What markedness marks: the markedness problem with direct objects

Åshild Næss

University of Nijmegen, PO Box 9103, Nijmegen 6500 HD, The Netherlands

Received 2 July 2003; accepted 2 July 2003
Available online 13 December 2003

Abstract

This paper discusses a number of problems associated with the widely accepted analysis of differential object marking (DOM) as reflecting the semantic markedness of highly individuated (definite and/or animate) direct objects. Firstly, such an account is in conflict with established notions of transitivity which take a typical object to be highly affected, since affectedness can be shown to correlate with a high degree of individuation. Secondly, the notion of markedness reversal, which is employed as a means of providing a unified account of differential marking of subjects and of objects (e.g. Aissen, 2000), cannot unproblematically be applied to the kinds of oppositions typically involved in DOM. Finally, the predictions made for subject marking only partly correspond to attested linguistic data.

An alternative analysis is proposed which takes accusative case-marking to be a marker of a high degree of affectedness in objects. By exploiting the association between affectedness and a high degree of individuation (definiteness/animacy), such an analysis can account for the DOM data, while avoiding the difficulties inherent to the approach which takes individuated objects to be semantically marked.

© 2003 Elsevier B.V. All rights reserved.

Keywords: Markedness; Markedness reversal; Differential object marking; Affectedness; Transitivity; Accusative case

1. Introduction

It has long been assumed in linguistic theory, and particularly in functional typology, that there is a correlation between certain properties of a noun phrase and the syntactic function it fulfills. More specifically, it is assumed that in the prototypical or unmarked case, subject
noun phrases are definite and refer to animate entities, while object noun phrases are typically indefinite and have inanimate referents (e.g. Croft, 1988; Comrie, 1989). The assumption is that this opposition of prototypical subject and object properties aids clause processing and comprehension by facilitating the identification of the NPs of a bivalent clause with their respective syntactic functions: where one NP is definite and animate, and the other indefinite and inanimate, the former can generally be assumed to be the subject, and the latter the object. When a clause deviates from this patterning, additional devices such as case-marking may be required in order to clarify which grammatical relation should be ascribed to which participant.

In this paper I will examine these assumptions, particularly with respect to what is known as differential object marking (DOM). I will argue that the analysis of DOM as reflecting a high degree of individuation in objects is inconsistent not only with accepted theories of transitivity and objecthood, but also with some basic premises of the very theory on which it builds, namely markedness theory. I will then present an alternative analysis which attempts to resolve these inconsistencies, based on the event-semantic properties of control and affectedness rather than on properties of noun phrases and their referents, and briefly outline how this analysis may be employed to account for a broader range of phenomena such as split case-marking systems in general.

2. The markedness of objects: differential object marking

Functional typology assumes a “natural” correlation between a high degree of individuation—that is, animacy and definiteness—and (transitive) subjects, on the one hand, and between a low degree of individuation and transitive objects on the other. In the words of Comrie (1989: 128), “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”.

The phenomenon known as DOM—that is, language-internal variation in object case-marking correlated with certain properties of the object—appears to illustrate this point perfectly. DOM occurs across a wide range of geographically and genetically diverse languages, and it is generally triggered by properties such as definiteness, specificity, or animacy of the object NP or NP referent. Languages that show DOM may, for instance, case-mark definite objects but not indefinite ones, animate objects but not inanimate ones, or only human objects but not objects referring to non-human animates or inanimates:

(1) **HEBREW** (Hopper and Thompson 1980:256):

   a. David natan matana lərīna
      D. gave present to.Rina
      ‘David gave a present to Rina’

   b. David natan et ha-matana lərīna
      D. gave OBJ DEF-present to.Rina
      ‘David gave the present to Rina’
Hebrew case-marks definite objects but not indefinite ones (1), while in Catalan, only human objects (in the standard written language, only pronoun objects) take case-marking (2).

Functional typology takes these case-marking patterns to be a reflex of the markedness of highly individuated or prominent objects. This assumption stems from a more general theory of case-marking which assumes that the basic function of case-marking on core arguments is to distinguish subject from object: in order for the language user to be able to identify the respective functions of the two arguments of a bivalent clause, one of them receives overt case marking. There is a fair amount of evidence for this analysis; for instance, Finnish, an accusative language, only employs overt accusative case on objects if the clause also contains a subject NP, otherwise, for instance in imperatives, the object is unmarked (Comrie, 1975); the Papuan language Waskia has an agent/subject marker ke which is used only when there might otherwise be some doubt as to which argument of a bivalent clause is the subject, such as when the subject is inanimate, or when it is displaced from its usual sentence-initial position (Ross and Paol, 1978); in the Australian split-ergative language Yuwaalaraay the ergative case-marker can be omitted “where the context serves to disambiguate the sentence”, for instance when the object is pronominal and therefore carries accusative case (Williams, 1980; for similar examples see also Dixon, 1994: 58–9).

It is argued that this need to overtly mark one argument for purposes of identification is greater in some cases than in others. If the distinction can be made through real-world knowledge alone, case-marking may not be necessary. We know that subjects are typically instigating participants, and that therefore they tend to be definite and animate. If a clause has only one NP with these properties, while the other is indefinite and/or inanimate, associating these NPs with the correct syntactic function is a straightforward matter: in the sentence Peter kicked the ball, anyone with a basic understanding of how the world works can figure out who did the kicking and who got kicked. The problem arises when both NPs have similar properties. In a sentence like Peter kicked John, word-order clues aside, the association is much less obvious; in principle, either participant could be the aggressor or the victim. Only in such cases, then, is there a real need for additional clues to identify grammatical functions; in other words it is only when the object of the clause is also highly individuated, that is, when it resembles a potential subject, that case-marking is necessary. It is “subject-like” properties of objects that triggers DOM: definiteness in the case of Hebrew, animacy/humanness in the case of Catalan, and so on.

Markedness theory assumes an iconic relation between form and meaning: a morphologically marked form is taken to reflect a somehow marked meaning. In the case of DOM,
the formal opposition is between the presence of overt case-marking versus the absence of any such marking. Since the overt marker occurs on definite and/or animate objects, this is taken to indicate that such objects are marked with respect to indefinite/inanimate ones; in other words, that definiteness and animacy are marked properties for objects (Aissen, 2000: 3). The unmarked object is taken to be one which is easily distinguishable from the subject in terms of degree of individuation, and so “subject-like” objects receive an overt morphological marker as an indication of their semantic markedness.


The functional-typological analysis of DOM has been formalised by Aissen (2000) within the framework of Optimality Theory (OT). Her model will be discussed here in some detail since it is a clear formulation of the assumptions and predictions involved in this type of analysis and so provides a good starting-point for discussion. The discussion below will therefore refer mainly to Aissen, but should be taken to hold for related approaches as well.

Aissen analyses the DOM phenomenon in terms of markedness reversal: “exactly what is marked for objects is unmarked for subjects” (Aissen, 2000: 3). She sets up two prominence scales relating to the dimensions of animacy and definiteness, where objects that rank higher on these scales are more likely to take overt case-marking than lower-ranking objects:

(3) PROMINENCE SCALES (Aissen 2000:2):

- Animacy scale: Human > Animate > Inanimate
- Definiteness scale: Personal pronoun > Proper noun > Definite NP > Indefinite specific NP > Non-specific NP

These scales are meant to capture the following generalisation: if in some language a direct object at some rank can be case-marked, then higher-ranked objects in the language can be case-marked, but not necessarily lower ones.

