0. Preliminaries: Two-term Case Systems in General.

Two-term case systems (henceforth 2-case systems) constitute a cross-linguistically valid type many of whose properties are due to the necessity that the language divide the universal semantic space of case meanings just into two subparts.

Languages where 2-case systems occur:
- Asia: Indo-Iranian
- Africa: Berber, Cushitic, Nilotic, Ethio-Semitic
- America: Salish, Uto-Aztecan
- Other: Yimas (New Guinea), Nias (Austronesian, Sumatra), Rumanian, Old French, English

Principal parameters of variation among 2-case systems:
- the “case zone”: the range of functions covered in a particular language by cases proper (vs. adpositions etc.);
- the distribution of functions from the “case zone” among the two cases.

A general typology of 2-case systems according to these parameters:
1. “narrow” systems, where the “case zone” includes only the core semantico-syntactic relations (e. g., Amharic, [Titov 1991]);
2. “broad” systems, where the “case zone” includes both core and peripheral functions (the overwhelming majority);
   2.1. “distributing” systems, where both cases have core as well as peripheral functions (examples below);
   2.2. “dividing” systems, where (almost) all peripheral functions are attributed to a single case (usually Oblique), which may also have a core function (the overwhelming majority).
Other important parameters of variation are applicable to other case systems as well; the one to be considered here in greater detail: alignment (accusative, ergative etc.) of the core functions and actance variations (“split case marking”, cf. [Moravcsik 1978a], [Comrie 1978], [Dixon 1979, 1994], [DeLancey 1981], [Tsunoda 1981], [Lazard 1994]).

2-case systems display various dominant alignments:
- accusative: Old-French, Berber, Nilotic, Uto-Aztecan;
- ergative: Nias;
- neutral: Romanian, Yimas, Salish;

However, most languages have coding alternations for the core relations that obscure the dominant alignment and, as I will argue below, even make the notion itself vacuous. The alternations especially prominent in the Indo-Iranian languages are the Tense-Aspect (TAM) Split and the Differential Object Marking (DOM, [Comrie 1979], [Lazard 1984, 1994], [Bossong 1985], cf. a recent formal account in [Aissen 2003]).


The 2-case systems attested in the Indo-Iranian languages show great variability for the parameters of “case zone” and “distribution”:
- the “case zone” ranges from just core relations (e.g., Wakhi, [Paxalina 1975]) to a long range including many peripheral functions (e.g., Kurdish dialects, [McKenzie 1961] or some Dardic languages, [Edelman 1983]); the set of peripheral functions included into the “case zone” also differs substantially across languages;
- most of the languages belong to the “dividing” type; it is noteworthy that it is the Indo-Iranian group where almost all the languages with the “distributing” 2-case systems are attested.

However, the group is consistent in that most of its members have both the TAM-split and the DOM.


The Indo-Iranian 2-case systems originate from the reduced systems with many cases (for the detailed account of this process see [Rastorgueva (ed.) 1975], [Edelman 1990]); thus the formal complexity such as allomorphy, the cumulative expression of case and number, non-segmental markers of case, cf. Kurdish (Kurmanji dialect, [Bakaev 1966: 63 — 264])

<table>
<thead>
<tr>
<th></th>
<th>‘father’</th>
<th>‘wife’</th>
<th>‘flour’</th>
<th>‘brother’</th>
<th>‘miller’</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dir</strong></td>
<td>bav</td>
<td>ʒən</td>
<td>ar</td>
<td>bəɾa</td>
<td>aʃvan</td>
</tr>
<tr>
<td><strong>Obl</strong></td>
<td>bav-e</td>
<td>ʒən-e</td>
<td>er</td>
<td>bəɾe</td>
<td>aʃven</td>
</tr>
</tbody>
</table>
However, there are some cases of “secondary agglutination” due to diachronic reanalysis, cf. Khowar (Dardic, [Edelman 1983: 212]) where both -o and -ān are former cumulative suffixes:

<table>
<thead>
<tr>
<th></th>
<th>‘brother’</th>
<th>‘son’</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sg</td>
<td>Pl</td>
</tr>
<tr>
<td>Dir</td>
<td>brār</td>
<td>brār-gini</td>
</tr>
<tr>
<td>Obl</td>
<td>brār-o</td>
<td>brār-gini-ān</td>
</tr>
</tbody>
</table>

