrie 1986; Croft 1988; Comrie 1989).

² Person is also a relevant dimension. The distinction between the local persons (1st and 2nd) and the 3rd can be articulated at the top end of the definiteness scale in (4b). Person-driven case is extensively discussed in Silverstein (1976) and Blake (1977), and in the other references of fn. 3. An analysis of such cases in terms like those of this paper is given in Aissen (1999b).

Hebrew, only objects high in prominence on the definiteness hierarchy are case marked; and in Romanian, it is objects which are high in prominence on both hierarchies that must be case marked.

Within this general conception of DOM, there is still substantial room for language-particular variation. First of all, although DOM is very common, there are languages which mark *all* direct objects; in these cases, object marking is not *differential*. Second, even among those that do have DOM, languages differ according to which dimensions are relevant. This is clear from our initial examples. And finally, even when DOM is determined in two languages by the same dimension, languages vary with respect to the ‘cut-off’ point. Turkish, for example, like Hebrew, has DOM determined by the definiteness hierarchy. But while DOM distinguishes definites from indefinites in Hebrew, it distinguishes specifics from non-specifics in Turkish (Enç 1991). The challenge then is to develop a theory of DOM which expresses the generalization in (3), and at the same time allows for the various ways in which DOM can be implemented in particular languages.

An intuition which recurs in the literature on DOM is that it is those direct objects which are most in need of being distinguished from subjects that get overtly case marked. This intuition is sometimes expressed as the idea that the function of DOM is to disambiguate subject from object. There may be cases in which DOM is motivated precisely by the need to disambiguate, but it is also clear that DOM is required in many instances where the absence of case marking could not possibly lead to ambiguity. In a weaker form, the intuition can be understood in the following terms: the high prominence which motivates DOM for objects is exactly the prominence which is unmarked for subjects. Thus, it is those direct objects which most resemble typical subjects that get overtly case marked. Because of the association of subjects with agenthood on the one hand, and topicality on the other, animacy and definiteness are unmarked properties for subjects (Keenan 1976). But they are not unmarked properties for objects. In fact, they seem to be marked properties for objects, in part perhaps because of the pressure to maximally differentiate subject and object.³ Thus, exactly what is marked for objects is unmarked for subjects, and vice versa -- an instance of what has been termed *markedness reversal* (Battistella 1990; Croft 1990; Battistella 1996). Comrie (1979, 19) puts it this way:

...in natural languages, certain grammatical relations tend to be characterized by certain features, in particular [that] subjects tend to be definite, animate, and topic (thematic); while direct objects tend to be indefinite, inanimate, and rhematic.

In later work, Comrie suggests that there is a relation between markedness reversal and markedness of structure:⁴

...the most natural kind of transitive construction is one where the A is high in animacy and definiteness, and the P is lower in animacy and definiteness; and any deviation from this pattern leads to a more marked construction. (Comrie 1989, 128)

This suggests a conception of DOM which is fundamentally iconic: nominals which are marked *qua* objects are morphologically more complex than ones which are unmarked *qua* objects. Functionally, the overt marking of atypical objects facilitates comprehension where it is most needed, but not elsewhere. DOM systems are thus relatively economical.

³Many explanations have been proposed for the association of high prominence elements with subjecthood, and low prominence elements with objecthood. Some are based on agency or the potential for control (Dixon 1994); others are based on viewpoint or empathy (Kuno 1976; Kuno and Kaburaki 1977; DeLancey 1981), or the related notion of predication perspective (Silverstein 1981). Yet others are based on discourse structure: topicality (Wierzbicka 1981; Givón 1994) and information flow (Comrie 1989, 128).

⁴ Comrie uses “A” and “P” as syntactic terms to refer to subject of transitive verb and object of transitive verb, respectively (Comrie 1989, 111).

Given a characterization of the relevant prominence scales, (3) makes clear cross-linguistic predictions. If any inanimate objects are case marked in a language with DOM, then at least some animate objects will be case marked; if any indefinite objects are case marked, then at least some definite objects will be marked, etc. Given the difficulty of discovering absolute linguistic universals, it is remarkable that the generalization in (3) seems to be without exceptions (Bossong 1985, p. 8). This cannot be attributed to the accident of a small number of languages with DOM, for DOM is extremely widespread, found within the Indo-European family (especially in Indo-Iranian and Romance), in Pama-Nyungan, in Dravidian, Uralic, Afro-Asiatic, and elsewhere.

In view of this, one must ask why there have been essentially no satisfactory treatments of DOM in the generative literature. A number of early generative studies of DOM provided quite comprehensive accounts, but do not succeed in expressing the generalization in (3) (e.g. (Isenberg 1968; Farkas 1978)). More recently, there have been very interesting discussions of the relation between DOM and the interpretation of indefinites (Enç 1991; Diesing 1992; Dobrovie-Sorin 1994; de Hoop 1996). But these discussions do not generalize to the broader understanding of DOM which has come out of functionalist and typological work, one which takes into account not only the way DOM distinguishes specific from non-specific indefinites, but also the way it distinguishes pronouns from non-pronouns (in some languages), humans from non-humans (in others), definites from indefinites (in yet others), and complex combinations of these.

A major stumbling block to a formal generative treatment of DOM is the need to refer to prominence scales. These have never been formally integrated into generative syntax, and as a consequence generalizations which depend on them cannot be fully integrated. The fact that DOM is characterized in many languages by a great deal of apparent fuzziness has perhaps reinforced the feeling that the principles underlying DOM are not part of core grammar. However, the exclusion of DOM from core grammar comes at a high cost, since it means that there is no account forthcoming from formal linguistics for what appears to be an excellent candidate for a linguistic universal.

Optimality Theory (OT) provides a way, I believe, to reconcile the underlying impulse of generative grammar to model syntax in a precise and rigorous fashion with a conception of DOM which is based on prominence scales. The purpose of the present paper is to develop an approach to this phenomenon within OT which is at the same time formal *and* expresses the functional-typological understanding of DOM. The analysis developed here succeeds in expressing the generalization in (3), while at the same time allowing for the range of language-particular variation which is observed. Furthermore, it does so in a way which relates DOM to markedness reversal and iconicity.⁵

2 Harmonic Alignment and Markedness Reversal

If markedness reversal lies at the heart of DOM, then the ideal situation would be one in which markedness reversal somehow entered into the formal account of DOM. It is quite unclear how this could be achieved in current conceptions of generative syntax. But developments in OT *phonology*, some general, and some quite specific, provide the conceptual and technical tools to build such an account. The essential conceptual innovation is the idea

⁵Hopper and Thompson (1980) suggest a different interpretation of DOM (see also Magier (1987)). In their account, DOM registers a high degree of clausal transitivity on one of the arguments of the clause. Various factors contribute to degree of transitivity, but one of them is individuation of the object. High rank on the prominence scales of (4) corresponds to a high degree of individuation. This account and the one proposed here (which is based conceptually on the approaches cited in the text) will lead to similar predictions about DOM, but they are conceptualized quite differently. The Hopper and Thompson approach is not related to markedness reversal and therefore does not, as far as I can tell, make predictions about discriminate *subject* marking systems (see section 6 below).

that constraints are universal but violable. The essential technical tool is HARMONIC ALIGNMENT, a formal operation for deriving constraints which express markedness reversal.⁶

2.1 Harmonic Alignment

The OT account of DOM requires, first of all, constraints which characterize the relative markedness of various associations of grammatical function with animacy and definiteness. Harmonic alignment provides the right kinds of constraints (Prince and Smolensky 1993, Ch. 8). Harmonic alignment operates on pairs of scales, aligning each element on one with each element on the other, and it generates constraint subhierarchies which express the relative markedness of each such association. Harmonic alignment is proposed in Prince and Smolensky as part of their account of the relation between sonority and syllable structure. The sonority hierarchy and the various prominence scales relevant to DOM belong to distinct ‘interface’ systems (the acoustic/articulatory system, the system of meaning), but the structural roles that these scales play in language design are strikingly parallel. The formal operation of harmonic alignment provides the means to express this parallelism between phonology and syntax.

The treatment of sonority and syllable structure in Prince and Smolensky is based on the idea that there are two kinds of structural positions in the syllable, the Peak and the Margin (Margin subsumes Onset and Coda), and that the unmarked situation is for the Peak to be filled by a relatively sonorous element, ideally a vowel, and the Margin by a relatively less sonorous one, ideally an obstruent. This is an instance of markedness reversal in the domain of phonology.

Harmonic alignment operates on two scales, one of which must be binary.⁷ It associates the high-ranking element on the binary scale with each of the elements on the other scale, left to right. It associates the low-ranking element on the binary scale with each of the elements on the other scale, right to left. Applied to the scale on syllable position (5) and the scale on sonority (6), it yields the two markedness hierarchies in (7) (where “ $x > y$ ” means “ x is more prominent than y ”, and “ $x \dots y$ ” means “ x is less marked than/more harmonic than y ”):

- (5) Peak > Margin
- (6) Vowel > Sonorant > Obstruent
- (7)
 - a. Peak/Vowel ... Peak/Sonorant ... Peak/Obstruent
 - b. Margin/Obstruent ... Margin/Sonorant ... Margin/Vowel

According to (7a), vowels are the least marked type of peak, and obstruents are the most marked type; according to (7b), obstruents are least marked type of margin, while vowels are the most marked type. Harmonic alignment provides then for the formal representation of markedness reversal.

The hierarchies in (7) express important markedness relations underlying syllable structure. A compelling feature of OT is that these hierarchies can be interpreted as constraint hierarchies and thereby constitute the

⁶ For the use of harmonic alignment in syntax, see for example (Artstein 1998; Aissen 1999b; Asudeh 1999; Lee 1999; Sharma 1999).

⁷ Alignment. Suppose given a binary dimension D_1 with a scale $X > Y$ on its elements $\{X, Y\}$, and another dimension D_2 with a scale $a > b \dots > z$ on its elements. The harmonic alignment of D_1 and D_2 is the pair of harmony scales:

$H_X : X/a > X/b > \dots > X/z$

$H_Y : Y/z > \dots > Y/b > Y/a$

The constraint alignment is the pair of constraint hierarchies:

$C_X : *X/z \gg \dots \gg *X/b \gg *X/a$

$C_Y : *Y/a \gg *Y/b \gg \dots \gg *Y/z$

[Prince and Smolensky 1993, 136]

substance from which grammars are constructed. Constraint hierarchies are derived from (7a,b) by inverting their ordering, and interpreting the ranked elements as avoid constraints:

- (8) a. *Margin/Vowel » ... » *Margin/Obstruent
 b. *Peak/Obstruent » ... » *Peak/Vowel

The high-ranking constraints in (8a,b) penalize vowel margins and obstruent peaks, respectively. Crucially, (8a,b), and more generally, constraint hierarchies derived by harmonic alignment, are universal subhierarchies. Their rankings are fixed, and thus cannot be the locus of language-particular difference. In the case of syllable structure, the fixed rankings in (8a) and (8b) express the linguistic universals underlying the distribution of sonority in the syllable.

2.2 Harmonic Alignment and DOM

Harmonic alignment can likewise be used to characterize the linguistic universals which underlie DOM. In the domain of DOM, the scales which are aligned are the scale on grammatical functions (the *relational scale*) and the scales on animacy and definiteness.

2.2.1 Animacy

Animacy plays a significant role in DOM. Following a great deal of work in diverse grammatical traditions, I assume the relational scale in (9), as well as the animacy scale from (4a), repeated below as (10):

- (9) Relational Scale: Su(bject) > Ob(ject)
 (10) Animacy Scale: Hum(an) > Anim(ate) > Inan(imate) (Croft 1988)

Harmonic alignment of grammatical function with animacy yields two hierarchies, one on subjects and one on objects; each expresses the relative markedness of possible associations with the various degrees of animacy.

- (11) a. Su/Hum ...Su/Anim ...Su/Inan
 b. Oj/Inan ...Oj/Anim ...Oj/Hum

The scale on subjects, (11a), says that human subjects are less marked than animate ones, which are in turn less marked than inanimate ones; the scale on objects, (11b), says the reverse. Again, these markedness hierarchies can be implemented as constraint hierarchies by reversing their rankings and interpreting the elements as avoid constraints:

- (12) a. *Su/Inan » *Su/Anim » *Su/Hum
 b. *Oj/Hum » *Oj/Anim » *Oj/Inan

According to these hierarchies, if any clauses are to be avoided because of the animacy of the subject and/or object, it will be clauses with inanimate subjects and/or human objects.