Furthermore, Aissen claims that precisely the opposite is the case for subjects: the properties that are high on the prominence scales are unmarked properties for subjects. This constitutes what is known as markedness reversal: the reversal of markedness values of the terms of an opposition within certain specific contexts. In order to formalise this markedness reversal, Aissen assumes a relational scale on which subject outranks object, and sets up a set of harmonic alignment hierarchies (4). Harmonic alignment is a process which operates on pairs of scales, where one of the two scales must be binary. The high-ranking element of the binary scale is associated with each of the elements on the other scale, left to right, and the low-ranking element of the binary scale—in this case, the Subject > Object scale—with each of the elements on the other scale, right to left. This gives rise to hierarchies representing the relative markedness of combinations of elements from the two scales, where the most unmarked association is either of the highest-ranked
element of one scale with the highest-ranked element of the other, or of the lowest-ranked
element of one scale with the lowest-ranked element of the other:

(4) HARMONIC ALIGNMENT HIERARCHIES (Aissen 2000:6-8):

Animacy:
Su/Hum ⊃ Su/Anim ⊃ Su/Inan (⊃ = “more unmarked than”)
Oj/Inan ⊃ Oj/Anim ⊃ Oj/Hum

Definiteness:
Su/Pro ⊃ Su/PN ⊃ Su/Def ⊃ Su/Spec ⊃ Su/NSpec
Oj/NSpec ⊃ Oj/Spec ⊃ Oj/Def ⊃ Oj/PI ⊃ Oj/Pro

These hierarchies are then integrated into an OT-style model by reinterpreting the elements
as constraints and reversing their rankings:

(5) *Su/Inan >> *Su/Anim >> *Su/Hum
    *Oj/Hum >> *Oj/Anim >> *Oj/Inan

The same, of course, applies to the definiteness hierarchies. The prediction that arises from
the constraint hierarchies in (5) is that to the extent that a language avoids or marks clauses
on the basis of the animacy of the subject and/or object, it will be clauses with inanimate
subjects and/or human objects (Aissen, 2000: 6). The corresponding hierarchies involving
definiteness predicts clauses with indefinite subjects and definite objects to be the most
highly marked ones.

The appeal of this analysis lies not least in its simplicity and its integration of subject and
object properties into one single system: while accounting for case-marking properties of
objects, it produces a model which also makes predictions about properties of subjects.
However, the approach has two major problems: It clashes with established notions of
transitivity and the properties of transitive clauses; and it assumes a notion of markedness
and of markedness reversal which is problematic with respect to some basic assumptions of
markedness theory.

4. The problem of affectedness

4.1. Affectedness and individuation

Aissen mentions briefly in a footnote the by now classical article by Hopper and
Thompson (1980) on transitivity, where typical objects, contrary to the functional-
typological analysis, are considered to be highly individuated. She notes that Hopper
and Thompson’s approach makes similar predictions to her own where DOM is concerned,
but that it is based on a very different conceptualisation of the phenomenon; and as it makes
no predictions about properties of subjects, and such predictions do follow from her own
analysis, she considers her approach to be more satisfactory.
The fact is, however, that inherent in Hopper and Thompson’s and similar approaches there lies a direct contradiction to the functional-typological analysis as presented above, and as formalised by Aissen—a contradiction which to my knowledge has not been addressed by any of the accounts under discussion.  

The problem revolves around the property of affectedness. A basic property which is almost universally ascribed to direct objects is that of being somehow affected by the verbal action; in fact, the property of affectedness is regularly appealed to in attempts to define the notion of direct object. Where there is more than one non-subject argument in a clause, it is apparently always the argument which is construed as most affected which is encoded in direct object position (e.g. Fillmore, 1977; Dixon, 1994). Affectedness, then, is relevant for objecthood; and affectedness can be shown to correlate with relatively high individuation. Hopper and Thompson (1980) take affectedness to depend on degree of individuation, because “an action can be more effectively transferred to a patient which is individuated than one which is not; thus a definite O is more completely affected than an indefinite one”. Crosslinguistic studies show that accusative case-marking on objects tends to be used for objects which are definite, human or animate, and highly affected (Moravcsik, 1978). In other words, there is a correlation between affectedness and high individuation.

In fact, clauses that do not have two fairly highly individuated arguments are less likely to be encoded as transitive, and so to have a direct object at all. For instance, objects that are low in individuation show a crosslinguistic tendency to be encoded in incorporation constructions, which are morphosyntactically intransitive (Mithun, 1984):

(6) TONGAN (Mithun 1984):
   a. Na’e inu ‘a e kavá ‘é Sione
      PAST drink ABS CONN kava ERG John
      ‘John drank the kava.’

   b. Na’e inu kava ‘a Sione
      PAST drink kava ABS Sione
      ‘John drank kava’

In (6b), the indefinite object kava is incorporated into the verb phrase, and the subject Sione takes absolutive rather than ergative case; the sentence is formally intransitive.

A different type of detransitivisation with low-individuation objects is found in Chamorro (Topping, 1973). Chamorro does not have incorporation, but a sentence with an indefinite object is formally similar to intransitive clauses in a number of ways. The subject pronoun in a clause with an indefinite object belongs to a set of postverbal pronouns which are also used for subjects of intransitive verbs (as well as for pronoun objects of transitive verbs), while in transitive clauses, a different set of preverbal subject pronouns is

---

1 Croft (1988) does discuss the problem of whether definite and animate objects are marked or unmarked, with reference to Hopper and Thompson. He does not, however, address the problem of the correlation between affectedness and high individuation, and in fact ends up characterising typical objects as both affected and low in individuation (see 4.2). For the reasons to be discussed in this section, this is not a satisfactory solution.
used. In addition, the verb in a clause with an indefinite object takes a special “indefinite object” prefix which does not occur in fully transitive clauses:

(7) CHAMORRO (Topping 1973):
   a. Hu li’e i lebłø   b. Manli’e’ yo’ lebłø
      1SG see DEF book  IO.see 1SG book
      ‘I saw the book’      ‘I saw a book’
   c. Macho’cho’ yo’
      work 1SG
      ‘I worked’

In other words, in many languages objects that are low in individuation simply are not objects from a formal point of view—they are encoded syntactically in intransitive constructions, which by definition do not have direct objects. The claim that the most typical direct objects are those which are low in individuation is therefore problematic, since such NPs often are not encoded as direct objects at all. In a study of noun incorporation in Cushitic languages, Sasse (1984) maintains that “only individuated items are candidates for Grammatical Relations due to their inherent pragmatic prominence”. In other words, unless a participant is relatively high in prominence, it is unlikely that the language will encode it as a separate argument. Thus relatively high individuation or prominence is in fact a prerequisite for being encoded as an object at all.

4.2. Claims and contradictions

This leaves us with a difficult problem. Either we have to assume that typical objects are highly affected and low in individuation at the same time; this is a view that is occasionally found in the literature, for instance in Croft (1988), where it is claimed that “the natural correlation of direct objects is with low animacy, low definiteness, and highly affected objects (i.e. genuine patients)”. But as the studies cited above show, affectedness in fact seems to correlate with high individuation, and low-individuation objects often are not treated formally as objects at all.

Alternatively, one might consider typical objects to be morphologically marked—taking overt case marking as opposed to no marking—while at the same time being unmarked members of their class. If typical objects are affected, and affectness correlates with high individuation, then since highly individuated objects are the ones which receive case-marking in DOM, the case-marked objects must be the semantically unmarked ones. But this is in direct conflict with established markedness theory, and so does not appear to be a desirable solution either.

