# Case marking strategies

Helen de Hoop & Andrej Malchukov[1]

Radboud University Nijmegen

DRAFT January 2006

**Abstract**

Two strategies of case marking in natural languages are discussed. These are defined as two violable constraints whose effects are shown to converge in the case of differential object marking but diverge in the case of differential subject marking. The strength of the case bearing arguments will be shown to be of utmost importance for case marking as well as voice alternations. The strength of arguments can be viewed as a function of their discourse prominence. The analysis of the case marking patterns we find cross-linguistically is couched in a bidirectional OT analysis.

## 1. Assumptions

In this section we wish to put forward our three basic assumptions:

(1) In ergative-absolutive systems ergative case is assigned to the first argument *x* of a two-place relation *R(x,y)*.

(2) In nominative-accusative systems accusative case is assigned to the second argument *y* of a two-place relation *R(x,y)*.

(3) Morphologically unmarked case can be the absence of case.

The first two assumptions deal with the linking between the first (highest) and second (lowest) argument in a transitive sentence and the type of case marking. For reasons of convenience, we will refer to these arguments quite sloppily as the subject and the object respectively, although we are aware of the fact that the labels *subject* and *object* may not be appropriate in all contexts, dependent on how they are actually defined.

In many languages, ergative and accusative case are assigned only or mainly in transitive sentences, while in intransitive sentences ergative and accusative case are usually not assigned. In that sense, we may call ergative and accusative 'dependent' cases, following Marantz (1991), since ergative and accusative crucially *depend on* the presence of another (core) argument (direct object and subject, respectively) in the clause. Unlike nominative case which is closely connected to the grammatical function of subject and which can combine with different thematic roles, ergative and accusative case are thematically more restricted in that they are mainly used for agents and patients, respectively. This cannot be reversed, since agents of intransitive sentences do not get ergative case in many languages, nor do intransitive subjects that fulfill the role of patient get accusative usually (although exceptions exist, as will be exemplified below).

That brings us to the third assumption. In Chomsky (1981) and subsequent work, a notion of abstract case (usually referred to as *Case* with a capital *C*) is used in connection to the inviolable Case Filter, that requires every lexically realized DP to bear case. Since many languages have little or no case morphology, the postulation of abstract Case guarantees DPs in structural case positions to be saved from the Case Filter. In this article, we take an Optimality Theoretic perspective (Prince and Smolensky 2004). Therefore, we assume that linguistic constraints such as the Case Filter, are violable in nature. Hence, we do not need the postulation of abstract Case. Some level of abstraction may still be needed, however, but in the absence of explicit evidence we simply assume that 'what you see is what you get'. Thus,

we interpret the absence of morphological case-marking as the absence of case (Aissen 1999; Aissen 2003). This also means that what is referred to as *nominative* or in other contexts *absolutive* case, can sometimes be viewed as the absence of case. This holds for example for the nominative case-marked subject in Hindi in sentence (4) below, but not for the nominative case-marked subject in Japanese in (5):

(4) Raam-∅ ek bakre-ko bec-taa hae.
Ram-NOM one goat-ACC selling is
'He sells the goat.'

(5) Boku -ga tomodati-ni hana-o ageta.
I-NOM friend-DAT flowers-ACC gave
'I gave flowers to my friend.'

In (4) the name *Ram* is glossed as being in the nominative, despite the fact that there is no such thing as a nominative marker in Hindi, the proper noun in (4) is in its unmarked, uninflected form. By contrast, there is a clear case marker, *viz. –ko*, on the direct object in (4). The uninflected word form is called nominative, but in fact, nominative case in Hindi can be viewed as the absence of case. By contrast, the nominative case in Japanese is expressed by a real case marker, and therefore we cannot equate nominative case with the absence of case in Japanese, unless the case marker gets dropped which is often the case in colloquial speech (Fry 2001; Lee 2002). In ergative case systems, the absolutive case is often unmarked and we will similarly assume that in the absence of morphological case marking, absolutive case is in fact the absence of case as well. This is illustrated in sentence (6) from Yup'ik.

(6) Angut-em tangrr-aa arnaq-∅.
Man-ERG sees woman-ABS
'The man sees the woman.'

## 2. Case marking: two basic functions

In this section we argue, following functional-typological insights, that two basic functions of case marking can be distinguished, the *identifying* function and the *distinguishing* one (Mallinson and Blake 1981; Kibrik 1985; Comrie 1989; Song 2001).

The identifying strategy makes use of case morphology to encode specific semantic/pragmatic information about the nominal argument in question. We say that case morphology is used to identify semantic or pragmatic properties. Lexical (inherent, oblique) as well as semantic cases are obvious examples of the identifying strategy (Butt and King 2003; Butt and King 2004). A clear example are the locative cases in Finnish:

(7) talossa house-INESSIVE 'in the house'
talosta house-ELATIVE 'out of the house'
talon house-ILLATIVE 'into the house'
talolla house-ADESSIVE 'on the house'
talolta house-ABLATIVE 'off the house'
talona house-ESSIVE 'as a house'

The lexical cases above clearly contribute to the meaning of the noun phrase, comparable with the meanings the prepositions in the English translations contribute. The identifying function of case is not restricted to lexical or semantic cases, however. In fact, structural or

grammatical cases identify certain semantic/thematic properties to a certain degree as well. For example, accusative case in direct object position can be argued to identify patienthood. Dative case in some languages is a structural case as well, yet it is clearly associated with thematic roles such as goal and experiencer. Ergative case is in many languages associated with 'true' agents. In Manipuri, for example, the ergative case on the agent in (8) marks high agentivity (the agent is in control, volitional), while a decrease in agentivity is signalled by the lack of ergative case in (9) (Bhat and Ningomba 1997).

(8) əy-nə tebəl-də theŋŋi.
I-ERG table-LOC touched
'I touched the table (volitionally).'

(9) əy tebəl-də theŋŋi.
I table-LOC touched
'I touched the table (involuntarily).'

In Manipuri, all and only true agents get ergative case. Thus, we can say ergative case in Manipuri identifies agentivity. We introduce a general constraint which states that ergative case identifies *strong subjects* (which we will write as *A*).

(10) IDENTIFY A/ERG: Ergative case identifies strong subjects (A ↔ ERG).

Clearly, the constraint IDENTIFY must be conceived of as a family of constraints and not a sole constraint. The notion *strength* will be elaborated upon in section 3.

The distinguishing strategy is a more specific strategy that is used for distinguishing between the two core arguments of a transitive clause, i.e. the subject and the object. The

intuition behind the distinguishing function is quite clear. When a two-place predicate *R(x,y)* is used to describe an event involving two participants, usually an agent and a patient, it is of utmost importance to avoid ambiguity as to which noun phrase corresponds to the first argument *x* (the agent) and which to the second argument *y* (the patient). For this purpose, case can be used to mark one of the arguments. If one argument is case marked, this already suffices for the purpose of disambiguation. Thus, from the distinguishing perspective, there is no need to case mark both arguments. Neither would it be necessary to case mark the one and only argument of a one-place (intransitive) predicate. Indeed, it has been argued that in many nominative-accusative case systems only the *y* is case marked (with accusative case) while the *x* remains morphologically unmarked. This view is in accordance with our assumption (3) presented above. When nominative case is the unmarked (uninflected) case form, we interpret it as the absence of case. Similarly, in pure ergative-absolutive systems only the *x* is case marked, while the *y* remains morphologically unmarked (absolutive). The only argument of an intransitive verb is unmarked as well, and although it is labeled absolutive or nominative, it can often be seen as the absence of case as well.

While Manipuri is an example of a radically identifying language, there are also languages which can be characterized as radically distinguishing. In Awtuw the object is obligatorily marked with accusative case if the object is equally high or higher than the subject in the animacy hierarchy (Feldman 1986):

(11) Tey tale-re yaw dæli
3FS woman-ACC pig bit
'The pig bit the woman.'

(12) Tey tale yaw dæli.
3FS woman pig bit
'The woman bit the pig.'