The constraint subhierarchies in (12a,b) will form the basis for the key constraints in our account of animacy-sensitive DOM, but it is important to recognize that they are independently motivated. The high-ranking constraint in (12a) is relevant in languages like Jakaltek (Craig 1977) and Halkomelem (Gerdt 1988), which exclude inanimate subjects (in transitive clauses) altogether. These constraints (or ones very much like these) are

also relevant in languages like Tzotzil and Chamorro which exclude transitive clauses in which the subject is inanimate and the object is human (Chung 1981; Cooreman 1987; Chung 1989; Aissen 1997; Chung 1998).⁸

2.2.2 Definiteness

In many languages DOM is determined entirely, or in part, by the dimension of definiteness. The definiteness scale, repeated below as (13), extends the familiar hierarchy on definiteness (definite > specific > non-specific) to names and pronouns. The short form of the hierarchy is assumed in some discussions of DOM (Comrie 1986; Croft 1988). The motivation for extending it comes from DOM itself: while it is necessary to distinguish pronouns and names from other definites, they are treated as the high end of the same scale. Thus, there are languages which case mark only personal pronoun objects, or only pronoun and proper noun objects, or pronoun, proper noun and definite objects, but there appear to be none which case mark just definites, excluding personal pronouns and names. I assume then that the definiteness scale has the form in (13), with abbreviations for each element indicated underneath:

(13) Definiteness Scale

Pronoun > Name > Definite > Indefinite Specific > NonSpecific
(Pro > PN > Def > Spec > NSpec)

If this scale is to be part of a cross-linguistically valid approach to DOM, it is necessary that the elements which it orders be defined in a way independent of language-particular morphosyntax. Following ideas and suggestions of Farkas (1997) I suggest that the relevant scale has to do with the ‘extent to which the value assigned to the discourse referent introduced by the noun phrase is fixed’.⁹ In the case of 1st and 2nd person pronouns, the value is fixed by the speech situation. The value of 3rd person is limited to a salient non-participant, while that of a name is fixed by convention. In the case of definite descriptions, the hearer is not free to assign just any value to the discourse referent introduced by the noun phrase. Definites are subject to a familiarity requirement, meaning that the value is determined by previous discourse. Indefinites are subject instead to a novelty requirement, but the degree to which the value assigned to the discourse referent is fixed or free can vary. It is more fixed in the case of partitives like ‘two of the girls’, where the value must be chosen from a familiar set, than it is in the case of ‘free choice’ any (e.g. ‘you can bring any friend’). It is also relatively fixed when used in combination with ‘certain’, (e.g. ‘you should bring a certain friend’). ‘certain’ suggests that the choice of value is restricted or that the choice of value matters. The restriction is often present in the form of a modifier (e.g. ...a certain friend who understands you’).

The degree of freedom available in fixing the value of the discourse referent introduced by an indefinite indeed appears to be highly relevant in determining DOM for indefinites. Enç (1991) identifies two types of

⁸ I say “or ones very much like these” because this prohibition on transitive clauses with high prominence objects and low-prominence subjects is restricted to clauses in which both arguments are 3rd person. In other work (Aissen 1997; Aissen 1999a), I have analyzed these as involving the dimension of obviation (which is restricted to 3rd persons), so that the constraints involved would penalize inanimate proximates, for example, rather than inanimate subjects.

⁹ A scale with very similar form is motivated on different grounds, and discussed in (Davison 1984). Davison (p. 614) characterizes the basis for the scale as follows:

the scale of NP types...ranks them as to how clearly or directly the contents of the NP can be used to pick out a particular individual in the discourse context.

This characterization is clearly related to the notion of individuation which has been linked to transitivity, cf. especially Hopper and Thompson (1980).

indefinites which carry the accusative suffix in Turkish: one is partitives like ‘two of the girls’; the other is indefinites which translate ‘a certain N’. She characterizes these as *specific* indefinites. In a discussion of DOM in Persian, Lazard identifies exactly these two classes of indefinites as the ones which obligatorily take the object suffix *-râ* (Lazard 1982). Along related lines, Isenberg (1968) argues for Spanish that DOM on indefinites forces the hearer to link the indefinite to an already present discourse referent. I will follow Enç in assuming that the class of specific indefinites includes partitive indefinites, but assume that it also includes indefinites which denote a novel discourse referent that is pragmatically salient.

Harmonic alignment between the relational scale and the definiteness scale produces two markedness hierarchies, one on subjects and one on objects:

- (14) a. Su/Pro ...Su/PN ...Su/Def ...Su/Spec ...Su/NSpec
b. Oj/NSpec ...Oj/Spec ...Oj/Def ...Oj/PN ...Oj/Pro

According to (14a), personal pronouns are the least marked type of subject, and non-specific indefinites are the most marked. This is reversed for objects, (14b): personal pronouns are the most marked type of object, and non-specific indefinites, the least marked. From these markedness hierarchies, the constraint subhierarchies in (15) are derived, again by inverting the ranking and interpreting the elements as avoid constraints:

- (15) a. *Su/NSpec » *Su/Spec » *Su/Def » *Su/PN » *Su/Pro
b. *Oj/Pro » *Oj/PN » *Oj/Def » *Oj/Spec » *Oj/NSpec

The constraints in (15a,b) will form the basis for our account of DOM which is entirely, or in part, determined by definiteness. Like the constraints on subject and object animacy in (12a,b), those in (15a,b) are independently motivated.

The two high-ranking constraints in (15a) penalize indefinite subjects, non-specific and specific, respectively. It is known that there are languages which exclude non-specific indefinite subjects, in compliance with the highest constraint in (15a). Diesing and Jelinek (1995) observe that Egyptian Arabic excludes indefinites in subject position ([Spec, IP]) unless they are interpreted partitively. Tagalog likewise has a well-known preference for definite subjects (Foley and Van Valin 1984; Kroeger 1993). It excludes indefinite non-specific subjects, per (15a), but apparently allows indefinites when specific (Kroeger, p. 15). There may be languages which exclude both specific and nonspecific indefinite subjects. Givón (1978, 295) claims that exclusion of indefinite subjects is a categorical constraint in most languages and a generalization which holds at the level of the text-count in others. He cites Bemba as a particular example. Keenan (1976) cites Malagasy. Neither source distinguishes specific from non-specific indefinites. The top constraints in (15a) may also be relevant in languages like Hindi where there is no definite article, and where the default interpretation for subjects (but not objects) is as definite (Singh 1994).

The high-ranking constraint in (15b) penalizes personal pronoun objects. It is relevant in languages like Chamorro (Chung 1984; Chung 1998), Mam (England 1983) and Halkomelem (Gerds 1988), all of which exclude transitive clauses in which the object is a personal pronoun and the subject is a non-pronoun. All these languages resort to constructions other than simple transitive clauses to express the combination of non-pronoun agent and pronoun patient. The top three constraints in (15b) penalize definite objects of various types, and might be relevant in a language like Tagalog, where definite patients cannot be realized as objects, but must be subjects (Foley and Van Valin 1984; Kroeger 1993).

3 Iconicity and Economy

The subhierarchies derived in the previous section penalize marked configurations more forcefully than unmarked ones. Thus, high-ranking constraints from those subhierarchies will be relevant in languages where transitive clauses

involving marked configurations are avoided. However, DOM arises precisely when these marked configurations are *not* avoided. Such configurations are tolerated, presumably because of higher-ranked constraints that penalize clauses in which the patient is not realized as object (cf. (Aissen 1999b)). In these cases, the marked transitive configuration is tolerated but the object is morphologically marked. Hence, what is needed at this point are constraints which characterize the relation between morphological complexity and markedness.

In thinking about how this should be implemented, it is important to focus on the morphology of DOM. In principle, there are various ways in which a language could morphologically distinguish between high- and low-prominence objects. It could mark the low-prominence ones and leave the high-prominence ones unmarked; or it could mark both, but with different morphemes. But this is not the way DOM works. *Overwhelmingly*, DOM is implemented by overtly marking the marked class of objects, and leaving the unmarked ones with no morphological mark (Bossong 1985, 125). Thus, Spanish examples like *Veo la casa* ‘I see the house’ vs. *Veo a la mujer* ‘I see the woman’, are typical: the low prominence (inanimate) object is unmarked (and thus identical to the subject form), while the high-prominence (human) object must be marked.

The morphology of DOM then is *privative*: zero expression contrasts with audible expression. We can understand this opposition in a principled way by assuming that zero expression involves no specification for morphological case, while audible expression involve a positive specification. More specifically, I assume that the representation of nominal arguments may include a value for the feature CASE, e.g. ACCUSATIVE, GENITIVE, DATIVE. Such specification will normally have an audible exponent. But CASE may also be left with no value, in which case there can be no audible exponent. We want then to penalize the absence of case specification most forcefully for high prominence objects. Following a suggestion of Paul Smolensky’s (5/97, p.c.), I propose that the absence of case specification is penalized by the constraint in (16), $\ast\emptyset_C$ (read: *Star Zero*), subscripted with *C* for CASE (see Aissen (1999b)).

(16) $\ast\emptyset_C$ ‘Star Zero’: Penalizes the absence of a value for the feature CASE.

This constraint does not directly force audible case, but since morphological case generally has an audible exponent, it enforces it indirectly.¹⁰

Finally we want to compel case specification (henceforth: case marking) on the most marked objects most forcefully. Since we already have constraints which characterize the relative markedness of objects, namely the subhierarchies in (12) and (15), the question is how to link $\ast\emptyset_C$ to those hierarchies. The operation of local conjunction provides the right connection (Smolensky 1995).¹¹ If we assume that local conjunction of $\ast\emptyset_C$ with the constraint subhierarchies in (12) and (15) preserves the ranking of those subhierarchies, then this operation yields new subhierarchies which characterize the relative markedness of zero case marking for objects of different types (see Aissen (1999b)). Local conjunction yields the subhierarchies shown in Figure 1, arranged vertically for the sake of clarity.

¹⁰ I assume that case marking is not present in input; hence, enforcing its presence in outputs cannot fall to a faithfulness constraint. Case marking signals grammatical function, and its motivation should be sought in the listener-oriented functional principle MINIMIZATION OF PERCEPTUAL CONFUSION (Boersma 1998).

¹¹ The local conjunction of C_1 and C_2 in domain *D*, $C_1 \& C_2$, is violated when there is some domain of type *D* in which both C_1 and C_2 are violated (Smolensky 1995). For the constraints of Figure 1 (below), the relevant domain is the nominal constituent itself.

Local conjunction of * $\emptyset$ with the subhierarchy on object animacy (12b)	Local conjunction of * $\emptyset$ with the subhierarchy on object definiteness (15b)
*Oj/Hum & * $\emptyset_C$ » *Oj/Anim & * $\emptyset_C$ » *Oj/Inan & * $\emptyset_C$	*Oj/Pro & * $\emptyset_C$ » *Oj/Name & * $\emptyset_C$ » *Oj/Def & * $\emptyset_C$ » *Oj/Spec & * $\emptyset_C$ » *Oj/NSpec & * $\emptyset_C$

Figure 1: Local conjunction of * $\emptyset$ with object-oriented subhierarchies

The high-ranking constraint of the left-hand subhierarchy penalizes the absence of case on human-referring objects. Hence, in a DOM system based on animacy, if there are *any* objects that are case marked, that set will include human objects. Likewise, the high-ranking constraint on the right penalizes the absence of case on object personal pronouns, guaranteeing that if any objects are case marked in a definiteness-based DOM system, personal pronouns will be. The effect of local conjunction here is to link markedness of content (expressed by the markedness subhierarchies) to markedness of expression (expressed by Star Zero). That content and expression are linked in this way is a fundamental idea of markedness theory (Jakobson 1939; Greenberg 1966). In the domain of DOM, this is expressed formally through the constraints in Figure 1. Thus they are ICONICITY CONSTRAINTS: they favor morphological marks for marked configurations.

If we say nothing more, the subhierarchies of Figure 1 will force case on *all* objects. Since this is precisely what does *not* happen in DOM systems, some constraint must penalize case. I assume that the relevant constraint is an ECONOMY condition, a version of *STRUC which penalizes the specification of morphological CASE.

(17) *STRUC_C: penalizes a value for the morphological category CASE

*STRUC_C can be interpolated at any point in the subhierarchies in Figure 1, ‘turning off’ case marking of all object types mentioned in the dominated constraints. However, there is no way that a less prominent object can be case marked if more prominent ones are not case marked. Hence, it is the interaction of the constraint hierarchies in Figure 1 with *STRUC_C that constitutes the core of this account. It is this interaction which expresses the generalization in (3): *The higher in prominence a direct object, the more likely it is to be overtly case marked.*

From this perspective, DOM involves the tension between iconicity and economy, a tension which has been extensively discussed in connection with other domains in morphosyntax in Haiman (1985a,b). Iconicity favors the morphological marking of syntactically marked configurations; economy penalizes it. The tension between these two principles is resolved in particular languages through constraint ranking. Section 4 surveys the extent to which the factorial typology defined by the interaction of the constraints in Figure 1 with *STRUC_C is in fact realized. In section 5, I turn to systems in which animacy and definiteness combine to determine DOM.