The only way to make the conflict go away would seem to be to abandon the idea that being affected is a fundamental property of direct objects: If we cannot at the same time claim that typical objects are affected and that they are low in individuation, perhaps we should give up the claim about affectness. But this would require some rather radical redefinition of the notion of object, which is closely associated with the concept of
affectedness in linguistic theory. Furthermore, it is not at all obvious how one should go about defining “object” if one cannot make reference neither to the formal property of case-marking—since only exceptional objects are taken to be case-marked in DOM, this cannot serve as a criterion for objects in general—nor to the semantic property of affectedness.

Unless this apparently unresolvable contradiction should be taken simply as a fact of human language—a highly unsatisfactory position from a scientific point of view—there must obviously be a flaw somewhere in the reasoning that produces it. Thus, the only way to attempt to find a solution is to carefully examine the claims and predictions made by the analyses in question, and try to uncover precisely where the problem arises.

There are two main claims made by the analysis of DOM as outlined above and formalised in Aissen’s paper. The first is that it is marked for direct objects to be definite and animate, as opposed to indefinite and animate; in other words, definiteness and animacy are marked properties in the context of syntactic objecthood. The second claim is that there is a relation of markedness reversal between subjects and objects with respect to these properties; in the words of Aissen, “exactly what is marked for objects is unmarked for subjects”.

These claims are based on markedness theory and the functional theory of case-marking, respectively. The first claim falls out of the form-meaning iconicity assumed in markedness theory: marked forms represent marked meanings. If definite and animate objects receive case-marking while indefinite and inanimate objects do not, then the first type of object is assumed to be marked with respect to the second.

The second claim is a somewhat more complex one which is derived from the theory of case-markers as discriminatory devices. If case-marking is only necessary when there is a possibility of ambiguity concerning which argument is subject and object in a bivalent clause, then an overt marker on an object must mean that it might be possible to mistake it for a subject; in other words, a marked object possesses “subject-like” properties. Since these properties are generally present in subjects, an overt marker is needed only when they are also present in objects, and so properties which are unmarked for subjects become marked for objects, and vice versa.

Note that this second claim is in principle independent of the first in that it only states that there are some properties in terms of which subjects and objects should ideally differ, but does not specify which “subject-like” properties give rise to possible ambiguities when applied to objects. This is a central question which will be addressed in more detail in 5.1. In conjunction with the first claim, however, the second claim is interpreted as meaning that the markedness reversal between subjects and objects holds specifically at the level of individuation, that is, for the properties of definiteness and animacy.

In order to evaluate these claims we have to take several things into consideration. Firstly, we must examine the notion of markedness and its content, to establish precisely which kind of definition of the term “markedness” underlies these claims. Secondly, we must relate this definition to the notion of markedness reversal, to see what exactly is entailed by the claim that “exactly what is marked for objects is unmarked for subjects”. Finally, we must consider the question of exactly which properties must be taken to be “unmarked” for subjects and objects, respectively; and we must take care to distinguish between properties of direct objects in general, independently of their expression in any
particular language or linguistic system, and the way these properties are encoded within specific formal systems of language.

5. Markedness and what it marks

5.1. Formal and semantic markedness

The notion of markedness is a complex one whose precise content remains a subject of debate. In particular, the concept of semantic markedness and its relation to markedness of form raises a number of questions. What does it mean for one meaning to be “marked” relative to another? Which criteria should one employ to determine which member of a semantic opposition is the marked one?

There is in fact a range of criteria generally assumed to be relevant for the notion of markedness (for a discussion, see, e.g. Croft, 1990). Rather than enter into a discussion of the various possible criteria here, I will focus on examining which criteria for markedness are being employed in the analysis under discussion: How is it determined which properties are “marked for objects”, a decision which is crucial to the analysis of DOM in terms of markedness reversal?

The studies cited above essentially employ two types of criteria for determining the markedness of subject and object properties: frequency and morphology (structural markedness). Comrie (1989) states that “in actual discourse there is a strong tendency for the information flow from A to P to correlate with an information flow from more or less animate and from more to less definite”—essentially an argument from frequency, although no actual frequency data are cited. Croft (1988), on the other hand, refers to a text count by Givón (1979) where 50% of direct objects were found to be indefinite as opposed to only 10% of subjects. According to Croft, this “suggests indefinite direct objects are unmarked” (p. 170). Note, however, that Givón himself is reluctant to take these percentages as evidence for clauses with indefinite objects being “unmarked” or “neutral”; although direct objects are more frequently indefinite than subjects are, they are not more frequently indefinite than definite.

The second type of criterion, and the one most frequently appealed to, is that of morphological structure: “unmarked values are morphosyntactically less complex than marked ones” (Croft, 1988: 160). That is, if one value of a semantic opposition tends to be expressed by a greater number of morphemes, or by a more complex morphosyntactic construction, than the opposite value, then the former is marked with respect to the latter. In the case of differential object marking, this criterion translates into presence versus absence of case marking: if one semantically defined class of direct objects take an overt case marker while another class does not, then the semantic feature characterising the first class is taken to be marked relative to that characterising the second class.

It is not entirely clear to what extent these criteria succeed in establishing the semantic markedness of definite/animate direct objects as opposed to their formal markedness. Indeed, as the semantic markedness of the pairing of direct objects and definiteness/animacy is being offered as an explanation for the case-marking pattern, it seems dubious
to appeal to these same case-marking patterns as an indication of this semantic markedness. There may exist other, independent criteria for establishing this semantic markedness, but aside from frequency, as discussed above, it is not clear from any of the studies cited what these would be.

The underlying justification for the analysis lies in the idea that it is most “natural” for subject and object NPs to have inherent properties that are somehow “opposite” in nature, so that these NPs can be unambiguously identified with their respective syntactic functions without the aid of any overt grammatical marking. This raises the question, however, of exactly which kinds of properties should be taken to enter into this opposition. By itself, the statement that “exactly what is marked for objects is unmarked for subjects” is completely meaningless; its validity depends crucially on which properties it refers to. For instance, it is extremely unmarked for a subject to be a noun phrase, but this obviously cannot be taken to imply that it is unmarked for objects not to be noun phrases. Any postulated opposition between subject and object properties must refer to a specific set of properties; and how can we be sure that exactly the properties of definiteness and animacy are the ones to participate in such an opposition?

This brings us to a second, and in a sense more fundamental, problem concerning the presence or absence of overt morphemes as a criterion for markedness. In the words of Croft (1990: 73): “The important and sometimes difficult question that must be answered in finding evidence for structural markedness is whether or not the morphemes being counted really are there to signal the category whose markedness is under examination” (my emphasis). In other words, how can we be sure which “marked meaning” is being represented by a “marked form”?

It is this question which will be the main concern of the rest of this paper: given the alternation in a number of languages between overt case-marking of some direct objects and the absence of such marking on other direct objects, precisely which semantic property is it that correlates with the formal marker? Put somewhat differently, what exactly is being marked by the accusative case? The claim of functional typology is that in languages with DOM, the accusative case is essentially a marker of high individuation in objects, represented by definiteness as opposed to definiteness, or animacy as opposed to inanimacy. However, given the problems with this account that were presented in (4) above, we should consider the possibility that “the morpheme being counted”, that is, the accusative case-marker, may in fact reflect some other property associated with certain kinds of objects. This possibility is further supported by the fact that the notion of markedness reversal, crucial to the analysis as presented by Aissen, is in fact difficult to apply to properties such as definiteness and animacy. This further suggests that if there is indeed an opposition between typical subject and object properties, this opposition does not operate at the level of definiteness and animacy.