In Fore, another Papuan language, it is the subject that is marked with ergative case if the object is higher in the animacy hierarchy than the subject (Scott 1978):

(13) Yagaa-wama wá aegúye.
pig-ERG man hit
'The pig hits the man.'

(14) Yagaa wá aegúye.
pig man hit
'The man hits (or kills) the pig.'

In the Fore example in (14), the man is higher in the animacy hierarchy than the pig, and that is why *man* is interpreted as the subject, even though the canonical SOV word order is overruled. If the speaker wants to express that the pig hit the man, then the subject needs to be explicitly marked as the subject (as in (13)).

The distinguishing function of case can be characterized as a global constraint like in Fore (i.e., the relative animacy of subject and object are measured) or as a local one as in classical cases of the markedness effects in differential object marking, where for instance the animacy or definiteness of the object is evaluated independent of the animacy or definiteness of the subject (see de Swart 2003 for more discussion and more examples of global distinguishability). To put the general motivation behind this type of case marking, whether locally or globally applied, in a constraint (de Swart 2003; de Hoop and Lamers to appear):

(15) DISTINGUISHABILITY: The two arguments of a transitive clause should be distinguishable.

Case marking is a way to distinguish between the subject and the object and hence to satisfy the above constraint. However, if the subject and object are otherwise distinguishable (like when in Fore the subject outranks the object in animacy), then case marking is not necessary to satisfy DISTINGUISHABILITY. But if the object is more 'subject-like' (absolutely or relatively), i.e., if it equals the (general or actual) subject in animacy/definiteness, the subject and the object can no longer be distinguished on the basis of these animacy/definiteness properties. In order to satisfy the constraint DISTINGUISHABILITY and to avoid *potential* ambiguity, case marking can apply.

Cross-linguistically, a merely distinguishing function of case is rare. This could be explained by the fact that there are alternative strategies for disambiguation between the two arguments of a transitive predicate, such as for instance the use of subject agreement, word order restrictions, context and/or intonation (Keenan 1978; Bouchard 2001; de Hoop and Lamers to appear). For example, when in Fore the two arguments are equal in animacy, solely word order determines what is the subject and what is the object: the first noun phrase will then be interpreted as the subject

Obviously, the identifying and the distinguishing function are not orthogonal, but overlap considerably. In fact, if case is systematically used to identify the subject or the object in a transitive clause, then of course we get differentiation 'for free'. Therefore, case systems that are completely based on the function of identification must be richer in case morphology than the mainly distinguishing ones. But as we will see, both functions are needed to account for the various case marking patterns we find across languages.

## 3. Case and Optimality Theory

In the previous section we have introduced two basic case marking constraints, DISTINGUISHABILITY and IDENTIFY, that we will use in the remainder of this article to account in an Optimality Theoretic fashion for cross-linguistic case marking patterns. But first, we will briefly discuss how our account relates to some previous Optimality Theoretic analyses of case marking.

A well studied paradigm of differential case marking is differential object marking: in many languages DPs higher in animacy or definiteness are marked, while those lower may be not (Bossong 1985; Aissen 2003). For example, in Turkish specific objects are case marked with accusative case, while non-specific objects remain without case (Enç 1991):

(16) Ali bir kitab-ı aldı.
Ali one book-ACC bought
'Ali bought a certain book.'

(17) Ali bir kitap aldı.
Ali one book bought
'Ali bought some book or other.'

An explanation of differential object marking in terms of distinguishability is straightforward and elegant (Aissen 1999, 2003). In a canonical transitive construction, the object is lower than the subject in animacy/definiteness, and thus when the object is animate/definite it is *marked* (for an object) which means it should be (case-)marked. Aissen crucially combines markedness of form with markedness of meaning within her constraints. Thus, she uses constraints such as "Avoid a specific object which is not case marked" (*Oj/Indef Spec &

*$Ø_C$) and which is universally ranked higher than "Avoid a non-specific object which is not case marked" (*Oj/Indef Nspec & *$Ø_C$). By inserting an economy constraint such as "Avoid case marking" (*$Struc_C$) in between these two constraints, Aissen accounts for the Turkish pattern in which specific indefinite objects are case marked while non-specific indefinite objects are not. The ranking Aissen proposes for Turkish differential object marking is therefore: *Oj/Indef Spec & *$Ø_C$ >> *$Struc_C$ >> *Oj/Indef Nspec & *$Ø_C$. We will briefly discuss three problematic aspects of Aissen's approach.

The first problem, pointed out by de Swart (2003), is that we encounter cases where distinguishability is globally evaluated, rather than locally. In those cases, markedness applies to the *relation* between the subject and the object in that the subject should outrank the object in animacy, and not just to the features of the object (inanimate) or the subject (animate) itself. De Swart discusses the following two examples from Malayalam. In this language, usually *only animate* direct objects are case marked. This could straightforwardly be accounted for in Aissen's framework by a constraint ranking such as "Avoid an animate object which is not case marked" (*Oj/Animate & *$Ø_C$) >> "Avoid case marking" (*$Struc_C$) >> "Avoid an inanimate object which is not case marked"(*Oj/Inanimate & *$Ø_C$). However, when both the subject and the object are inanimate ánd there is a danger of ambiguity, the *inanimate* object gets case marked. This is illustrated below (Asher and Kumari 1997):

(18) Kappal tiramaalakaɭe bheediccu.
ship waves-ACC split
"The ship broke through the waves."

(19) tiramaalakaɭ kappaline bheediccu.
waves ship-ACC split
"The waves split the ship."

Thus, the inanimate objects in (18) and (19) are case marked, but only in sentences like these, in order to distinguish them from the inanimate subjects. However, this cannot be accounted for in terms of constraints such as Aissen's. In her framework we can only rank the economy constraint above the constraint that requires marking of inanimate objects and then *no* inanimate objects would be case marked, or we can rank it below, in which case *all* inanimate objects would be case marked. Hence, unlike local differential case marking (e.g., case marking of all and only animate objects) instances of global differential case marking (case marking of the object depending on the properties of the object *and* the subject) cannot be accounted for by Aissen's approach (see section 4 below for further discussion of differential case marking in Fore and Awtuw, driven by global distinguishability).

The second problem for Aissen's analysis of differential case marking, pointed out by de Swart (2003) and Woolford (2001; Woolford 2004), is that DISTINGUISHABILITY would predict that differential subject marking would mirror differential object marking in the sense that inanimate and indefinite subjects (which are more 'object-like') would be case-marked. This prediction is actually borne out in some languages, such as Qiang (a Tibetan language), where the subject in a transitive clause only takes agentive case when it is inanimate (Lapolla 2003).

(20) MoVu-wu qa datuəZ.
wind-AGT 1sg knocked-down
'The wind knocked me down.'

But in fact, such examples are rare. More often, DISTINGUISHABILITY effects in subject case marking manifest themselves in a split between nouns and pronouns. In many split-ergative languages, (first/second person) pronouns do not get ergative case marking when they are the

subject of a transitive verb, while nouns do (Silverstein 1976; Aissen 1999). By contrast, (first/second person) pronominal direct objects receive accusative case marking while nominal objects do not. A well-known example of this pattern is Dyirbal (Dixon 1979):

| Case marking in Dyirbal | 1st, 2nd person pronoun | 3rd person pronoun | noun |
|---|---|---|---|
| Transitive subject | ∅ | -ŋgu | -ŋgu |
| Transitive object | -na | ∅ | ∅ |

Table 1: Dyirbal

Although such a pattern can readily be explained in terms of (local) DISTINGUISHABILITY, since first and second person DPs more often occur as subjects than third person DPs, it does not account for the rarity of the animate/inanimate subject split. This was also pointed out by de Hoop and Narasimhan (de Hoop and Narasimhan 2005; de Hoop and Narasimhan submitted) in relation to the case alternations found in Hindi. While the object case alternation in Hindi correlates with animacy and/or specificity of the object (accusative case marks an animate or specific object), the subject case alternation is mainly triggered by the perfectivity of the transitive sentence. That is, ergative case is assigned in the context of a perfective verb phrase. Nominative case is the unmarked (morphologically zero) case which functions as the elsewhere case for both the subject and the object (see the sentences below).