4 One-Dimensional DOM

The points of possible interpolation of *STRUC_C into the subhierarchies of Figure 1 define the set of possible DOM systems based on one dimension: animacy or definiteness. Here, I consider the extent to which this typology is instantiated by known languages, starting with definiteness.

4.1 Definiteness

Figure 2 shows the points where *STRUC_C can be interpolated among the subhierarchy on object definiteness

(Figure 1, right-hand side), and cites one language for each of the possible DOM types.

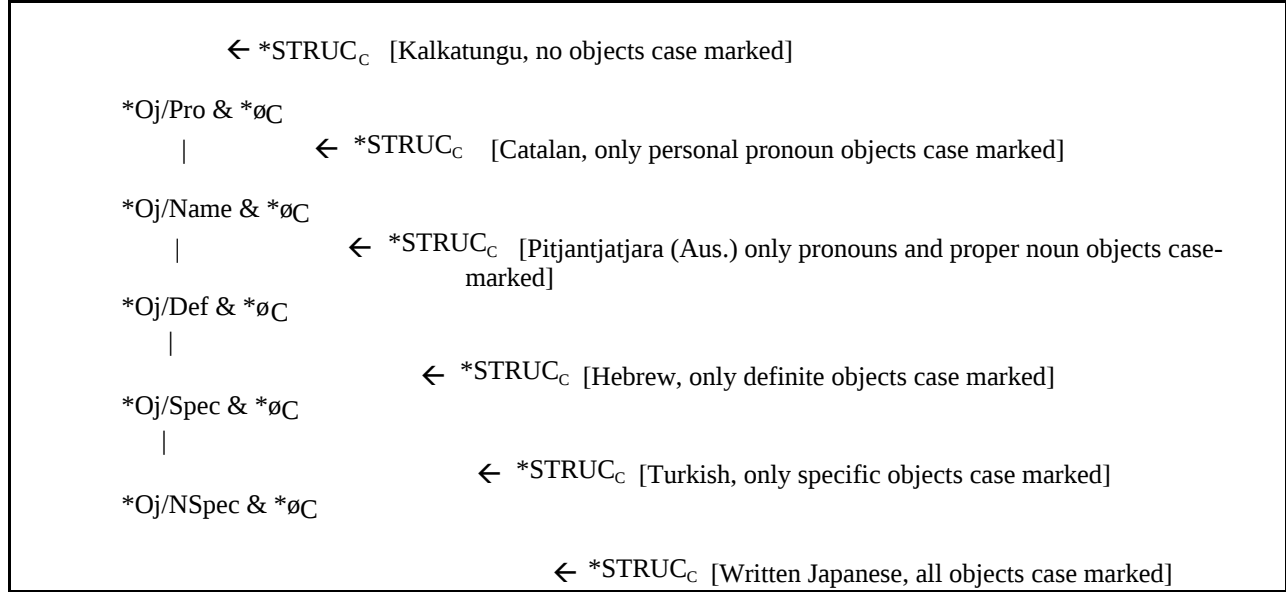


Figure 2: Interpolation of *STRUC_C in the subhierarchy on object definiteness

If *STRUC_C dominates all the constraints on object definiteness, then case marking is penalized for all objects. An example is a language like Kalkatungu (Pama-Nyungan), in which no objects at all are case marked (Blake 1979). Kalkutungu is a language with ergative case marking in which *all* transitive subjects are case marked, but no objects or intransitive subjects. Examples (18a,b) show zero marking of the object and overt case marking of the subject.

- (18) a. maɾapai-tu caa kunka t̪umaji-ŋa.
 woman-ERG here stick break-PAST
 The woman broke the stick. (Blake 1979, 27)
- b. ŋa-tu ŋaŋa maɾapai...
 I-ERG saw woman
 I saw the woman. (Blake 1979, 28)

Even very high prominence objects like the 1st person pronoun are unmarked in object function, as shown in (19) (compare with the case marking for the 1st person subject in (18b)).

- (19) l̪i-ji kuu-ŋku ŋai ŋaŋt̪amaji t̪iinta.
 3s-ERG rain-ERG me find in.the.middle
 The rain caught me in the open. (Blake 1979, 32)

The pronominal form *ŋai* is the citation form, as well as the form used in object and intransitive subject function. Thus I assume it has no specification for case.

If *STRUC_C is ranked one step lower, below the top-ranked constraint, it penalizes case marking on all objects except personal pronouns. An example is Catalan, where (strong) personal pronouns must be preceded by *a*.^{12,13} (Here and below, I gloss the morphology associated with DOM by ACC.)

- (20) a. No m' havien vist *a* *mi*.
 NEG CL they.have seen ACC 1SG
 They had not seen me. (Comrie 1979, 15)
- b. É! te telefonará *a* *ti*.
 he CL will.phone ACC 2SG
 He'll phone you. (Rigau 1986, 161)
- c. A *ell* no el vull.
 ACC 3SG.MASC NEG CL I.want
 Him, I don't want. (Vallduví 1992, 76)

Other definite objects, including proper nouns and definite human common noun phrases, are not preceded by *a* (whether preposed or not):

- (21) No havien vist *l'alcalde*.
 NEG they.have seen the mayor
 They had not seen the mayor. (Comrie 1979, 15)
- (22) ...docs *el Joan* el veiem ben poc.
 since the John CL we.see very little
 since John we see very little of. (Vallduví 1992, 90)

Demotion of the economy condition *STRUC_C below *Oj/Name & *ø_C penalizes case marking of all objects other than personal pronouns and names, for which case marking is forced by higher ranked constraints. According to the description of Bowe (1990), Pitjantjatjara, a Pama-Nyungan language of Australia, exemplifies this case. In Pitjantjatjara, pronouns and proper noun objects are case marked with the suffix *-nya*.¹⁴

- (23) Tjitji-ngku *Billy-nya/ngayu-nya* nya-ngu.
 child-ERG Billy-ACC/1SG-ACC see-PAST
 The child saw Billy/me.

Other objects, including definite, human-referring ones, are not case marked.

¹²Comrie (1979) cites Catalan as a language in which DOM is driven by person (1st and 2nd versus 3rd). But examples like (20c), and the contrast with (22), show that the split is between personal pronouns (of all three persons) and all other nominals.

¹³Examples (20a-c) all contain an obligatory clitic pronoun, which doubles the strong pronominal object. This raises the possibility that it is the obligatory presence of the clitic which motivates the presence of *a* before the strong pronouns. But this cannot be the whole story. In the left-detachment construction illustrated by (20c) and (22), a clitic always resumes a left-detached direct object. If the detached object is a strong pronoun, as in (18c), it must be preceded by *a*; if it is a non-pronoun, as in (22), *a* does not occur. Thus, while it is true that a strong pronoun object always induces an object clitic (and *a*), it is not true that a clitic always induces *a*.

¹⁴Bowe glosses the case marker on the object ABS (for absolutive). I am systematically using ACC for this function.

- (24) a. Billy-lu tjitji nya-ngu.
Billy-ERG child see-PAST
Billy saw the child.
- b. NgayuluṅAtju puṅu kati-ngu.
1ST.NOM.REFL wood bring-PAST
I brought the wood all by myself.

Further demotion of *STRUC_C below *Oj/Def & *ø_C describes a language in which pronouns, names, and definite objects are overtly case marked, but indefinites are not. Hebrew is such a language (Givón 1978). In Hebrew, the preposition *ʔet* is obligatory with definite objects, including inanimates, and does not occur with indefinites.^{15,16}

- (25) a. Ha-seret herʔa ʔet-ha-milxama.
the-movie showed ACC-the-war
The movie showed the war.
- b. Ha-seret herʔa (*ʔet-) milxama.
the movie showed (ACC-)war
The movie showed a war.

Example (25a) is ungrammatical without *ʔet*, while (25b) is ungrammatical with it. That the relevant property is definiteness rather than specificity is shown by the fact that (26b) disallows *ʔet* even though the object refers to a specific doctor.¹⁷

- (26) a. Hu mexapes ʔet-ha-rofe.
he is-looking ACC-the-doctor
He's looking for the doctor.
- b. Hu mexapes (*ʔet) rofe exad.
he is-looking (ACC) doctor one
He's looking for a certain doctor.

Turkish is like Hebrew in that all definite objects are obligatorily case marked, whether they are personal pronouns, names, or definite common nouns, and irrespective of animacy:

- (27) a. Zeynep Ali-yi /on-u /adam-ı. /o masa-yı. gördü.
Zeynep Ali-ACC /him-ACC /man-ACC /that table-ACC saw
Zeynep saw Ali/him/the man/that table.

¹⁵Thanks to Edit Doron for these examples.

¹⁶ There is one exception to this which involves the dimension of animacy. The interrogative pronoun “who” is obligatorily case marked with *ʔet*. The inanimate interrogative is not marked; neither are the pronouns “someone” and “noone”. Yehuda Falk and Hagit Borer each brought this to my attention.

¹⁷ Specificity is marked by *xad*, the “gender inflected numeral ‘one’” (Givón 1978, 303). This element is not limited to objects.

- b. *Zeynep Ali / o / adam / o masa gördü. (Enç 1991, 9)

Turkish differs from Hebrew, though, in that some indefinites are case marked. Enç (1991) characterizes this class as *specific*. Notably, the class includes partitive indefinites. Thus, the sentence *Several children entered my room* can be followed by (28), in which the object may be case marked or not.

- (28) İki kız/kız-ı tanıyordum. (Enç 1991, 6)
two girl/girl-ACC I.knew
I knew two girls.

The version with the case marker is ‘about two girls who are included in the set of children, established by the [previous, JA] utterance...’, while the version without ‘is about two girls who are excluded from the original set of children’ [Enç 1991, 6]. The class of specifics also includes cases in which the speaker has a specific referent in mind and the identity of that referent matters. Enç cites the examples in (29a,b), with the translations indicated:

- (29) a. Ali bir kitab-ı aldı. (Enç 1991, 5)
Ali one book-ACC bought
A book is such that Ali bought it.
- b. Ali bir kitap aldı. (Enç 1991, 5)
Ali one book bought
Ali bought some book or other.

The difference between Turkish and Hebrew can be characterized by the ranking of *STRUC_C relative to *Oj/Spec & *ø_C. In Hebrew, *STRUC_C is ranked above it, favoring economy, i.e. no case marking on specific indefinites; in Turkish, it is ranked below it, favoring iconicity, i.e. case marking on specific indefinites.

At this point, it might be helpful to show some tableaux which illustrate this analysis. The tableaux in (30)-(31) contrast the treatment of indefinite, specific direct objects in Hebrew and Turkish. The input is the same in both cases: the nominal is specified for semantic role and for the relevant semantic and pragmatic properties. Here, I consider only those candidates in which the nominal is realized as direct object, and the top three constraints on object definiteness are encapsulated and represented by *Oj/Def & *ø_C. In each case, there are two candidates, one with a value for CASE specified (as ACC) and one with no value specified. In both cases, the top three constraints are irrelevant. In Hebrew, *STRUC_C eliminates the candidate specified for CASE, leaving the unspecified candidate as winner.

- (30) Hebrew specific, indefinite patient

Patient: specific, indefinite	*Oj/Def & *ø _C	*STRUC _C	*Oj/Spec & *ø _C	*Oj/NSpec & *ø _C
Oj: specific indefinite CASE: ACC		*!		
☞ Oj, specific, indefinite CASE:			*	*

In Turkish, the candidate with no value specified for CASE is eliminated by *Oj/Spec & *ø_C, leaving the candidate specified for CASE as winner (see (31)). The winner violates the economy condition *STRUC_C but this violation is necessary in order to ensure compliance with *Oj/Spec & *ø_C, a higher ranked constraint.

(31) Turkish specific, indefinite patient

Patient: specific, indefinite	*Oj/Def & *ø _C	*Oj/Spec & *ø _C	*Struc _C	*Oj/NSpec & *ø _C
☞ Oj: specific, indefinite CASE: ACC			*	
Oj: specific indefinite CASE:		*!		*

Finally, ranking *STRUC_C below the full subhierarchy yields a language in which all objects are overtly case marked, i.e., one without differential object marking. Written Japanese is apparently such a language, as is Dhalandji (Australia) (Austin 1981). Example (32), from Dhalandji, shows overt case marking for both primary and secondary object, one high prominence (the personal pronoun), the other low prominence (indefinite, probably non-specific) (PURP.DS = purposive, different subject):

(32) ...wantha-rpuka ngatha-nha papa-nha.
 give-PURP.DS me-ACC water-ACC
 [I am waiting for the whiteman] to give me water. (Austin 1981, 223)

This account predicts then that if a language case marks any objects, it will case mark definite ones. A language may mark specific objects, and leave non-specific ones unmarked. But no language will case mark specific indefinites, but not definites. These predictions appear to be borne out.