5.2. Types of markedness relations

An analysis in terms of markedness reversal presupposes a clear understanding of the markedness relations involved. If definite and animate objects are marked, an important question to ask is: marked with respect to what? Over which kinds of semantic opposition is this markedness relation to be defined?
Andrews (1990) distinguishes between two possible types of markedness relations: *inclusive asymmetry* and *polar opposition*. Inclusive asymmetry is defined for semantic oppositions in the same way as purely formal markedness: as the presence versus absence of a given feature (a *privative* opposition in the terminology of Trubetzkoy, 1939). This corresponds to Jakobson’s (1932) conception of semantic markedness: as the signalisation of some feature A—the marked term of the opposition—versus the nonsignalisation of A, the unmarked term. Note that this definition gives rise to a double interpretation of the unmarked term: it can be taken to represent the opposite of the marked term, “not A”; or it can be interpreted as being merely unspecified for property A. In any case, a relation defined in this way unambiguously identifies the marked versus the unmarked term; the very nature of the opposition is such that there is something present in one of the terms that is absent in the other, and the marked term is by definition the one in which the feature in question is present.

The notion of polar opposition, on the other hand, refers to the presence of a feature A on one term versus the presence of the opposite feature B on the other. Such a relation can be illustrated, for instance, by the masculine–feminine opposition. This type of opposition contains no inherent markedness relation; rather, the markedness of one term relative to the other is imposed by contextual, cultural, or cognitive factors.

The question, then, is this: Which kind of markedness relation is taken to hold for the marked properties of objects in a DOM system? If definiteness is a marked property, is it marked with respect to *indefiniteness* (a positive property representing not the absence of definiteness, but the opposite of definiteness, making this relation one of polar opposition), or to *non-definiteness* (the absence of definiteness, i.e. a relation of inclusive asymmetry)?

Evidence from a number of languages suggests that what is encoded by DOM is a relation of the latter kind. In Sinhala, for instance, animate objects are optionally marked with the accusative case, while inanimate objects never take accusative case-marking. This means that any object with an overt accusative marker is unambiguously animate, but an unmarked object might be either animate or inanimate; in other words the value of the unmarked term is that of being unspecified for animacy rather than being specifically inanimate.

(8) **Sinhala** (Gair and Paolillo 1997):

a. mamə minïha dëkkka
   *I saw the man*

b. mamə minïha-wə dëkkka
   *I saw the man*

  ‘I saw the man’

  ‘I saw the man’

c. mamə ee potə(*-wə) kiyewwa
   *I read that book(*-ACC)* read

  ‘I read that book’

In Hindi, the situation is somewhat more complex, as DOM here depends on both animacy and definiteness. While all human-referring objects take case-marking, whether they are definite or indefinite, inanimate object NPs are optionally case-marked if they are definite, but never marked when they are indefinite. An unmarked inanimate object, then, might be
either definite or indefinite, while marked inanimate objects are always definite:

(9) **Hindi** (Mohanan 1994 quoted in Aissen 2000):
   a. Ravi-ne kacca kele kaata
      Ravi-ERG unripe banana cut
      ‘Ravi cut the unripe banana’
   b. Ravi-ne kacca kele-ko kaataa
      Ravi-ERG unripe banana-ACC cut
      ‘Ravi cut the unripe banana’

The unmarked term is ambiguous; the object in (9a) may be read either as definite or as indefinite. The accusative-marked object in (b), on the other hand, is unambiguously definite. The contrast encoded by the opposition between accusative-marked and unmarked objects in Hindi, then, is that between specifically definite objects on the one hand and objects *unspecified for definiteness* on the other, rather than between specifically definite and specifically indefinite objects.

Similar patterning is found in many but not all other languages with DOM. There are cases where no optionality occurs, and so no such interpretation of the unmarked term as being unspecified for the property in question is possible. Note, however, that as mentioned above, the unmarked term of a relation of inclusive asymmetry has a double interpretation: it may be taken as being either unspecified for property A, or as representing the opposite of A. One might argue, then, that the latter interpretation is implied when there is an exact correlation between case-marking and the positive property of the opposition; that is, if all definite objects are case-marked, then unmarked objects are necessarily interpreted as indefinite. This is a possible interpretation of an inclusive asymmetry relation and would allow for a unified account of all instances of DOM; the alternative would be to say that cases such as (8–9) above involve inclusive asymmetry, whereas in other DOM languages, such as e.g. Hebrew, the relation is rather one of polar opposition.

There is no doubt, however, that in a significant number of cases the relation involved is one of inclusive asymmetry, as illustrated by (8–9). This is highly significant for the further analysis of such cases in terms of markedness reversal.

5.3. **Markedness reversal**

If what is encoded by DOM is a markedness relation of inclusive asymmetry—specification versus nonspecification of a property—then this has consequences for the notion of markedness reversal which is appealed to by Aissen as characteristic of the relationship between subject and object properties. Reversal is defined in Shapiro (1983) as “the phenomenon whereby a marked context reverses the normal markedness values of the terms of an opposition”. In other words, the essence of markedness reversal is that what is marked in one context—the context of syntactic objecthood, in this case—may be unmarked in a different context, such as that of syntactic subjecthood.
The problem is that markedness reversal presupposes a notion of markedness that is dependent on context in the first place. For context to be able to influence the markedness values of an opposition, markedness must be defined as a value imposed by factors external to the opposition as such rather than being inherent to the opposition itself. In other words, markedness must be defined on a polar opposition.

However, as Andrews (1990) points out, relations of inclusive asymmetry cannot be reversed by context, because the relative markedness of the terms of the opposition is not imposed by context in the first place. The marked term of such an opposition is defined as marked because it possesses a specific feature that the unmarked term lacks, and this markedness relation does not change with context; the term that lacks specification for the feature in question is by definition unmarked, regardless of the context in which it occurs. This means that if the markedness relation encoded by DOM is one of inclusive asymmetry, as suggested in 5.2, then an analysis in terms of markedness reversal is problematic, because markedness reversal cannot be defined in a straightforward manner over such a relation.

The only way to let markedness reversal operate on a relation of inclusive asymmetry is to assume two pairs of asymmetric features, where the positive term of one opposition is the polar opposite of the positive term of the other (Battistella, 1996: 58). This creates a set of four features:

A vs. non-A  
non-B vs. B

On this kind of system, says Battistella, reversal can be defined as “the situation in which a category is sometimes characterized as A and sometimes as non-B (or, equivalently, sometimes as B and sometimes as non-A)”. That is, the same category would be marked with respect to the A versus non-A opposition, but unmarked with respect to the B versus non-B opposition.

If the category in question is definiteness, such an approach would mean operating with on the one hand an opposition between definite and nondefinite (in the sense of unspecified for definiteness), and on the other an opposition between indefinite (a positive value and the polar opposite of definiteness) and nonindefinite (specified for indefiniteness). An NP would then, for instance, have to be characterised as (specifically) definite in a context where indefiniteness is the norm, namely in object position, but as unspecified for definiteness in a context where definiteness is the norm, that is in subject position. In this way, the same NP would be marked in object position, since it is given a positive and therefore marked value of definiteness in this context, but unmarked in subject position, where it is characterised in terms of the unmarked value of the indefinite–nonindefinite opposition.