(21) Wo ek laD.kaa / ek laD.ke-ko dekhtaa hae.
he one boy / one boy-ACC seeing is
'He sees a boy / the boy.'

(22) Us-ne ek laD.kaa / ek laD.ke-ko dekhaa.
he-ERG one boy / one boy-ACC saw
'He saw a boy / the boy.'

In Aissen's (1999, 2003) approach, sometimes the function of case-marking is to mark those subjects and objects that are less typical. Thus, ergative case should be assigned to agents which are somehow marked (in the grammatical role of subject) or weak (in terms of prominence) and hence confusable with patients. Overtly marking a weak subject or a strong object helps to tag the role of one argument with respect to the other. However, Hindi presents a serious counter-example for Aissen's predictions for subject case assignment, since the ergative case-marking of the subject of a perfective transitive clause in Hindi is clearly not triggered by a decrease in prominence of that subject. That is, subjects of perfective clauses are not less prominent nor in any other way more similar to objects than subjects of imperfective clauses. Following Aissen's account, we would expect atypical subjects (i.e. those subjects which are weak or low in prominence, e.g. non-volitional, inanimate) to receive overt case-marking. By contrast, subjects of actions which (successfully) bring about an event to completion (as in perfective contexts) are usually characterized as being typically agentive (Dowty 1991), or high on the transitivity scale (Hopper and Thompson 1980), hence strong or more prominent, and not weak or less prominent as predicted by Aissen. Further evidence that the ergative case marker *ne* in Hindi marks properties associated with strong or prominent subjects rather than with weak ones can be seen in constructions where *ne* is used optionally. Even though there is no second argument in intransitive predicates, case-marking may be used to resolve a different type of ambiguity in the subject itself. In Hindi, the single argument of a small set of primarily *bodily emission* verbs is optionally ergative. In such cases, ergative case-marking is argued to be associated with an interpretation of the agent as being volitional,

i.e., in *control* of the event the predicate refers to and as such more like a typical agent (Butt and King 1991; Mohanan 1994).

(23) Raam-ne chIIk-aa.
Raam-ERG sneezed
'Raam sneezed (purposefully).'

(24) Raam chIIkaa-aa.
Raam sneezed.
'Raam sneezed.'

Hence, as pointed out by de Hoop and Narasimhan (2005, submitted), Hindi provides clear evidence against the fact that differential case-marking on subjects is always motivated by the need to disambiguate subjects from objects in transitive clauses. Rather, ergative case-marking in Hindi seems to fulfill an *identifying* role, identifying strong or prominent subjects. So, it seems that we cannot account for the case marking patterns in Hindi only by appealing to the distinguishing role of case. De Hoop and Narasimhan (2005, submitted) posit an additional motivation for overt case-marking, which is to identify strong or prominent arguments. The determination of argument strength is made at a level of mapping between semantic features to argument strength. We assume that languages can vary in which features determine argument strength. In Hindi, perfectivity and sometimes also volitionality can make a subject strong, while animacy/specificity can make an object strong.

The third and last problem we wish to discuss in relation to Aissen's OT approach to cross-linguistic case marking becomes clear when we consider for example the subject case alternation found in Lezgian, which is triggered by volitionality. If the transitive subject is

non-volitional, this is a 'marked' feature for a subject, as would be predicted by Aissen. However, the feature of non-volitionality on the subject is marked with an oblique case, while ergative case is maintained for the volitional subject (Haspelmath 1993):

(25) Ajal-di get'e xa-na.
child-ERG pot broke
'The child broke the pot.'

(26) Zamiira.di-waj get'e xa-na.
Zamira-OBL pot break-AOR
'Zamira broke the pot (accidentally/involuntarily).'

So, here we see an instantiation of differential subject marking that reflects features of the subject (in particular, volitionality versus non-volitionality). Both the weak and the strong subject are case marked and thus, neither remains without case. Up to now we discussed varieties of differential case marking where one form was case marked form and the other form was the unmarked or null form (the absence of case marking). Clearly, this type of (asymmetrical) differential case marking can be analysed as the result of the interaction between DISTINGUISHABILITY and economy (Aissen 1999, 2003). However, case alternations between two overt cases (such as ergative – oblique in the Lezgian examples) cannot be due to the interaction between DISTINGUISHABILITY and economy which is also why these patterns cannot be accounted for in the framework of Aissen (1999, 2003). Indeed, in cases of symmetrical differential case marking, of the type observed for Lezgian, DISTINGUISHABILITY would be vacuously satisfied as both ergative and oblique case on the subject suffice to distinguish between the subject and the object. Hence, whenever we find symmetrical case

alternations (i.e., two types of morphological case) instead of asymmetrical ones (overt case marking versus no case marking) we will suggest an analysis in terms of IDENTIFY.

While Aissen (1999, 2003) uses the distinguishing function of case marking as the basis of her OT account of differential case-marking patterns, the constraints that Legendre *et al.* (1993) use in their earlier OT account of case-marking are quite different. Two constraints are in fact pure instantiations of the identifying function of case marking, although they do not play a very important part in their approach. These constraints associate agents with ergative (or nominative) case and patients with accusative (or absolutive) case. Note that they follow syntactic analyses which equate nominative and ergative case (both called *C1*), arguing that both cases are the highest case in the syntactic structure and behave alike with respect to binding, whereas accusative and absolutive (called *C2*) are situated lower (Bobaljik 1993; Laka 1993). Furthermore, one of their constraints, one that is of utmost importance in accusative languages, requires high-prominent arguments to receive nominative or ergative case.

In their account, Legendre *et al.* (1993) only consider agent and patient roles as input (A, P), and they label the arguments already in the input in terms of *prominence in the discourse* (X=high-prominent; x=low-prominent). They stipulate that if a predicate has only one argument, this argument is always high-prominent in the discourse. Hence, intransitives are always either *A* or *P* in the input. They also stipulate that ordinary transitive constructions are always *AP* in the input. That is, both arguments of a two-place predicate are assumed to be high-prominent in the discourse. Crucially, if one of the two arguments of a two-place predicate is low-prominent, this should lead to the use of a passive construction (of which the input is assumed to be *aP*) or an antipassive (of which the input is assumed to be *Ap*). This view has certain problems, as argued by de Hoop (1999), but we can safely ignore these as we will restrict ourselves to a discussion of ordinary transitive clauses within this paper. For an

Optimality Theoretic account of the relation between case and voice alternations, we refer the reader to Malchukov (to appear) and to section 5 below.

One general shortcoming of the approach of Legendre *et al.* (1993) is that their characterization of case-marking patterns as (nominative-) accusative and ergative(-absolutive) systems is made on the basis of whether the intransitive subject is marked like the transitive subject or like the transitive object. However, it is not clear that case-marking is based on any kind of perceived similarity between the two types of subjects in an accusative system, or the intransitive subject and the transitive object in an ergative one. Rather, these are all arguments in the morphologically unmarked form, and only one of the two arguments of a two-place predicate is being assigned a specific case marker. In our proposal, lack of overt marking as we often find with nominative and absolutive case, is treated as a lack of case, as we already explained above.

An obvious problem for the account of Legendre *et al.* arises when in a language more than one case-marker is available for assignment to a particular argument, that is, when the language allows for differential object marking or differential subject marking, or even both, as Hindi does (de Hoop and Narasimhan 2005). It turns out that the constraints on case-assignment are often sensitive to additional semantic restrictions in terms of contextually restricted faithfulness constraints (Stiebels 2000; Woolford 2001) that specify the conditions under which one or the other case-marker is associated with the argument. This can be witnessed for transitive sentences, whose input should always be *AP* according to Legendre *et al.*, but the problem is even obvious with intransitive inputs.

