Alena Witzlack-Makarevich*, Taras Zakharko,
Lennart Bierkandt, Fernando Zúñiga and Balthasar Bickel

Decomposing hierarchical alignment: Co-arguments as conditions on alignment and the limits of referential hierarchies as explanations in verb agreement

Abstract: Apart from common cases of differential argument marking, referential hierarchies affect argument marking in two ways: (a) through hierarchical marking, where markers compete for a slot and the competition is resolved by a hierarchy, and (b) through co-argument sensitivity, where the marking of one argument depends on the properties of its co-argument. Here we show that while co-argument sensitivity cannot be analyzed in terms of hierarchical marking, hierarchical marking can be analyzed in terms of co-argument sensitivity. Once hierarchical effects on marking are analyzed in terms of co-argument sensitivity, it becomes possible to examine alignment patterns relative to referential categories in exactly the same way as one can examine alignment patterns relative to referential categories in cases of differential argument marking and indeed any other condition on alignment (such as tense or clause type). As a result, instances of hierarchical marking of any kind turn out not to present a special case in the typology of alignment, and there is no need for positing an additional non-basic alignment type such as "hierarchical alignment". While hierarchies are not needed for descriptive and comparative purposes, we also cast doubt on their relevance in diachrony: examining two families for which hierarchical agreement has been postulated, Algonquian and Kiranti, we find only

*Corresponding author: Alena Witzlack-Makarevich, Institute for Scandinavian Studies, Frisian and General Linguistics, Department of General Linguistics, Christian-Albrechts-Universität zu Kiel, Olshausenstraße 40, 24098 Kiel, Germany, E-mail: awitzlack@isfas.uni-kiel.de
Taras Zakharko, Department of Comparative Linguistics, University of Zürich, Plattenstrasse 54, CH-8032 Zürich, Switzerland, E-mail: taras.zakharko@uzh.ch
Lennart Bierkandt, Institut für Anglistik, Friedrich-Schiller-Universität Jena, Ernst-Abbe-Platz 8, 07743 Jena, Germany, E-mail: post@lennartbierkandt.de
Fernando Zúñiga, Department of Linguistics, University of Bern, Laenggassstrasse 49, CH-3000 Bern 9, Switzerland, E-mail: fernando.zuniga@isw.unibe.ch
Balthasar Bickel, Department of Comparative Linguistics, University of Zürich, Plattenstrasse 54, CH-8032 Zürich, Switzerland, E-mail: balthasar.bickel@uzh.ch

weak and very limited statistical evidence for agreement paradigms to have been shaped by a principled ranking of person categories.

Keywords: morphosyntactic alignment, hierarchical agreement, Algonquian, Kiranti

1 Introduction

Alignment is understood here as the way in which the generalized semantic argument roles S, A, and P[1] are treated alike by case assignment, agreement marking, and other morphosyntactic operations. The three roles S, A, and P yield five logically possible alignment types: nominative-accusative (subsetting roles as {S, A} vs. {P}), ergative-absolutive ({S, P} vs. {A}), tripartite ({S} vs. {A} vs. {P}), horizontal ({S} vs. {A, P}), and neutral ({S, A, P}). These five basic types have figured prominently in typological research and haven proven useful for descriptive purposes as well.

At the same time, it has often been noted that not all cases of argument marking can be straightforwardly characterized in terms of these basic alignment types because the way in which case or agreement aligns argument roles is often subject to various conditions. In many cases the complications created by such conditions, however, are more apparent than real: phenomena like differential subject and differential object marking, for example, simply suggest that alignments may hold only for certain referential types, e. g., one might find accusative alignment for first person referents and ergative alignment for third person referents. The characterization and analysis of such systems is straightforward and is not different in principle from well-established alignment conditions based on clausal properties such as tense, aspect, or subordination, which also frequently result in alignment splits.

By contrast, other ways in which referential properties affect alignments are far less straightforward and have given rise to the postulation of further alignment types such as "hierarchical alignment" (Mallinson and Blake 1981; Nichols 1992; Siewierska 1998) or "inverse alignment" and "inverse voice" (Gildea 1994) or even more specific types, such as "Austronesian alignment" (Aldridge 2012). None of these concepts has received general acceptance because (a) their

---

1 We understand the generalized argument roles S, A, and P as exclusively semantic notions, defined by the logical entailments of a predicate and not by its morphological marking or syntactic behavior – for details, see Bickel and Nichols (2009), Bickel et al. (2010), Bickel (2011), Witzlack-Makarevich (2011), and Bickel et al. (2014). On other uses of S, A, and P (or O) and a comparison of approaches to these notions, see Haspelmath (2011). For the sake of simplicity we focus here on one- and two-place predicates, for the application of the suggested method to three-place predicates and challenges thereof see van Lier et al. (this issue).

definitions do not follow the same conceptual logic of forming subsets of argument roles that underlies the five basic alignment types (cf. Creissels 2009), (b) they tend to be ad-hoc and not universally applicable, and (c) they tend to also contain traces of the five basic alignment types (Nichols 1992; Bickel 1995; Zúñiga 2006; Haude 2009). The key challenge of all these phenomena for alignment typology is that the relevant condition is not a simple referential feature but an entire constellation of such features.

Here we propose a new analysis of the effects that such constellations can have on alignment. The basic idea is that all such effects boil down to a sensitivity to co-arguments, and that such a sensitivity is not in principle different from well-established kinds of conditions on alignments, such as e. g., tense or aspect of a clause. We argue that this new analysis challenges the assumption that a notion of referential hierarchy or scale is part of the universal inventory of analytical concepts in grammar. By contrast, there is a series of phenomena that (we argue) cannot be successfully analyzed without the notion of co-argument sensitivity. This notion is therefore a necessary ingredient of the universal inventory of analytical concepts.

In the following we first describe in more detail the challenges posed by various types of feature constellations: situations where argument roles seem to compete for expression or marking so that only the "highest" role is selected (Section 2), and situations where expression or marking of one argument depends on properties of another argument (Section 3). In Section 4 we introduce our response to these two kinds of challenges, showing that hierarchies are not needed for descriptive or comparative purposes. Specifically, we argue that any descriptive generalizations and simplifications that emerge from using hierarchies are more apparent than real and that not much is gained by using them. In Section 5 we examine the extent to which hierarchies may be needed at least for the purpose of explaining the diachronic development of specific paradigms. We analyze whether, and if so to what extent, referential hierarchies have shaped the evolution of agreement paradigms in two language families for which hierarchical agreement has been postulated, viz. Algonquian and Kiranti. We find that the statistical evidence for such effects is very limited, never affecting entire systems and in the best case of evidence affecting only highly specific aspects of individual sub-paradigms. Section 6 summarizes the paper and suggests general conclusions.

2 Role competition

In some languages, case or agreement marking can be conditioned by the referential properties of all arguments that are involved and represented in the clause.

The phenomenon comes in two flavors: (i) in some systems, discussed in this section, arguments of a clause compete for a particular agreement slot or for a particular case marker. It is normally assumed that in order to provide an account for such a system, it is necessary to posit a referential hierarchy (or scale) of a certain form (language-specific or universal). Then, one can say that only the argument that ranks higher on the hierarchy than other arguments of the same clause gets access to a particular agreement slot or case marker. Such cases underlie the traditional label "hierarchical agreement" or recently "hierarchical indexation" (Rose 2009), as well as what we will refer to as "hierarchical case marking". (ii) In the other systems, discussed in detailed in Section 3, argument marking also depends on the whole constellation of the arguments in a clause, or what one might call its "scenario". However, in contrast to the first type, in such systems it is impossible to account for the distribution of markers in terms of a unified referential hierarchy because the relevant conditions determining their distribution involve several variables at once (e. g., 'assign accusative to the P argument if the A argument is second person singular and nowhere else').

Hierarchical agreement can be illustrated by the following examples from the Tupian language Emerillon. Emerillon has two sets of agreement prefixes. Whereas one set marks S and A arguments, the other set marks P arguments. In the case of one-place verbs, the prefix slot shows agreement with the S argument, as in (1a) and (1b):

(1)   Emerillon (French Guiana, Tupi-Guaraní)[2]
    a. ***ere**-zaug-tanẽ-po?*
       2sS/A-bathe-DESID-INTER
       'Do you want to bathe?'
       (Rose 2011: 296)
    b. *Polo   **o**-manõ.*
       Paulo **3S/A**-die
       'Paulo died.'
       (Rose 2011: 24)

In the case of two-place verbs, there is a competition between the two arguments for the same agreement slot. The following person hierarchy determines the access to this agreement slot:

---

2 Abbreviations: Abbreviations used in the glosses follow Leipzig Glossing Rules with the following additions and modifications: CONTR contrastive, d dual, DESID desiderative, e exclusive, G goal, INTER interrogative, INV inverse, ns nonsingular, p plural, s singular, SAP speech act participant.