4.2 Animacy

Interpolation of *STRUC_C into the subhierarchy on object animacy (Figure 1, left-hand side) predicts two types of languages with animacy-sensitive DOM: languages in which only objects referring to humans are case marked, and languages in which only objects referring to animates are case marked. As noted at the outset, there are languages in which only animate-referring direct objects can be overtly case marked. Sinhalese is an example, though in Sinhalese, case marking is optional. It is difficult, however, to find languages in which all and only human-referring objects are case marked. The reason is that in individual languages, animacy-sensitive DOM frequently extends beyond, or retracts within, the human category. Interestingly, it does so in ways that are clearly culturally determined.

In Yiddish, DOM is restricted to humans, but does not cover the entire category. Among common nouns, overt case marking is restricted to eight common nouns, most of which denote individuals worthy of respect. This set contains three masculine nouns which are obligatorily case marked (*taty* "father", *zaidy* "grandfather" and *reby* "teacher"), two masculines which are optionally case marked (*Iid* "Jew" and *ménc* "person") and three feminine nouns, which are optionally case marked (*mamy* "mother", *buby* "grandmother", *mjmy* "aunt").¹⁸ In both Ritharngu and Hindi, DOM leaks across the human-animate boundary, but in different ways. In Ritharngu, all human-referring direct objects are case marked but so are a few nouns referring to 'higher animals' like kangaroos, dogs, and emus. In contrast, nouns referring to fish and raccoons are not case marked in object function. (Heath 1980). In Hindi, DOM likewise extends beyond the human-animate boundary, but according to Mohanan (1993, 28), the higher animals include elephants and lions, and excludes peacocks and mice. Leakage across the animate-inanimate boundary is also found. In Bayungo, all animate direct objects (including humans) are overtly case marked, but so

¹⁸As predicted by the definiteness scale, personal pronouns are case marked. So, generally, are names. According to (Katz 1987), the likelihood of object marking on names correlates with the degree of familiarity with the name. For the Yiddish data, I have relied on (Katz 1987) and (Birnbaum 1979). Thanks to Ellen Prince for bringing the Yiddish data to my attention, and to David Perlmutter for discussion. Perlmutter (p.c.) has suggested that the optionality of overt case with the two masculines mentioned in the text might be linked to definiteness.

are two nouns which English-speakers would consider inanimate, *murla* ‘meat’ and *thanuwa* ‘vegetable food’ (Austin 1981).

There are two ways to analyze these cases. It might be that the three basic categories HUMAN, ANIMATE, and INANIMATE are understood differently in particular languages. Or it might be that they are understood in basically the same way, but that there is further language-particular ranking within the basic categories. On this view, DOM would make the cut somewhere within the basic category, rather than at its boundary. I assume this is the case. First of all, it seems unlikely that speakers of any language do not distinguish between humans and animals. Second, there is a parallel within the category of PERSON. While all languages apparently rank the local persons (1st and 2nd) over 3rd, languages vary in ranking of the local persons themselves. In some languages, 1st person outranks 2nd, in others, 2nd outranks 1st. Language-particular articulation within the category of LOCAL PERSON must then be permitted, and this constitutes a precedent for permitting language-particular ranking within the categories of HUMAN, ANIMATE, and INANIMATE.

Figure 3 then represents this conception. Interpolation of *STRUC_C at various points yields different DOM systems based on animacy. Languages which instantiate the various types are indicated, but I forgo examples. See references cited above and in fn. 19.¹⁹

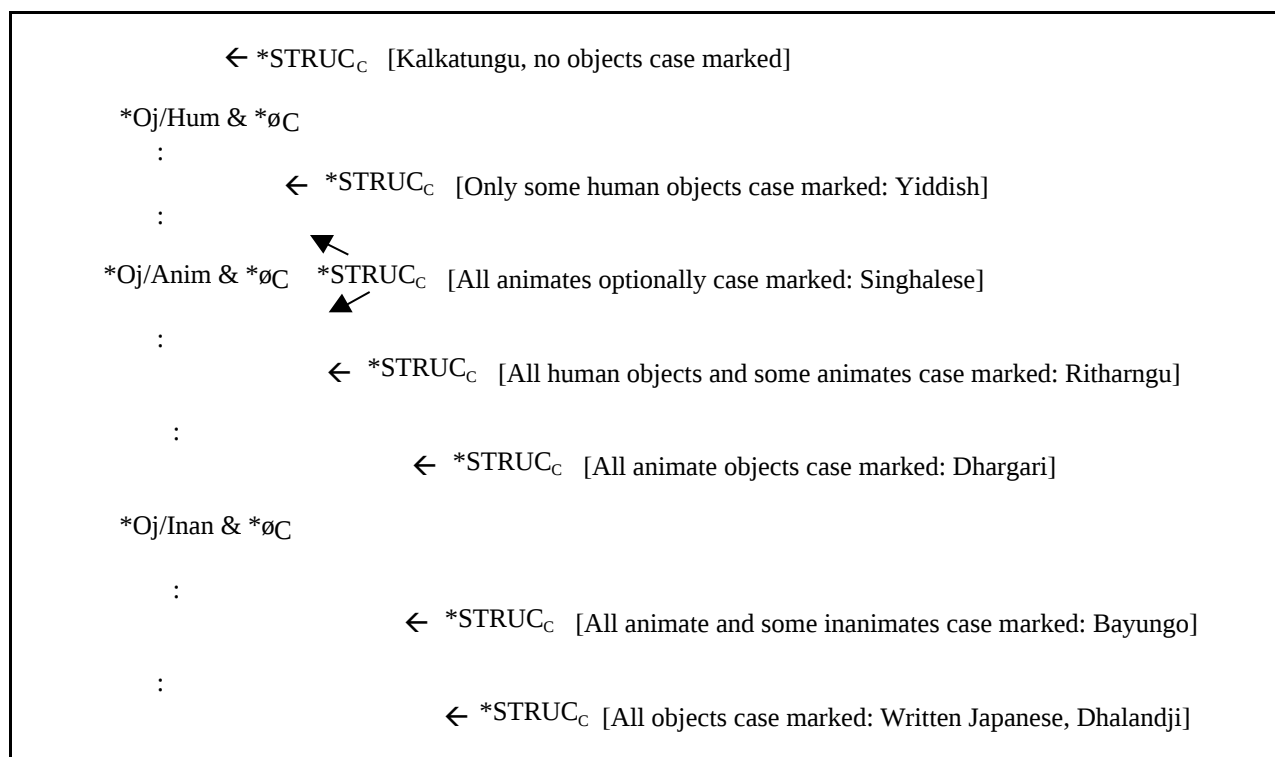


Figure 3: Interpolation of *STRUC_C in the subierarchy on object animacy

The languages cited in Figure 3 are all consistent with the generalization that case marking of relatively low prominence objects entails the possibility of case marking of objects higher in prominence. This accounts predicts

¹⁹ For Bayungo, Dhargari, and Dhalandji, see Austin (1981).

that the reverse is not found, e.g. languages in which only inanimates are case marked, or only non-humans. To my knowledge, no such languages exist.

5 Two-Dimensional DOM

So far we have focused on languages in which DOM is determined by a single dimension of prominence, either animacy or definiteness. However, it is common for DOM to be determined by both. In Romanian, the possibility of DOM is largely restricted to objects which are both animate-referring *and* specific. However, within this space, distinctions in degree of definiteness are relevant to whether DOM is obligatory or optional: it is obligatory for pronouns or names, but optional for definites and indefinite specifics. In Hindi, both animate and inanimate objects can be case marked, but case marking of inanimates is possible only for definites, while case marking of humans (and some animates) is possible for indefinites as well. Further, case marking for human (and some animates) is generally obligatory, while case marking of inanimates is generally optional. Characterizing these systems then requires reference both to degree of animacy and degree of definiteness.

The most straightforward approach to two-dimensional DOM involves the ranking of a set of composite properties, formed by crossing the animacy scale and the definiteness scale. The result is shown in Figure 4, which is constructed by taking the cross-product of those two scales.

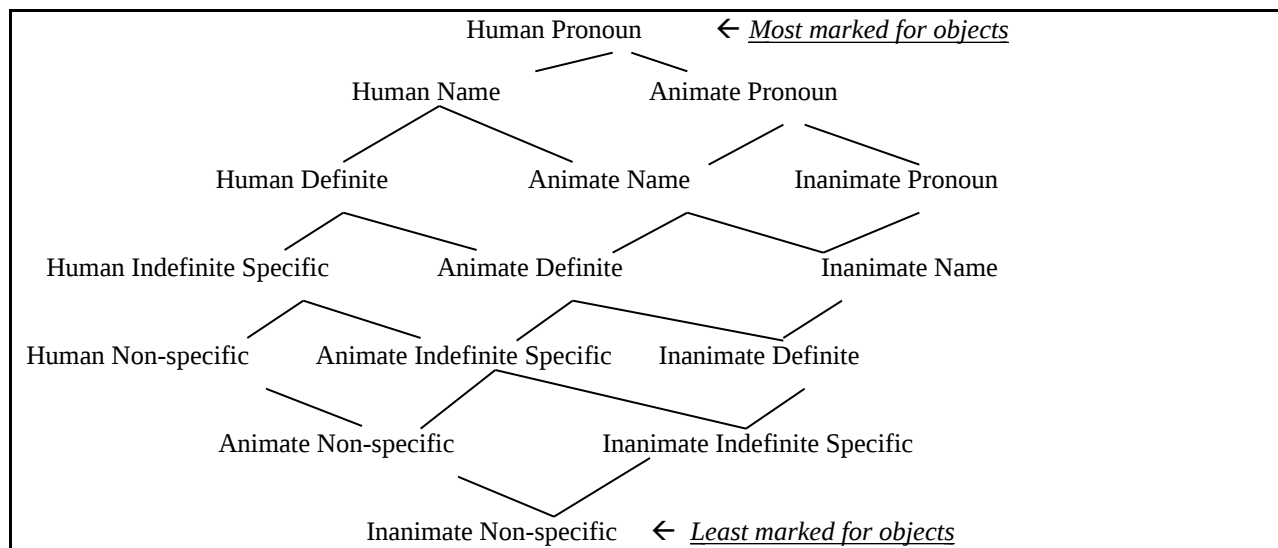


Figure 4: Relative markedness wrt the dimensions of animacy and definiteness

Recalling the basic principles underlying DOM from (3) (*the higher in prominence a direct object, the more likely it is to be overtly case marked*), Figure 4 can be interpreted as a hypothesis about the distribution of two-dimensional DOM. Higher (dominating) elements are more likely to be case marked *qua* objects than lower (dominated) elements. Intuitively, DOM should flow from the top of this structure down. Human pronouns outrank all other elements and should be the most susceptible, cross-linguistically, to DOM; inanimate non-specifics are outranked by all other elements, and should be least susceptible. More generally and precisely,

- (33) If in Figure 4, **a** dominates **b**, then:
- a. if an object of type **b** may be case marked, then all objects of type **a** may be case marked.
 - b. if an object of type **b** must be case marked, then all objects of type **a** must be case marked.
 - c. if no object of type **a** can be case marked, then no object of type **b** can be case marked.

Note that the structure in Figure 4 involves a partial ranking. Elements at the same level horizontally have no fixed relative ranking, e.g. human definites and inanimate pronouns. Thus, Figure 4 predicts that there could be two dimensional DOM systems in which human indefinites are case marked, but inanimate-referring names are not (Romanian is an example), and as well as systems in which inanimate names are case marked, but human indefinites are not (Hebrew is a near candidate, if we see Hebrew as a two-dimensional system per fn.16).

This account (i.e. (33)) makes a general claim about the distribution of case marking in two-dimensional DOM systems: if an inanimate object at some degree of definiteness (e.g. pronoun, name, definite) can be case marked, then so can a human and so can an animate at that same degree. In other words, case marking of inanimate definites entails case marking of *all* definites. However, case marking of a human object at some degree of definiteness does not entail case marking of non-human object at the same degree. Thus, Figure 4 predicts a situation which in fact seems to be common, one in which *all* objects at some degree of definiteness can be case marked, but only human (or animate) objects at lesser degree of definiteness. This holds in Romanian, where case marking for humans and animates extends down to indefinite specifics, but only as far as pronouns for inanimates. In Hindi, case marking for humans and animates extends to indefinites, but for inanimates, only to definites .