As pointed out by de Hoop and Narasimhan (2005), arguments that are definite and animate can be seen as stronger, more prominent in the discourse, or as typical full-fledged arguments, independent of whether they are the subject or the object of a transitive clause. Cross-linguistically, the strength of nominal constituents seems to influence case marking

(a.o., de Hoop 1996). This is also clear in the account of Legendre *et al.* (1993) who emphasize the role of discourse prominence for case marking and use Optimality Theoretic constraints such as "High-prominence arguments receive subject case-marking" and "Low-prominence arguments do not receive subject or object case marking". Also thematic properties contributing to agentivity or patienthood of the core arguments, seem to contribute to the strength of an argument (Dowty 1991). Importantly, different perspectives on the strength of DPs point in the same direction: animate and specific DPs are usually high-prominent in the discourse and they are also often realized as real syntactic arguments (thus qualify as better agents and patients). We will refer to these DPs as *strong* DPs in the remainder of this article, basically following de Hoop and Narasimhan (2005). Note by the way that this notion of *strength* for arguments only slightly differs from the notion of *strength* used in de Hoop (1996) to account for the role of DP semantics in (abstract) differential case marking. De Hoop (1996) focusses on the difference between quantificational versus predicative types of noun phrases but this difference can also be captured under the notion of strength as we use it here.

Thus, we can define the identifying function of case as identifying/marking the strong arguments, both subjects and objects. What counts as a strong DP, that is the cut-off point between strong and weak, may vary from language to language. It should be kept in mind that similarly to DISTINGUISHABILITY (or markedness on Aissen's approach) IDENTIFY is a family of constraints rather than a unitary constraint, subsuming several subhierarchies (see for example the complex pattern of case marking in Aranda in (40) below, resulting from the fact that relative marking of both DISTINGUISHABILITY and IDENTIFY constraints with respect to economy varies with the ranking of the DP in the animacy hierarchy). Besides, the identifying function does not always apply to strong DPs in general, it can also be applied to identify/mark more specific properties of certain DPs, such as volitionality of subjects.

## 4. A bidirectional OT approach to case marking

As we have pointed out above, the strength of noun phrases seems to influence case-marking. In languages with differential case-marking, noun phrases that are strong are likely to be overtly case-marked (de Hoop 1996). However, this does not always hold. In fact, sometimes the weak rather than the strong arguments receive overt case marking (cf. Aissen 1999). We claim that cross-linguistic variation in case marking patterns can be analysed in terms of differences in the relative strengths of the two basic case marking constraints, IDENTIFY and DISTINGUISHABILITY, in relation to economy. In this section we will present our bidirectional OT analysis of the resulting case patterns.

Recall the pattern in Manipuri, repeated below for convenience. The ergative case on the subject marks high agentivity while a decrease in agentivity of the subject is signalled by the lack of ergative case:

(27) əy-nə tebəl-də theŋŋi.
I-ERG table-LOC touched
"I touched the table (volitionally)."

(28) əy tebəl-də theŋŋi.
I table-LOC touched
"I touched the table (involuntarily)."

We have a minimal pair of two form-meaning pairs here, where one form (ergative) corresponds to one meaning (the strong subject that we will write as *A*) and the other form (without case) to the other meaning (the weak subject, written as *a*).

A pattern as such with two related forms and two related meanings and a one-to-one mapping between the forms and the meanings suggests an analysis in terms of bidirectional OT (Blutner 2000). The markedness principle states that an unmarked form goes with an unmarked meaning, and a marked form with a marked meaning (Horn 1984), and it is known to follow from bidirectional optimization (from form to meaning and from meaning to form)..

Let us briefly illustrate the main characteristics of bidirectional OT. Assume that we have two forms $f_1$ and $f_2$ and two meanings $m_1$ and $m_2$. We stipulate that the form $f_1$ is less marked than the form $f_2$, which means that for a given meaning $m$, form $f_1$ will be the optimal form. Furthermore, interpretation $m_1$ is less marked than the interpretation $m_2$, which means that for a given form $f$, meaning $m_1$ will be the optimal meaning. Thus, the following ordering relation between form-meaning pairs can be derived, represented in an arrow diagram, where the arrows point to the preferred pair (Blutner 2000; Dekker and van Rooij 2000).

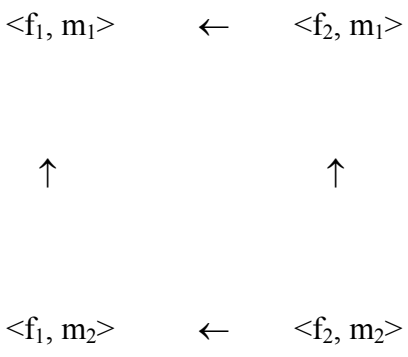


Figure 1: Bidirectional optimization

In Blutner's (2000) framework a form-meaning pair $\langle f, m \rangle$ is called *super-optimal* if and only if there is no other super-optimal pair $\langle f', m \rangle$ such that $\langle f', m \rangle$ is more harmonic than $\langle f, m \rangle$ and there is no other super-optimal pair $\langle f, m' \rangle$ such that $\langle f, m' \rangle$ is more harmonic than $\langle f, m \rangle$. The reader may verify that according to this definition, there are two super-optimal pairs in the diagram in Figure 1, namely $\langle f_1, m_1 \rangle$ and $\langle f_2, m_2 \rangle$. Indeed, although $f_2$ is

not an optimal form itself and $m_2$ not an optimal meaning, the pair $<f_2, m_2>$ is super-optimal, because there is no super-optimal pair that blocks it. That is, the two candidates $<f_1, m_2>$ and $<f_2, m_1>$ are not super-optimal themselves because they are each blocked by the other super-optimal pair $<f_1, m_1>$ and that is why they cannot block $<f_2, m_2>$.

Thus, bidirectional OT provides us with two super-optimal form-meaning pairs, one linking the unmarked form to the unmarked meaning, and one linking the marked form to the marked meaning. This is in accordance with the markedness principle that is a general principle in natural language. Let us now turn to our bidirectional analysis of the Manipuri case marking pattern. In Manipuri, subject case marking is completely determined by IDENTIFY, so all and only strong subjects get ergative case (Bhat and Ningomba 1997). Thus, we get two super-optimal form-meaning pairs for the subject as illustrated in Tableau 1:

| Subject | IDENTIFY (A/ERG) | ECONOMY |
|---|---|---|
| ✌ [ERG, A] | | * |
| [ERG, a] | * | * |
| [∅, A] | * | |
| ✌ [∅, a] | | |

Tableau 1: Subject case marking in Manipuri

The bidirectional OT tableau can be read as follows. The fourth candidate form-meaning pair is super-optimal because it does not violate any of the constraints. This pair combines a weak subject meaning with a null form. The second and the third candidate form-meaning pair are both blocked by this fourth candidate. The second pair has the same meaning as the fourth, but its form (ergative case) is less economical (suboptimal), while the third pair has the same form as the winning fourth candidate, but it is linked to a less harmonic meaning (as it induces

a violation of IDENTIFY). Therefore, the second and third form-meaning pairs are blocked by the fourth pair. The first pair, however, cannot be blocked by the super-optimal pair, since it differs in both form *and* meaning from that one. Hence, it is not blocked by a super-optimal pair at all (as the other two candidates are not super-optimal themselves). That is why the first candidate comes out as super-optimal as well. Hence, as a result we have two winning form-meaning pairs, one which assigns no case to a weak subject, and another which assigns ergative case to a strong subject.

Above we presented Fore as an example where DISTINGUISHABILITY rather than IDENTIFY governs differential subject marking. The relevant pattern is repeated below.

(29) Yagaa-wama wá aegúye.
pig-ERG man hit
'The pig hits the man.'

(30) Yagaa wá aegúye.
pig man hit
'The man hits (or kills) the pig.'

In Fore, the subject receives ergative case marking when it is (relatively) weak, in the sense that it is less prominent than the object (i.e., the object outranks the subject in the animacy hierarchy). In order to distinguish this weak subject from the object, the weak subject is case marked. Hence, when the subject is relatively strong (as animate or more animate than the object) it remains without case. This is illustrated in the bidirectional OT tableau below:

| Subject | DISTINGUISHABILITY | ECONOMY |
|---|---|---|
| [ERG, A] | | * |
| ✌ [ERG, a] | | * |
| ✌ [∅, A] | | |
| [∅, a] | * | |

Tableau 2: Subject case marking in Fore

The third candidate form-meaning pair clearly is a super-optimal pair as it does not violate any of the constraints. Hence, the third pair blocks both pairs that differ from the super-optimal pair in either form (the first candidate) or meaning (the fourth candidate). This leaves the second candidate which differs from the third one in both form and meaning, as the second super-optimal form-meaning pair.