(2)   1/2 ≻ 3

As (3a) and (3b) show, if there is a first or second person argument in a clause, the verb agrees with this argument, no matter whether it is an A or a P. The verb agrees with the third person A argument only if there is no first or second person argument in the clause, as in (3c).[3]

(3)   Emerillon (French Guiana, Tupi-Guaraní)
       a. *nõde-i-a-ne*   ***ere-****zika*.
          1INCL.POSS-mother-a-CONTR **2sS/A**-kill
          'You killed our mother.'
          (Rose 2011: 25)
       b. *zawar* ***de-****suʔu.*
          dog   **2sP**-kill
          'A dog killed you.'
          (Rose 2011: 70)
       c. *kija*   ***o-****baʔe*   *pariru-pe.*
          hammock   **3S/A**-make son.in.law-for
          'She makes a hammock for her son-in-law.'
          (Rose 2011: 122)

Another example of hierarchical agreement comes from Plains Cree. In the case of one-place verbs, the prefix slot shows agreement with the S argument, as in (4a) and (4b). In the case of two-place verbs, the situation is more complex: if there is a second person argument, the verb agrees with it, no matter whether it is an A or a P, as (4c) and (4d) show. Only if there is no second person argument, the verb agrees with the first person argument, as in (4e) and (4f).

(4)   Plains Cree (Canada, Algonquian)
       a. ***ki-****pimipahtā-n.*
          2-run-sSAP
          'You run.'
          (Dahlstrom 1986: 31)
       b. ***ni-****pimipahtā-n.*
          1-run-sSAP
          'I run.'
          (Dahlstrom 1986: 31)

---

3 In first and second person ("local") scenarios the verb always shows agreement with the A argument.

    c. ***ki-*pēhtaw-i-n.**
      2-hear-2-1→sSAP
      'You hear me.'
      (Dahlstrom 1986: 28)

    d. ***ki-*pēhtaw-iti-n.**
      2-hear-1-2→sSAP
      'I hear you.'
      (Dahlstrom 1986: 28)

    e. ***ni-*wāpaht-ē-n.**
      1-see-TR-sSAP
      'I see it.'
      (Dahlstrom 1986: 36)

    f. ***ni-*pakamahw-ikw-w.**
      1-hit-INV-3
      'He hits me.'
      (Dahlstrom 1986: 29)

The access to prefixal agreement is usually described as being governed by the well-known Algonquian hierarchy (cf. Dahlstrom 1986; Zúñiga 2006):

(5)   2 > 1 > 3

    A similar situation is found in the domain of case marking and is particularly common in some Western Malayo-Polynesian languages. Parallel to the established term of hierarchical agreement, one can refer to such cases as hierarchical case marking. For instance, in Tagalog (Schachter and Otanes 1972; Kroeger 1993), only one noun phrase of a clause is marked by *ang=*, glossed here as "nominative".[4] The choice of which noun phrase is marked with *ang=* depends on the referential properties of arguments and adjuncts of the clause: the marker is assigned to whichever element in the clause is most "prominent" in terms of referentiality, definiteness, and further discourse-pragmatic factors (cf. Himmelmann 2005).[5] The following examples illustrate this:

4 The exact status of *ang*-marked noun phrases is still under debate in Austronesian linguistics where it is referred to as "topic", "focus", "pivot", "nominative", "subject" or "specific article", see Himmelmann (2002) for discussion and references.

5 Competitive access to a single category is what makes Tagalog similar to the distribution of Algonquian proximative-obviative case marking, but there are of course many differences. The proximative marks arguments typically expressing the person from whose point of view events are described, such as the protagonist in narratives; all other arguments are obviative. And unlike in Tagalog, a clause can also be without any proximative argument (e. g., Cowell and

(6)   Tagalog (Austronesian; Southeast Asia)
     a. *bumili*   **ang**=*lalake ng*=*isda*   *sa*=*tindahan*.
      PFV.A.buy **NOM**=man OBL=fish LOC=store
      'The man bought fish at the/a store.'
      (Kroeger 1993: 26)
     b. *binili*   *ng*=*lalake* **ang**=*isda*   *sa*=*tindahan*.
      PFV.P.buy OBL=man **NOM**=fish LOC=store
      'The/a man bought the fish at the/a store.'
      (Kroeger 1993: 26)
     c. *binilhan*   *ng*=*lalake ng*=*isda*   **ang**=*tindahan*.
      PFV.G.buy OBL=man OBL=fish **NOM**=store
      'The/a man bought fish at the store.'
      (Kroeger 1993: 26)

Similar to Emerillon, one can formulate the rule of the distribution of *ang* = NOM as being governed by a hierarchy like the following:

(7)   more prominent > less prominent

     Phenomena like those in Emerillon, Plains Cree, and Tagalog could all be typologized as representing "hierarchical alignment". But if we allow for it, hierarchical alignment would have a special position among alignment types. As mentioned in the introduction, the three roles S, A, and P yield only five logically possible alignment types: accusative, ergative, tripartite, horizontal, and neutral. Hierarchical alignment would be an additional, *non-basic* type. Such an extension incurs several problems: In contrast to the mere comparison of argument roles that underlies the basic types, the idea of hierarchical alignment is based on a different principle: the concept merely refers to the existence of some additional conditioning factors of argument marking, but it does not specify what the grouping of argument roles is. This makes the theory of alignment inconsistent, as already noted by Zúñiga (2007) and Creissels (2009). Along the same lines, if hierarchical alignment indeed exists as a type of its own, we would expect it to be on a logical par with the other types and, in particular, we would expect that the existence of hierarchical alignment in one construction or form would exclude any other kind of alignment in the same

Moss 2008: 350 on Arapaho); selection of a proximative argument is strictly enforced only when there are both third and first or second persons involved in the same clause.

construction or form. Yet this prediction is not warranted: Within the same agreement system, one often – perhaps indeed mostly – finds "hierarchical" principles along with other alignment types (cf. Nichols 1992: 68). To illustrate this point, consider again the Emerillon prefixes illustrated in (1) and (3). In these examples, the prefixes show accusative alignment, for instance, *ere-* 2sS/A in (1a) and (3a) is used to mark the S and A arguments of the first person inclusive, whereas the prefix *de-* 2sP is used to mark the P argument, as in (3b). Similar observations have been made for Algonquian languages (Bickel 1995; Zúñiga 2006: 126).

How can we then capture agreement systems like that of Emerillon? As "hierarchical with a trace of accusative alignment"? As a combination of "hierarchical", "neutral" and "accusative"? These choices seem all unattractive. What we want to capture is both the accusative alignment of prefixes and the effects of arguments' referential values on the distribution of agreement markers. A notion of "hierarchical alignment" does not seem to generate insights here.

3 Co-argument sensitivity

Not every system with effects of referential values on argument marking can be analyzed as governed by a hierarchy. An example is prefixal agreement in the Sino-Tibetan language Puma (Bickel et al. 2007), which in the singular is limited to two markers: *tʌ-* for second and *pʌ-* for third person. Given verb forms like in (8), one is reminded of Cree and Emerillon: the third person agreement prefix *pʌ-* appears only if there is no second person agreement (*tʌ-*), as in (8a), and for the second person, the argument role does not seem to matter – *tʌ-* can refer to P (8b), A (8c)–(8d) or S (8e) arguments:

(8) Puma (Nepal, Sino-Tibetan)
  a. ***pʌ-cind-oŋ***
     3sA-teach-1sP.PST
     's/he (A) taught me (P)'
  b. ***tʌ-cind-a***
     2-teach-PST
     's/he (A) taught you (P)'
  c. ***tʌ-cind-i***
     2-teach-3sP.PST
     'you (A) taught him/her (P)'

    d. ***tʌ-cind-oŋ***
       2-teach-1sS/P.PST
       'you (A) taught me (P)'
    e. ***tʌ-phind-a***
       2-jump-PST
       'you (S) jumped'

However, a statement like "the prefix marks whichever argument ranks higher on a 2 ≻ 3 hierarchy" would not correctly account for the distribution of the two prefixes because their appearance also depends on the referential status of the other argument in a transitive clause, i. e., of the co-argument. This becomes clear once we add the following forms, where the prefixes fail to appear even if the person that they denote are the highest on the 2 ≻ 3 hierarchy:

(9)   Puma (Nepal, Sino-Tibetan)
    a. *cin-na-a*
       teach-1>2-PST
       'I (A) taught you (P)'
    b. *cind-i*
       teach-3sP.PST
       's/he taught him/her'

In both cases, the prefixes are blocked not by a higher-ranking prefix – and not by a ∅- exponence prefix either, since this would paradoxically have to denote a first person in (9a) and a third person in (9b) – but by the properties of the co-argument: in (9a), *tʌ*- '2' is blocked because its co-argument is a first person A argument; in this case, the verb shows no prefix, but instead a suffix *-na* '1 acting on 2'. In (9b), the third person prefix *pʌ*- is blocked because its co-argument is a third person, in which case there is no third person agreement marker at all. This can only be captured by saying that *pʌ*- '3sA' marks a third person if and only if its co-argument is a first person.[6]

    Co-argument sensitivity that is not based on a hierarchy is also found with case marking. This can be illustrated by the following data from Ik. In Ik, the P argument can be either in the nominative, as in (10a) and (10b), or in the accusative case, as in (10c) and (10d):

---

6 The patterns are even more complex, and even less susceptible to an analysis in terms of hierarchical agreement, once one includes dual and plural forms. See Bickel et al. (2007) and Sharma (2014) for full paradigms and analyses.