A 2-case system of a genuinely recent origin is attested only in the Northern Tati dialects ([Grjunberg 1963, 1966]), where the Oblique case marker -(r)æ is traced back to the postposition *=rādiy. That it is a new formation is clearly seen from the comparison of pronominal paradigms of Northern and Southern Tati dialects ([Grjunberg 1966: 288], [Yar-Shater 147 — 149]):

<table>
<thead>
<tr>
<th></th>
<th>Northern Tati</th>
<th>Southern Tati</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sg 1Sg 2Sg</td>
<td>Sg 1Sg 2Sg</td>
</tr>
<tr>
<td>Dir men</td>
<td>ty</td>
<td>az</td>
</tr>
<tr>
<td>Obl men-æ</td>
<td>ty-æ</td>
<td>cemæ(n)</td>
</tr>
</tbody>
</table>

The recent development of the Oblique case in Northern Tati has parallels in its distribution: it marks only the individuated U (henceforth U’) and sometimes Poss, whereas Southern Tati Oblique has a variety of functions.

3. Marking of the Core Relations: Global Constraints vs. Particular Argument Marking Rules.

A common view on different type of core argument alignment (accusative, ergative, neutral etc.), advocated by Comrie [1978], Dixon [1979, 1994], Kibrik [1979, 1992, 1997] is based on the assumption that the primary function of case-marking is to distinguish arguments from each other. Thus, those coding strategies which successfully discriminate between A and U (viz. accusative and ergative) are more “iconic” than those which do not make such a distinction (viz. neutral). Under this assumption some logically possible strategies have been claimed to be impossible and non-occurring due to their “unnaturalness”, viz. the following one I call “quasi-neutral”: S vs. A + U (it distinguishes the intransitive argument from the transitive ones while failing to differentiate among the latter).

Also, Comrie [1979] claims that the main function of DOM is to distinguish those Us from As that are animate and definite and therefore have more properties of the prototypical As.

However, these views are subject to objections of theoretical (cf. [Moravcsik 1978b]) as well as of factual nature. If they have been correct, there would have been no such examples as the following:

Yaghnobi ([Xromov 1987: 663], [Bossong 1985: 18]):

(1) šumox tim čupīn iyot
    you.Dir also shepherds.Dir.Sg were.2Pl
‘Have you (S) been shepherds, too?’

(2) \( \text{man} \quad \text{awi} \quad \text{adih-} \quad \text{im} \)
I.Obl he.Obl hit- 1Sg

‘I (A) hit (U) him’

Hindi, [Junghare 1983: 43]:

(3) \( \text{ek} \quad \text{gā́v=} \quad \text{mē} \quad \text{ek} \quad \text{kisā́n} \quad \text{thā} \)
one village=in one farmer.Dir was

‘In one village there was a peasant (S)’

(4) \( \text{ek} \quad \text{din} \quad \text{kisā́n=} \quad \text{ne} \quad \text{laṛkī=} \quad \text{ko} \quad \text{khet=} \quad \text{mē} \quad \text{bhejā} \)
one day farmer.Obl=Agt daughter.Obl=Obj farm= in sent

‘One day the peasant (A) sent his daughter (U) to the farm’

The Hindi and Yaghnobi examples seem to be entirely different from the point of view of the ‘discriminatory’ theory of case-marking: while (2) fails to distinguish A from U (thus exemplifying the allegedly non-existent “quasi-neutral” alignment; cf. [Payne 1979, 1980]), (4) “over-marks” them, which results in a “three-way” alignment. Also, both (2) and (4) contradict Comrie’s generalization: in (2) both core arguments are definite and animate, but nevertheless their case-marking is similar (that they are distinguished by verb agreement is not important here); in (4) it would have sufficed to mark the A leaving the U unmarked, since the alleged “*A=U” constraint would have been satisfied, however it is ungrammatical not to mark definite and animate Us in Hindi even in the Past tense where the A is marked as well.