Now, we have illustrated how we can account for differential case marking patterns on the basis of IDENTIFY as well as those on the basis of DISTINGUISHABILITY, both in relation to a general principle of economy. Strikingly, whereas IDENTIFY in relation to economy results in ergative case marking of the *strong* subject, DISTINGUISHABILITY in relation to economy results in ergative case marking of the *weak* subject. This explains the variation in differential subject marking patterns that is found cross-linguistically and that was presented as a problem for the previous OT account of case marking by Aissen (1999, 2003).

Another problem that was pointed out in the previous section is the phenomenon of *symmetrical* differential case marking, where two case forms alternate rather than one case form alternating with no case. The relevant examples are repeated in (31) and (32) (Haspelmath 1993).

(31) Ajal-di get'e xa-na.
child-ERG pot broke
'The child broke the pot.'

(32) Zamiira.di-waj get'e xa-na.
Zamira-OBL pot break-AOR
'Zamira broke the pot (accidentally/involuntarily).'

This pattern can be accounted for straightforwardly in a bidirectional OT approach as well. In Lezgian ergative case switches to oblique if the subject is weak (in this case non-volitional), hence we get the following pattern:

| Subject | IDENTIFY (A/ERG) | ECONOMY |
|---|---|---|
| ✌ [ERG, A] | | * |
| [ERG, a] | * | * |
| [OBL, A] | * | * |
| ✌ [OBL, a] | | * |

Tableau 3: Subject case marking in Lezgian

Things are different for asymmetrical differential case marking which is often triggered by DISTINGUISHABILITY. This can be illustrated by the well-known pattern of differential case marking when different types of nominal constituents select different cases (for example, nouns are marked differently from pronouns in Australian split ergative languages), as in the example of Dyirbal discussed above. This type of differential case marking is due to DISTINGUISHABILITY in interaction with economy. This is illustrated for differential subject marking in Dyirbal:

| Subject | DISTINGUISHABILITY | ECONOMY |
|---|---|---|
| [ERG, A] | | * |
| ✌ [ERG, a] | | * |
| ✌ [∅, A] | | |
| [∅, a] | * | |

Tableau 4: Subject case marking in Dyirbal

In Dyirbal, first and second person pronouns count as strong subjects (*A*'s) and they do not need case marking in order to distinguish them from objects. This is in accordance with the fact that the same categories (i.e., first and second person pronouns) count as strong objects (*p*'s), hence they do receive accusative case marking when they are the object of a transitive clause. The same correlation between case marking strategies and differential case marking patterns (symmetrical versus asymmetrical) which we observed for differential subject marking can be observed for differential object marking too. That is, in a split ergative language like Dyirbal, differential case marking of the object is asymmetrical and clearly due to DISTINGUISHABILITY.

On the other hand, symmetrical differential object marking must be due to IDENTIFY. Differential object marking in Finnish serves to illustrate this analysis. In Finnish, weak objects get partitive case while strong objects get accusative case. In Finnish, a strong object is obtained when the predicate is bounded (that is, non-homogeneous) (Kiparsky 1998).

(33) Ammuin karhu-a.
shot.1sg bear-PART
"I shot at a/the bear."

(34) Ammuin karhu-t.
shot.1sg bear-ACC
"I shot a/the bear."

So, in Finnish, IDENTIFY does the job all by itself. Note that the economy is vacuously violated here. And again, DISTINGUISHABILITY would be vacuously satisfied, as both accusative and partitive case would suffice to distinguish the object from the subject.

| Object | IDENTIFY (P/ACC) | ECONOMY |
|---|---|---|
| ✌ [ACC, P] | | * |
| [ACC, p] | * | * |
| [PART, P] | * | * |
| ✌ [PART, p] | | * |

Tableau 5: Object case marking in Finnish

So, while asymmetrical differential case marking can sometimes be explained by DISTINGUISHABILITY, symmetrical differential case marking should be due to IDENTIFY. This is consistent with the observation by Woolford (2004) that differential case marking patterns are heterogeneous, and some are determined by an alternation in argument structure. Yet, she does not relate the semantic types of differential case marking to the syntactic patterns in the way suggested here. We think the two different strategies of case marking provide us with a straightforward explanation for the attested correlations.

Unlike in differential subject marking, in the case of asymmetrical differential object marking the effects of IDENTIFY and DISTINGUISHABILITY converge. This explains the cross-linguistic consistency of differential object marking patterns as compared to differential

subject marking. In order to illustrate this, we will briefly discuss differential object marking in Hindi.

In differential object marking in Hindi, there are again two forms (one without case and one with accusative case) and two meanings (a weak and a strong object) and this leads to two super-optimal form-meaning pairs, as illustrated in the tableau:

| Object | IDENTIFY (P/ACC) | ECONOMY |
|---|---|---|
| ✌ [ACC, P] | | * |
| [ACC, p] | * | * |
| [∅, P] | * | |
| ✌ [∅, p] | | |

Tableau 6a: Object case marking in Hindi

However, as we pointed out above, we cannot make a principled distinction between the identifying and the distinguishing function of case here, since both functions predict the *strong* objects to be case marked. Indeed, DISTINGUISHABILITY would give rise to exactly the same two super-optimal pairs as IDENTIFY:

| Object | DISTINGUISHABILITY | ECONOMY |
|---|---|---|
| ✌ [ACC, P] | | * |
| [ACC, p] | | * |
| [∅, P] | * | |
| ✌ [∅, p] | | |

Tableau 6b: Object case marking in Hindi

Although we have seen that the two constraints lead to the same type of differential object marking (namely, marking the strong objects), we may wonder whether we can still distinguish between differential object marking triggered by IDENTIFY and differential object marking triggered by DISTINGUISHABILITY. That is, does the convergence of the effects of the two constraints mask two different types of differential object marking after all? We think that indeed it does and that some patterns are in fact better explained by IDENTIFY, some others by DISTINGUISHABILITY. For example, animacy effects in differential object marking are usually due to DISTINGUISHABILITY, as is obvious in languages of the Awtuw type, repeated below for convenience.

| | | | | |
|---|---|---|---|---|
| (35) | tey | tale-re | yaw | dæli |
| | 3FS | woman-ACC | pig | bit |
| | 'The pig bit the woman.' | | | |
| (36) | tey | tale | yaw | dæli. |
| | 3FS | woman | pig | bit |
| | 'The woman bit the pig.' | | | |

In Awtuw, the object is marked with accusative case if it outranks the subject in animacy. Because we are dealing with global (relative) differential object marking here, the pattern must be due to DISTINGUISHABILITY. In the following tableau, the strong object *P* stands for an object which outranks the subject in animacy.

| Object | DISTINGUISHABILITY | ECONOMY |
|---|---|---|
| ✌ [ACC, P] | | * |
| [ACC, p] | | * |
| [∅, P] | * | |
| ✌ [∅, p] | | |

Tableau 7: Object case marking in Awtuw

However, this is not the best explanation for the differential object marking pattern in Central Pomo, where patientive case is locally (that is, independent of the case of the subject) assigned to human objects only (Mithun 1991):

(37) M'u·tu ?a·hk'úm.
he.PAT I.killed
'I killed him.'

(38) Mu·l ?a·hk'úm.
he I.killed
'I killed it (the bee).'

This usual differential object marking pattern could be explained by DISTINGUISHABILITY. But strikingly, it carries over to differential subject marking:

(39) Q'alá·w m'u·tu.
died he.PAT
'He died.'