Figure 4 provides the right kind of structure for characterizing two other features of two-dimensional DOM systems, a property we will exploit below. One is the fact that these systems generally define three categories of objects: those for which case marking is obligatory, those for which it is optional, and those for which it is impossible. These categories can be mapped onto zones in Figure 4: the first class occupies some upper region of that structure, the third class occupies some lower region, and the middle, optional, class occupies a middle zone. This is the case in Romanian, in 12th century Spanish, and in Hindi (see below for details on the latter two languages). And second, even in zones where DOM is optional, it is often the case that the probability of case marking varies depending on the particular properties of the object. Again, Figure 4 provides the right structure: higher (dominating) elements are more likely to be case marked than lower (dominated) elements. We will see below how these properties of Figure 4 play a role in the evaluation of candidate sets.

Assuming that DOM can ‘cut off’ at any point consistent with (33), the structure in Figure 4 defines a very large set of possible DOM case systems. It is premature to assess the extent to which these possibilities are actually realized, but some examples of how DOM structures the space defined by the structure in Figure 4 will illustrate the nature of two-dimensional DOM systems. The following sections treat the DOM systems of 12th century Spanish (as attested in the epic poem *Cantar de Mío Cid*, Hindi, and Persian – a set of languages which vary in interesting ways with respect to the class of objects which can be case marked. With very few exceptions, 12th c. Spanish permits case marking only of animates, but definiteness plays a role in restricting its scope. In Hindi, DOM is ‘driven’ primarily by animacy, but there is significant incursion into the class of inanimates. In Persian, DOM is primarily driven by definiteness, but animacy also plays a role.

A caveat is necessary here: the data around two-dimensional DOM systems is complex and relevant literature is sometimes incomplete or contradictory. The following discussions are based on my understanding of the source material, but no doubt they are inaccurate and/or oversimplified in some respect(s). Hence, they should be regarded as approximations, or sketches. Corrections and clarifications will be welcome.

5.1 Spanish of El Cantar de Mío Cid

Like Modern Spanish, the Spanish of the *El Cantar de Mío Cid* (henceforth *CMC*) marked some, but not all, direct objects with the preposition *a* (the same preposition used for marking indirect objects), i.e. it had DOM. This stage

of the language distinguishes three categories of direct object: (i) those which must be case marked, (ii) those which cannot be case marked, and (iii) those which are optionally case marked. Obligatorily marked are those direct objects which are (strong) personal pronouns and names of both persons and animals (Melis 1995).²⁰ The category of optionally marked objects comprises common noun phrases referring to humans, especially definite ones, but also indefinites (Melis 1995). Names of geographic places are also optionally marked (Monedero Carrillo de Albornoz 1978). I treat the latter as inanimate-referring names. Inanimate-referring common noun phrases are not case marked. With some assumptions addressed below, this is mapped onto Figure 4 as shown in Figure 5:

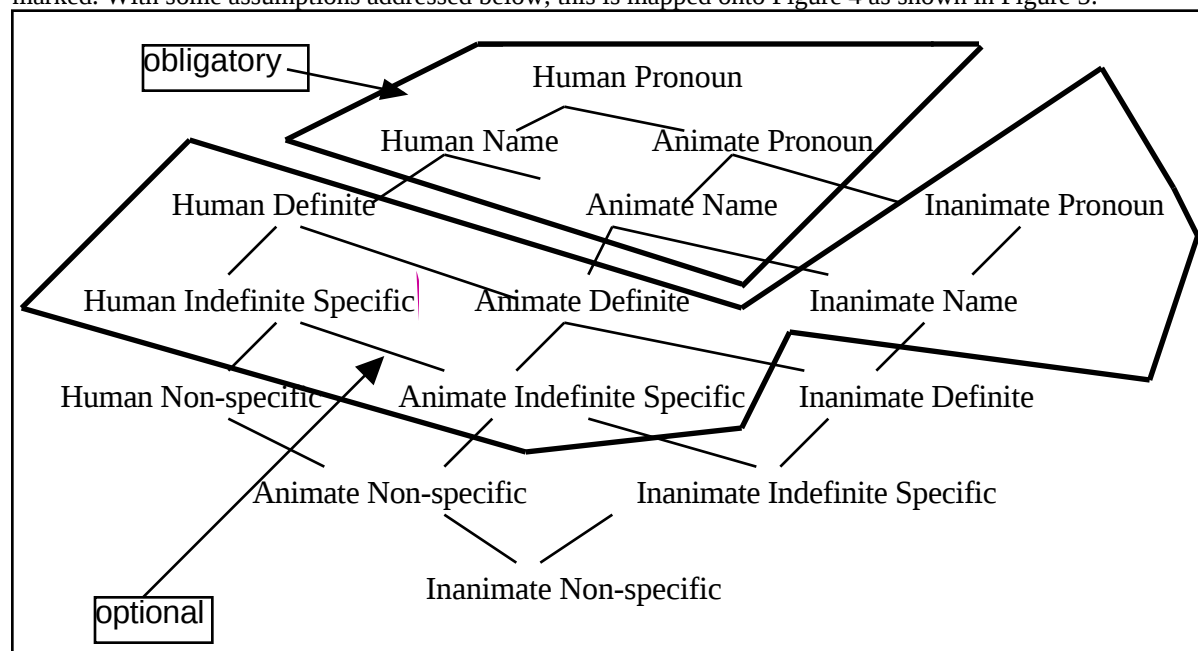


Figure 5: Spanish of Cantar de Mío Cid

Figure 5 goes beyond what was described above in two respects. First, it includes inanimate pronouns in the class of optionally marked object directs. In fact, the strong personal pronoun cannot refer to an inanimate in direct object function at all, so there is no way to know whether it would be case marked. Our account predicts that if it occurred, it would be case marked. Second, Figure 5 distinguishes specific indefinites from non-specific indefinites, and asserts that the latter are never case marked. According to García and van Putte (1995, 124), only 22% of human-referring indefinites in the CMC are marked with *a*, and marking of indefinite singulars (35%) is more likely than marking of indefinite plurals (18%). The former are more likely to be specific than the latter. Further, *a* is more likely in 12th c. Spanish with objects denoting discursively prominent arguments (Melis 1995). Cutting the data this way is also in line with the situation in Modern Spanish, where specific indefinites, but not non-specifics, occur with *a*. Not explicitly represented in Figure 5 is that case marking is more likely for a human-

²⁰ There is a slight idealization here. Personal pronoun objects are marked with *a* 100% of the time (13 instances). Personal names are marked 98% of the time (there are 52 instances, and 51 of them are marked). See García and van Putte (1995) for discussion.

referring common noun phrase which is definite (46%) than for one which is indefinite (22%), i.e. the probability of DOM decreases as one descends the structure in Figure 5.

This situation can be characterized by the constraint structure represented in Figure 6.

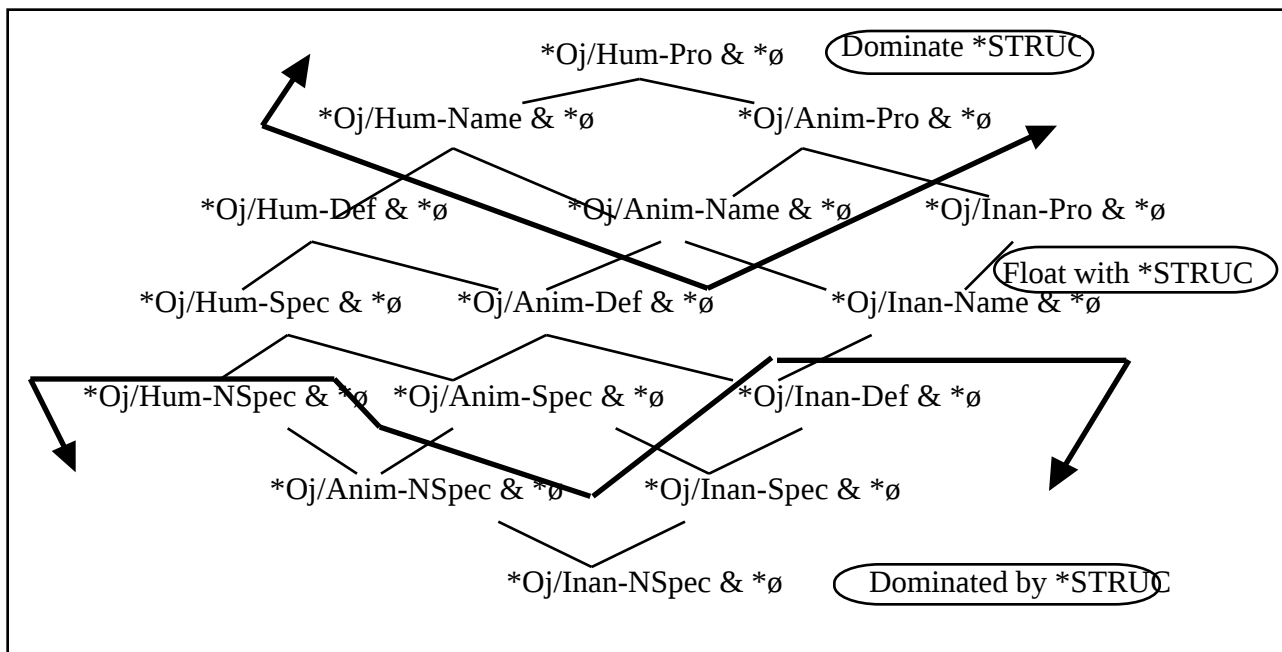


Figure 6: Characterizing DOM in the *Cantar de Mío Cid*

The constraints in this structure are derived by local conjunction of the two subhierarchies on object markedness, (12b) and (15b), followed by local conjunction with $*\emptyset_C$. (For example, local conjunction of $*Oj/Hum$ and $*Oj/Pro$ yields $*Oj/Hum \& *Oj/Pro$, abbreviated in Figure 6 as $*Oj/Hum-Pro$.) The resulting subhierarchy is iconic to the structure in Figure 5. The constraints are partially ranked: if one constraint dominates a second, then the first outranks the second.

In the Spanish of the CMC, the four constraints at the top which favor case marking of high prominence objects are strictly ranked above $*STRUC_C$. This guarantees overt case marking for human and animate pronouns and names. The swathe of constraints at the bottom which favor case marking for low-prominence objects are all dominated by $*STRUC_C$. This precludes casemarking for such objects. The constraints in the middle band govern case marking of intermediate prominence objects. For these, case marking is optional. This can be modeled by allowing $*STRUC_C$ to ‘float’ (rerank) with respect to each of these constraints. For each such constraint, there will be evaluations in which it outranks $*STRUC_C$, and evaluations in which it is outranked by $*STRUC_C$. Evaluations of the first type will yield winning candidates with case marking, those of the second type will yield winners without case marking. For the language as a whole, the result is the appearance of optionality. Further, it may be possible to characterize the diminishing frequency of optional case as one descends the structure in Figure 4 if the grammar is equipped with an algorithm for computing the probability of particular constraint rankings. The approach of Boersma and Hayes (1999), which involves probabilistic constraint ranking, seems to give the right kinds of results

for two-dimensional DOM, and can account both for the optionality of case marking in the middle zone, and for the decreasing likelihood of overt case as one descends the hierarchy.²¹

The system of DOM attested in 12th c. Spanish is very close to that of modern Romanian, which, however, uses an entirely distinct preposition to mark the direct object (*pe*, from Latin *per*) (Farkas 1978; Dobrovie-Sorin 1994). As in 12th c. Spanish, the direct objects for which DOM is obligatory are the human- and animate-referring pronouns and names; the ones for which it is optional are the human and animate definites and indefinite specifics. Some inanimate pronouns can be marked with *pe*, but not inanimate (geographic) names, or any other inanimates. Both 12th century Spanish and Romanian then largely restrict DOM to humans and animates, with very limited incursion into the class of inanimates, and with definiteness playing an important role in restricting the extent of DOM.²²

5.2 Hindi²³

Hindi marks some direct objects, but not all, with the postposition *ko* (again, the same postposition used for indirect objects). In general, Hindi permits more overt case marking of objects than 12th c. Spanish (or modern Romanian) does. On the one hand, it requires overt marking in places that 12th c. Spanish simply allows it; on the other, it permits overt marking in domains where 12th c. Spanish disallows it. These differences should be describable by demoting *STRUC_C among the constraints in Figure 6.

Hindi requires extensive case marking for human-referring objects. As in 12th c. Spanish, case marking is obligatory with human-referring personal pronouns and names (see refs. of fn. 23).

- (34) Ram ne Ravi ko piiṭa.
 Ram ERG Ravi ACC beat
 Ram beat Ravi. (Mohanani 1994a, 183)

But it is also obligatory with human-referring definite descriptions, and appears to be the rule for indefinites as well. Mohanani (1994a, 80) cites (35) below showing that a human object must be case marked, whether it is interpreted as definite or indefinite; Comrie (1989, 133) cites (36) to show that case marking is preferred in this situation.

