

On the scope of the referential hierarchy in the typology of grammatical relations

Balthasar Bickel
University of Leipzig

1. Introduction

In the late seventies, Bernard Comrie was one of the first linguists to explore the effects of the referential hierarchy (RH) on the distribution of grammatical relations (GRs). The referential hierarchy is also known in the literature as the animacy, empathy or indexibility hierarchy and ranks speech act participants (i.e. first and second person) above third persons, animates above inanimates, or more topical referents above less topical referents. Depending on the language, the hierarchy is sometimes extended by analogy to rankings of possessors above possessees, singulars above plurals, or other notions. In his 1981 textbook, Comrie analyzed RH effects as explaining (a) differential case (or adposition) marking of transitive subject ('A') noun phrases in low RH positions (e.g. inanimate or third person) and of object ('P') noun phrases in high RH positions (e.g. animate or first or second person), and (b) hierarchical verb agreement coupled with a direct vs. inverse distinction, as in Algonquian (Comrie 1981: Chapter 6). The relevant effects are captured by what I call here the Marking-based RH Hypothesis:

(1) *The Marking-based RH Hypothesis:*

- a. For A arguments, the odds for zero case-marking correlate positively with the rank of the argument on the referential hierarchy.
- b. For P arguments, the odds for zero case-marking correlate negatively with the rank of the argument on the referential hierarchy.
- c. For transitive direction marking, the odds for zero direction-marking are higher for argument scenarios where A ranks higher than P ('direct' scenarios) than for scenarios where P ranks higher than A ('inverse' scenarios).

Because nominatives and absolutes tend to have zero exponence, the most common case systems that are consonant with the hypothesis are those that reserve accusative alignment for upper and ergative alignment for lower segments of the referential hierarchy. In such systems, high-ranking A arguments receive the same zero-marking as S (intransitive subject) arguments, whereas low-ranking A arguments are assigned an overt marker ('ergative'); high-ranking P arguments receive an overt marker ('accusative') while low-ranking P arguments receive the same zero-marking as S.

The link with alignment typology motivates a more general approach, where the referential hierarchy is expected to determine whether GRs are aligned ergatively or accusatively regardless of whether the alignment manifests itself in case or any other construction, such as verb agreement or word order, and regardless of the concrete morphological exponence pattern (zero vs nonzero marking). This

generalization, first proposed explicitly by Silverstein (1976:112f), is captured by what I call here the Alignment-based RH Hypothesis:

(2) *The Alignment-based RH Hypothesis:*

- a. For a given grammatical relation in any kind of construction, the odds for accusative alignment correlate positively with the rank of the argument in that relation on the referential hierarchy.
- b. For a given grammatical relation in any kind of construction, the odds for ergative alignment correlate negatively with the rank of the argument in that relation on the referential hierarchy.

In this chapter I submit these two hypotheses to empirical testing against typological data. I first test two predictions from the Alignment-based RH Hypothesis: the prediction that the distribution of alignment patterns in verb agreement reflects the referential hierarchy (Section 2), and the prediction that the referential hierarchy determines case alignment regardless of actual marking patterns (Section 3). I show that the Alignment-based RH Hypothesis fails with regard to verb agreement, and that it has no better support than the Marking-based RH Hypothesis with regard to case-marking. However, unlike the Alignment-based RH Hypothesis, the Marking-based RH Hypothesis makes no predictions on other grammatical relations than those identified by case or inverse marking, and in Section 4 I discuss possible alternative accounts that may capture the full range of RH impacts on grammatical relations. Section 5 summarizes the results.

2. The referential hierarchy and verb agreement

According to the Alignment-based RH Hypothesis, the referential hierarchy is expected to determine alignment patterns in any kind of grammatical relation. I test this prediction on data from verb agreement. The hypothesis predicts that if agreement morphology shows an alignment split, accusative alignment is expected to be limited to higher RH positions and ergative alignment is expected to be limited to lower RH positions. In order to test this prediction, I searched Bakker & Siewierska's (2006) verb agreement database ($N = 402$) for languages coded as having any kind of RH-sensitive alignment splits, excluding cliticized pronouns (which are expected to behave like pronouns).¹ Table 1 summarizes the results.

INSERT TABLE 1 ABOUT HERE

The Alignment Hypothesis only makes predictions about the location of accusative and ergative alignments on the hierarchy. Tripartite alignments, where S, A, and P are all treated differently, share properties of both alignments: there is a distinct form for A as in ergative alignment, and there is a distinct form for P as in accusative alignment. This excludes Ainu, Chácobo, Comox, Kamass, and Yimas from the testing ground of the hypothesis. One of the remaining languages, Paumarí, distributes alignment types on non-contiguous segments of the hierarchy (third

person plural, but not third person singular or dual, pattern with first and second person). Under one reading of the Alignment Hypothesis, this runs counter the prediction. A more friendly reading of the hypothesis, however, restricts the predictions to languages with ‘clean’ splits, assuming that non-contiguous splits are due to other factors of paradigm development. I follow this reasoning here and exclude languages like Paumarí from the testing ground. What is left, then, are six languages (Seri, Maung, Nez Perce, Washo, Tepehua, and Maricopa). All of these show accusative alignment on higher and ergative alignment on lower ranks of the RH, and this at first sight supports the Alignment Hypothesis.