(10) Ik (Kuliak, Uganda)

    a. *en-í-a   nk-a   wík-**a***
       see-1s-A I-NOM children-**NOM**
       'I see the children.'
       (König 2009: 151)

    b. *en-es-íd-a   bi-a   wík-**a**.*
       see-IRR-2s-A you-NOM children-**NOM**
       'You (sg) will see the children.'
       (König 2009: 151)

    c. *en-es-uɠot-a   wík-á   njíní-**ka**.*
       see-IRR-AND-A children-NOM 1pINCL-**ACC**
       'The children will see us (incl.).'
       (König 2009: 150)

    d. *en-es-át-a   ńt-a   ceki-**ka**.*
       see-IRR-3p-A they-NOM woman-**ACC**
       'They will see the woman.'
       (König 2009: 159)

What determines the distribution of the case markers on the P argument is the nature of the A argument: the P argument is in the accusative case only if the A argument is a third person, otherwise the P argument is in the nominative case. This is made explicit in Table 1, where subscripted numbers indicate person (e. g., S_1 refers to the first person S argument).

Table 1: Distribution of case marking in Ik.

marking	marked argument	co-argument
NOM	S_1	n.a.
NOM	S_2	n.a.
NOM	S_3	n.a.
NOM	A_1	any
NOM	A_2	any
NOM	A_3	any
NOM	P_1	A_2
NOM	P_2	A_1
NOM	P_3	A_1
NOM	P_3	A_2
ACC	P_1	A_3
ACC	P_2	A_3
ACC	P_3	A_3

Similar to Puma, and in contrast to the Emerillon, Cree or Tagalog examples, it is impossible to account for differential argument marking here by formulating a referential hierarchy of any kind, whereas the reference to the nature of co-arguments is unavoidable. For example, one could consider an analysis of Ik based on a hierarchy that ranks first and second person above third person (1/2 ≻ 3). Then, if the P argument is higher or equal to the A argument on this hierarchy, it is in the accusative case. This hierarchy would correctly predict the distribution of the accusative and nominative cases in the scenarios involving third persons, but it would overgenerate in that the analysis would also predict the accusative P argument in scenarios involving only first and second person ("local" scenarios).

The examples from Puma and Ik show that for some systems of agreement and case marking, it is unavoidable to make reference to specific constellations of person features. In principle, however, it seems that reference to such constellations is not fundamentally different from reference to single referential properties. Reference to single referential properties is something that we need anyway for any type of differential object or differential subject marking, i. e., alignment types need to be stated relative to referential contexts anyway (Bickel 2011). This paves the way towards a unified analysis, to which we turn in the following.

4 Analysis of co-argument sensitive argument marking

Whereas hierarchical marking systems are typically squeezed into alignment typology by introducing one or more non-basic alignment types, non-hierarchical cases of co-argument sensitivity are often either ignored in the discussion of alignment, or they are classified as instances of more familiar hierarchical alignment (e. g., in the case of Limbu in Siewierska 2005). To include these systems, one could in principle consider expanding the typology of alignment systems with yet another type, "co-argument sensitive alignment". But as discussed above, this would obscure the notion of alignment even more. The alternative that we propose is to analyze both co-argument sensitivity and hierarchical agreement in terms of basic alignments that are subject to, or conditioned by, specific referential effects (Zúñiga 2007; Bickel 2008, 2011). Such an analysis would expand the range of known splits of alignment patterns by subsystems in a language (as surveyed already by e. g., Dixon 1994), but would keep the five basic alignment types (ergative, accusative, neutral, tripartite, and horizontal)

clearly identifiable within each subsystem. And the analysis would allow the characterization and analysis of both hierarchical marking systems and systems with non-hierarchical co-argument sensitivity parallel to cases with well-established alignment conditions based on clausal properties such as tense, aspect, subordination, or polarity (Witzlack-Makarevich 2011).

But there is a fundamental challenge when one tries to spell out the basic idea: alignments can only be formulated for sets of roles under a single, general condition, e. g., {S, A} vs. {P} for all third person arguments. Now, co-argument sensitivity means that the relevant conditions are by definition *not* general across all arguments, e. g., the condition may be split between a third person P co-occurring with a first A as opposed to a third person P co-occurring with a third person A (as is indeed the case for instance in Ik, cf. the P_3 rows in Table 1). As a result, there is no immediate sense in which one could compare third person P arguments to third person S and A arguments.

In response to this challenge we suggest to compute what we call "exhaustive alignments", i. e., we retrieve all possible alignment patterns for each referential type under the condition of every possible co-argument (cf. Witzlack-Makarevich 2011). For this we first build all possible triads of the three argument roles, whereby for the A and P arguments we list all co-arguments that these arguments can occur with. In Ik, for example, a first person A argument (A_1) can co-occur with a second person and a third person P argument (P_2 and P_3); a first person P argument (P_1) can co-occur with a second person A and a third person A argument (A_2 and A_3). Combined with first person S arguments (S_1), this results in four triads of the three argument roles:[7]

(11)   Comparative triads for the first person argument in Ik
   a. $\langle S_1, A_1$ [with P_2], P_1 [with A_2]$\rangle$
   b. $\langle S_1, A_1$ [with P_2], P_1 [with A_3] $\rangle$
   c. $\langle S_1, A_1$ [with P_3], P_1 [with A_2] $\rangle$
   d. $\langle S_1, A_1$ [with P_3], P_1 [with A_3] $\rangle$

The full set of comparative triads in Ik is given in the first three columns of Table 2.

In a second step, we determine whether argument marking is the same or different for the three arguments in each triad, e. g., whether first person arguments receive the same marking when comparing S_1, A_1 [with P_2], and P_1 [with A_2]. This results in alignment statements for each triad (given in the last column of Table 2). Aggregating by referential category, we can then deduce proportions of individual basic alignment types for each such category. In the Ik

---

7 We ignore reflexive scenarios, such as $A_1{\rightarrow}P_1$, which are treated as intransitive in Ik (cf. König 2009: 178).

Table 2: Comparative triads for determining the alignment of Ik case marking.

Comparative triads			Alignment
S argument	**A argument**	**P argument**	
S_1	A_1 [with P_2]	P_1 [with A_2]	neutral
S_1	A_1 [with P_2]	P_1 [with A_3]	accusative
S_1	A_1 [with P_3]	P_1 [with A_2]	neutral
S_1	A_1 [with P_3]	P_1 [with A_3]	accusative
S_2	A_2 [with P_1]	P_2 [with A_1]	neutral
S_2	A_2 [with P_1]	P_2 [with A_3]	accusative
S_2	A_2 [with P_3]	P_2 [with A_1]	neutral
S_2	A_2 [with P_3]	P_2 [with A_3]	accusative
S_3	A_3 [with P_1]	P_3 [with A_1]	neutral
S_3	A_3 [with P_1]	P_3 [with A_2]	neutral
S_3	A_3 [with P_1]	P_3 [with A_3]	accusative
S_3	A_3 [with P_2]	P_3 [with A_1]	neutral
S_3	A_3 [with P_2]	P_3 [with A_2]	neutral
S_3	A_3 [with P_2]	P_3 [with A_3]	accusative
S_3	A_3 [with P_3]	P_3 [with A_1]	neutral
S_3	A_3 [with P_3]	P_3 [with A_2]	neutral
S_3	A_3 [with P_3]	P_3 [with A_3]	accusative

example, this means that for the first and second person, i. e., in all four triads $\langle S_1, A_1..., P_1...\rangle$ and all four triads $\langle S_2, A_2..., P_2...\rangle$, we get 50 % neutral and 50 % accusative alignment. For the third persons, i. e., all nine triads with $\langle S_3, A_3..., P_3...\rangle$, we obtain 66 % neutral and 33 % accusative alignment.

The same analysis can be generalized to hierarchical marking and – with some adjustment – to agreement. We will illustrate this by considering a system of hierarchical agreement in Plains Cree, covering both aspects. In order to show how the generalization works, we first demonstrate in the following that all cases of hierarchical marking can be analyzed as cases of co-argument sensitivity without any appeal to a hierarchy. Note that the reverse is not true: as we found in Section 3, not all cases of co-argument sensitivity can be analyzed in terms of hierarchical marking. Once hierarchical argument marking is analyzed in terms of co-argument sensitivity, its alignments can be computed in exactly the same way as we did for co-argument sensitive case marking in Ik.