The natural way to account for these patterns of case-marking (which are quite widespread among the languages in question, see [Bossong 1985]) is to regard them as arising through the application of different rules marking particular arguments. Thus, in Yaghnobi there is a rule that requires that A be marked by Obl in the Past tense (“Past-A”); and also a rule that marks with Obl the individualized Us (“Indiv-U”). These rules apply independently of each other which follows from the very fact that they are motivated by different functional factors. The corresponding rules in Hindi differ only in the following respects: (i) the contexts for their application are slightly different (e. g., Perfective aspect instead of Past tense for the “Past-A”); (ii) the case-markers they introduce are adpositions rather than suffixes, and they are different rather than similar.

Such an approach to case-marking has the following advantages over the “discriminatory” approach:

• it does not presuppose that any arguments in a clause must be distinguished, thus leaving the responsibility for justifying such claims for the advocates of the “discriminatory” approach;
• it does not state that the Yaghnobi and Hindi patterns are in any way “unnatural”, which the “discriminatory” approach does;
• it does not miss the generalization that the Yaghnobi and Hindi examples, contrary to their superficial differences, are motivated by the same functional factors;
• it is based on the assumption that cases and case-markers have genuine meanings, and are not merely diacritics for distinguishing arguments (cf. [Wierzbicka 1981, 1983]).

There are also other minor patterns of the case-marking of the core relations that speak in favour of the approach presented here. Thus, in Wakhi in addition to “Past-A” and “Indiv-U” which mark both arguments with Obl, there is a rule which requires Obl on the S argument, too, just in case it is a pronoun. Thus we have the following examples [Paxalina 1975: 54] which demonstrate the ‘neutral’ alignment:

(5) \[ \text{maž} \ yān \ zaqlay \ tu \]

I.Obl then small was

‘I (S) was a small child then’

(6) \[ \text{taw} \ maž \ fand δṣt \]

you.Obl I.Obl deceived

‘You (A) have deceived me (U)’

Another interesting feature of Indo-Iranian is absence of clear cases of the so called ‘nominal split’, where pronouns follow the accusative and nouns the ergative alignment (cf. [Silverstein 1976], [Moravcsik 1978], [Comrie 1979], [Dixon 1979, 1994]). This fact is perhaps partly due to the TAM-split, and perhaps to some diachronic matters. Taleshi, however, presents a subtle pattern of the interaction of TAM and nominal splits: personal pronouns in this language follow the accusative alignment regardless of tense, cf. the following examples [Pirejko 1966: 305 — 308]:

(7) \[ \text{hъrdan-} *(i) \ \text{kitob} \ \text{dъrya=} \ \text{še} \]

child- Obl/*Dir book.Dir tore= 3Sg

‘The child (A) tore a book (U)’

(8) \[ \text{аž}/*mъni} \ \text{gatъ=} \ \text{me} \]

I.Dir/*Obl took= 1Sg

‘I (A) took <it (U)>’

(9) \[ \text{mъni}/*аž} \ \text{ba} \ \text{dъs dano=} \ \text{še} \]

I.Obl/*Dir deceived= 3Sg

‘He (A) deceived me (U)’

However, in order to capture such a pattern, in the approach advocated here it is necessary only to restrict the application of “Past-A” to nouns, leaving other rules unchanged (for instance, “Indiv-U” is responsible for the Obl marking of pronominal U in the Past tense as well as in the Present).
Another piece of evidence for the proposed approach to argument marking give the instances of “non-canonical” marking, when certain verbs or constructions mark their core arguments differently from what is required by the “global” rules. Consider, for example, the “indirect subject construction”, universally characterized as having non-volitional meaning ([Klaiman 1981], [Onishi 2001]; here it is obligation), attested in the Nuristani language Kati. It consists of a non-finite form of the verb, a copula with the “impersonal” 3Sg agreement, and of the full range of the verb’s arguments, of which the S/A must receive the Oblique case ([Grjunberg 1980: 184, 257]):

(10) \(dalk’eř \ ye \ Ktivī \ estuk= \ asi\)

    tomorrow I.Obl Ktivi.Dir go.Ger= Cop.3Sg

    ‘Tomorrow I must go to Ktivi’

(11) \(ye \ davo \ yustuk= \ asi\)

    I.Obl medicine.Dir eat.Ger= Cop.3Sg

    ‘I (A) must take the medicine (U)’

Here the only way to capture the “non-canonical” argument marking is by postulating a construction-specific case-marking rule; this, however, is in line with the approach I present, where actually all case-marking rules (in particular “Past-A”) may be regarded as construction-specific (although motivated by “global” functional principles).