However, the support is very weak for but one case (Seri), indicated in Table 1 by question marks. For Maung, Curnow (1999) has offered an alternative analysis that does not involve any alignment pattern in the system, and his analysis has better data coverage. Four other languages in Table 1 are marked by ‘?’ because they contain traces of contradictory evidence: Nez Perce has plural prefixes (*pe-* ‘1/2/3pS/A’ and *nees-* ‘1/2/3pP’) which align accusatively and are also used with third person arguments; in addition there is a prefix *e-* dedicated to third person singular P arguments (Crook 1999). Both these facts suggest that the third person shows at least some traces of accusative alignment, in conflict with the prediction. Similarly, in Washo, there is an intransitive third person marker *?-* ‘3S’, which also covers 3>1 (though not 3>2 and 3>3) scenarios (Jacobsen 1964). This again represents a minor pattern of accusative for third persons in conflict with the prediction. Tepehua, too, has traces of accusative alignment in the third person: while third person singular has neutral alignment as a result of zero marking, plural arguments trigger *ta-* when in S or A function and *lak-* when in P function (Watters 1988). In Maricopa, the evidence for accusative alignment in the agreement morphology rests on the first person prefix *?-* which occurs in both intransitive and 1>3 transitive forms. However, the same prefix also occurs in 2>1 scenarios, where it marks a P argument — in fact it even disambiguates the role of another first person marker (*ny-*) which is role-neutral: in *?-ny-m-aham-m* [1-1-2-hit-REAL] ‘you hit me’, *?-* signals that *ny-* ‘1’ is in P role. Without the prefix, *ny-* covers both A and P roles, as in *ny-aham-m* [1-hit-REAL] ‘I hit you’ or ‘he hit me’. Second person is marked by a role-neutral prefix and third person is zero-marked in all roles (Gordon 1986).²

If the Alignment-based RH Hypothesis is a genuine universal of human language, we expect it to have strong and frequent effects on the way alignment is split in verb agreement. The present findings, however, suggest that only one single language (Seri) supports the hypothesis — and this only insofar as accusative alignment for first person plural and second person (singular and plural) is opposed to zero-marked agreement for third persons in any role. The zero exponence could also be analyzed as mere absence of agreement, and this would further weaken the relevance of Seri for the hypothesis. In return and in support of the Alignment Hypothesis, however, one could add Chinook, one of the languages on which Silverstein based the formulation of the hypothesis in the first place. Further research might detect a couple of more cases, or remove a couple of question marks in Table 1, or both. But in any of these events, half a dozen languages out of 400 is a very narrow testing ground for genuine universals, and not much can be inferred from such a small number.

Despite this small number, one could still note that Table 1 contains no straightforward counterexample to Silverstein’s predictions — i.e. there is no language with ergative alignment on higher and accusative or neutral alignment on lower RH positions. Yet this may just as well be chance:³ even if there were seven

languages supporting the prediction (including those with question marks and also Chinook), and none contradicting it, such a score (7:0) only reaches a .015 probability level on a χ^2 -based one-sample randomization test (Janssen et al. 2006), i.e. there is only weak evidence to reject the null hypothesis of a chance distribution. The actually attested score of relatively clear cases (2:0) is not significant at all. And if one were to discover only one clear counterexample, i.e. a language that distributes ergative and accusative alignment in a way exactly opposite to Seri or Chinook, the evidence for rejecting the null hypothesis of chance distribution would dissipate even under the best-case scenario for the hypothesis, where all question marks in Table 1 were proved wrong, or one were to discover an equal number of other languages in line with the hypothesis (a 7:1 score, with an associated p -level of .077). In fact, field research over the past few years has unearthed robust counterexamples.

These counterexamples come from the Kiranti branch of Sino-Tibetan (Bickel 2000, Siewierska 2004). The basic pattern found in a number of languages of this family is shown in Table 2.

INSERT TABLE 2 ABOUT HERE

Here, the first person, or at least the first person singular, aligns ergatively, the second person is coded neutrally, and the third person aligns accusatively — i.e. the exact opposite of what is expected from the Alignment-based RH Hypothesis. Table 3 exemplifies the pattern in Puma, a language of the Southern subgroup of Central Kiranti (Bickel et al. 2007a). The Puma verb agrees with S, and in transitives, with both A and P arguments.

INSERT TABLE 3 ABOUT HERE

While many person-marking affix strings are specialized for specific scenarios defined by both A and P roles (represented as ‘>’ in Table 3), there is a consistent pattern of ergative alignment in the first person singular and plural. This pattern is identified by dark grey shading in the table. Further down the referential hierarchy, and with first person dual arguments, we find neutral (no shading) or accusative alignment (light grey shading). To be sure, in most of these cases there are also some counter-examples, but they all involve specific markers selected by global argument marking rules in the sense of Silverstein 1976 (a typical feature of Kiranti morphology in general: see Bickel 1995; Bickel et al. 2007b). For example, there is dedicated A-marking for third person singular arguments (p_{Λ} -), but it is limited to forms with concomitant first person P-arguments. The more general pattern aligns A with S marking in the third person.

The forms in (3) and (4) illustrate the contrast between first and third person alignment with past indicative forms of the intransitive verb *puks-* ‘to go’ and the transitive verb *pukq-* ‘to take’; the relevant affixes are underlined.⁴

- (3) a. 1sS *puks-oŋ*
 go-1sS/P.PST
 ‘I went.’
- b. 1sP *p Λ -pukq-oŋ*
 3S/A-take-1sS/P.PST
 ‘S/he took me.’

- c. 1sP *tʌ-pukɔ-ɔŋ*
2-take-1sS/P.PST
'You took me.'
- d. 1sA *pukɔ-u-ŋ*
take-3sP-1sA
'I took him/her.'
- e. 1sA *puk-na-a*
take-1>2-PST
'I take you.'
- (4) a. 3sS *∅-puks-a*
3sS/A-go-PST
'S/he went'
- b. 3sP *∅-pukɔ-i*
3sS/A-take-3sP
'S/he took him/her.'
- c. 3sP *tʌ-pukɔ-i*
2-go-3sP
'You took him/her.';
- d. 3sP *pukɔ-u-ŋ*
took-3sP-1sA
'I took him/her.'
- e. 3sA *∅-pukɔ-i*
3sS/A-take-3sP
'S/he took him/her.'
- f. 3sA *∅-tʌ-pukɔ-a*
3sS/A-2-take-PST
'S/he took you.'
- g. 3sA *pʌ-pukɔ-ɔŋ*
3S/A-take-1sS/P.PST
'S/he took me.'