- (35) Ilaa-ne bacce-ko/*baccaa uṭ^haayaa.
 Ilaa-ERG child-ACC/child lifted
 Ila lifted a/the child.

²¹ In the Boersma and Hayes approach, constraints are ranked on a continuous scale, rather than discretely. Hence two constraints may be relatively close or relatively distant. The grammar is also stochastic: at each evaluation, the constraint ranking is slightly perturbed. This has the consequence that when constraints are sufficiently close, alternative rankings are possible, resulting in variable outputs. In the present case, *STRUC_C needs to be close enough in ranking to the entire set of constraints in the ‘optional’ zone that it can rerank with each of them. Under the assumptions of Boersma and Hayes, this entails that the constraints in the ‘optional’ zone must also be close to one another. A potential worry is that they might then rerank. However, to the extent that the rankings among these constraints are fixed universally (via harmonic alignment), they cannot rerank. If necessary, the algorithm itself can enforce this, cf. Boersma and Hayes (1999, 23).

²² Spanish was discussed here rather than Romanian because of the very close relation in Romanian between DOM and clitic doubling. It appears to be true that every direct object which is doubled by a clitic is also marked with *pe*, though the reverse is not true. Hence, the description of DOM in Romanian might be parasitic on a description of clitic doubling.

²³ To my knowledge, there is no systematic study of DOM in Hindi. However, there are a number of sources which deal with aspects of it, especially (Masica 1982; Junghare 1983; Magier 1987; Butt 1993; Mohanani 1993; Mohanani 1994a; Mohanani 1994b). The grammars of (Kellogg 1938; McGregor 1972) are also informative.

- (36) ?Aurat baccā bulā rahī hai.
 woman child calling PROG is
 The woman is calling a child.

Mohanani's discussion implies that even non-specific (referential) indefinites require *ko* when they refer to humans. However, Masica (1982) cites text examples of this sort without *ko*, so I will assume that case marking is obligatory with specific indefinites, but optional with non-specifics, when these are human-referring.

The two languages contrast sharply in their treatment of inanimates. 12th c. Spanish does not permit overt case marking for inanimates except for names. In Hindi, overt marking is an *option* for inanimates, but *only when definite*.²⁴ Thus, both (37a,b) are possible, and while the object in (37a) can be interpreted as either indefinite or definite, the case marked one in (37b) must refer to an already mentioned banana:

- (37) a. Ravii-ne kaccaa kelaa kaaṭaa.
 Ravi-ERG unripe banana cut
 Ravi cut the/an unripe banana.
- b. Ravii-ne kaccaa kele-ko kaaṭaa.
 Ravi-ERG unripe banana-ACC cut
 Ravi cut the unripe banana. (Mohanani 1994a, 87-88)

Object marking has quite a different function then for inanimates than it has for humans: in the case of inanimates only, it marks *definiteness*. The claim that *ko* is possible with indefinite animates, but not indefinite inanimates, is supported by Singh's observation that the indefinite marker *ek* may occur with *ko* marked objects if they refer to animates (38a), but not inanimates (38b). Analogous data is cited in Mohanani (1994a, 79).

- (38) a. LaRke-ne aaj subah ek laRkii-ko dekhaa.
 boy-ERG today morning one girl-ACC see-PERF
 The boy saw a girl this morning.
- b. *LaRke-ne aaj ek phuul-ko dekhaa
 flower-ACC
 (The boy saw a flower this morning.) (Singh 1994, 227)

Figure 7 characterizes the distribution of *ko*.

²⁴ Hindi sources vary considerably on the status of overt object marking on inanimates. Most suggest it is restricted to *definites*. However, Butt (1993) claims that it is possible for specific indefinite inanimates as well, and cites several examples. In contrast, Singh (1994) implies that some speakers disallow *ko* with inanimates altogether. I will assume this represents dialect differences. Whether this is true needs to be clarified. All of these systems can be described by reranking of *STRUC_C among the constraints in Figure 7.

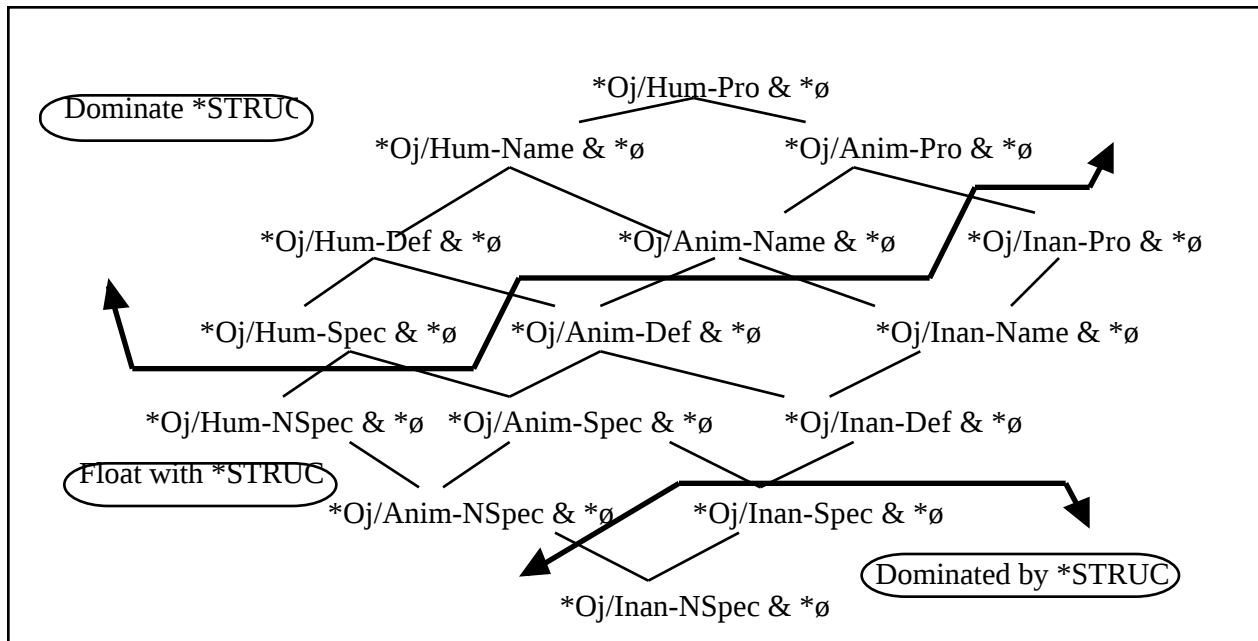


Figure 7: Characterizing DOM in Hindi

Figure 7 defines three zones for overt case marking in Hindi: an obligatory zone (defined by the constraints in the upper third of the structure), an optional zone (defined by the constraints in the middle zone), and an excluded zone (the two constraints at the bottom). All human-referring constraints fall in the obligatory zone, except for the one which involves non-specifics, which falls in the optional zone. No inanimate-referring constraints fall in the obligatory zone. Those which refer to definite inanimates fall in the optional zone, and the others fall in the excluded zone. The facts for animates have been less discussed in the literature, so the treatment in Figure 7 is somewhat arbitrary. As noted earlier, Mohanan reports that some animals are treated like inanimates, in the sense that the presence of *ko* entails definiteness, while others (e.g. lions and elephants) are treated like humans. I have assumed that if animals are referred to by personal pronouns or names, that they will assimilate to the human class and require overt case marking; animate definites are included in the optional zone, perhaps an oversimplification.

As noted at the outset, on the whole, Hindi allows for more overt case marking of objects than does 12th c. Spanish. Obligatory marking is found with a larger class of objects in Hindi than in 12th c. Spanish: it is optional with human definites in 12th c. Spanish, and obligatory in Hindi. This is reflected by the ranking of *STRUC_C in Figures 6 and 7. In Hindi, the constraint which penalizes zero marking of human definite objects (*Oj/Hum-Def & *ø_C) is among those which strictly dominate *STRUC_C. In 12th c. Spanish, it is among those which rerank with *STRUC_C. Furthermore, case marking is extended to definite inanimates in Hindi, but not in 12th c. Spanish. Hence, the constraint which penalizes zero-marking of inanimate definites strictly outranks *STRUC_C in 12th c. Spanish, but the same constraint reranks with *STRUC_C in Hindi. Despite the differences, both languages are consistent with the basic hypothesis: if overt marking is possible with direct objects with property *a*, then it is possible with direct objects with property *b*, where *b* dominates *a* in Figure 4.

5.3 Persian

Lazard (1984) observes that Persian and Hindi are languages in which both definiteness and animacy play a role in determining DOM, but he suggests that the two languages differ in how each of the dimensions is weighted: animacy is ‘stronger’ than definiteness in Hindi, while in Persian, definiteness is ‘stronger’ than animacy. If this is correct, it should correspond to a difference in constraint ranking. We are fortunate in having Lazard (1982), a thorough and thoughtful discussion of DOM in Persian (see also Bosson (1985, pp. 57ff)).

Persian uses the suffix *-râ* to mark some but not all direct objects. Lazard distinguishes three values for definiteness, which correspond to what I am calling DEFINITE, SPECIFIC, and NON-SPECIFIC. He observes that all definites are obligatorily suffixed with *-râ*, regardless of animacy, e.g.

- (39) Ketâb-râ xâdam.
book-ACC I.read
I read the book. (Lazard 1982, 181)

Indefinites are optionally marked, but according to Lazard, specific indefinites are, as a rule, marked. He notes two classes of specific indefinites: those which have a partitive sense, and those with the sense of *a certain*. Both classes require the suffix *-râ*, exactly as both classes require the accusative suffix in Turkish.

- (40) Yeki az ân ketâbhâ-râ xâdam.
INDEF of DEM books-ACC I.read
I read one of these books. (Lazard 1982, 183)

- (41) (Yek) ketâb-i-râ xâd ke...
INDEF book-INDEF-ACC he.read which
He read a certain book which... (Lazard 1982, 183)

In contrast to both Spanish (12th c. and modern) and Hindi, this is so irregardless of the animacy of the object. Persian, like Turkish then, requires case marking for all specific objects.

With respect to other non-specific indefinites, however, Persian appears to be different from Turkish. For the class as a whole, case marking is optional, but within the class of non-specifics, the distribution of case marking is determined by animacy. According to Lazard, the probability of case marking decreases sharply as one moves from human to inanimate. The result is that marking is found generally with humans or animates, but not with inanimates. Lazard cites three minimal pairs as evidence (Lazard 1982, 185-6), all examples in which the objects which seem to be equivalent with respect to definiteness, but not with respect to animacy. In (42), Lazard notes a *preference* to include *-râ* when the object is human and a *preference* to omit it when the object is inanimate.

- (42) a. Mard-i-râ did.
man-INDEF-ACC he.saw
He saw a man.

b. Medâd-i xarid.
pencil-INDEF he.bought
He bought a pencil.

Example (43) conflates two examples, which differ only with respect to animacy of the object:²⁵

²⁵ ‘EZ’ in (43) glosses the clitic *-e* (the *ezafe*), which occurs between a nominal head and post-head modifiers.

- (43) Çah-e tork nazd-e emperâtur-e rum *qâsed-i-râ* / *payâm-i*
king-EZ Turk before-EZ emperor-EZ Rome messenger-INDEF-ACC/ message-INDEF

ferestâd tâ bâ u hamdast çavad.

he.sent so with him ally he.became

The king of the Turks sent a messenger /a message to the Roman emperor in order to ally with him. (Lazard 1982, 185)

In the first version of (43), with a human-referring object, *-râ* could be omitted, but it is preferably present; in the second, the suffix ‘ne pourrait guère accompagner *payâm-i*’ [could scarcely accompany *payâm-i*, JA]. The final example is one in which case marking the object alters the sense in a direction that correlates with animacy:

- (44) a. xarguç-râ dust dâram
rabbit-ACC liking I.have
I like rabbits.
- b. xarguç dust dâram.
rabbit liking I.have
I like rabbit.