Cases of hierarchical agreement can be analyzed in terms of co-argument sensitivity as follows. Take as an example prefixal agreement in Plains Cree, as illustrated in (4) above. Table 3 lists all logically possible scenarios (first column). The table also indicates which prefix is used with the respective scenario

Table 3: Co-argument sensitive distribution of Plains Cree verbal agreement prefixes; the argument marked with a prefix is printed in bold face.

	Scenario	Prefix	Marked argument	Unmarked co-argument(s)
(a)	**S_1**	*ni-* '1'	S_1	n. a.
(b)	**S_2**	*ki-* '2'	S_2	n. a.
(c)	S_3	none	none	n. a.
(d)	A_1→**P_2**	*ki-* '2'	P_2	A_1
(e)	**A_1**→P_3	*ni-* '1'	A_1	P_3
(f)	**A_2**→P_1	*ki-* '2'	A_2	P_1
(g)	**A_2**→P_3	*ki-* '2'	A_2	P_3
(h)	A_3→**P_1**	*ni-* '1'	P_1	A_3
(i)	A_3→**P_2**	*ki-* '2'	P_2	A_3
(j)	A_3→P_3	none	none	A_3 & P_3

(second column) and explicitly states the marked argument and its unmarked co-argument (third and fourth column, respectively). The table shows that prefix marking is not restricted to a particular role (S, A, and P can all be marked with the same prefix) and also makes clear that whether the arguments of one and the same referential type are marked or unmarked depends on their co-argument. Thus, a first person argument in a two-argument scenario is marked in rows (e) and (h), but unmarked in rows (d) and (f).

In Table 4 exactly the same procedure for establishing comparative triads as in Ik example in Table 2 is applied to Plains Cree prefixal agreement. For the second person we obtain neutral alignment for all triads (the four bottom rows in Table 4). For the first person, we find four different alignment types (the four top rows in Table 4). The third person is not marked in the prefix position and is therefore excluded here (but we eventually include it when analyzing all other argument markers, including suffixes).

Table 4: Comparative triads of the Plains Cree verbal agreement prefixes.

S argument	A argument	P argument	Alignment
S_1 (*ni-*)	A_1 (∅) [with P_2]	P_1 (∅) [with A_2]	horizontal
S_1 (*ni-*)	A_1 (∅) [with P_2]	P_1 (*ni-*) [with A_3]	ergative
S_1 (*ni-*)	A_1 (*ni-*) [with P_3]	P_1 (∅) [with A_2]	accusative
S_1 (*ni-*)	A_1 (*ni-*) [with P_3]	P_1 (*ni-*) [with A_3]	neutral
S_2 (*ki-*)	A_2 (*ki-*) [with P_1]	P_2 (*ki-*) [with A_1]	neutral
S_2 (*ki-*)	A_2 (*ki-*) [with P_1]	P_2 (*ki-*) [with A_3]	neutral
S_2 (*ki-*)	A_2 (*ki-*) [with P_3]	P_2 (*ki-*) [with A_1]	neutral
S_2 (*ki-*)	A_2 (*ki-*) [with P_3]	P_2 (*ki-*) [with A_3]	neutral

In many languages analyzed as showing hierarchical agreement, it is common to find many argument indexes on the verb (e. g., in the Algonquian languages): some index only person of one of the arguments, others only number, whereas still others combine the indexes of both arguments in a portmanteau morpheme. This complexity does not change the mechanism of determining alignments presented above: For each slot separately, all argument indexes on the verb can be listed in the same fashion as the two indexes in the Plains Cree prefix slot in Table 3, building comparative triads in the same way as in Table 4. The complete analysis of Plains Cree indexes includes one prefix slot (illustrated above), as well as four suffixal slots (Dahlstrom 1991). They all get a separate notation parallel to the one in Table 3 and Table 4 each. This mechanism equally applies to the indexes which only code number, animacy or obviation, as their distribution can be conditioned by co-arguments.

The representation of Plains Cree prefixal agreement in terms of co-argument sensitivity raises the question whether we do not lose a possibly important generalization, viz. the relevance of the 2 ≻ 1 ≻ 3 hierarchy for the marking in the prefix slot, and whether we do not make a seemingly simple picture unnecessarily complex. Under closer inspection however, it is unclear to what extent hierarchies are genuine generalizations and whether analyses in terms of hierarchies are indeed less complex than analyses in terms of co-argument sensitivity: The generalization of a 2 ≻ 1 ≻ 3 hierarchy in Plains Cree quickly loses its descriptive appeal once one moves away from prefixes and considers the entire agreement system. In order to account for the full distribution of agreement markers one needs to postulate at least three, partially conflicting hierarchies:[8]

(12)   Plains Cree hierarchies
    a. Plains Cree hierarchy I:
       2/1pINCL ≻ 1 ≻ 3
    b. Plains Cree hierarchy II:
       1p ≻ 1pINCL/2p ≻ 3 animate ≻ sSAP ≻ 3 inanimate
    c. Plains Cree Hierarchy III:
       SAP ≻ 3 proximate ≻ 3 obviative (≻3 further obviative)
       (Zúñiga 2006: 85–86)

These hierarchies are not only language-specific but specific to parts of the paradigm, e. g., to individual slots. Such idiosyncrasies are fairly common in the

---

8 Also see Macaulay (2005, 2009) for similar conclusions with respect to a range of further Algonquian languages and Mithun (2012) for a wider survey of hierarchical systems in genealogically unrelated languages of Northern California.

languages categorized as possessing hierarchical systems. Another example is provided by Aguaruna (Jivaroan, Peru): in order to analyze the distribution of case markers in this language Overall (2007, 2009) needs to posit the following Aguaruna-specific hierarchy:

(13)   1s ≻ 2s ≻ 1p/2p ≻ 3

Hierarchies like the ones in Plains Cree and Aguaruna capture the facts, but they do not explain anything more than an analysis in terms of co-argument sensitivity. As a result, the seeming generalization does not lead to more insights into the inner workings of these languages.

Further, once hierarchies are language-specific, it is difficult, often impossible, to compare them for reconstruction purposes or when exploring areal or universal distributions. Alignment statements, by contrast, are well-established notions for comparative purposes, they make it possible to compare proportions (or degrees) of, say, "ergativity" along with the effects of specific referential categories (for some recent applications, see Bornkessel-Schlesewsky et al. 2008; Bickel and Nichols 2009; Bickel et al. 2013, 2015a, 2015b, 2015c).

5 Case study: Hierarchies as determinants in Algonquian and Kiranti diachrony?

The previous discussion has shown that (a) systems of hierarchical marking can be represented in terms of co-argument sensitivity and that (b) this representation allows for consistent notions of alignment and for systematic comparison of alignments across referential categories and across languages. An immediate consequence of this is that we in fact do not need hierarchies for either descriptive or comparative purposes. Any generalizations that hierarchies seem to capture turn out to be more apparent than real because they are fraught with idiosyncrasies and peculiarities that leave little or no room for explanatory insights.

However, while hierarchies are not needed for descriptive or comparative purposes, it may still be possible that hierarchies play a systematic role in shaping agreement paradigms over time. Evidence for this would come from the observation that individual markers compete for the same slot and that there is a trend across an entire genealogical unit (e. g., a family like Algonquian) that all such competitions are resolved by a uniform hierarchy (e. g., the 2 ≻ 1 ≻ 3 hierarchy). If there is such a trend, it would be plausible to assume that this hierarchy played a shaping role in the history of the family and that it therefore reflects a deeper,

perhaps cognitive, reality. If there is no detectable systematic trend, this would suggest that hierarchies are obsolete artifacts not only for the purpose of describing and comparing languages, but also for explaining their diachronies.

We address this issue by way of a case study of Algonquian and Kiranti verbal paradigms because they are rich in co-argument sensitive markers that could be, and indeed sometimes are, analyzed in terms of hierarchies (Hockett 1966; DeLancey 1981; Michailovsky 1988; Dahlstrom 1991; Ebert 1991; Bickel 1995; LaPolla 2003; Zúñiga 2006). We focus on person categories, ignoring number and the inclusive/exclusive distinction because only person categories are equally relevant for both families and indeed recur in virtually all work on the referential hierarchies.

5.1 Data

Our sample includes 8 Algonquian languages and 16 Kiranti languages. For lack of sufficient data we limited the Algonquian survey to the paradigms of what is called the independent order, i. e., main clause forms. For Kiranti we limited the survey to the indicative mood paradigms and included both past and non-past tenses for those languages which have different person and/or number marking in the two paradigms (though this does not necessarily result in different rankings of person values, as we will see).