Since in many languages definiteness (and sometimes also indefiniteness) is morphologically marked on NPs, it does not appear reasonable to claim that such NPs in some contexts have the semantic value of “definite”, but in other contexts rather the value “non-indefinite”; this would imply that the marker in question would have both these readings. This is certainly not the claim that is being made by functional typology with
The concept of “markedness reversal” that is taken to hold for subjects versus objects is of a somewhat different kind: Definite objects are claimed to be marked with respect to other types of objects, while indefinite subjects are taken to be marked with respect to other types of subjects; the “markedness reversal” lies in the relation of polar opposition between the marked values for objects versus subjects, respectively. It is marked for objects to be definite, while it is marked for subjects to be indefinite; reversely, the idea is that it is unmarked for objects to be indefinite, and unmarked for subjects to be definite, and this is what is meant by markedness reversal. We have seen, however, that DOM does not encode a definiteness-indefiniteness opposition but rather a definiteness—nondefiniteness opposition. In other words, the relationship between the marked objects and the “other types of objects”, and the unmarked subjects and the “other types of subjects” must be defined as an inclusive asymmetry: definiteness is marked for objects as opposed to non-definiteness, whereas indefiniteness is marked for subjects as opposed to non-indefiniteness. Figure (10) illustrates the oppositions involved:

The crucial point here is that the marked terms of each opposition, indefiniteness for subjects and definiteness for objects, are marked, not with respect to each other (which would be a markedness relation of polar opposition), but with respect to their unspecified counterpart. Examples (8–9) above showed this to be the case for objects: where a marked object is definite, the corresponding object is not necessarily indefinite, rather, is it nondefinite, unspecified for definiteness. Given that this is how DOM works in many languages, this is the only way that the statement that “what is marked for objects is unmarked for subjects” can be interpreted: what is marked for objects is being specified for definiteness, whereas what is marked for subjects is being specified for indefiniteness. The corresponding unmarked instances consist of being unspecified for definiteness and indefiniteness, respectively. Thus the unmarked property for a subject would be that of being non-indefinite, that is, not specifically indefinite, while for an object the unmarked property is being nondefinite. A similar set of properties would be needed for animacy: animate objects would be marked with respect to nonanimate ones, while inanimate subjects would be marked with respect to non-inanimate ones.

The acceptability of this analysis hinges on the plausibility of the oppositions involved. The proliferation of terms needed to rescue the markedness reversal analysis is no problem if the oppositions in question can be shown to be relevant for the description of the languages involved. However, I am not aware of any languages which operate with a distinction between an overt marker for indefiniteness and an absence of marking which
would designate “unspecified for indefiniteness”, not to mention a double opposition of on the one hand a marked definite contrasted with an unmarked term unspecified for definiteness, and on the other an indefinite—unspecified for indefiniteness contrast. A similar argument could be made with respect to animacy versus non-animacy and inanimacy versus non-inanimacy. If languages do not operate with such contrasts, it would be preferable not to have to postulate them for the sake of supporting a specific analysis.

5.4. Predictions for subjects

One of the advantages claimed for Aissen’s model is that it not only accounts for DOM patterns, but also makes predictions about subject case marking. More specifically, “if any subjects are case marked, it will be common noun subjects, and within that class, indefinite, inanimate ones” (Aissen, 2000: 28). Figure (10) suggests that the parallel to DOM based on definiteness would be a system where specifically indefinite subjects were case-marked, whereas unmarked subjects might be either indefinite or non-indefinite. Something similar would be predicted for animacy, with overt marking of specifically inanimate subjects.

Overt marking of indefinite as opposed to nonindefinite (or even specifically definite) subjects appears not to be reliably attested (Comrie, 1989: 130). Marking of only inanimate subjects, or only common noun subjects as opposed to pronouns, or only 3rd but not 1st and 2nd person pronoun subjects, all of which may be considered to represents portions of an animacy hierarchy, is fairly widespread. Interestingly, however, it seems almost only to occur in split-marking systems, that is, in systems that also show differential marking of objects; this is certainly true for all the cases cited by Aissen as having this type of marking (Guugu Yimidhirr, Punjabi, Dyirbal). In the most straightforward cases, exactly the types of NPs that take no ergative marker in subject position are the ones that take accusative marking in object position, such as in Dyirbal (Dixon, 1994: 85–86); other languages show an overlap so that some types of NPs show both ergative marking in subject position and accusative marking in object position. It seems, then, as though “differential subject marking” of the type predicted by Aissen rarely occur independently of differential object marking.

There are, however, other languages where only some subjects take case-marking, but where the patterning is quite different. In one variety of the Iranian language Wakhi, for instance, only 1st and 2nd person subjects, the highest on the animacy hierarchy, can take overt case-marking (Bashir, 1986). In Mayan, only the pronominal system shows ergative marking (Givón, 2001). In the Tibeto-Burman language Meithei, the Papuan Folopa (Podopa), and a number of other languages, overt case-marking of subjects does not depend on the animacy or definiteness hierarchies, but rather occurs only on subjects that are perceived to be in control of the verbal action (Dixon, 1994: 29ff).

Aissen comments with respect to the cases that do follow her predictions that “D[ifferential] S[ubject] M[arking] is widely known as ‘split ergative’ case marking,

---

2 Sinhala has an overt marker of indefiniteness contrasting with zero, but it is not clear to me whether the zero-marked form has a possible “unspecified” interpretation as opposed to being specifically definite (Gair and Paolillo, 1997: 17–18).
but this is a confusing misnomer, since it has nothing to do in any significant sense with ergativity” (Aissen, 2000: 28). This may well be; but if it is indeed the case that such subject case-marking patterns only arise in combination with differential object marking, then the label “split”, at least, would appear to be adequate. The other type of split Aissen predicts, that based on definiteness, is unattested for subjects, while other types of differential subject marking which do occur do not function along the parameters indicated by Aissen.

5.5. Conclusions

The above discussion has attempted to establish the difficulties associated with an analysis of DOM in terms of definiteness and animacy, and with the concept of markedness reversal with respect to these parameters. If the markedness relation found with DOM is indeed one of definiteness versus non-definiteness (rather than indefiniteness), or animacy versus non-animacy (rather than inanimacy), that is, a relation of inclusive asymmetry, and if markedness reversal cannot be defined on such a relation, then we are left with two possibilities. Either there is no markedness reversal involved in the phenomenon under discussion, and the statement that “exactly what is marked for objects is unmarked for subjects” is simply false. Or else this markedness reversal is defined over some other property than definiteness or animacy. This would mean that the answer to the question posed in 5.1 above—“What is being marked by the accusative case?”—is not high individuation of objects, but something else. If there is some kind of markedness reversal involved in the differential marking of subjects and objects, then the property ultimately encoded by DOM has to be something other than definiteness and animacy; it must be some property for which the notion of markedness reversal is relevant.

6. Affectedness revisited

6.1. Expressing affectedness

On the face of it, the suggestion that DOM might encode some property other than definiteness and/or animacy seems highly unlikely. If DOM correlates systematically with definiteness, or with animacy, or with both, then surely definiteness and/or animacy is what is being encoded by the accusative case? What possible grounds could there be for postulating some additional, invisible feature which is “really” what DOM marks?

The problem is, as we have seen, that there are great difficulties with the account that assumes DOM to be a reflex of the markedness of definite and animate objects. We have seen that it leads to seemingly unresolvable contradictions when confronted with theories of transitivity and facts concerning the encoding of arguments in transitive clauses; and what is perhaps even more problematic, it runs into difficulties with its own theoretical basis, that of markedness theory. This all seems to suggest that some other analysis might be more appropriate.
It is not only the case-marking of objects that correlates with definiteness and animacy. We have already seen that a fundamental property of objects, namely affectedness, is also to a large extent dependent on these properties. Recall the quote from Hopper and Thompson (1980) cited above: “a definite O is more completely affected than an indefinite one”.

I propose, then, that what is being marked by the accusative case is not a high degree of individuation, but a high degree of affectedness. In other words, the objects that get case-marking in a DOM system do so not because they are definite and animate, but because they are affected. But affectedness is a rather abstract property which tends to co-vary with individuation.