As can be seen when comparing (3a) with (3b-c), the first person is coded by the same suffix *-ɔŋ* in both S and P functions. When in A function, the first person is marked differently: by *-ŋ* when the P is third person (3d) and as part of the portemanteau morpheme *-na* when the P is second person (3e). In contrast to this, third person shows accusatively-aligned agreement: whereas in S function, third person singular has zero exponence (4a), in P function it is marked by *-i* (4b-c) or its pre-nasal allomorph *-u* (4d). In A function, third person singular is coded like S, i.e. by zero, when the P argument is third or second person (4e-f).

If the P argument is first person, as in (4g), a third person A argument is coded (as noted above) by the prefix *pʌ-*, which at first sight suggests a tripartite pattern (*∅* for S, *-i* for P, and *pʌ-* for A), but a closer look at the paradigm shows that the overall distribution of *pʌ-* follows again an accusative pattern: the prefix also marks third

person S in the dual (though not in the singular and the plural), but it is never found marking any argument in P function:

- (5) a. \emptyset -*puks-a*.
 3sS/A-go-PST
 ‘S/he went.’
- b. $p\Lambda$ -*puks-a-ci*.
 3S/A-go-PST-d
 ‘They (two) went.’
- c. $m\Lambda$ -*puks-a*.
 3pS/A-go-PST
 ‘They (three or more) went.’

This pattern of a ‘reverse ergativity split’ is frequent in Kiranti languages, with a higher concentration in the Central and Eastern branches (which are likely to form a unified taxon at some very early level) than in the Western branch of the family. Table 3 shows the distribution of reverse splits according to genealogical subgroups established by regular sound change (based on work by van Driem 1990, 2001, Michailovsky 1994, Opgenort 2004b, 2004a, and my own ongoing research).

INSERT TABLE 4 ABOUT HERE

On one possible count, Table 4 provides twelve cases of a reverse split. Earlier we found that there are no more than two cases of a RH-predicted split. This distribution would suggest a trend in the opposite direction from what is predicted by the Alignment-based RH Hypothesis, or, more likely, that plain chance is at work. However, it is possible, perhaps indeed likely, that the twelve languages share a reverse split because they inherited it from Proto-Kiranti: although only about two thirds of the entire family is at present sufficiently documented and represented in Table 4, reverse splits appear to be the default pattern, and some exceptions demonstrably represent innovations (e.g. in Belhare, where first person P markers are recent developments of generic nouns, cf. Ebert 1991, Bickel and Gaenszle 2005). But even under this scenario, the Kiranti facts speak against a universal principle disfavoring reverse splits (the Alignment-based RH Hypothesis): even if Kiranti counts only as a single genealogically independent case, the total score (two cases following the hypothesis, one contradicting it) would not suggest any significant trend. But the absence of statistical evidence can never demonstrate the absence of an underlying principle, especially if the sample size is so extremely small as it is here.

Stronger evidence against the Alignment-based RH Hypothesis comes from the following: the Kiranti family is a fairly heterogeneous group of about 30 languages, and even on the most conservative estimate, the family is as old as Germanic. Therefore, if it is inherited, the reverse split pattern would need to have been diachronically stable to a degree that is unexpected if there was a strong universal principle acting against it — at each generation, for a few thousand years. Now, the survival of patterns that are otherwise disfavored by universal principles are sometimes strengthened by areal connections (Nichols 2003). But this is an unlikely scenario in the Kiranti context. First, the distribution of reverse ergative splits is

not geographically contiguous: the reverse-split languages Bahing and Hayu, for example, are completely surrounded by languages lacking the pattern (Jero, Wambule, Dumi, Camling, Thulung, and Nepali, the Indo-Aryan lingua franca of the region).⁵ Second, the distribution of ergativity splits cross-cuts the few sociological units that could arguably bring about some areal integration and language contact: one such unit that is well-established ethnographically by intermarriage patterns, a shared mythology and a traditional ethnonym (*Rāi*; cf. Gaenszle 1991: Chapter 3.2; Gaenszle in press) contains both languages with reverse splits (e.g. Puma) and languages without (e.g. Camling); and some languages with splits are part of the unit (e.g. Puma), while others are not (e.g. Limbu).

Whether the Kiranti pattern consists of several independent cases of parallel innovations, or whether it represents a single but diachronically stable case (or a combination of these two possibilities), it undermines the evidence for a universal principle in verb agreement that would favor ergative alignment in the lower and accusative alignment in the upper parts of the referential hierarchy. And, as we have seen before, there is no good statistical evidence for such a principle on a world-wide scale.

3. The referential hierarchy and case marking

There is no empirical support for the Alignment-based RH Hypothesis because it already fails in one central domain outside case, viz. split alignment in verb agreement. But what is the empirical support in the domain of case marking? For those languages where nominatives or absolutes are zero-marked, the Alignment-based RH Hypothesis makes the same predictions as the Marking-based RH Hypothesis. For these cases of zero-marking we expect higher-ranking arguments to favor accusative and lower-ranking arguments to favor ergative alignment because such an alignment distribution ensures that higher-ranking As and lower-ranking Ps are zero-marked (being assigned ‘unmarked’ nominatives and absolutes, respectively), while lower-ranking As and higher-ranking Ps are marked by special case morphology (ergative and accusative or dative, respectively). But for languages with alignment splits and nonzero (‘marked’) nominatives or absolutes, the two hypotheses make opposite predictions: under the Alignment-based Hypothesis, a nonzero nominative behaves like an zero-marked nominative, and so we expect it to be favored by high-ranking arguments; the Marking-based Hypothesis, by contrast, predicts that the nonzero nominative is favored by lower-ranking A arguments. With regard to nonzero absolutes, the Alignment-based Hypothesis predicts that the absolute is favored by low-ranking arguments; the Marking-based hypothesis predicts that it is favored by high-ranking P arguments. Table 5 shows the predictions in detail.