In general, there is more obligatory object marking in Persian than in either 12th c. Spanish or Hindi. This is reflected in Figure 8 by the fact that *STRUC_C is outranked by a significantly larger set of constraints than in either of the other two languages. Since overt case marking is required for specific inanimates, *STRUC must be dominated by *Oj/Inan-Spec & *∅. This means that it is dominated by every constraint which dominates *Oj/Inan-Spec & *∅

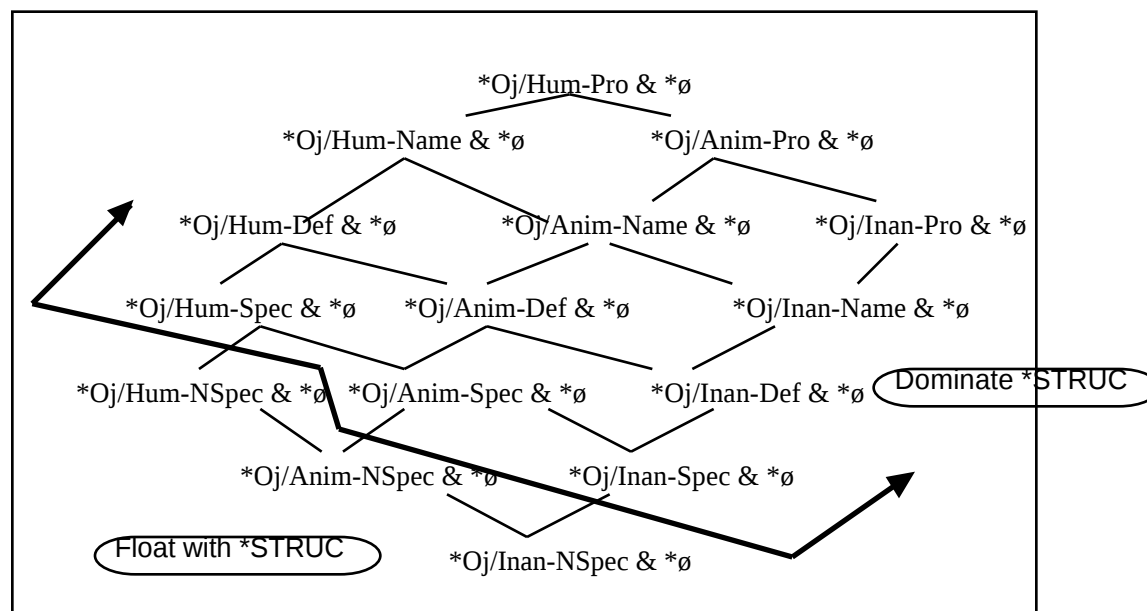


Figure 8: Characterizing DOM in Persian

Animacy comes into play, according to Lazard's description, only with low-salience indefinites, where case marking is optional, but where the probability of case marking declines as one descends from humans to inanimates. In Figure 8, the relevant constraints are the three below the line, which make up a subhierarchy which is strictly ranked. Each of the three constraints floats with *STRUC_C, but the probability of dominating *STRUC_C sharply decreases as one descends the subhierarchy, a fact which can be expressed in the Boersma and Hayes approach by fixing appropriately the relative closeness of the four constraints involved.

5.4 Last point: the spread of DOM

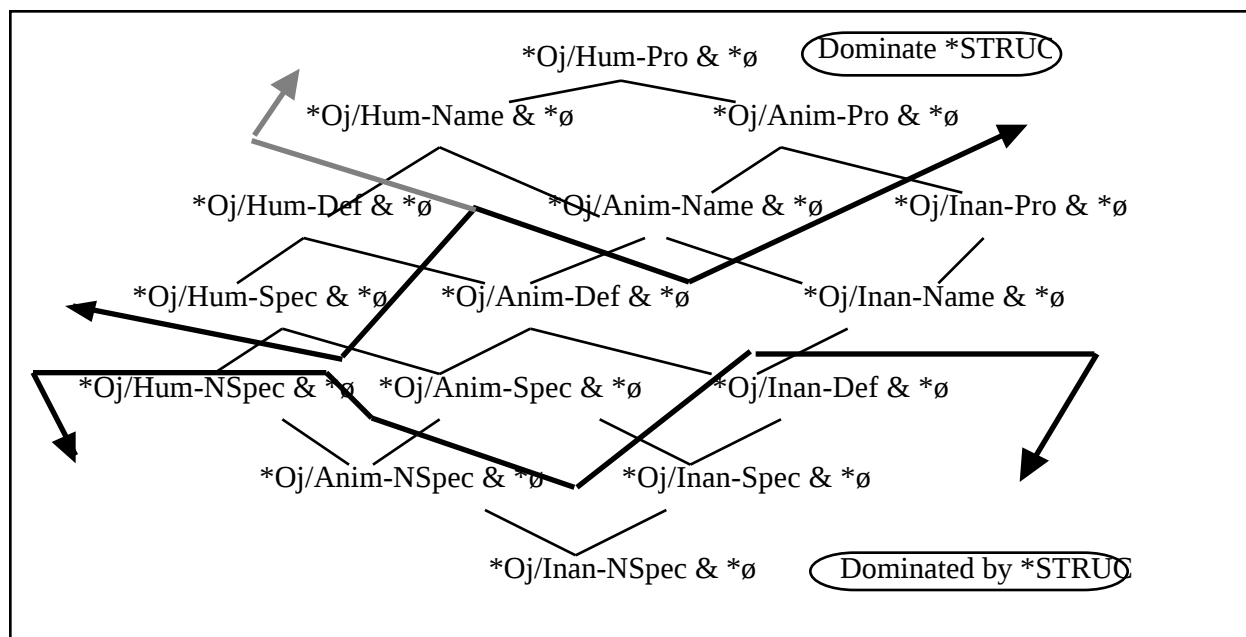
Figure 4 (and the corresponding constraint hierarchy in Figure 6, ignoring the specific positioning of *STRUC_C) provides the basis for an account of two-dimensional DOM. Direct objects are more likely to be overtly case marked if they are positioned high in that figure than low. This likelihood is expressed in various ways. Case marking may be possible (obligatory or optional) for a high prominence object but not a lower one. Or it may be obligatory for a high prominence object but only optional for a lower one. Or it may be possible for both, but with significant differences in attested frequency which correlate with degree of prominence.

This account also generates some predictions about the historical expansion (and retraction of DOM). If DOM is correctly characterized in terms of the constraint hierarchy in Figure 6 and the ranking of *STRUC_C with respect to that hierarchy, then DOM should expand (and retract) along the paths defined in that figure. While the actual conditions on DOM may vary from one period to another, the basic generalization that more prominent direct objects are always more likely to be case marked than ones of lower prominence should hold at all stages. Spanish and Persian are both languages in which the expansion of DOM has been documented, and each is consistent with the scenario sketched above.

DOM occurs widely in Romance today, but it was not a feature of Latin. The preposition *a* first appeared as a direct object marker with the personal pronouns and with proper names (García and Putte 1995; Melis 1995; Pensado 1995). As noted above, in 12th c. Spanish, it was obligatory with these categories of direct object, and clearly optional with human/animate common nouns. The situation in Modern Spanish is complex, but one difference is clear: *a* is now essentially obligatory with definite and indefinite-specific, human-referring direct objects.²⁶

The difference in constraint ranking between these two periods of the language can be (partly) characterized as demotion of the economy constraint *STRUC_C below the constraints which penalize zero marking for definite and specific human-referring objects. This is shown in Figure 9 (the lighter line shows the ranking posited earlier for the Spanish of the CMC).

²⁶ On DOM in Spanish, see (Isenberg 1968; Kliffer 1982; Weissenrieder 1985; Weissenrieder 1990; Weissenrieder 1991; Laca 1995; Pensado 1995) as well as standard textbooks like (Ramsey and Spaulding 1894/1956).



Although Figure 9 oversimplifies in some respects (e.g. inanimate common noun objects are in some contexts case marked (see especially Weissenrieder (1985)), it does characterize fairly accurately one important change in the distribution of DOM in Spanish.

DOM was first documented in Persian in documents of the 9th and 10th centuries; it was not a feature of Middle Persian (Bossong 1985, pp. 58ff.). In its first documentation, DOM was found only with definite objects, and it was optional. At present, DOM is obligatory with definite objects, and furthermore, it has extended into the class of indefinite objects. For some of these, it is apparently obligatory; for others, it is optional. In Persian too, then, the historical expansion of DOM can be expressed as demotion of the economy condition, *STRUC_C, among the constraints in Figure 6.

6 Conclusion

This analysis developed in this paper is based on a conception of DOM which is well-developed in the functional and typological literature cited earlier. In that conception, DOM involves the morphological marking of those nominals which are most marked as objects. The goal here has been to show that this conception can be expressed quite directly within a formal approach to grammar based on Optimality Theory. Central to the account is the operation of *harmonic alignment* which derives constraints that express the relative markedness of the various nominal types in the grammatical *object* function. However, by definition (see fn. 7), harmonic alignment does more than this: it automatically derives a set of constraints which express the relative markedness of exactly the same nominal types in the grammatical *subject* function. Hence, this account predicts that there should be case marking systems in which some subjects are marked, but not all and it predicts that the factors that favor differential subject marking (DSM) will be the mirror image of those that favor DOM. That both types of systems exist and that they are

in a mirror image relation was the main point of Silverstein's original work (1976; 1981). Hence, harmonic alignment succeeds in capturing Silverstein's generalization.

Figure 4 predicts that if any subjects are case marked, it will be common noun subjects, and within that class, inanimate, indefinite ones. Least likely to be case marked are human-referring personal pronouns, especially 1st and 2nd person. These predictions are realized in languages like Guugu Yimidhirr (Pama-Nyungan) in which all transitive subjects are overtly case marked *except* personal pronouns (Haviland 1979), and Punjabi (Indo-European) and Dyirbal (Pama-Nyungan), where all subjects *except* 1st and 2nd person pronouns are case marked (Dixon 1972; Bhatia 1993). (DSM is widely known as 'split ergative' case marking, but this is a confusing misnomer, since it has nothing to do in any significant sense with ergativity (on this point, see especially Goddard (1982)).

Woolford (to appear) suggests that DSM does not exist in the way that DOM does because it does not make reference to multiple dimensions, e.g. both animacy and definiteness. While it is true that complex systems of the sort found in Spanish, Hindi, and Persian are hard to find for DSM, it does not seem to be true that they do not exist. In Fore, for example, neither personal pronouns *nor* names may be marked in subject function, but inanimates must be (Scott 1978; Donohue 1999). These correspond respectively to the least marked and the most marked subjects (see Figure 4). Elements between the two extremes may, and under some circumstances must, be marked as subjects. The present account predicts that such systems should exist. Central to their description will be the constraint hierarchies on *subject* definiteness and *subject* animacy, conjoined with Star Zero.

David Pesetsky (Pesetsky 1997; Pesetsky 1998) has suggested that Optimality Theory is a theory of constraint interaction in *phonology*, and that it is relevant to syntax only in the domain of pronunciation. The results achieved in the present analysis are possible only because it is articulated within a theory in which constraints are universal and violable. The question is whether this analysis deals with what is essentially pronunciation. It is true that DOM involves always the choice between audible case marking and no case marking, clearly a matter of pronunciation. But the analysis depends crucially on markedness constraints which have to do with aspects of syntax which are much deeper. The initial constraint hierarchies derived by alignment of prominence scales play the key role in this analysis, and they are motivated by phenomena related to transitivity and voice (see section 2.2 above), phenomena which are not simply a matter of pronunciation.

The relevance of hierarchy alignment to both syntax and to phonology highlights a deep parallelism between these components. Many of the basic elements of phonology and syntax are different. Articulatory and acoustic parameters, syllable structure, and other low levels of prosodic organization are relevant to phonology but not syntax; grammatical relations, definiteness, and animacy are relevant to syntax but not to phonology. However, at a more abstract level there are fundamental similarities in the way that these elements associate. The basic principle is that prominent structural positions attract elements which are prominent on other dimensions. In phonology, the peak of the syllable is more prominent than the margins, and it attracts elements which are relatively high in sonority, leaving elements of lesser sonority to function as margins. In syntax, the subject is the most prominent argument position, and it attracts elements which are relatively prominent on one of various dimensions, e.g. semantic role, animacy, definiteness, person, topicality, leaving less prominent elements to function as non-subjects. In both domains, there is alignment of prominent substance with prominent structure.

REFERENCES

- Aissen, J. (1997). On the syntax of obviation. *Language* 73(4): 705-750.
Aissen, J. (1999a). Agent focus and inverse in Tzotzil. *Language* 75(3): 451-485.
Aissen, J. (1999b). Markedness and subject choice in optimality theory. *Natural Language and Linguistic Theory* 17(4): 673-711.
Artstein, R. (1998). Hierarchies. <http://www.eden.rutgers.edu/~artstein>.
Asudeh, A. (1999). Linking, optionality, and ambiguity in Marathi: an optimality theory analysis. <http://www.stanford.edu/~asudeh>.
Austin, P. (1981). Case marking in southern Pilbara languages. *Australian Journal of Linguistics* 1: 211-226.
Battistella, E. (1990). *Markedness: the evaluative superstructure of language*. Albany: SUNY Press.
Battistella, E. L. (1996). *The logic of markedness*. Oxford: Oxford University Press.
Bhatia, T. K. (1993). *Punjabi. A cognitive-descriptive grammar*. London: Routledge.