How do we measure affectedness? How do we assess whether one participant is more or less affected than another? If object NPs are to be marked as being more or less affected, then there has to be some way of measuring the extent of their affectedness. But affectedness can take many different forms and is not easily quantified. So how can we impose some kind of classification onto this rather hazy concept?

It seems to me that there are two main ways of assessing the affectedness of an entity. One is in terms of part-whole relations. A cup with only its handle broken off is rather “less broken” than one which is completely smashed to bits; in general, an entity where only a subpart is affected is less affected than one which is affected as a whole.

In order to express such partial affectedness of entities we often resort to linguistic expressions of indefiniteness. We say, for instance, I drank some milk to express the affectedness of only a subpart of the available milk; the milk would imply affectedness of the whole rather than of a part. We killed a pig implies that there were several pigs and only one of them, a subpart of the whole, was subjected to being killed; the pig would mean that the only available pig was killed, that is, the totality rather than a part would be affected. In other words, expressions of definiteness are used as expressions of total affectedness.

The other dimension along which we assess degree of affectedness is a rather vaguer one, namely salience. From a human point of view, which presumably is what linguistic expressions encode, some types of effects are both more easily perceptible and of more interest than others. More specifically, effects on some types of entities are more salient to human perception and interests than others. An effect on a human participant is more likely to impinge directly on the lives of both the human in question and those surrounding him than an effect on an inanimate object; the situation described by Peter killed John is far more likely to have profound effects on all parties involved than, say, Peter broke the pot. In short, effects on human or animate entities are perceived as more dramatic, more significant, than effects on inanimates.

A case-marking system which reflects the link between affectedness, animacy and salience is found in the Tibeto-Burman language Dolakha Newari (Genetti, 1997). In this language the case-marker = ta, glossed “dative” is found on all recipient arguments (“indirect objects”) of ditransitive verbs, as well as on some patient arguments of monotransitive verbs. Case-marking of transitive patients is assigned as follows: (1) human patients are case-marked if they are given or accessible in the discourse; (2) non-human animate patients are case-marked when they are topical, that is, in contexts where “the manipulation of these animate patients [is] crucial to the plot resolution” (Genetti, 1997: 44); (3) inanimate patients are never case-marked.
The Dolakha Newari system case-marks highly salient objects. Recipients of ditransitive verbs are always animate in Dolakha Newari, which makes them inherently salient; furthermore, such arguments, being the ultimate goal of the action, are highly salient with respect to the action denoted by the ditransitive verbs. For patient arguments, salience is assessed by a combination of animacy and topicality. For human patients, which are high on the animacy scale, case-marking assignment depends only on givenness or accessibility. Nonhuman animate patients, on the other hand, may also receive case-marking, but only when the fact of their affectedness is central in the context—“crucial to the plot resolution”. Inanimate patients, however, are never considered salient enough to receive case-marking, even if they are highly topical. In other words, animate objects are considered to be more saliently affected than inanimate ones.

It appears, then, that affectedness is to a large extent encoded in language in terms of properties such as definiteness and animacy. On this basis we might re-cast the DOM phenomenon as follows: The tendency to case-mark objects that are high in definiteness and animacy is in fact a reflection of the accusative case as marking objects which are construed as being highly affected. But affectedness is as it were operationalised in terms of definiteness and animacy; when a language has to decide which kinds of objects are “affected enough” to receive case-marking, it may make this categorisation on the basis of more easily measurable properties on which affectedness depends. It is a simple matter to decide whether an object is definite (and so more affected than an indefinite object) or animate (and so more affected than an inanimate object), and so definiteness and animacy provide the yard-sticks by means of which affectedness may be measured. In this way we can capture the correlation between definiteness/animacy and affectedness which was discussed above, and also explain the tendency for accusative case to appear on definite animate objects.

6.2. Affectedness and accusative case

There are languages where the accusative case quite clearly reflects high affectedness of objects. For example, in Finnish, the accusative alternates with the partitive, the latter encoding partly affected objects:

(11) **Finnish** (Kittilä 2002, p.c.):
   - a. Hän jo-i maito-n
     - s/he drink-PAST.3SG. milk-ACC
     - ‘S/he drank (all) the milk.’
   - b. Hän jo-i maito-a
     - s/he drink-PAST.3SG. milk-PART
     - ‘S/he drank (some of the) milk.’

Although this is not a case of an accusative marker alternating with lack of case marking, it nevertheless a form of “differential object marking”, and one that correlates with degree of affectedness rather than with definiteness or animacy.

The Tibeto-Burman language Meithei shows an interesting link between animacy-governed DOM and the property of affectedness. The patient marker –pu occurs only on human or animate objects (12a–b). However, it is homophonous with an “adversative marker” (12c) which “signals that the –pu marked noun phrase is ill-fated in being
acted upon or that the verb is unexpected, unanticipated, or unfortunate” (Chelliah, 1997: 117):

(12) MEITHEI (Chelliah 1997):
   a. əy part léy-ru-khi-ni
      I parts buy-ADIR-STILL-COP ‘I will buy spare parts there’
   b. əŋáŋ-si má-bu īl-í
      child-PDET he-PAT push-N1HP ‘This child pushed him’
   c. əy-bu hi hon-bə həy-t-e-ne
      1-ADV boat row-NOM proficient-NEG-ASRT-SI ‘(But unfortunately). I don’t know how to row boats.’

A unified account of the semantics of this marker is possible under the assumption that it functions as a marker of affectedness, where animate or human objects are taken to be highly affected.

The examples above suggest that in some languages at least, accusative case is clearly a marker of high affectedness of objects, and that even where the marking is apparently triggered by animacy of the object referent there may still be a connection between the presence of the marker and the affectedness of the participant in question. If the accusative case functions as a marker of affectedness in some languages, then, one should consider the possibility that it might do so in languages in general, albeit not always as transparently as in (11–12).

6.3. Affectedness and markedness reversal

6.3.1. The control-affectedness opposition

If the possibility of markedness reversal is to provide an argument for affectedness rather than individuation as the property encoded by DOM, then there has to be a plausible polar opposition involving the property of affectedness, over which markedness reversal can be defined. That is, there has to exist a positive property which might be considered the “opposite” of affectedness, and which in the unmarked case is present with subjects, but not with objects.

Affectedness is a property of participants in verbal events; for an entity to be affected is generally defined in terms such as “undergoing a change of state as a direct result of an event” (see, e.g. Lakoff, 1977; Langacker, 1991; Dowty, 1991). In the case under discussion here, that is, affectedness of syntactic direct objects, it is, furthermore, a property of participants in transitive events. While transitivity is a notoriously difficult phenomenon to pin down in a precise manner, it is fairly generally agreed that the prototypical transitive situation is one where a volitional, controlling agent performs an action which has a physical and perceivable effect on a patient (see, e.g. Lakoff, 1977; DeLancey, 1985, 1987).

Affectedness, in other words, constitutes one “end” of a transitive event; the other, in the prototypical case, consists of the volitional instigation of the event. I will refer to such volitional instigation as control over the event.
Control and affectedness are not opposites in a strictly logical sense. They may, for instance, co-occur within one entity; it is perfectly possible for a participant which controls an event to also be affected by it. They have, however, been characterised as *causal opposites* ([Croft, 1988]3 meaning that they represent opposite poles of an event of causation. In the context of transitive semantics, then, there is a sense in which control and affectedness function as “opposite” properties.