INSERT TABLE 5 ABOUT HERE

There is no large-scale survey available that allows extensive testing of any of these hypotheses. What is possible to test, however, is one special effect on pronouns and nouns that is predicted by the hypothesis (Comrie 2005). First and second person pronouns always rank higher on the RH than lexical nouns (barring perhaps some lexical honorifics), and it is likely that also third person pronouns

more often rank higher than lexical nouns. In Comrie's (2005) sample of 172 languages there are 20 languages in line with the Marking-based hypothesis, and 2 in conflict with it (the Cushitic language Iraqw, where accusative alignment is limited to nouns, and pronouns are neutral; and the Austronesian language Chamorro, where ergative alignment is limited to pronouns, and nouns are neutral). The only language in the database with an alignment split and a nonzero nominative or absolutive is Middle Atlas Berber (Afroasiatic), and in this language, a nonzero nominative in accusative alignment is limited to lexical nouns; pronouns have no case-marking. This violates the prediction of the Alignment-based hypothesis, but it is a single instance, and nothing universal can be inferred from it. There are a few further languages with a differential use of a nonzero S/A case marker, but in the instances I am aware of, the distribution is at least in part governed by focus structure (e.g. in the Papuan languages Fore or Kâte, cf. Donohue and Donohue 1997, Suter 2006, respectively), and this requires nonzero marking on an argument regardless of what one would expect from its position on the referential hierarchy.

The 20:2 score for languages with zero-marked nominatives and absolutives allows slightly more inference to universal trends, but it still falls short of the sample size required to postulate robust universals. The score itself is unlikely to be due to chance but it is possible – indeed likely – that other factors than the RH also affect the distribution of case markers on nouns and pronouns: for example, Comrie (2005) notes that 4 of the supporting languages (i.e. 20%) are from Australia, and we cannot exclude therefore areal (or genealogical) confounding factors. Also specific etymologies and paradigm structures are often demonstrably relevant for the distribution on case markers. For example, if an ergative develops from an instrumental, a limitation of the ergative to low-ranking As is to be expected just because animate nouns may never have had an instrumental form to begin with (Garrett 1990). Specific etymologies can also work against the prediction from the RH: for example, an ergative system might survive in pronouns, which often are generally more conservative morphologically, while lexical NPs lose case or develop new accusative marking, and this might result in a distribution that reverses what is predicted by the RH – such as happened in a number of Dardic and other Indo-Aryan languages (Filimonova 2005).

Thus, when limited to the special case of pronouns vs. nouns, the Marking-based RH Hypothesis receives some, albeit not very strong quantitative support, and we cannot rule out that such factors as areal influence, genealogical stability and individual etymologies might in the end account for the distributional findings equally well or better. But the Marking-based RH Hypothesis reaches far beyond this special case of pronoun vs. noun declension, and there is a rich literature on differential subject and object marking demonstrating that in many languages from many different parts of the world, the referential hierarchy is directly responsible for the distribution of case markers in discourse: there are many languages like Nepali (Indo-Aryan), where the odds for overt ergative case-marking have a distributional peak for A arguments that are nontopical, nonspecific, or inanimate, and the odds for dative case have a distributional peak for P arguments that are definite, specific and animate (Bickel 2007; Pokharel 2054).

4. Beyond case marking

The preceding two sections suggest that overall the Marking-based RH hypothesis has better empirical support than the Alignment-based RH hypothesis. A likely reason for this is that the Marking-based Hypothesis can be reduced to standard Zipfian effects: we know from discourse studies that A arguments are more frequently topical, i.e. filled by referents higher on the hierarchy, while P arguments are more frequently NPs with referents lower on the hierarchy, especially with rhematic and new referents (see, e.g. DuBois et al. 2003 or Jäger 2007 for statistical evidence of this in various languages). As per Zipf's Law, more frequent patterns generally tend to be less overtly marked and therefore, higher-ranking As and lower-ranking Ps are more likely to be zero-marked than lower-ranking As and higher-ranking Ps.

However, once the Marking-based hypothesis is reduced to standard Zipfian effects, there is no inherent reason that it should be limited to case marking, and so one would expect its scope to extend to other kinds of grammatical relations. For example, one would expect a universal trend for nonzero agreement morphology to be split in such a way that A-agreement is limited to inanimate or third person arguments; or P-agreement to animate, or first and second person arguments. This prediction is explored by Siewierska (2004), who finds equivocal evidence. There is a substantial number of languages where P agreement is limited to animate or human referents (e.g. in the Papuan language Hua: Haiman 1980), or where it is favored by such referents (e.g. Swahili: Seidl and Dimitriadis 1997, among others). But Siewierska (2004: 150) also identifies quite a few languages with the opposite pattern, where verbs show P-agreement only with third persons but not with speech act participants, so that a full assessment of the prediction must await further statistical testing with careful control for genealogical and areal confounding factors. The same is true for the prediction on A-agreement. Although there are languages like English where nonzero agreement is limited to third person A (and S) arguments, there is also a substantial — and probably much larger — number of language where A-agreement is limited to or favored by positions higher up on the referential hierarchy (e.g. Kiowa: Watkins and McKenzie 1984; Zúñiga 2006).