We considered only overt markers because consistent coding of zero morphemes is extremely complex when person markers interact with other categories such as number, tense, aspect, negation etc. as they regularly do in the analyzed data. (The Plains Cree prefix pattern, where simple person markers are in opposition to a zero morpheme is highly exceptional.) There tend to be a great number of possible analyses of zeros, and working through all options would constitute a separate research project. Furthermore, the choice between analytical options often directly depends on whether or not one assumes a hierarchy, the very issue we aim to test. In any event, if hierarchies have shaped paradigms diachronically, one would expect them to leave a systematic trace already among overt markers. Also, we excluded all portmanteau markers from our analysis, as they mark two arguments at the same time and therefore do not provide any evidence for a particular ranking between these arguments. Our analyses largely follow the language specialists' morpheme analyses in the literature, although we took the liberty of occasionally unifying and harmonizing different analyses. The agreement paradigms of these languages were coded in the AUTOTYP database of grammatical relations (Witzlack-Makarevich et al. 2011).[9]

---

9 The database and all analyses and the algorithm used to obtain the calculations reported below are available as Supporting Online Material at http://www.autotyp.uzh.ch/available.html.

For hypothesis testing we make assumptions about the topology of the Algonquian and Kiranti trees. These assumptions are made explicit in Figures 1 and 2. The Kiranti tree is based on the summary in Bickel and Gaenszle (2015). For Algonquian we test our hypothesis both on a flat tree that is traditionally posited (e. g., Mithun 1999) and a more structured tree that we construct on the basis of shared innovations that have developed in a west-to-east cline (Goddard 1994; Oxford 2015). In each tree, all branches have length 1 between the nodes that are posited in the literature even when the nodes are not relevant for our sample. For example, there is evidence for Kulung being a descendant of a "Khambu" node covering several closely related languages that are not included in our sample. Therefore, there is one additional node between "Central Kiranti" and the Kulung paradigms we selected here. This results in a total length of 2 between Central Kiranti and Kulung, the same length as between Central Kiranti and Bantawa, which develops via the node "Southern Kiranti".
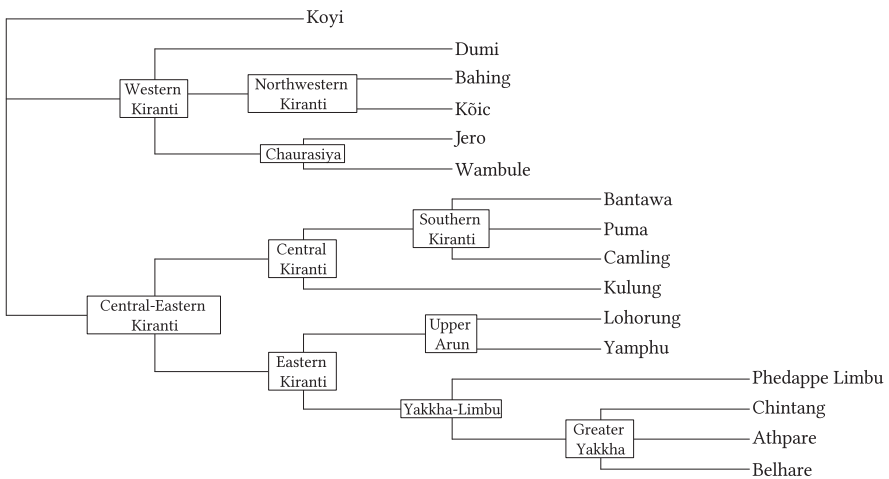

Figure 1: Genealogy of the Kiranti languages in the sample (Bickel and Gaenszle 2015).

5.2 Methods

For detecting hierarchical rankings of persons (first, second, and third) within each slot we apply the following algorithm:[10]

---

10 The algorithm is implemented as Python script and available as Supporting Online Material 5.

(a)

```
                                                            Passamaquoddy

                   ┌─────────────────┐
                   │Eastern Algonquian├──────────────────── Micmac
                   └─────────────────┘
                                                            Munsee

                                                            Eastern Ojibwa

                                                            Plains Cree

                                                            Arapaho

                                                            Blackfoot

                                                            Cheyenne
```

(b)

```
                                                            Passamaquoddy

                              ┌─────────────────┐
                              │Eastern Algonquian├───────── Micmac
                              └─────────────────┘
                      ┌──┐                                  Munsee
                      │3 │
                      └──┘                                  Eastern Ojibwa
                ┌──┐
                │2 │
                └──┘                          Cheyenne
          ┌──┐
          │1 │                                Plains Cree
          └──┘
                                              Arapaho

                                    Blackfoot
```
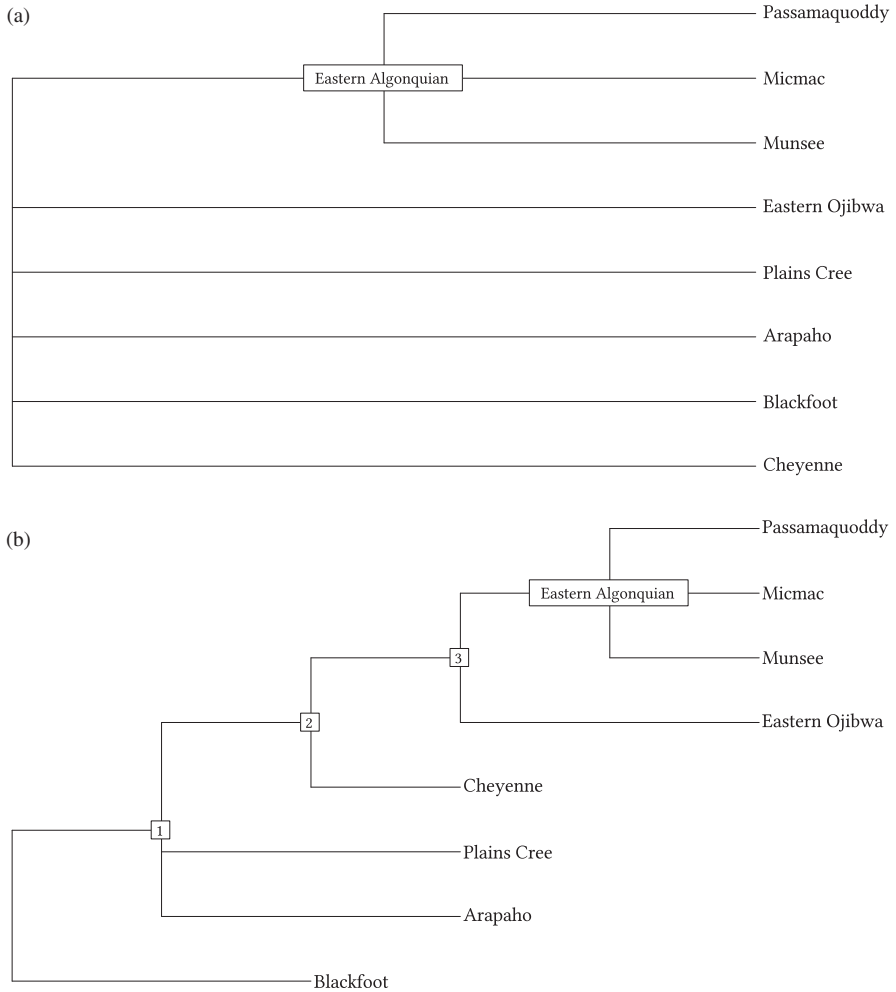
Figure 2: Two possible genealogies of Algonquian. (a) Traditional genealogy of the Algonquian languages in the sample (Mithun 1999). (b) Alternative genealogy of the Algonquian languages in the sample, derived from the west-to-east cline of shared innovations observed by Goddard (1994) and Oxford (2015). Subgroups are labeled by the numbers specifying innovations in Oxford (2015).

1. The input to the algorithm is a set of occurrences of agreement affixes with the indication of their respective slots and their co-arguments in the way illustrated for the Plains Cree prefixes *ki-* and *ni-* in Table 3 above. (For expository purposes, Table 3 only represents the person feature, but the

complete data set also includes the number, clusivity, and – for Algonquian languages – animacy features and obviation status, as these features can also condition the presence or absence of an agreement affix.)

2. Because evidence for hierarchical order means that markers compete for the same slot, the algorithm first compiles sets of markers which could potentially occur in a given slot in accordance with their person feature. This is done by putting together sets of person, number, clusivity, and obviation markers occurring in the same slot and then extracting the information about the person feature that they express from these sets. In the case of the Plains Cree prefix slot in Table 3 only two prefixes occur in this slot and thus only these two compete with each other. As the two markers express first and second person, only the ranking of these two person categories can be established in the prefix slot unlike in, say, Eastern Ojibwa, where the third person prefix also competes for the prefix slot. The remaining person markers (in our example, those expressing third person) cannot enter the slot and are hence irrelevant for the competition. Therefore, the algorithm ignores all contexts of occurrence which refer to these features, e. g., in Plains Cree, we do not need to consider the contexts where the co-argument is third person (rows (a), (c), (d), and (f) in Table 5).