2-case systems in the Indo-Iranian languages reveal great diversity with regard to both parameters of the typology outlined above. It is important that much of this variation does not strictly correspond to any genetic or areal divisions among the languages surveyed. The main dividing line lies perhaps between the “core” Iranian (and also Dardic and Nuristani) languages, such as Kurdish, Pashto, Kati etc. and the “core” Indian languages, such as Hindi and especially Nepali (with Braj [Liperovskij 1988] and Avadhi [Liperovskij 1997] as exceptions), which differ profoundly in how they treat their case systems. Indian languages in general tend to supplant it by a newly grammaticalized system of postpositions, which have been introduced even into the domain of core argument marking (see [Zograf 1976] for an extensive discussion). The case paradigm is considerably reduced; for instance, Nepali merely distinguishes DirSg from a common form for OblSg and Pl, and with a class of masculine nouns only ([Koroljov 1965: 46], ‘son’):

<table>
<thead>
<tr>
<th></th>
<th>Sg</th>
<th>Pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dir</td>
<td>choro</td>
<td></td>
</tr>
<tr>
<td>Obl</td>
<td>chorā</td>
<td></td>
</tr>
</tbody>
</table>

The functional load of cases in the Indian languages is also reduced: whereas in Hindi Obl still can mark the goal of motion when used without postpositions ([Mohanan 1994: 88]), in Nepali and
other languages it is used only before postpositions, and somewhere (e.g. Gujarati [Savel’jeva 1965: 24 — 25] and especially Tsigani [Ventcel’ 1964: 52 — 74]) it is even better to regard those as full-fledged case suffixes.

Other languages, on the contrary, make use of their two cases to a much greater extent (although here some exceptions also exist). It is common among the Iranian, Dardic and Nuristani languages to use Obl at least in the functions of A (in the Past tense), U’ (usually in all tenses), Adr and Poss. This association of functions is a peculiar feature of these languages, mentioned already in [Bossong 1985: 116 — 121], who even postulates and implicational universal:

if in a given language A and U’ are marked by the same case then this case also marks either Adr or Poss

This generalization is violated only by Pashto where Obl seems not to have the last two functions (both marked by prepositions, [Grjunberg 1987]). Such a non-trivial (and presumably non-attested anywhere else) pattern of polysemy of a case is best explained in diachronic terms (see [Rastorgueva (ed.) 1975: 180 — 191]): the Adr and Poss functions are inherited by the Iranian Obl from its ancient prototype GenDat; the Agentive function arose later when a “passive” construction was generalized (for different opinions on this issue see [Benveniste 1952], [Pirejko 1968, 1979], [Anderson 1977], [Klaiman 1978, 1987], [Butt 2001]); and the U’ function arose through an extension of the Adr, which is a common phenomenon ([Bossong 1985: 109 — 111; Lehmann 1995: 97 — 105]).

However, some languages, for instance, the Pamir Munji [Grjunberg 1972] employ pre- or postpositions to mark U’, which is a relatively recent development (cf. what was said above about the two groups of Tati dialects). Some Dardic languages, e.g. Shina ([Anderson 1977], [Edelman 1983: 283; 295 — 296], mark A with a postposition, too. Keeping in mind that the Central Iranian languages (Persian, Tajik, etc.) have lost their original case systems altogether and use postpositions for marking U’ ([Bossong 1985]), it is possible to outline a global tendency of transfer from a 2-case system to a more complex multi-layer system where both old and newly grammaticalized case markers encode some core functions. This diachronic development has reached farthest in the Indo-Aryan languages, while some Iranian and Dardic languages show intermediate stages or even stay at the beginning of the process.