In (39) the subject of the intransitive clause is a patient and when it is human it gets the same case marking as the object of the transitive clause in (37). This can only be explained by IDENTIFY as it is a strong P-argument that gets marked, but in an intransitive clause, thus in the absence of another argument (Malchukov 2005). Hence, the case marking in (39) cannot be explained by DISTINGUISHABILITY. Therefore, the differential object marking pattern in this language is also better explained by IDENTIFY.

| Object | IDENTIFY (P/PAT) | ECONOMY |
|---|---|---|
| ✌ [PAT, P] | | * |
| [PAT, p] | * | * |
| [∅, P] | * | |
| ✌ [∅, p] | | |

Tableau 8: Object case marking in Central Pomo

To sum up, taking into account two functions of case marking, DISTINGUISHABILITY and IDENTIFY, we can explain the asymmetries between differential subject marking and differential object marking. While the two constraints give rise to the same type of differential object marking (marking the strong object), the two constraints diverge in the case of differential subject marking (DISTINGUISHABILITY predicts the case marking of the weak subject in order to distinguish it from the object, while IDENTIFY predicts the case marking of the strong subject). But even in the case of differential object marking, we can find examples that seems to be triggered by the need to satisfy DISTINGUISHABILITY, such as Awtuw, as well as examples where the key constraint seems to be IDENTIFY, as in Central Pomo.

Our approach can also account for violations of the Silverstein's hierarchy constraints on case patterns, such as the use of ergative case in Aranda (Silverstein 1976, discussed by

Woolford 2004). In Aranda, surprisingly both first person pronouns (strongest in the person hierarchy) and inanimate nouns (weakest in the animacy hierarchy) get assigned ergative case:

(40) ERG: 1st > 2nd > 3rd > human > animate > inanimate.

This pattern results from the fact that both IDENTIFY and DISTINGUISHABILITY interact with the animacy hierarchy (Silverstein 1976; Aissen 2003) but interact in the opposite way. Recall that DISTINGUISHABILITY compels differential case marking on the weakest (inanimate) subjects. Hence, the constraint penalizing inanimate subjects unmarked for ergative case is the strongest and the marking of inanimate subjects in Aranda results from the fact that this constraint is stronger than economy considerations which penalize morphological case, while other segments of the animacy hierarchy are dominated by the economy constraints. On the other hand, IDENTIFY penalizes case marking of subjects lower in the animacy hierarchy: only the strongest subjects, i.e., the first person pronouns, are identified/marked with ergative case. Hence, this constraint outranks the economy constraint as well. Thus, the effects of both IDENTIFY and DISTINGUISHABILITY are visible in Aranda: the ergative case on the first person pronoun (strong subject) is due to IDENTIFY, while the ergative case on the inanimate noun (weak subject) satisfies DISTINGUISHABILITY. The other types of subjects hold an intermediate position between weak and strong, and we will write them as *A-a* (in this representation, we abstract away from the fact that this is not a unitary type, as it relates to different positions in the animacy hierarchy).

| Subject | DISTINGUISHABILITY | IDENTIFY(A/ERG) | ECONOMY |
|---|---|---|---|
| ✌ [ERG, A] | | | * |
| [ERG, A-a] | | * | * |

| | | | |
|---|---|---|---|
| ✌ [ERG, a] | | * | * |
| [∅, A] | | * | |
| ✌ [∅, A-a] | | | |
| [∅, a] | * | | |

Tableau 9: Subject case marking in Aranda

The fifth candidate form-meaning pair which combines no case with an intermediate subject (not strong, not weak) comes out as super-optimal. This one blocks the candidates that differ from it only in form or in meaning. However, the first and third candidate pairs differ both in form *and* in meaning from the super-optimal pair. Hence, they come out as super-optimal as well. One might think that the third candidate pair would be blocked by the first one, since these two only differ in meaning but not in form. However, this is not the case. First, there is no other case form available, hence there is no other form for the other meaning (the weak, inanimate subject) and we may expect ambiguity. But more importantly, a difference in form is not really necessary here, because the first person pronoun (a strong subject) can never be inanimate (a weak subject) at the same time. That is, ergative case is not ambiguous here between marking a strong subject and marking a weak subject.

## 5. More on economy

So far we have worked with a very general economy constraint (simply called ECONOMY in the OT tableaux) that interacted with our two basic case marking constraints DISTINGUISHABILITY and IDENTIFY. In order to account for a few more cross-linguistic generalizations in the domain of case marking, we need to refine this constraint. That is, we will replace ECONOMY by Malchukov's (to appear) constraint PAIP, which was originally an

abbreviation of "Primary Actant Immunity Principle". In our formulation PAIP penalizes case marking an (otherwise) unmarked argument. Thus, in general, PAIP penalizes morphological case marking of the absolutive argument in ergative languages and of the nominative argument in nominative-accusative languages. Of course, this can be seen as a reformulation of a constraint that states that one argument should always bear the unmarked case, or to put it differently, that the unmarked case (either nominative or absolutive) is obligatorily present in every sentence of a language. This constraint thus resembles Tsunoda's (1981) "Unmarked Case Constraint" as well as Bobaljik's (1993) "Obligatory Case Parameter".

(41) PAIP: Avoid (case) marking of the unmarked argument.

PAIP thus penalizes 'marking the unmarked'. As has been argued by Malchukov (to appear), the potential conflict between IDENTIFY and PAIP can explain the striking fact that differential object marking is normally found in nominative-accusative languages, while differential subject marking is usually found in ergative languages (Bossong 1985; Drossard 1991).

In nominative-accusative languages, where the subject of a transitive sentence is the unmarked argument, differential object marking does not violate PAIP which is satisfied by the nominative subject. But in ergative languages, where the object of a transitive verb is the unmarked argument, object marking would induce a violation of PAIP. In those languages, we often see that a weak object leads to the use of an antipassive construction, while a strong object remains in the unmarked (absolutive) case. Take for example the alternation between (42) and (43) from Greenlandic Eskimo (Bittner 1988):

(42) Jaaku-p arnaq tuqut-p-aa.
Jacob-ERG woman kill-IND-3sERG/3sNOM
'Jacob killed the woman.'

(43) Jaaku arna-mik tuqut-si-v-uq.
Jacob woman-INSTR kill-AP-IND-3sNOM
'Jacob killed a woman.'

Note that in (43) the object or the *y* argument is non-specific, whereas it is specific in (42). Yet, only (42) is a true transitive construction with ergative case on the subject and both subject and object agreement on the verb, whereas (43) is in fact an intransitive, more specifically an antipassive, construction and its only true argument (the subject) is therefore unmarked for case, whereas the *y* argument is marked with oblique (instrumental) case.

Note that marking of the weak objects in an antipassive construction might appear problematic since the constraints DISTINGUISHABILITY and IDENTIFY both predict preferential marking of strong objects. The contradiction is only apparent, though. As noted above, in ergative languages the constraint IDENTIFY (P/ACC) cannot be saisfied due to the lack of accusative case. Hence, the strong object in a canonical transitive construction will remain unmarked. However, a weaker version of the IDENTIFY constraint, requiring differential marking of weak and strong objects, can be satisfied indirectly, through marking of the weak object with an oblique case, that is, differently from strong objects. In section 5 we shall provide more examples showing how a weaker version of IDENTIFY interacting with economy can explain some unusual patterns of differntial case marking.

In other words, weak objects may cause the shift to antipassive in ergative languages. This can be explained as the avoidance of a PAIP violation. If the alternation would have been marked on the (absolutive) object, that would violate PAIP which states that the unmarked

argument should not be messed around with. By using the antipassive construction, the subject becomes the unmarked argument (in the absolutive case), which means that PAIP as a requirement that the clause must have an unmarked argument is fulfilled.