What is more, there is evidence that these properties are “opposite” in a more fundamental sense within the context of prototypically transitive clauses. While control and affectedness can logically co-occur within a single argument, and sometimes do, the kinds of clauses where such co-occurrence is found tend, crosslinguistically, to deviate from the formally transitive prototype. One example is found with the so-called “ingestive verbs” ([Masica, 1976], that is, verbs such as “eat” which involve incorporating some substance or property into the agent’s own body or mind—in other words they involve a degree of *affectedness* of the controlling agent. One need not look farther than English to establish that such verbs behave in a manner that deviates from the transitive prototype—they have traditionally been characterised as “pseudo-intransitive”, meaning that they may occur equally felicitously with or without a direct object, unlike core transitive verbs such as *break* or *kill*. Other languages show other types of transitivity deviations with these verbs; for instance, in the Brazilian language isolate Trumai, the verb “eat” takes an oblique object ([Guirardello, 1999]); in the Tibeto-Burman language Chepang, the object of “eat” lacks the case-marker normally found on direct objects ([Caughley, 1982]); in the Papuan language Amele “eat” is among the verbs that may occur without an object clitic, while objects of most other transitive verbs are obligatorily cross-referenced ([Roberts, 1987]), and so forth. It is also not uncommon for ingestive verbs to be allowed in causative constructions which otherwise only apply to intransitive verbs ([Amberber, 2002]).

Conversely, if an affected object is construed as somehow being in control of the event that affects it (even if it is not itself the instigator of the event), languages may encode it with formal means that differ from those found in core transitive clauses. In Icelandic, for instance, an object which shows some degree of control over its own affectedness takes the dative rather than the accusative case:

(13) **ICELANDIC** Barðdal (2001):

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Hann klóraði mig</td>
<td>b. Hann klóraði mér</td>
</tr>
<tr>
<td>he,NOM scratched me,ACC</td>
<td>he,NOM scratched me,DAT</td>
</tr>
<tr>
<td>‘He scratched me(ACC)’</td>
<td>‘He scratched me(DAT)’</td>
</tr>
</tbody>
</table>

In (13a), the scratching is interpreted as a violent and painful act, probably intended to hurt me. (13b), on the other hand, means that I had an itch and the subject participant helped out by scratching me, perhaps in a place I could not reach myself. In other words, in (13b) the object is affected by its own volition, it is in *control* of its own affectedness; and it is therefore encoded in the dative case, rather than the accusative which is reserved for “normal” objects which are not in control.

---

3 Croft is here referring to volitionality, which is included in my definition of control.
What these examples show is that in prototypical transitive clauses, subjects are not affected, and objects are not in control. Failure to fulfill this requirement leads to failure to encode the clause as fully transitive in many languages. In other words, prototypical transitivity requires an opposition between subject and object with respect to the properties of affectedness and control: in the prototypical transitive clause, control and affectedness do not co-occur on the same argument (see also Testelec, 1998). In this sense, then, control and affectedness function as opposite properties in transitive clauses.

6.3.2. Reversal

It follows that there is indeed a type of markedness reversal involved in the subject–object opposition: In the unmarked case, subjects are controlling and nonaffected, while objects are affected and noncontrolling—“unmarked” in the sense that this distribution of control and affectedness over the core arguments will yield a basic transitive clause, while any other configuration of these properties may lead to constructions that deviate from the transitive prototype. This opposition is consistent with Battistella’s (1996) definition of reversal: we have two oppositions of inclusive asymmetry, affected versus nonaffected and controlling versus noncontrolling, where the positive terms of each opposition are in a relation of polar opposition to each other. The unmarked subject can then be characterised “sometimes as A and sometimes as non-B” (Battistella, 1996: 58), that is, sometimes as controlling and sometimes as nonaffected, depending on which opposition is relevant for the marking pattern in question. The properties of typical subjects and objects can be set out like in (14):

\[
\begin{array}{ll}
\text{SUBJECT} & \text{OBJECT} \\
\text{controlling} & \text{affected} \\
\text{nonaffected} & \text{noncontrolling}
\end{array}
\]

In the case of DOM, the relevant property is affectedness: objects that are positively specified for affectedness take the accusative case, while objects that are nonaffected—and subjects, which are also nonaffected—take no marking. The oppositions involved are illustrated in (15):

\[
\begin{array}{ll}
\text{SUBJECT} & \text{OBJECT} \\
\text{controlling} & \text{affected} \\
\text{noncontrolling} & \text{nonaffected}
\end{array}
\]

\[
\text{polar opposition}
\]

\[
\begin{array}{ll}
\text{inclusive} & \text{inclusive} \\
\text{asymmetry} & \text{asymmetry}
\end{array}
\]

It is significant that this double opposition—affectedness versus nonaffectedness, control versus noncontrol—is not a stipulation constructed to save an analysis, but is supported by independent evidence, unlike the definite versus nondefinite/indefinite versus nonindefinite
feature set which would be required to define markedness reversal with respect to the property of definiteness. We have seen from the examples in the previous section that, quite independently of any analysis in terms of markedness and reversal, in a core transitive clause subjects are indeed controlling and nonaffected, and objects are affected and noncontrolling, and that both sets of properties have consequences for the encoding of both argument types. The double opposition is necessary for a coherent analysis of subjects and objects in transitive clauses; the fact that it can also form the basis for an analysis in terms of markedness reversal is just a welcome bonus.

7. Disentangling the levels

7.1. Semantics versus case-marking

So far, we have shown that differential object marking is better accounted for in terms of the affectedness of object participants than in terms of properties of individuation of NP referents such as definiteness and animacy. We have not, however, solved the central problem: the apparent contradiction with Hopper and Thompson-style transitivity models as outlined in 4. If the accusative case marks highly affected objects, then what we appear to have shown is that affected objects are marked with respect to nonaffected ones, since highly affected objects would be the ones to take overt case-marking; and this is in even more direct contradiction with an account which explicitly states that affected objects are the most prototypical ones, or at least that prototypical transitive clauses have affected objects. Thus it also contradicts the statement in 6.3.2 that affectedness and noncontrol are unmarked properties of objects.

The apparent contradiction in fact stems from a failure to keep apart two distinct levels of analysis. There are two different, though related, things under discussion here: transitive semantics on the one hand, and case-marking phenomena on the other. The former is to do with general properties of transitive objects as opposed to subjects, irrespective of the formal system these subjects and objects are encoded in. The latter, on the other hand, concerns the way in which objects which do or do not show these properties are formally encoded within a specific kind of structural system, namely accusative case-marking systems.

7.2. Transitive semantics

Transitive semantics, as mentioned earlier, is taken to involve a volitional, controlling agent performing an act which has a physical and perceivable effect on a patient. Schematically, this might be represented as follows:

(16) CONTROLLING
    AGENT ----------- action ------------ AFFECTED
    PATIENT

In this schema, which makes no statement about the relative markedness of the participants involved, or of participants with different properties, being affected is a fundamental
characteristic of the patient participant. This is in line with the common understanding of what a transitive object is; the typical object is a patient, and a patient is, by definition, a participant affected by the verbal action.

(16) is a semantic abstraction, independent both of the properties of the constituents that instantiate the participants in a given clause (e.g. the definiteness of an NP or the animacy of its referent) and of the properties of the structural systems that encode this schema in actual language. Case-marking patterns such as DOM are such structural properties, which arise when (16) is to be given formal expression in a linguistic system.