Turning to the typological parameters themselves, it is necessary to say that although all three main types of system (“narrow”, “distributing” and “dividing”) are attested in the Indo-Iranian languages, their distribution follows the general pattern: “narrow” and “distributing” systems are rare, while “dividing” constitute the major type. Wakhi has a typical “narrow” system, marking only the core relations by the cases proper, while employing various adpositions for all peripheral functions. “Broad” systems vary with respect to the number of functions represented: some Indian
languages, such as Panjabi [Cummings, Bailey 1925], select a single peripheral function (such as location in space or time) for their Obl, Iranian Taleshi [Pirejko 1966] and Shugnani [Paxalina 1969] have Poss as the only non-core relation they mark with a bare case; other languages possess a much richer “case zone”, which may include Poss and Adr as well as both locative relations (Goal and Location). Such “broadest” systems are attested in some Dardic and Nuristani languages (Edelman 1983), in Yaghnobi [Xromov 1987] and Kurdish ([McKenzie 1961]).

One of the most interesting features of the Indo-Iranian 2-case systems are the “distributing” systems, which to my knowledge are extremely rare if not non-existent outside this region. Let us compare the range of functions of Dir and Obl in a typical “dividing” system (Kurdish dialect Mukre [McKenzie 1961]) and in a “dividing” system (Nuristani Kati, [Grjunberg 1980]):

<table>
<thead>
<tr>
<th></th>
<th>Mukre</th>
<th>Kati</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dir</strong></td>
<td>Pred, S, A, U</td>
<td>Pred, S, A, U, Goal, Loc</td>
</tr>
<tr>
<td><strong>Obl</strong></td>
<td>A, U’, Adr, Poss, Goal, Loc, Temp</td>
<td>A, U’, Adr, Poss</td>
</tr>
</tbody>
</table>

It may be clearly seen that while in a “dividing” system it is possible to say that Dir is a “core” case while Obl is a “peripheral” case, such a distinction is very difficult to make in a “distributing” system where both cases have functions from both domains. Consider other “distributing” systems:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dir</strong></td>
<td>Pred, S, A, U, Goal, Temp</td>
<td>Pred, S, A, U, Goal</td>
<td>Pred, S, A, U, Goal, Loc, Temp</td>
</tr>
<tr>
<td><strong>Obl</strong></td>
<td>U’, Adr, Loc, Temp</td>
<td>A, U’, Adr, Poss, Loc, Temp</td>
<td>A, U’, Adr, Poss</td>
</tr>
</tbody>
</table>

A common pattern arises: in “distributing” systems it is the Goal function which is always aligned to the core relations, whereas Adr and Poss are uniformly treated as peripheral; other roles (Loc, Temp) seem to “float” somewhere in between, being marked by the Dir in some languages and the Obl in other. This tendency may again be explained in diachronic terms: the Goal function was marked by the Proto-Indo-Iranian Acc, which collapsed with Nom and gave rise to Dir of modern languages ([Rastorgueva (ed.) 1975: 128]), while Adr and Poss have been among the functions of the former peripheral cases whereof the modern Obl is the reflex (see above). Loc and Temp are, quite probably, innovative functions, just as A and U’. If such an account really holds, then the overwhelming majority of the Iranian 2-case systems, which belongs to the “dividing” type, must have undergone diachronic restructuring and redistribution of functions; it is among the primary goals of further research to find factual evidence for such changes. If “distributing” systems are found to be diachronically unstable in comparison to “dividing” ones, it may serve as an explanation for the fact that it is the latter type which is the commonest in the languages of the world that have 2-case systems.
5. Conclusions.

I do hope to have shown that 2-case systems constitute an important phenomenon by themselves, revealing patterns of typological variation relevant for case systems in general. The Indo-Iranian is of particular interest here, since this is the group of languages where 2-case systems show greatest diversity coupled with a relatively well-documented history. The data presented by these languages bear important evidence on such issues as diachronic origins of 2-case systems or functional rationale of the types of core argument marking.

To sum up the discussion, I re-state the main points of my presentation:

- different types of marking of core arguments which co-exist in a language are best viewed as arising through complex interaction of rules responsible for case marking of individual actants, rather than from the need to distinguish them; such an account, for example, correctly predicts both “quasi-neutral” and “three-way” techniques in languages with both DOM and TAM-split;
- 2-case systems in Indo-Iranian show a diachronic tendency for being supplanted by a newly grammaticalized layer of adpositions, although in different languages this process has advanced to different degrees;
- the Indo-Iranian languages conform to cross-linguistic trends, favouring “broad” and “dividing” systems against “narrow” and “distributing” ones; the latter seem to reflect the more conservative stages of case systems, retaining the ancient distribution of case functions.

Abbreviations


References