Similarly, we see evidence for PAIP in the nominative-accusative languages as well. While in ergative languages a change in the strength of the subject can affect the form of the subject exclusively (differential subject marking), since this does not violate PAIP, in nominative-accusative languages, on the other hand, a weak agent subject regularly leads to passivization. Thus, a passive applies when the subject is indefinite, non-specific or not important in the discourse. Similarly, in some languages passive forms are used to indicate non-volitionality of the subject (cf. (Masica 1991) on Sinhala and Dhivehi). In ergative languages, on the other hand, a non-volitional subject may lead to a differential subject marking pattern, as observed above. Thus, features that trigger differential subject marking in ergative languages, may cause the use of a passive construction in nominative-accusative languages. Again, this can be straightforwardly explained by PAIP. If a change in meaning would have been marked on the (nominative) subject of a transitive clause, that would violate PAIP which states that the unmarked argument should not be marked. By using the passive construction, the object is promoted to the function of subject and hence becomes the unmarked argument (in the nominative case), thus satisfying PAIP.

As is expected if PAIP is a violable constraint, it may be violated as well in certain circumstances. An example of a violation of PAIP comes from Warlpiri (Hale 1973):

| | | | |
|---|---|---|---|
| (44) | Njuntulu-lu | npa-tju | ŋatju. |
| | 2SG-ERG | 2SG-1SG | speared.1SG |
| | 'You speared me.' | | |

(45) Njuntulu-lu npa-tju-la ŋatju-ku.
2SG-ERG 2SG-1SG-la speared.1SG-DAT
'You speared at me."/ "You tried to spear me.'

In (44) the object is in the unmarked case, the absolutive, while the subject is in the ergative case. In (45), however, we assume that the object *me* is weak as the action is attempted but not necessarily accomplished and this weakness is marked by dative case on the object. In other ergative languages, as we have seen above, a weak object may result in the use of an antipassive construction (cf. Legendre *et al.* 1993). But in Warlpiri, both arguments of the transitive clause are actually case marked in (45) and this means that PAIP is indeed violated. We can argue that in those examples again IDENTIFY and PAIP are conflicting, but IDENTIFY outranks PAIP.

To sum up, we have seen harmonic cases, where IDENTIFY and PAIP reinforce each other. That is, both constraints can be satisfied in case of differential object marking in nominative-accusative languages and differential subject marking in ergative languages. When the two constraints are in conflict, we commonly see that a voice alternation is a way to resolve the conflict. As predicted, passivization is applied when there is need to encode a subject alternation in a nominative-accusative language, while antipassivization applies when an object alternation must be encoded in an ergative language (Malchukov, to appear). Note that this was also observed by Legendre *et al.* (1993) who argue that passives occur when the input is *aP* (with a weak subject), while antipassives are the result of an *Ap* input (with a weak object). However, Legendre *et al.* do not account for the fact that passives are found more often in nominative-accusative languages, while antipassives are found more often in ergative

languages. In our approach this is straightforwardly explained by the interaction between two conflicting constraints, as an attempt to satisfy PAIP.

We have also seen the case in Warlpiri where PAIP is clearly violated, as it is outranked by IDENTIFY. Hindi which was discussed above is another example of a language where PAIP is clearly ranked lower than IDENTIFY. Hindi shows both differential subject marking (based on a split between perfective and imperfective tenses) and differential object marking (based on the features animacy and specificity of the object). The two case marking alternations are for the greater part independent of each other in Hindi, which means that an ergative-accusative pattern can be found in a perfective context with an animate and/or specific object, as in the following example:

| | | | | |
|---|---|---|---|---|
| (46) | Us-ne | ek | laD.ke-ko | dekhaa. |
| | he-ERG | one | boy-ACC | saw |
| | 'He saw the boy.' | | | |

This means that the constraint PAIP is ranked low in Hindi (cf. de Hoop and Narasimhan 2005). In Kashmiri, which also has differential subject and differential object marking, animate objects are *not* marked in perfective contexts, i.e. when ergative case is assigned to the subject (Klaiman 1987). In (47) with an imperfective verb the first person pronominal object gets accusative case, while it remains caseless in (48) with a perfective verb (Wali and Koul 1997).

| | | | | |
|---|---|---|---|---|
| (47) | Su | chu | me | parına:va:n. |
| | he | is | 1sgACC | teaching |
| | 'He is teaching me.' | | | |

(48) Nana-n roTus bɪ.
Nana-ERG caught 1sg
'Nana caught me.'

This means that PAIP is ranked higher in Kashmiri than in Hindi, and so in Kashmiri one argument must remain unmarked (if not the subject, than the object).

A final puzzle is presented by certain cases where an A-feature is marked on the object, or where a P-feature is marked on the subject, to satisfy PAIP. One example is the following Russian construction, where non-volitionality of the subject is marked by instrumental case on the object:

(49) On krutil rulj.
he rotate wheel-ACC
'He rotated the wheel (consciously).'

(50) On krutil ruljom.
he rotate wheel-INSTR
'He rotated the wheel unconsciously.'

Similarly, an example of a feature of the strong object (animacy) that is encoded on the subject is found in Pitjantjantjara (Rose 1996; Kittilä 2002):

(51) Ilyatjari-lu pony tati-nu.
Ilyatjari-ERG pony climbed
'Ilyatjari mounted the pony.'

(52) Nyantju wala winki puli tati-nu.
horse very quickly hill climbed
'The horse climbed the hill flat out.'

Thus, a marked semantic feature of the object (the object is animate) is not marked on the object but instead it is marked through ergative case on the subject.

In the examples above from Russian and Pitjantjantjara it seems that IDENTIFY is violated in order to satisfy PAIP. In Russian, which is a nominative-accusative language, non-volitionality of the subject is encoded on the object in order to leave the unmarked nominative subject unmarked. In Pitjantjantjara, which is an ergative language, it is the animacy of the object that gets encoded on the subject in order to leave the unmarked absolutive object untroubled. In those two cases we could argue that PAIP outranks IDENTIFY. But in fact, IDENTIFY is not really violated here. In Russian, the alternation between a weak and a strong subject *is* identified by case after all, albeit a case alternation on the object. Similarly, in Pitjantjantara the strong object is marked, but by case marking the subject. Malchukov (to appear) argues that IDENTIFY is satisfied after all, even by the marking of the 'wrong' argument. We therefore assume that in these languages PAIP outranks a specific constraint such as IDENTIFY (A/ERG) or IDENTIFY (P/ACC) which universally outranks a more general version of IDENTIFY (Woolford 2001). While satisfaction of PAIP in these cases necessarily implies the violation of the specific constraint IDENTIFY (A/ERG) or IDENTIFY (P/ACC) the more general version of IDENTIFY can still be satisfied by identification of the relevant features on the wrong arguments.

## 6. Conclusions

We have distinguished two main strategies of case marking, DISTINGUISHABILITY and IDENTIFY, and argued that they both play a role in the languages of the world. DISTINGUISHABILITY requires case marking to *disambiguate*, i.e. to distinguish between the arguments of a transitive clause. IDENTIFY is a strategy that requires case marking for the *identification* of specific semantic-pragmatic information. In this article we have investigated the two strategies of case marking and analysed them as violable constraints. We have shown that they converge in the domain of differential object marking whereas they diverge in the domain of differential subject marking. We have analysed these patterns within a bidirectional OT approach. We have also argued that while asymmetrical differential case marking can sometimes be explained by DISTINGUISHABILITY, symmetrical differential case marking must always be due to IDENTIFY. The variety of patterns that we predict find across languages is predicted on the basis of these two case marking constraints in relation to a principle of economy. Finally, we have further specified the economy constraint that interacts with DISTINGUISHABILITY and IDENTIFY, in order to account for attested case-voice correlations as well as certain instances of case marking of the 'wrong' argument.

## References

Aissen, J. (1999). "Markedness and subject choice in optimality theory." *Natural Language and Linguistic Theory* 17: 673-711.

Aissen, J. (2003). "Differential Object Marking: Iconicity vs. Economy." *Natural Language and Linguistic Theory* 21: 435-483.

Asher, R. E. and T. C. Kumari (1997). *Malayalam*. London/New York, Routledge.

Bhat, D. N. S. and M. S. Ningomba (1997). *Manipuri grammar*. München/Newcastle, Lincom.

Bittner, M. (1988). *Canonical and noncanonical argument expressions*, University of Texas, Austin.

Blutner, R. (2000). "Some aspects of optimality in natural language interpretation." *Journal of Semantics* 17: 189-216.