7.3. Argument encoding: accusative versus ergative systems

Assuming that the basic function of core case-marking is a discriminatory one, as discussed in Section 2, a given language must, in encoding (16) formally, attempt to ensure that the two participants in the event—the controlling agent and the affected patient—are easily distinguished and identified with the correct grammatical relation. Given that the simplest solution to this is to overtly mark one of the participants while leaving the other unmarked, there are, then, two logical possibilities: (1) mark the instigating participant; (2) mark the affected participant. If the language chooses to formally mark the affected participant, we have what is known as an accusative case-marking system.

In other words, within an accusative system, affectedness functions as a marked property in the sense that it triggers overt case-marking. This is the key to understanding the apparent conflict between the two analyses discussed above which seemed to imply that individuated objects are unmarked and marked at the same time: They are unmarked at the level of transitive semantics, because a transitive clause typically has a highly affected and individuated object; but they are marked within an accusative case-marking system because such a system treats affectedness as the marked property in the opposition between the controlling initiator of the event at one end, and the entity affected by the event at the other. Note that this is a polar opposition where the relative markedness of the terms involved is determined entirely by context—in this case, by the case-marking system in question.

In terms of the schema of control and affectedness presented in (16), the accusative-ergative distinction can be conceived of as follows: Accusative and ergative languages differ in the way that they construe transitive situations; more specifically, in which phase of a transitive situation is construed by the case-marking system as marked relative to the other. In accusative languages, the effect of the action on the patient is treated as marked, being singled out with the accusative case, while the initiation of the action is the unmarked phase. Conversely, in ergative languages it is the initiation phase which is considered to be marked, taking ergative case, while the action’s effects are the unmarked feature of the situation. In the words of Halliday (1968: 185), the distinction is between “an ‘ergative’ form of organization, based on causation, where the question is whether the cause is external to the action or not” and “a ‘transitive’ [accusative, ÂN] form of organization, based on extension, where the question is whether the action extends beyond the actor or not”.

What this means is that in accusative languages, situations are characterised mainly by their initiation phase. The unmarked S argument is grouped together with the A, and so all
intransitive situations are treated as if they had a controlling initiator, even when the S participant is in fact semantically patientive. When we describe a situation in an accusative language such as English, the characteristic that all situations, transitive or intransitive, have in common, is an initiating participant: accusative languages as it were characterise situations in terms of who did X, while the effects on a participant distinct from the “doer” is the additional element which marks a clause as transitive as opposed to intransitive. In this sense, objects are “marked” participants with respect to subjects, in that they only occur in certain types of clauses whereas subjects are always present; and the more distinct such an object participant is from the unmarked subject participant, the more likely the clause is to be encoded as transitive.

7.4. Subject marking and split ergativity

The analysis just outlined predicts differential marking of subjects to work in terms of control: controlling subjects should take overt (ergative) case, while less controlling subjects may be unmarked, or take other forms of marking (cf. Figure 15). That this is so for a number of languages was mentioned in 5.3; languages whose case-marking system functions on a “semantic basis” (Dixon, 1994) tend to mark controlling as opposed to non-controlling subjects, and many languages that are usually labeled ergative reserve the ergative case for arguments which exert some measure of volitional control. Similarly, in split-marking systems where the split is conditioned by verb semantics (“active-inactive” or “split-S” and “fluid-S” languages), the relevant parameter tends to be the controllability of the action in question (Mithun, 1991; Dixon, 1994).

Other kinds of ergative splits, I would suggest, should rather be analysed as marking deviations from certain pre-defined patterns of control and affectedness. In such systems, the ergative case is employed in contexts where control is presented a marked property, not in opposition to affectedness, but with respect to certain expected patterns of natural events. We analyse events according to certain expectations towards the world: the various participants in an event, and the event itself, are attributed certain properties representing the “normal” way in which entities and events interact. When actual events violate these expectations, languages may choose to mark this in the structural encoding of these events. This may happen, for instance, when the subject, generally a controlling entity, is low on the animacy hierarchy and so not normally expected to be in control; when the clause is negated, counterfactual or has irrealis mood, in which cases there is no real-world event that could be controlled (see, e.g. the data on Cavineña in Dixon, 1994); or when the clause has past tense or perfective aspect, in which case there is no longer a possibility of controlling the event, as it has already been brought to an end. Such tense/aspect splits reflect perfectly the characterisations of accusative versus ergative systems laid out in 7.2; in past tense or perfective aspect, the most salient feature of the event in question is its result, as this result may still be perceivable, while the action that produced it is no longer in evidence. The event is characterised mainly by its effect, and so takes ergative marking. In the present tense or imperfective aspect, however, the effect may not yet be achieved; it is “not yet fully accessible to an observer” (DeLancey, 1987: 61). Since the effect is unavailable, the situation must rather be described in terms of its initiation by a controller, which fits our description of an accusative system.
This analysis of split-marking systems furthermore explains why the type of “differential subject marking” predicted by Aissen almost only occurs in combination with differential object marking. Such marking of subjects is part of an overall system which marks, as it were, unexpected distributions of control and affectedness. Thus not only deviations from expected control patterns, but also deviations from expected affectedness patterns trigger overt marking—a human-referring pronoun, for instance, might be marked with accusative case when it is in object position, since according to the underlying “expectations” of the system such a pronoun is most likely to be a controller rather than an affected entity. This kind of system is fundamentally different from one which just marks all instances of one of the terms of the control-affectedness opposition—either all affected objects or all controlling subjects. In the split-marking case, both types of properties may trigger marking, but only when they occur on participants not normally expected to show them. In “pure” ergative or accusative systems (to the extent that such things exist), on the other hand, only one of the two properties triggers marking, but it does so consistently, assigning case-marking to all participants judged to be controlling or affected, whether or not this is consistent with the “expected” patterning.

8. Conclusion

The assumption made by functional typology that there is a natural opposition between subjects and objects at the level of the inherent properties of the noun phrases that instantiate them (definiteness and animacy) does not conform very well to observable language data, as it is in many cases a prerequisite for an argument’s being encoded as an object that it is relatively highly individuated, as is the typical subject. Furthermore, the notion of markedness reversal which underlies this analysis is not in fact easily applicable to the kinds of oppositions involved. The relevant opposition, rather, is between the nature of the individual noun phrases’ involvement in the verbal event. The statement that “exactly what is marked for objects is unmarked for subjects” holds only within the domain of properties which define the relation of participants to the verbal action: an unmarked subject is controlling and nonaffected, an unmarked object is affected and noncontrolling, and deviations from this distribution is frequently reflected in the formal encoding of such clauses in ways other than with a prototypical transitive construction. However, properties of case-marking systems—notably, the need to unambiguously identify the participants in the clause with their respective grammatical relations—makes it necessary to single out one of the positive properties involved as marked relative to the other. Thus, accusative case-marking systems single out affectedness as the marked term in the control-affectedness opposition, and mark affected objects accordingly, while ergative systems treat control as the marked term and so overtly mark controlling subjects. The control-affectedness opposition can function in this way because it contains no inherent markedness relation; the extent to which one of these terms is marked with respect to the other is not imposed by the nature of the opposition itself, but arises entirely from the context in which it is embedded.

Of course, most case-marking systems work in terms of generalisation and abstraction. It is not implied here that all accusative case-marking systems independently assess the
degree of affectedness of each individual object of each individual clause and assign case-marking accordingly. Rather, such systems classify a broad range of objects as resembling the affected prototype, and case-mark them accordingly. But assuming the basic semantic content of the accusative case to be “affected participant” allows for a coherent account of a set of phenomena for which this paper has attempted to show that currently available analyses are inadequate.

References


Caughley, R.C., 1982. The syntax and morphology of the verb in Chepang. Pacific Linguistics Series B No. 84. Research School of Pacific Studies, Australian National University, Canberra.