A-agreement limited to higher RH positions contradicts the expectations from the Marking-based RH Hypothesis. However, this kind of agreement rule is likely to reflect a competing universal trend: a general trend for reserving grammatical relations, such as those that trigger agreement, to the topmost segments of the referential hierarchy. I call this the GR-based RH Hypothesis:

(6) *GR-based RH Hypothesis*

If grammatical relations (of any construction, with any alignment) have restricted access based on the referential hierarchy, the odds for this access correlate positively with the rank of an argument on the referential hierarchy.

The hypothesis follows from grammaticalization theory: under this theory, grammatical relations are taken to represent grammaticalized topicality assignments, and since higher positions in the referential hierarchies are intrinsically more likely to be topical, they will have better access to grammatical relations than lower positions (Givón 2001).

Although again full statistical testing must await better surveying, the GR-based hypothesis has independent support from the many languages with strictly hierarchical agreement, i.e. where access to the agreement-triggering GR is based on the position of an argument on the referential hierarchy rather than its semantic role (see DeLancey 1981; Ebert 1987; Siewierska 2004; Zúñiga 2006, among others). One example is the Tibeto-Burman language Gyarong (lCog-rtse rGyal-roñ): when a speech act participant (i.e. a high-ranking referent) co-occurs with a third person (i.e. a lower ranking referent), the speech act participant triggers agreement, no matter what its role is. Thus, both ‘I give him’ and ‘he gives me’ trigger first person agreement (-ŋ), resulting in *wuŋ* and *wuwuŋ*, respectively. The meanings are differentiated by an inverse-marking prefix *wu-* (Nagano 1984, 2003; Bickel 1995)

Beyond agreement, the evidence for the GR-based hypothesis is uncertain. With regard to relative constructions, for example, there are both languages where the relativizable GR favors higher-ranking arguments and languages where the same GR favors lower-ranking arguments. Available survey work does not yet allow an assessment of what, if anything, is more common.

An example of a GR preference to high-ranking arguments is Tagalog, where only *ang*-NPs, which are topical and/or specific, can be the target of relativization (Schachter 1976; Kroeger 1993). The role of the *ang*-NP is signaled by the verb morphology as, e.g. A (marked by the infix *-um-*) or P (unmarked):

- (7) a. *ang=lalaki=ng b<um>asa ng=diyaryo*
 NOM=man=ATTR <A>read OBL=newspaper
 ‘the man who read a newspaper.’
- b. *ang=diyaryo=ng b<in>asa ng=lalaki*
 NOM=newspaper=ATTR [P-]<PFV>read OBL=man
 ‘the newspaper that the man read’
- c. **ang=lalaki=ng b<in>asa ang=diyaryo*
 NOM=man=ATTR [P-]<PFV>read NOM=newspaper
Intended: ‘the man who read a newspaper.’

Relativization on an A argument necessitates the A voice, as in (7a), because this ensures that the A argument is in the relativizable *ang*-function. Relativization on P, by contrast, requires the P voice, as in (7b), so as to ensure that now the P argument is in the *ang*-function. (7c) is ungrammatical because the relativized argument (‘the man’) is not in *ang*-function.

The opposite pattern is found in Movima, an isolate of Bolivia (Haude 2006) where only lower-ranking arguments in transitive clauses can be targets of relativization. Lower-ranking arguments of transitive clauses — called ‘ARG2’ by Haude — are formally characterized by, among other properties, (i) optional realization (in contrast to ARG1 arguments, which are obligatory), (ii) a lack of first and second person clitic pronouns, (iii) obviative case-marking (under some conditions), and (iv) the fact that they refer to P arguments if the verb is marked as direct and to A arguments if the verb is marked as inverse. Arguments lower on the referential hierarchy are assigned to ARG2, arguments higher on the hierarchy to ARG1, in either case regardless of their role. Relativization is possible only on ARG2 (obviative) arguments (or on the sole argument of intransitives):

- (8) a. [isos wa:ka [di' chik<a>ye=is neyru=s beñ'i]]
 ART.pPST cow REL find<DIR>=PL.ABSENT here=DET grassland
 'the cows which they had found in this grassland'
- b. [us ney juyeni [di' alwani-kaya=y'i]]
 ART.M here person REL talk-INV=1p
 'that person who had spoken to us'

Relativization on a P argument, as in (8a), necessitates direct verb morphology (infix *-a*), relativization on a A argument, as in (8b), inverse verb morphology (suffix *-kaya*). In both cases, the remaining argument in the relative clause is ARG1, i.e. the higher-ranking argument (the clitic pronouns *is* 'they (absent)' in (8a) and *y'i* 'we' in (8b). This is in direct contrast to the Tagalog pattern observed earlier in (7).

5. Conclusions

Empirical assessments of referential hierarchy effects on grammatical relations are hampered by the lack of large-scale typological databases beyond person categories in case marking and verb agreement. Where statistical testing is possible, we find no support for a general trend linking accusative alignment with high RH positions and ergative alignment with low RH positions. What is empirically supported is a more specific correlation between zero case exponence and high-ranking A or low-ranking P arguments, although the dataset that allows testing this is so far limited to two dozen languages with case splits based on a pronoun vs. noun distinction (Comrie 2005). If this correlation nevertheless reflects a genuine universal, it is possible that it extends to verb agreement. The evidence accumulated so far is equivocal (Siewierska 2004), however, and one reason for this could be that any possible trend towards zero verb agreement morphology for high-ranking A arguments is cancelled out by a counteracting principle that favors agreement with high-ranking arguments across any semantic role (assuming such a principle would have audible, nonzero, effects.) However, although there is some tentative evidence for such a principle, it cannot be taken as a given. It is also possible that the referential hierarchy has a different impact on case (or adposition) and verb agreement because these two ways of marking arguments have a fundamentally different relationship to referential distinctions: the fact that normally case is realized on NPs implies that is much more closely tied to the active, on-line management of reference and information than agreement, and this causes perhaps a more direct impact of referential distinctions on the realization of case than on agreement.