Table 5: Filtering the Plains Cree verbal agreement prefixes.

	Marker	Argument	Co-argument	Filter output
(a)	ni- 1	A_1	P_3	ignore
(b)	*ki-* 2	A_2	P_1	consider for ranking
(c)	*ki-* 2	A_2	P_3	ignore
(d)	*ni-* 1	P_1	A_3	ignore
(e)	*ki-* 2	P_2	A_1	consider for ranking
(f)	*ki-* 2	P_2	A_3	ignore

3. The remaining contexts can be directly interpreted in terms of ranking. For instance, in our simplified Plains Cree example there are two remaining contexts (the contexts (b) and (e) in Table 5). In the context (b), a second person A acts on a first person P, and it is A which is marked (by the prefix *ki-* '2'). Thus, for this context, the prefix marking the second person wins over the prefix marking the first person and consequently this context provides a point in support of the 2 ≻ 1 ranking. In the second relevant context (e), a first person A acts on a second person P, and it is P which is marked (again by the prefix *ki-* '2'). This context provides another point in support of the 2 ≻ 1 ranking. Now, because in some languages number and

obviation status interact with the person ranking (cf. the hierarchies in [12]), the actual amount of the environments considered for ranking is higher and corresponds to all allowed person/number/obviation/animacy combinations, which in case of Plains Cree prefixes results in eight different contexts supporting the ordering 2 ≻ 1 in the prefix position.

Once all verb agreement slots of a language are evaluated in the way described above, we tabulate the counts of contexts per paradigm that support specific pairwise rankings (first and second, first and third, second and third). We then perform statistical analyses to determine whether or not a specific ranking (say, 1 ≻ 3) plays a significant role in shaping the paradigms of each family in diachrony.[11] We examine both scenarios in which each individual paradigm underwent its own independent history, and also scenarios where the paradigms are correlated and are simultaneously affected by a hierarchy principle. For assessing either scenario, we use two approaches.

The first approach, called here a *set-based approach*, takes each paradigm in a family to be the result of an independent diachronic trial (a particular instance of morphological evolution). As such, a paradigm has a certain number of contexts that support a given ranking hypothesis, e. g., there may be 6 out of 8 relevant contexts that support a 1 ≻ 3 ranking. We then test for each ranking whether the number of its supporting contexts significantly exceeds what can be expected under the null hypothesis of a plain chance process, i. e., that there are about as many contexts that support the ranking as there are contexts that contradict the ranking.[12] This results in *p*-values for each ranking hypothesis, i. e., probabilities that the observed counts result from the chance process assumed by the null hypothesis. These values are then collected from all paradigms in the family. Such a collection amounts to multiple testing of the same hypothesis, inflating the risk of false positives. In response to this, we correct the *p*-values using the Holm-Bonferroni correction (Holm 1979). Finally, we compute the proportion of paradigms that significantly support a given ranking hypothesis and take this proportion as an index of the extent to which the paradigms were shaped in their development by the given person ranking. Any proportion higher than 50 % would at least be suggestive; for robust evidence one would

---

11 See Supporting Online Material 6 for a richly annotated R script implementing the analysis and Supporting Online Material 7 for a PDF version of the script containing the annotations and the output.

12 Technically, our null hypothesis assumes a Bernoulli process with a 0.5 probability of success. Note that if there are no contexts in a paradigm that are relevant for a specific ranking, we assume that the null hypothesis is true.

expect the large majority of languages to show a significant preference for the same ranking.

The second approach, called here a *tree-based approach*, adapts methods from computational phylogenetics for estimating the extent to which the evolution of a trait (here rankings) is biased towards a certain optimum value, i. e., adapts to this value as the result of selection pressure (following Hansen 1997; Butler and King 2004 and using the model fitting algorithms implemented in Clavel et al. 2015). We take the proportion of counts that support a given ranking as a continuous trait that a paradigm can have, e. g., a value of $6/8 = 0.75$ in the example 8 above. We then fit two evolutionary models per family tree and per pairwise ranking hypothesis: first, a simple model of Brownian motion, where at each node in the tree, the proportion randomly changes in no specific direction, i. e., without a preference for or against a hypothesized ranking; the only parameter of this model is σ, the intensity of random diffusion in the process. The second model is what is known as an Ornstein-Uhlenbeck model and assumes that there is a preferred optimum proportion towards which paradigms evolve with certain strength in a family. In addition to σ, this model has parameters θ for the optimum value and α for the strength of attraction to this value. The parameters of both models are fitted via maximum likelihood.[13] The fitted models are then compared using likelihood ratio tests assuming a χ^2 distribution with degrees of freedom equal to the difference in the number of parameters, as well as with a sample-size corrected version of Akaike's An Information Criterion (AIC_c), which allows comparing the fit of the models by taking both their complexity and their likelihood into account: the model with the smallest AIC_c is the one that is simplest and at the same time still best fitting, hence preferable.

For model fitting, we set all lengths of branches between nodes in the trees to 1, as explained in Section 5.1. This assumes that each structural change in a paradigm requires the birth of a new language, i. e., a new node, and that rates of non-cladogenetic change elsewhere in the language (e. g., lexical change in cognate replacement) have no direct impact on the number of opportunities for restructuring of a paradigm. These assumptions are consistent with the research tradition in historical linguistics and also with more recent observations that structural change tends to be characterized by punctuated evolution (Dediu and Levinson 2012).

---

13 Since the parameter estimating procedures are not bounded to the [0, 1] interval that is inherent to the proportions we examine, we apply an arcsine square root transformation to the proportions before model fitting. See Supporting Online Materials 6 and 7 for details.

The two approaches have the same goal but differ in a number of ways. The set-based method has a relatively weak power of detecting relevant signals because the overall counts of relevant contexts are low and the method does not directly model diachronic trajectories. The advantage however is that the method does not assume that each diachronic process must have followed a tree but also allows for areal diffusion effects or wave models of a family (François 2014). The tree-based method has the advantage that it is much more powerful (because it depends less on the total counts) and that it directly models diachronic processes. The disadvantage is that the method only allows changes along a tree and not for example along the waves in a dialect landscape (unless such a landscape is modeled as a set of non-congruent trees).

Given this mix of advantages and disadvantages we apply both methods and expect that if a given ranking has an effect, it is detectable by at least one of the methods, and if the effect is strong, by both methods. Also, if the effects are so strong as to motivate hierarchical agreement as a fundamental principle in the history of a family, we expect the effects to leave detectable statistical signals in all systems in a family, across all paradigms. We apply the two methods to the entire families but in the case of Kiranti, we also examine possible signals within each of the major sub-groups, Western Kiranti (including Koyi) and Central-Eastern Kiranti (Figure 1). We do this because there is no reason to assume that an effect can only be detected in the deeper time span that the entire family represents. It could also be a more shallow effect. For Algonquian our sample is too small for testing the hypothesis within shallower branches.

5.3 Results

The aggregated output of the algorithm described in Section 5.2 is shown in Table 6 for the Kiranti languages and in Table 7 for the Algonquian languages (See Supporting Online Materials 3 and 4). The numbers in the tables are the counts of relevant individual slots in what we call here "competition contexts", i. e., contexts where a particular ranking of person values can be observed because an agreement marker wins the competition for a particular slot against another marker or against other markers.

As the shape and size of paradigms vary across languages, the bare numbers in the tables cannot be directly compared. What can be directly compared, however, is the preference for a particular ranking of two person features within a language. For example, in Plains Cree in ten competition contexts there is evidence for 2 ≻ 1 ranking, whereas in four contexts there is evidence for 1 ≻ 2 ranking. Eight of the ten contexts evidencing the 2 ≻ 1 ranking correspond to the

Table 6: Pairwise ranking of person values in the Kiranti languages.

Language	TAM	1 > 2	2 > 1	1 > 3	3 > 1	2 > 3	3 > 2
Athpare	IND	0	4	2	1	4	0
Bahing	NPST.IND	4	0	8	0	2	0
	PST.IND	4	0	8	0	2	0
Bantawa	IND	0	0	1	0	1	0
Belhare	IND	6	1	0	0	1	0
Camling	IND	2	1	8	0	13	0
Chintang	NPST.IND	9	6	14	0	3	0
Dumi	NPST.IND	5	0	0	0	2	0
Jero	IND	3	0	1	0	4	0
Kõic	NPST.IND	4	4	5	4	4	3
	PST.IND	7	4	5	4	3	4
Koyi	NPST.IND	8	0	17	0	4	0
	PST.IND	8	0	14	3	4	0
Kulung	NPST.IND	3	1	4	0	4	0
	PST.IND	3	3	4	0	4	0
Limbu	NPST.IND	0	0	0	0	6	6
	PST.IND	0	0	0	0	6	6
Lohorung	NPST.IND	0	0	4	6	0	1
	PST.IND	0	0	4	6	0	1
Puma	NPST.IND	2	6	7	2	12	1
	PST.IND	2	6	7	2	12	1
Wambule	IND	1	3	5	0	3	0
Yamphu	IND	3	3	1	3	2	1

Table 7: Pairwise ranking of person values in the Algonquian languages.