- Birnbaum, S. A. (1979). *Yiddish. A survey and a grammar*. Toronto: University of Toronto Press.
- Blake, B. (1977). *Case marking in Australian languages*. Canberra: Australian Institute of Aboriginal Studies.
- Blake, B. (1979). *A Kalkatungu Grammar*. Canberra: Pacific Linguistics.
- Boersma, P. (1998). *Functional phonology*. Amsterdam: Faculteit der Geesteswetenschappen, Universiteit van Amsterdam.
- Boersma, P. and B. Hayes (1999). Empirical tests of the Gradual Learning Algorithm. <http://www.fon.hum.uva.nl/paul/gla/>
- Bossong, G. (1985). *Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Gunter Narr Verlag.
- Bowe, H. (1990). *Categories, constituents and constituent order in Pitjantjatjara*. London: Routledge.
- Butt, M. (1993). Object specificity and agreement in Hindi/Urdu. In *Papers from the 29th regional meeting of the Chicago Linguistic Society*, Chicago: Chicago Linguistic Society, pp. 89-103.
- Chung, S. (1981). Transitivity and surface filters in Chamorro. In J. Hollyman and A. Pawley, (eds.), *Studies in Pacific languages and cultures, in honour of Bruce Biggs*, Auckland: Linguistic Society of New Zealand, pp. 311-332.
- Chung, S. (1984). Identifiability and null objects in Chamorro. In *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*, Berkeley: Berkeley Linguistics Society, pp. 116-130.
- Chung, S. (1989). On the notion "null anaphor" in Chamorro. In O. Jaeggli and K. Safir, (eds.), *The null subject parameter*, Dordrecht: Kluwer Academic Publishers, pp. 143-184.
- Chung, S. (1998). *The design of agreement. Evidence from Chamorro*. Chicago: University of Chicago Press.
- Comrie, B. (1979). Definite and animate direct objects: A natural class. *Linguistica silesiana* 3: 13-21.
- Comrie, B. (1980). Agreement, animacy, and voice. In G. Brettschneider and C. Lehmann, (eds.), *Wege zur Universalien Forschung: Sprachwissenschaftliche Beiträge zum 60. Geburtstag von Hansjakob Seiler*, Tübingen: Gunter Narr, pp. 229-234.
- Comrie, B. (1986). Markedness, grammar, people, and the world. In F. Eckman, E. Moravcsik and J. Wirth, (eds.), *Markedness*, New York: Plenum Press, pp. 85-106.
- Comrie, B. (1989). *Language universals and linguistic typology*. Chicago: University of Chicago Press.
- Cooreman, A. (1987). *Transitivity and discourse continuity in Chamorro narratives*. Berlin: Mouton de Gruyter.
- Craig, C. (1977). *The structure of Jacalteco*. Austin: University of Texas Press.
- Croft, W. (1988). Agreement vs. case marking and direct objects. In M. Barlow and C. Ferguson, (eds.), *Agreement in natural language: Approaches, theories, descriptions*, Stanford, CA: CSLI, pp. 159-179.
- Croft, W. (1990). *Typology and universals*. Cambridge: Cambridge University Press.
- Davison, A. (1984). Syntactic markedness and the definition of sentence topic. *Language* 60(4): 797-846.
- DeLancey, S. (1981). An interpretation of split ergativity. *Language* 57: 626-657.
- Diesing, M. (1992). *Indefinites*. Cambridge: MIT Press.
- Diesing, M. and E. Jelinek (1995). Distributing arguments. *Natural Language Semantics* 3: 123-176.
- Dixon, R. M. W. (1972). *The Dyirbal language of North Queensland*. Cambridge: Cambridge University Press.
- Dixon, R. M. W. (1994). *Ergativity*. Cambridge: Cambridge University Press.
- Dobrovie-Sorin, C. (1994). *The syntax of Romanian*. Berlin: Mouton de Gruyter.
- Donohue, C. (1999). Optimizing Fore case and word order. Stanford, CA, Stanford University.
- Enç, M. (1991). The semantics of specificity. *Linguistic Inquiry* 22(1): 1-26.
- England, N. (1983). Ergativity in Mamean (Mayan) languages. *International Journal of American Linguistics* 49: 1-19.
- Farkas, D. (1978). Direct and indirect object reduplication in Romanian. In *Papers from the fourteenth regional meeting of the Chicago Linguistic Society*, Chicago: Chicago Linguistic Society, pp. 88-97.
- Farkas, D. (1997). Towards a semantic typology of noun phrases. Colloque de syntaxe et sémantique de Paris, Université Paris 7.
- Foley, W. and R. Van Valin, Jr. (1984). *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.
- Gair, J. (1970). *Colloquial Sinhalese clause structures*. The Hague: Mouton.
- García, E. and F. v. Putte (1995). La mejor palabra es la que no se habla. In C. Pensado, (ed.) *El complemento directo preposicional*, Madrid: Visor Libros, pp. 113-132.
- Gerdts, D. (1988). A nominal hierarchy in Halkomelem clausal organization. *Anthropological Linguistics* 30(1): 20-36.
- Gerdts, D. (1988). *Object and absolutive in Halkomelem Salish*. New York: Garland.
- Givón, T. (1978). Definiteness and referentiality. In J. Greenberg, (ed.) *Universals of human language*, Stanford, CA: Stanford University Press, pp. 291-330.
- Givón, T. (1994). The pragmatics of de-transitive voice. In T. Givón, (ed.) *Voice and inversion*, Amsterdam: John Benjamins, pp. 3-44.
- Goddard, C. (1982). Case systems and case marking in Australian languages: a new interpretation. *Australian Journal of Linguistics* 2: 167-196.
- Greenberg, J. (1966). *Language universals with special reference to feature hierarchies*. The Hague: Mouton.

- Haiman, J., Ed. (1985a). *Iconicity in syntax*. Amsterdam, John Benjamins.
- Haiman, J. (1985b). *Natural syntax: iconicity and erosion*. Cambridge: Cambridge University Press.
- Haviland, J. (1979). Guugu Yimidhirr. In R. M. W. Dixon and B. Blake, (eds.), *Handbook of Australian Languages. Volume 1*, Amsterdam: John Benjamins, pp. 27-180.
- Heath, J. (1980). *Basic materials in Ritharngu: Grammar, texts and dictionary*. Canberra.
- de Hoop, H. (1996). *Case configuration and noun phrase interpretation*. New York: Garland.
- Hopper, P. and S. Thompson (1980). Transitivity in grammar and discourse. *Language* 56: 251-299.
- Isenberg, H. (1968). *Das direkte Objekt im Spanischen*. Berlin: Akademie-Verlag.
- Jakobson, R. (1939). Signe zéro. In *Melanges de linguistique offerts a Charles Bally sous les auspices de la Faculte des lettres de L'Universite de Geneve par des collegues, des confreres, des disciples reconnaissants*, Geneve: Georg et cie, s.a., pp. .
- Junghare, I. (1983). Markers of definiteness in Indo-Aryan. In A. Dahlstrom, C. Brugman, M. Macaulay et al, (eds.), *Proceedings of the Ninth Annual Meeting of the Berkeley Linguistics Society*, Berkeley: Berkeley Linguistics Society, pp. 116-127.
- Katz, D. (1987). *Grammar of the Yiddish language*. London: Duckworth.
- Keenan, E. (1976). Towards a universal definition of "subject". In C. Li, (ed.), *Subject and topic*, New York: Academic Press, pp. 303-333.
- Kellogg, S. H. (1938). *A grammar of the Hindi language*. London: Routledge.
- Kliffer, M. (1982). Personal *a*, kinesis and individuation. In P. Baldi, (ed.) *Papers from the XII linguistic symposium on Romance languages*, Amsterdam: Benjamins, pp. 195-216.
- Kroeger, P. (1993). *Phrase structure and grammatical relations in Tagalog*. Stanford, CA: CSLI Publications.
- Kuno, S. (1976). Subject, theme, and the speaker's empathy—a reexamination of relativization phenomena. In C. Li, (ed.) *Subject and topic*, New York: Academic Press, pp. 417-444.
- Kuno, S. and E. Kaburaki (1977). Empathy and syntax. *Linguistic Inquiry* 8: 627-672.
- Laca, B. (1995). Sobre el uso del acusativo preposicional en español. In C. Pensado, (ed.) *El complemento directo preposicional*, Madrid: Visor, pp. 61-91.
- Lazard, G. (1982). Le morphème *râ* en persan et les relations actanciellles. *Bulletin de la société de linguistique de Paris* 73(1): 177-208.
- Lazard, G. (1984). Actance variations and categories of the object. In F. Plank, (ed.) *Objects: Towards a theory of grammatical relations*, London: Academic Press, pp. 269-292.
- Lee, H. (1999). The emergence of the unmarked order in Hindi. [lee-1999-1206] on the LFG Archive.
- Magier, D. (1987). The transitivity prototype: evidence from Hindi. *Word* 38(3): 187-199.
- Masica, C. (1982). Identified object marking in Hindi and other languages. In O. N. Koul, (ed.) *Topics in Hindi Linguistics*, New Delhi: Bahri Publications, pp. 16-50.
- McGregor, R. S. (1972). *Outline of Hindi grammar*. Oxford: Clarendon Press.
- Melis, C. (1995). El objeto directo personal en el Cantar de Mio Cid: Estudio sintáctico-pragmático. In C. Pensado, (ed.) *El complemento directo preposicional*, Madrid: Visor, pp. 133-163.
- Mohanan, T. (1993). Case alternation on objects in Hindi. *South Asian Language Review* 3: 1-30.
- Mohanan, T. (1994a). *Argument structure in Hindi*. Stanford: CSLI.
- Mohanan, T. (1994b). Case OCP: A constraint on word order in Hindi. In M. Butt, T. H. King and G. Ramchand, (eds.), *Theoretical perspectives on word order in South Asian languages*, Stanford, CA: CSLI, pp. 185-215.
- Monedero Carrillo de Albornoz, C. (1978). El objeto directo preposicional y la estilística épica. *Verba* 5: 259-303.
- Pensado, C., Ed. (1995). *El complemento directo preposicional*. Madrid, Visor.
- Pensado, C. (1995). El complemento directo preposicional estado de la cuestión y bibliografía comentada. In C. Pensado, (ed.) *El complemento directo preposicional*, Madrid: Visor, pp. 11-59.
- Pesetsky, D. (1997). Optimality theory and syntax: movement and pronunciation. In D. Archangeli and D. T. Langendoen, (eds.), *Optimality Theory. An overview*, Oxford: Basil Blackwell, pp. 134-170.
- Pesetsky, D. (1998). Some optimality principles of sentence pronunciation. In P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis and D. Pesetsky, (eds.), *Is the best good enough? Optimality and competition in syntax*, Cambridge: MIT Press, pp. 337-383.
- Prince, A. and P. Smolensky (1993). *Optimality theory: constraint interaction in generative grammar*. Piscataway, NJ, Rutgers University Center for Cognitive Science.
- Ramsey, M. M. and R. Spaulding (1894/1956). *A textbook of Modern Spanish*. New York: Holt, Rinehart and Winston.
- Rigau, G. (1986). Some remarks on the nature of strong pronouns in null-subject languages. In Ivonne Bordelois, Heles Contreras and Karen Zagona (eds.) *Generative studies in Spanish syntax*, Dordrecht: Foris, pp. 143-163.
- Scott, G. (1978). *The Fore language of Papua New Guinea*. Canberra: School of Pacific Studies.

- Sharma, D. (1999). Case marking and person hierarchy effects in Kashmiri. Stanford, CA, Stanford University.
- Silverstein, M. (1976). Hierarchy of features and ergativity. In R. M. W. Dixon, (ed.) *Grammatical categories in Australian languages*, Canberra: Australian Institute of Aboriginal Studies, pp. 112-171.
- Silverstein, M. (1981). Case marking and the nature of language. *Australian Journal of Linguistics* 1: 227-244.
- Singh, M. (1994). Thematic roles, word order, and definiteness. In M. Butt, T. H. King and G. Ramchand, (eds.), *Theoretical perspectives on word order in South Asian languages*, Stanford, CA: CSLI, pp. 217-235.
- Smolensky, P. (1995). On the internal structure of the constraint component Con of UG. Handout of talk, 3/95.
- Vallduví, E. (1992). *The informational component*. New York: Garland.
- Weissenrieder, M. (1985). Exceptional uses of the accusative *a*. *Hispania* 68: 393-398.
- Weissenrieder, M. (1990). Variable uses of the direct-object marker *a*. *Hispania* 73: 223-231.
- Weissenrieder, M. (1991). A functional approach to the accusative *A*. *Hispania* 74: 146-156.
- Wierzbicka, A. (1981). Case marking and human nature. *Australian Journal of Linguistics* 1: 1-42.
- Woolford, E. (to appear). Case patterns. In J. Grimshaw, G. Legendre and S. Vikner, (eds.), *Optimality-theoretic syntax*, Cambridge, MA: MIT Press.