References

- Allen, N. J. (1975), *Sketch of Thulung grammar*, Ithaca, N.Y.: Cornell University.
- Bakker, D. and Siewierksa, A. (2006), 'The agreement database', <http://www.lotschool.nl/Research/ltrc/agreement.htm>; accessed August 7, 2006.
- Bickel, B. (1995), 'In the vestibule of meaning: transitivity inversion as a morphological phenomenon', *Studies in Language*, 19, 73 – 127.
- (2000), 'Person and evidence in Himalayan languages', *Linguistics of the Tibeto-Burman Area*, 23, 1 – 12.
- (2003), 'Belhare', in G. Thurgood and R. J. LaPolla (eds.), *The Sino-Tibetan languages*, London: Routledge, 546 – 70.
- (2007), 'Grammatical relations typology', in J. J. Song (ed.), *The Oxford Handbook of Language Typology*, Oxford: Oxford University Press.
- Bickel, B. and Gaenszle, M. (2005), '1P antipassives and the political history of the Southern Kirant', Paper presented at the 11th Himalayan Languages Symposium, Bangkok, December 6 - 9, 2005; handout available at http://www.uni-leipzig.de/~ff/cpdp/frameset_publ.html.
- Bickel, B., Gaenszle, M., Rai, A., Rai, P. D., Rai, S. K., Rai, V. S., and Sharma (Gautam), N. P. (2007a), 'Two ways of suspending object agreement in Puma: between incorporation, antipassivization, and optional agreement', *Himalayan linguistics*, 7, 1 - 18.
- Bickel, B., Banjade, G., Gaenszle, M., Lieven, E., Paudyal, N., Rai, I., Rai, M., Rai, N. K., and Stoll, S. (2007b), 'Free prefix ordering in Chintang', *Language*, 83, 43 – 73.
- Comrie, B. (1981), *Language universals and linguistic typology*, Chicago: University of Chicago Press.
- (2005), 'Alignment of case marking', in M. Haspelmath, et al. (eds.), *The world atlas of language structures*, Oxford: Oxford University Press, 398 – 405.
- Crook, H. D. (1999), *The phonology and morphology of Nez Perce stress*, Ph. D. dissertation, University of California, Los Angeles.
- Curnow, T. J. (1999), 'Maung verbal agreement revisited: a reponse to Donohue (1998)', *Australian Journal of Linguistics*, 19, 141 – 59.
- Cysouw, M. (2002), 'Against implicational universals', *Linguistic Typology*, 7, 89-110.
- DeLancey, S. (1981), 'An interpretation of split ergativity and related patterns', *Language*, 57, 626 – 57.
- Dixon, R. M. W. (1979), 'Ergativity', *Language*, 55, 59 – 138.
- Donohue, C. and Donohue, M. (1997), 'Fore case marking', *Language and Linguistics in Melanesia*, 28, 69 – 98.
- Driem, G. van (1987), *A grammar of Limbu*, Berlin: Mouton de Gruyter.
- (1990), 'The fall and rise of the phoneme /r/ in Eastern Kiranti: sound change in Tibeto-Burman', *Bulletin of the School of Oriental and African Studies*, 53, 83 – 86.
- (1991), 'Bahing and the Proto-Kiranti verb', *Bulletin of the School of Oriental and Asian Studies*, 54 (2), 336 - 56.
- (1992), 'Le proto-kiranti revisité: morphologie verbale du lohorung', *Acta Linguistica Hafniensia*, 24, 33 –75.
- (1993), *A grammar of Dumi*, Berlin: Mouton de Gruyter.
- (2001), *Language of the Himalayas: an ethnolinguistic handbook of the Greater Himalayan Region, containing an introduction to the Symbiotic Theory of Language*, Leiden: Brill.
- DuBois, J. W., Kumpf, L. E., and Ashby, W. J. (eds.) (2003), *Preferred argument structure: grammar as architecture for function*, Amsterdam: Benjamins.
- Ebert, K. H. (1987), 'Grammatical marking of speech act participants', *Journal of Pragmatics*, 11, 473 – 82.
- (1991), 'Inverse and pseudo-inverse prefixes in Kiranti languages: evidence from Belhare, Athpare and Dungmali', *Linguistics of the Tibeto-Burman Area*, 14, 73 - 92.
- (1994), *The structure of Kiranti languages*, Zürich: ASAS Press.
- (1997), *A grammar of Athpare*, München: LINCOM Europa.
- Filimonova, E. (2005), 'The noun phrase hierarchy and relational marking: problems and counterevidence', *Linguistic Typology*, 9, 77 - 113.
- Gaenszle, M. (1991), *Verwandtschaft und Mythologie bei den Mewahang Rai in Ostnepal: eine ethnographische Studie zum Problem der "ethnischen Identität"*, Stuttgart: Steiner.
- (in press), 'Rai mythology', in K. Ebert and M. Gaenszle (eds.), *Rai mythology: Kiranti oral texts*, Harvard: Harvard University Press.
- Garrett, A. (1990), 'The origin of NP split ergativity', *Language*, 66, 261 – 96.
- Givón, T. (2001), *Syntax*, Amsterdam: Benjamins.