Language	1 > 2	2 > 1	1 > 3	3 > 1	2 > 3	3 > 2
Arapaho	0	6	21	9	20	0
Blackfoot	2	8	4	0	0	0
Cheyenne	2	12	28	2	16	2
Eastern Ojibwa	2	8	28	0	16	0
Micmac	10	2	56	30	48	6
Munsee	2	8	28	6	16	6
Passamaquoddy	0	8	24	12	16	12
Plains Cree	4	10	48	12	27	10

eight contexts of the paradigm where in the prefix slot the prefix *ki-* '2' wins over the prefix *ni-* '1', as illustrated by the data in (4) above. These are the four contexts with a second person (singular and plural) acting on a first person singular and first person plural exclusive and four contexts with the second

person (again, singular and plural) being acted upon by the first person singular and first person plural exclusive. Two more points for the 2 ≻ 1 ranking come from the distribution of the suffix *-nāwāw* '2p', which in the two contexts where the second person plural interacting with the first person singular wins over the competing suffix *-n* '1s/2s' (cf. Zúñiga 2006: 78). The four points in support of the 1 ≻ 2 ranking stem from the distribution of the suffix *-nān* '1pEXCL', which blocks the suffixes *-nāwāw* '2p' (in two contexts) and *-n* '1s/2s' (also in two contexts).

Tables 6 and 7 suggest that among first and second person either ranking (1 ≻ 2 or 2 ≻ 1) could show an effect while for first and third, and second and third persons only the ranking 1 ≻ 3 and 2 ≻ 3 are candidates for effects, as these rankings consistently outrank their inverse rankings. We therefore selected only the rankings 1 ≻ 2, 2 ≻ 1, 1 ≻ 3 and 2 ≻ 3 for further analysis. This is also in line with the common hypothesis that, if there is any ranking, it will rank first and second above third persons.

Table 8 summarizes the main results from both the set-based and the tree-based methods (for more detailed results, see the Supporting Online Materials 6 and 7). The results from models assuming combined or correlated evolution of paradigms in Kiranti are not included here because they never fitted better than models that separate the paradigms (see Supporting Online Material 7: Section 3 for the set-based and Sections 4.1.1.3 and 4.1.2.3 for the tree-based results). Also, the table omits the tree-based models with variable histories in Western vs. Central-Eastern Kiranti because they never fitted the data better than the simpler

Table 8: Summary of main results for each ranking where Tables 6 and 7 suggest possible evidence.

Ranking	Kiranti Non-Past		Kiranti Past		Algonquian	
	P	ΔAIC_c	*P*	ΔAIC_c	*P*	ΔAIC_c
1 ≻ 2	0 (W 0.17)	2.67	0 (W 0.20)	3.61	0	1.98 (9.42)
2 ≻ 1	0		0		0.25	
1 ≻ 3	0.12 (W 0.33, CE 0.20)	−4.89	0 (W 0.40, CE 0.11)	−0.07	0.62	5.60 (11.77)
2 ≻ 3	0.12 (W 0, CE 0.22)	−0.52	0.14 (W 0, CE 0.22)	6.22	0.62	5.60 (11.46)

P: Proportion of paradigms with significant preference for the respective ranking (set-based method); ΔAIC_c: Reduction in sample-size corrected AIC_c by a model with an optimum (Ornstein-Uhlenbeck) relative to a Brownian Motion model (positive values mean non-reduction, i. e., a Brownian Motion model fits better); "W" stands for the Western and "CE" for the Central-Eastern group. The values in brackets for Algonquian show the results from the alternative, more structured tree (Figure 2b).

models assuming a single history in the tree-based approach (see Supporting Online Material 7, Section 4.1.2). In the set-based approach, separately modeling the groups occasionally raises the evidence for a ranking, and whenever the proportion is different from a model without groups, the table reports the separate estimates.

The only robust evidence for any ranking is found for the 1 ≻ 3 ranking in Kiranti non-past paradigms ($\Delta AIC_c = -4.89$ and $LR = 11.6$, $df = 2$, $p < 0.01$), where we find selective pressure (with strength $\alpha = 0.78$) towards an optimum proportion ($\theta = 0.91$) of contexts showing this ranking. In past tense paradigms, a model with an optimum trend for this ranking does not fit better than a Brownian Motion model in terms of AIC_c reduction ($\Delta AIC_c = -0.07$), although a likelihood ratio tests suggests borderline evidence ($LR = 7.42$, $df = 2$, $p = 0.024$). This difference between the two paradigms is consistent with the observation that there is no evidence for models that assume correlated evolution between the paradigms (see Supporting Online Material 7, Section 4.1.1.3). Also, the evidence for the 1 ≻ 3 ranking in the non-past is not replicated by the set-based approach, where the ranking is supported only by 12–33 % of the languages in the family in the non-past and by 11–40 % in the past (and 44 % in the Western sub-group when modeling both tense paradigms together, see Supporting Online Material 7, Section 3). In no case does the proportion reach even half of the languages.

Apart from the 1 ≻ 3 ranking, there is no other candidate evidence for a particular ranking. Likelihood ratio tests suggest borderline evidence for 2 ≻ 3 rankings in the non-past paradigms ($LR = 7.23$, $df = 2$, $p = 0.027$), but the gain in model fit is negligible ($\Delta AIC_c = -0.52$). This is in line with the set-based approach, which does not reveal any further trends because no proportion exceeds 25 %.

For Algonquian, the set-based method suggests that 63 % of the paradigms show a significant preference for 1 ≻ 3 and 2 ≻ 3 rankings (Table 8). While this may point towards a diachronic bias, it could as well be a product of chance, since finding evidence in two thirds of 8 cases is far from significant (e. g., under a binomial test), and a closer inspection of Table 7 suggests that the results might be inflated by the higher overall count of competition contexts that are relevant for these two rankings. Consistent with this, there is no support for the two rankings from tree-based modeling, regardless of whether one assumes a flat or a structured tree topology (all $\Delta AIC_c > 0$ and all $p_{LR} > 0.1$, see Supporting Online Material 7, Section 4.2). The likelihood ratio tests approach significance for the 2 ≻ 1 ranking ($p = 0.057$ for the flat and $p = 0.064$ for the more structured tree) but AIC_c comparison does not support any evidence for the model with an optimum (see last column in Table 8).

5.4 Discussion

Overall, our results suggest that hierarchical rankings of person are unlikely to have systematically shaped the evolution of agreement paradigms in Kiranti or Algonquian. The only evidence we found for a ranking to have played a role is the 1 ≻ 3 ranking in Kiranti when using a tree-based approach. However, the ranking is not detectable with other methods, and this weakens the significance of these findings. More importantly, the evidence is limited to non-past tense paradigms; past tense paradigms do not support the ranking. This pattern is not what one would expect if a principle of hierarchical agreement would be a consistent force that shapes the morphosyntactic evolution of these languages.

6 Conclusion

In this paper we showed that apart from differential argument marking, referential properties affect argument marking in two ways: (a) through hierarchical marking, where markers compete for a slot and the competition is resolved by a hierarchy, and (b) through co-argument sensitivity, where the marking of one argument depends on the properties of a co-argument. While co-argument sensitivity cannot be analyzed in terms of hierarchical marking, hierarchical marking can be analyzed in terms of co-argument sensitivity. Hence, co-argument sensitivity is a notion that is needed for descriptive purposes while hierarchies are not needed. We further showed that once hierarchical effects are analyzed in terms of co-argument sensitivity, they allow examining alignment patterns relative to referential categories in exactly the same way as one can examine alignment patterns relative to referential categories in cases of other known hierarchy effects, such as differential argument marking. Hence, analyses in terms of co-argument sensitivity are suitable for comparative purposes as well. As a result, cases of hierarchical marking of any kind turn out not to present a special case in typologies of alignment, and there is no need for positing an additional non-basic alignment type like "hierarchical alignment".

While hierarchies are not needed for descriptive and comparative purposes we also cast doubt on their relevance in diachrony: for at least two families whose agreement systems are often cited as being based on hierarchies, Algonquian and Kiranti, we find only very limited evidence for the evolution of paradigms to be shaped by the way person categories are ranked. The only relatively robust evidence we find is for the ranking of first person over third person in Kiranti, but this evidence is limited to one method and one tense

paradigm and this makes it unlikely to reflect the effects of an underlying principle that would generally constrain diachrony towards hierarchical agreement in the family. This is consistent with the findings by Gildea and Zúñiga (this issue), who show that the evolution of apparent hierarchy effects in synchronic grammars follows a disparate set of reconstructable pathways and that there is no signal of a general hierarchy principle that would have shaped these pathways.