Bobaljik, J. D. (1993). Nominally Absolutive is not Absolutely Nominative. *Proceedings of the Eleventh West Coast Conference on Formal Linguistics*. J. Mead. Stanford, CSLI: 44-60.

Bossong, G. (1985). *Differentielle Objectmarkierung in den neuiranischen Sprachen*. Tübingen, Narr.

Bouchard, D. (2001). "The concept of 'universal' and the Case of Japanese." *Lingua* 111: 247-277.

Butt, M. and T. H. King (1991). Semantic case in Urdu. *Papers from the 27th Regional Meeting of the Chicago Linguistic Society*. Chicago, CLS: 31-45.

Butt, M. and T. H. King (2003). Case systems: beyond structural distinctions. *New perspectives on case theory*. E. Brandner and H. Zinsmeister. Stanford, CSLI.

Butt, M. and T. H. King (2004). The status of case. *Clause structure in South-Asian languages*. V. Dayal and A. Mahajan. Dordrecht, Kluwer academic publishers.

Chomsky, N. (1981). *Lectures on Government and Binding. The Pisa Lectures*. Dordrecht, Foris Publications.

Comrie, B. (1989). *Language universals and linguistic typology*. Chicago, University of Chicago Press.

de Hoop, H. (1996). *Case configuration and noun phrase interpretation*. New York, Garland.

de Hoop, H. (1999). Optimal case assignment. *Linguistics in the Netherlands 1999*. R. van Bezooijen and R. Kager. Amsterdam, John Benjamins: 97-109.

de Hoop, H. and M. Lamers (to appear). Incremental distinguishability of subject and object. *Case, valency, and transitivity*. L. Kulikov, A. L. Malchukov and P. de Swart. Amsterdam, John Benjamins.

de Hoop, H. and B. Narasimhan (2005). Differential case-marking in Hindi. *Competition and variation in natural languages: the case for case*. M. Amberber and H. de Hoop. Oxford, Elsevier: 321-345.

de Hoop, H. and B. Narasimhan (submitted). Ergative case-marking in Hindi. *Differential Subject Marking*. H. de Hoop and P. de Swart. Dordrecht, Kluwer.

de Swart, P. (2003). The case mirror. Department of Linguistics. Nijmegen, Radboud University Nijmegen.

Dekker, P. and R. van Rooij (2000). "Bi-directional Optimality Theory: an application of Game Theory." *Journal of Semantics* 17: 217-247.

Dixon, R. M. W. (1979). "Ergativity." *Language* 55: 59-138.

Dowty, D. (1991). "Thematic proto-roles and argument selection." *Language* 67: 547-619.

Drossard, W. (1991). Kasusmarkierung: zur Zentralität und Peripherizität von Paritizipanten. *Partizipation: das sprachliche Erfassen von Sachverhalten*. H. Seiler and W. Premper. Tübingen, Narr: 447-480.

Enç, M. (1991). "The Semantics of Specificity." *Linguistic Inquiry* 22(1): 1-25.

Feldman, H. (1986). *A grammar of Awtuw*. Canberra, Australian National University.

Fry, J. (2001). *Ellipsis and 'wa'-marking in Japanese conversation*, PhD dissertation, Stanford University.

Hale, K. (1973). Person marking in Warlpiri. *A festschrift for Morris Halle*. S. Anderson and P. Kiparsky. New York, Holt et al.: 308-344.

Haspelmath, M. (1993). *A grammar of Lezgian*. Berlin, Mouton.

Hopper, P. J. and S. A. Thompson (1980). "Transtivity in grammar and discourse." *Language* 56: 251-299.

Horn, L. (1984). Towards a new taxonomy of pragmatic inference: Q-based and R-based implicature. *Meaning, form, and use in context: linguistic applications*. D. Schiffrin. Washington, Georgetown University Press: 11-42.

Keenan, E. (1978). Language variation and the logical structure of universal grammar. *Language universals*. H. Seiler. Tübingen, Günther Narr Verlag: 89-123.

Kibrik, A. E. (1985). Toward a typology of ergativity. *Grammar inside and outside the clause*. J. Nichols and A. Woodbury. Cambridge, Cambridge University Press: 268-324.

Kiparsky, P. (1998). Partitive case and aspect. *The projection of arguments: lexical and compositional factors*. M. Butt and W. Geuder. Stanford, CSLI: 265-307.

Kittilä, S. (2002). *Transitivity: towards a comprehensive typology*. Turku, Abo Akademis Tryckeri.

Klaiman, M. H. (1987). Mechanisms of ergativity in South Asia. *Studies in ergativity*. R. M. W. Dixon. Amsterdam, North Holland: 61-103.

Laka, I. (1993). Unergatives that assign ergative, unaccusatives that assign accusative. *Papers on case and agreement*. J. D. Bobaljik and C. Philips. Cambridge, Mass., MITWPL.

Lapolla, R. (2003). *A grammar of Qiang*. Berlin, Mouton de Gruyter.

Lee, K. (2002). "Nominative case marker deletion in spoken Japanese: an analysis from the perspective of information structure." *Journal of Pragmatics* 34: 683-709.

Legendre, G., W. Raymond, et al. (1993). An Optimality-Theoretic typology of case and grammatical voice systems. *Proceedings of the 19th meeting of the Berkeley Linguistics Society*. Berkeley.

Malchukov, A. L. (2005). Animacy and asymmetries in differential case marking. Paper presented at PIONIER Workshop on Animacy. Nijmegen.

Malchukov, A. L. (to appear). Transitivity parameters and transitivity alternations: constraining co-variation. *Case, valency and transitivity: a cross-linguistic perspective*. L. Kulikov, A. L. Malchukov and P. de Swart. Amsterdam, John Benjamins.

Mallinson, G. and B. Blake (1981). *Language Typology: Cross-cultural studies in syntax*. Amsterdam, North-Holland.

Marantz, A. (1991). Case and licensing. *Proceedings of the Eighth Eastern States Conference on Linguistics (ESCOL '91)*. G. Westphal, B. Ao and H.-R. Chae. Ithaca, NY, CLC Publications: 234-253.

Masica, C. (1991). *The Indo-Arian languages*. Cambridge, Cambridge University Press.

Mithun, M. (1991). "Active/agentive case marking and its motivations." *Language* 67: 510-546.

Mohanan, T. (1994). *Argument structure in Hindi*. Stanford, Stanford University.

Prince, A. and P. Smolensky (2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford, Blackwell.

Rose, D. (1996). Pitjantjatjara processes. An Australian experiential grammar. *Functional descriptions: theory and practice*. Hasan and e. al. Amsterdam, John Benjamins.

Scott, G. (1978). *The Fore language of Papua New Guinea*. Canberra, Pacific Linguistics.

Silverstein, M. (1976). Hierarchy of features and ergativity. *Grammatical categories in Australian languages*. R. M. W. Dixon. Canberra, Australian Institute of Aboriginal Studies: 112-171.

Song, J. J. (2001). *Linguistic Typology: Morphology and Syntax*. London, Longman.

Stiebels, B. (2000). Linker inventories, linking splits and lexical economy. *Lexicon in Focus*. B. Stiebels and D. Wunderlich. Berlin, Akademie Verlag: 211-245.

Tsunoda, T. (1981). "Split case-marking in verb types and tense/aspect/mood." *Linguistics* 19: 389-438.

Wali, K. and O. Koul (1997). *Kashmiri: a cognitive-descriptive grammar*. London, Routledge.

Woolford, E. (2001). Case patterns. *Optimality-theoretic syntax*. G. Legendre, J. Grimshaw and S. Vikner. Cambridge, Mass., MIT Press: 509-545.

Woolford, E. (2004). Differential subject marking at argument structure, syntax, and PF. Paper presented at the PIONIER workshop on Differential Subject Marking. Nijmegen.

[1] We are grateful to the Netherlands Organisation for Scientific Research, NWO, for financial support (grants no. 220-70-003, 051.02.070 and DN 30-609). Jonathan Bobaljik, Peter de Swart, and Joost Zwarts kindly commented upon an earlier version of the manuscript.