- Gordon, L. (1986), *Maricopa morphology and syntax*, Berkeley: University of California Press.
- Haiman, J. (1980), *Hua: a Papuan language of the Eastern Highlands of New Guinea*, Amsterdam: Benjamins.
- Haude, K. (2006), *A grammar of Movima*, Ph. D. dissertation, Radboud University, Nijmegen.
- Jacobsen, W. H. J. (1964), *A grammar of the Washo language*, Ph.D. dissertation, University of California, Berkeley.
- Jäger, G. (2007), 'Evolutionary Game Theory and typology: a case study', *Language*, 83, 74 - 109.
- Janssen, D., Bickel, B., and Zúñiga, F. (2006), 'Randomization tests in language typology', *Linguistic Typology*, 10, 419 - 40.
- Kroeger, P. (1993), *Phrase structure and grammatical relations in Tagalog*, Stanford: CSLI.
- Lahaussois, A. (2003), 'Thulung Rai', *Himalayan Linguistics Archive*, 1, 1 - 25.
- Michailovsky, B. (1988), *La langue hayu*, Paris: CNRS.
- (1994), 'Manner vs. place of articulation in the Kiranti initial stops', in H. Kitamura, T. Nishida, and Y. Nagano (eds.), *Current issues in Sino-Tibetan linguistics*, Osaka: National Museum of Ethnology, 766 - 72.
- (1997), 'Catégories verbales et intransitivité duale en Limbu', *Studi Italiani di Linguistica Teorica e Applicata*, 26, 307 - 25.
- (2001), 'Limbu nous autres and First Person morphology', *Linguistics of the Tibeto-Burman Area*, 24, 145-56.
- Nagano, Y. (1984), *A historical study of the rGyarong verb system*, Tokyo: Seishido.
- (2003), 'Cogtse Gyarong', in G. Thurgood and R. J. LaPolla (eds.), *The Sino-Tibetan languages*, London: Routledge, 469 - 98.
- Nichols, J. (2003), 'Diversity and stability in language', in R. D. Janda and B. D. Joseph (eds.), *Handbook of Historical Linguistics*, London: Blackwell, 283 - 310.
- Opgenort, J. R. (2004a), 'Implosives and preglottalized stops in Kiranti', *Linguistics of the Tibeto-Burman Area*, 27, 1 - 27.
- (2004b), *A Grammar of Wambule*, Leiden: Brill.
- (2005), *A grammar of Jero, with a historical comparative study of the Kiranti languages*, Leiden: Brill.
- Pokharel, M. (2054 VS = 1998 AD), *Nepāli Vākya Vyākaraṇ*, Kāṭhmāṇḍū: Nepāl Rājakīya Pragya-Pratiṣṭhān.
- Rai, N. K. (1984), 'A descriptive study of Bantawa', Ph.D. dissertation, Deccan College Post-Graduate Research Institute, Pune (India).
- Rutgers, R. (1998), *Yamphu*, Leiden: CNWS Publications.
- Schachter, P. (1976), 'The subject in Philippine languages: topic, actor, actor-topic, or none of the above', in C. N. Li (ed.), *Subject and topic*, New York: Academic Press.
- Seidl, A. and Dimitriadis, A. (1997), 'The discourse function of object marking in Swahili', *Proceedings of the 33rd Annual Meeting of the Chicago Linguistics Society*, 373 - 89.
- Siewierska, A. (2004), *Person*, Cambridge: Cambridge University Press.
- Silverstein, M. (1976), 'Hierarchy of features and ergativity', in R. M. W. Dixon (ed.), *Grammatical Categories in Australian Languages*, New Jersey: Humanities Press, 112 - 71.
- Suter, E. (2006), 'The optional ergative in Kâte', Ms. to appear in a Festschrift.
- Toba, S. (1988), 'The pronominal affix system in Khaling', in D. Bradley, E. J. A. Henderson, and M. Mazaudon (eds.), *Prosodic analysis and Asian linguistics: to honour R.K. Sprigg*, Canberra: Pacific Linguistics (PL-C 104), 201 - 04.
- Tolsma, G. (1999), *A grammar of Kulung*, Ph. D. dissertation, University of Leiden.
- Watkins, L. J. and McKenzie, P. (1984), *A grammar of Kiowa*, Lincoln: University of Nebraska Press.
- Watters, J. K. (1988), *Topics in Tepehua grammar*, Ph.D. dissertation, University of California, Berkeley.
- Zúñiga, F. (2006), *Deixis and alignment: inverse systems in indigenous languages of the Americas*, Amsterdam: Benjamins.

Acknowledgments

The research reported here was funded by DFG Grant No. BI 799/3-1. I am grateful to Anna Siewierksa, Grev Corbett, and Mickey Noonan for helpful comments on a first draft.

Notes

¹ I am grateful to Dik Bakker who helped me perform the search in early August 2006. The database is available online at <http://www.lotschool.nl/Research/ltrc/agreement.htm>. Languages retrieved by the search but excluded from Table 1 because they involve clitics are Ngiyambaa, Yukulta, Warao, and Nadeb. My criterion for clitic status was (a) a phonological structure reflecting simple reduction of full pronouns, with near-identity in segments, and (b) a ban on cooccurrence of clitics with coreferential argument NPs in the same clause. Note that from the raw search results I excluded cases with missing values in the relevant alignment variable, i.e. I did not expand the database. But see the discussion below for possible additions.

² There is another second person prefix, *k-*, which covers only S and A. However, this is not an RH effect but results from the fact that *k-* is limited to imperative forms, and that in Maricopa, just like in most other languages, imperatives follow a near-universal principle of accusative alignment (see Dixon 1979, Comrie 1981, among others).

³ cf. Cysouw 2002 on the problematic nature of zeros as evidence for universals.

⁴ Complete Puma paradigms can be inspected at the DOBES archive (<http://www.mpi.nl/DOBES>) by searching for sessions of genre 'paradigm' with content 'verb'.

⁵ As Mickey Noonan reminds me, the distribution could have been very different in the past, when the split developed. But we have no independent evidence on this.

Table 1: Verb agreement with RH-sensitive alignment splits in Bakker & Siewierska (2006)

| <i>Language</i> | <i>Split, as coded in the database</i> | <i>Stock</i> | <i>Location</i> |
|-----------------|---|------------------|------------------|
| Ainu | 1s,2 accusative; 1p tripartite | (isolate) | E. Eurasia |
| Chácobo | 1s,2s,3s,1p tripartite; 2p,3p accusative; | Panoan | S. America |
| Comox | 1,2 accusative; 3s tripartite | Salish | N. America |
| Kamass | 1,2 accusative; 3 tripartite | Uralic | Eurasia |
| Paumarí | 1,2,3p accusative; 3s,3d ergative | Arauan | S. America |
| Seri | 1s tripartite; 2s,p accusative; 3 neutral (zero) | (isolate) | N. America |
| Yimas | 1,2 tripartite; 3 ergative | Lower Sepik | Papua New Guinea |
| ? Maricopa | 1,2 accusative; 3 neutral (zero) | Yuman | N. America |
| ? Maung | 1,2 accusative; 3 ergative | Iwaidjan | Australia |
| ? Nez Perce | 1,2 accusative; 3 ergative | Plateau Penutian | N. America |
| ? Tepehua | 1,2 accusative; 3 neutral | Totonac-Tepehuan | C. America |
| ? Washo | 1,2 accusative; 3 ergative | (isolate) | N. America |

Table 2. Split ergativity in Kiranti

| | A | S | P |
|---------------|---|---|---|
| 1 singular | | | |
| 1 nonsingular | | | |
| 2 | | | |
| 3 | | | |

Table 3. Split ergativity in Puma

| | A | S | P |
|-----------------|--|-------------------------|--------------|
| 1s | -ŋ (>3) -na (>2) | -ŋa (NPST) -oŋ (PST) | |
| 1d ^a | -ci~cɔ ^b | | |
| | ni- ci~cɔ (>2) | | |
| 1p ^a | -m (>3) ni- i~ni(n)~nɔ (>2) | -i~ni(n)~nɔ | |
| 2s | tɔ- | | -na (1s>) |
| 2d | tɔ- ci~cɔ | | -na-ci (1s>) |
| 2p | tɔ- i~ni(n)~nɔ | | |
| | tɔ- m (>3) | | |
| 3s | ø- | | -u ~ -i |
| | pɔ- (>1) | | |
| 3d | pɔ- ci~cɔ | | |
| | ni-pɔ- ci~cɔ (>1) ni- ci~cɔ (>2) | | |
| 3p | mɔ- | | |
| | pɔ- (>3s) ni-pɔ- (>1s) ni-pɔ- i~ni(n)~nɔ (>1ns) ni- i~ni(n)~nɔ (>2) | -ci | |

^a All first person nonsingular forms distinguish exclusive vs. inclusive forms, marked by -ka, but omitted here since this does not affect alignment

^b For the conditions regulating allomorphies (marked here by a tilde), see Bickel et al. (2007a)

Table 4. Reverse ergativity splits in Kiranti verb agreement

| <i>Branch</i> | <i>With reverse ergativity split (at least in part)</i> | <i>Without ergative alignment (i.e. with first person showing tripartite, accusative or neutral alignment)</i> |
|---------------|---|--|
| Eastern | Athpare (Ebert 1991, 1997) | Belhare (Bickel 2003) |
| | Limbu (van Driem 1987, Michailovsky 1997, 2001) | |
| | Mewahang (M. Gaenszle, p.c.) | |
| | Lohorung (van Driem 1992) | |
| | Chintang (Bickel et al. 2007b) | |
| | Yamphu (Rutgers 1998) | |
| Central | Puma (Bickel et al. 2007a) | Camling (Ebert 1991, 1994) |
| | Bantawa (Rai 1984; Ebert 1994) | |
| | Kulung (Tolsma 1999) | |
| Western | Hayu (Michailovsky 1988) | Wambule ^a (Opgenort 2004b) |
| | Khaling (Toba 1988, Ebert 1994) | Thulung (Allen 1975; Lahaussis 2003) |
| | Bahing ^b (van Driem 1991) | Jero (Opgenort 2005) |
| | | Dumi (van Driem 1993) |

^a Wambule (Opgenort 2004b) has *-ki* '1ns' covering both S and P, but it denotes inclusive in S and exclusive in P function. There are no other traces of ergative alignment in the paradigm.

^b In Bahing (van Driem 1991), alignment is also split on tense: first person shows S=A alignment in the nonpast and S=P alignment in the past. In this language, S=P alignment also extends to the second person (singular in the nonpast; all numbers in the past), but again not to third person, which has a thorough S=A alignment. An extension to second person is also found in Kulung.

Table 5. Comparing predictions^a

| <i>RH rank</i> | <i>In line with prediction:</i> | | <i>Against prediction:</i> | |
|----------------------------|--|---|--|--|
| | <i>high</i> | <i>low</i> | <i>high</i> | <i>low</i> |
| Marking-based Hypothesis | non-zero accusative and zero nominative; or non-zero absolutive and zero ergative; or zero neutral | non-zero ergative and zero absolutive; or non-zero nominative and zero accusative; or neutral | ergative and zero absolutive; or nonzero nominative and zero accusative; or zero neutral | accusative and zero nominative; or non-zero absolutive and zero ergative; or neutral |
| Alignment-based Hypothesis | accusative or neutral alignment | ergative or neutral alignment | ergative or neutral alignment | accusative or neutral alignment |

^a For consistency, I use the terminology of alignment typology also for the absence of overt case-marking, called here 'zero neutral'. If both pronouns and nouns are zero-marked, the predictions are vacuous.