Our case study results are also in line with what emerges from a larger typological survey study which shows that person hierarchies play no significant role in the evolution of agreement paradigms (Bickel et al. 2015c). Our results are furthermore consistent with recent findings that referential hierarchies are much less important for argument coding in general (through agreement or case) than previously assumed (see Filimonova 2005; Bickel 2008; Phillipps 2013; Fauconnier and Verstraete 2014; Bickel et al. 2015b).

References

Aldridge, Edith. 2012. Antipassive in Austronesian alignment change. In Dianne Jonas, John B. Whitman & Andrew Garrett (eds.), *Grammatical change: Origins, nature, outcomes*, 332–346. Oxford: Oxford University Press.

Bickel, Balthasar. 1995. In the vestibule of meaning: Transitivity inversion as a morphological phenomenon. *Studies in Language* 19(1). 73–127.

Bickel, Balthasar. 2008. On the scope of the referential hierarchy in the typology of grammatical relations. In Greville G. Corbett & Michael Noonan (eds.), *Case and grammatical relations: Studies in honor of Bernard Comrie*, 191–210. Amsterdam & Philadelphia: John Benjamins.

Bickel, Balthasar. 2011. Grammatical relations typology. In Jae Jung Song (ed.), *The Oxford handbook of language typology*, 399–444. Oxford: Oxford University Press.

Bickel, Balthasar, Martin Gaenszle, Arjun Rai, Prem Dhoj Rai, Shree Kumar Rai, Vishnu S. Rai & Narayan P. Sharma (Gautam). 2007. Two ways of suspending object agreement in Puma: between incorporation, antipassivization, and optional agreement. *Himalayan Linguistics* 7. 1–18.

Bickel, Balthasar & Martin Gaenszle. 2015. First person objects, antipassives, and the political history of the Southern Kirant. *Journal of South Asian Languages and Linguistics* 2. 63–86.

Bickel, Balthasar, Giorgio Iemmolo, Taras Zakharko & Alena Witzlack-Makarevich. 2013. Patterns of alignment in verb agreement. In Dik Bakker & Martin Haspelmath (eds.), *Languages across boundaries: Studies in memory of Anna Siewierska*, 15–36. Berlin & Boston: De Gruyter Mouton.

Bickel, Balthasar & Johanna Nichols. 2009. Case marking and alignment. In Andrej Malchukov & Andrew Spencer (eds.), *The Oxford handbook of case*, 304–321. Oxford: Oxford University Press.

Bickel, Balthasar, Manoj Rai, Netra Paudyal, Goma Banjade, Toya Nath Bhatta, Martin Gaenszle, Elena Lieven, Iccha Purna Rai, Novel K. Rai & Sabine Stoll. 2010. Ditransitives and three-argument verbs in Chintang and Belhare (Southeastern Kiranti). In Andrej Malchukov, Martin Haspelmath & Bernard Comrie (eds.), *Studies in ditransitive constructions. A comparative handbook*, 382–408. Berlin & New York: De Gruyter Mouton.

Bickel, Balthasar, Alena Witzlack-Makarevich, Kamal K. Choudhary, Matthias Schlesewsky & Ina Bornkessel-Schlesewsky. 2015a. The neurophysiology of language processing shapes the evolution of grammar: Evidence from case marking. *PLoS ONE* 10(8). e0132819.

Bickel, Balthasar, Alena Witzlack-Makarevich & Taras Zakharko. 2015b. Typological evidence against universal effects of referential scales on case alignment. In Ina Bornkessel-Schlesewsky, Andrej L. Malchukov & Marc Richards (eds.), *Scales and hierarchies: A cross-disciplinary perspective*, 7–43. Berlin & Boston: De Gruyter Mouton.

Bickel, Balthasar, Alena Witzlack-Makarevich, Taras Zakharko & Giorgio Iemmolo. 2015c. Exploring diachronic universals of agreement: alignment patterns and zero marking across person categories. In Jürg Fleischer, Elisabeth Rieken & Paul Widmer (eds.), *Agreement from a diachronic perspective*, 29–52. Berlin & Boston: De Gruyter Mouton.

Bickel, Balthasar, Taras Zakharko, Lennart Bierkandt & Alena Witzlack-Makarevich. 2014. Semantic role clustering: An empirical assessment of semantic role types in non-default case assignment. *Studies in Language* 38(3). 485–511.

Bornkessel-Schlesewsky, Ina, Kamal Kumar Choudhary, Alena Witzlack-Makarevich & Balthasar Bickel. 2008. Bridging the gap between processing preferences and typological distributions: Initial evidence from the online comprehension of control constructions in Hindi. In Andrej Malchukov & Marc Richards (eds.), *Scales (Linguistische ArbeitsBerichte 86)*, 397–436. Leipzig: Institut Für Linguistik. http://www.uni-leipzig.de/~asw/lab/lab86/LAB86_Bornkessel_et_al.pdf.

Butler, Marguerite A. & Aaron A. King. 2004. Phylogenetic comparative analysis: A modeling approach for adaptive evolution. *The American Naturalist* 164. 683–695.

Clavel, Julien, Gilles Escarguel & Gildas Merceron. 2015. mvMORPH: An R package for fitting multivariate evolutionary models to morphometric data. *Methods in Ecology and Evolution* 6. 1311–1319.

Cowell, Andrew & Alonzo Moss Sr. 2008. *The Arapaho language*. Boulder, CO: University Press of Colorado.

Creissels, Denis. 2009. Ergativity/accusativity revisited. Paper presented at the 8th Biannual Conference of the Association for Linguistic Typology, Berkeley, July 24–28 http://www.deniscreissels.fr/public/Creissels-ergativity.pdf

Dahlstrom, Amy. 1986. *Plains Cree morphosyntax.* Berkeley, CA: University of California, Berkeley dissertation.

Dahlstrom, Amy. 1991. *Plains Cree morphosyntax.* New York: Garland.

Dediu, Dan & Stephen C. Levinson. 2012. Abstract profiles of structural stability point to universal tendencies, family-specific factors, and ancient connections between languages. *PloS One* 7. e45198.

DeLancey, Scott. 1981. An interpretation of split ergativity and related patterns. *Language* 57(3). 626–657.

Dixon, R. M. W. 1994. *Ergativity.* Cambridge: Cambridge University Press.

Ebert, Karen H. 1991. Inverse and pseudo-inverse prefixes in Kiranti languages: Evidence from Belhare, Athpare and Dungmali. *Linguistics of the Tibeto-Burman Area* 14. 73–92.

Fauconnier, Stefanie & Jean-Christophe Verstraete. 2014. A and O as each other's mirror image? Problems with markedness reversal. *Linguistic Typology* 18. 3–49.

Filimonova, Elena. 2005. The noun phrase hierarchy and relational marking: Problems and counterevidence. *Linguistic Typology* 9. 77–113.

François, Alexandre. 2014. Trees, waves and linkages: Models of language diversification. In Claire Bowern & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 161–189. London: Routledge.

Gildea, Spike. 1994. Semantic and pragmatic inverse: Inverse alignment and inverse voice in Carib of Surinam. In Talmy Givón (ed.), *Voice and inversion*, 187–230. Amsterdam & Philadelphia: John Benjamins.

Goddard, Ives. 1994. The west-to-east cline in Algonquian dialectology. *Papers of the Algonquian Conference* 25. 187–211.

Hansen, Thomas F. 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution* 51. 1341–1351.

Haspelmath, Martin. 2011. On S, A, P, T, and R as comparative concepts for alignment typology. *Linguistic Typology* 15. 535–567.

Haude, Katharina. 2009. Hierarchical alignment in Movima. *International Journal of American Linguistics* 75(4). 513–532.

Himmelmann, Nikolaus P. 2002. Voice in western Austronesian: An update. In Fay Wouk & Malcolm Ross (eds.), *The history and typology of western Austronesian voice systems*, 7–16. Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, The Australian National University.

Himmelmann, Nikolaus P. 2005. Tagalog. In K. Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar*, 350–376. London: Taylor & Francis.

Hockett, Charles F. 1966. What Algonquian is really like. *International Journal of American Linguistics* 32. 59–73.

Holm, Sture. 1979. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* 6. 65–70.

König, Christa. 2009. Ik. In Gerrit Jan Dimmendaal (ed.), *Coding participant marking: Construction types in twelve African languages*, 141–172. Amsterdam & Philadelphia: John Benjamins.

Kroeger, Paul. 1993. *Phrase structure and grammatical relations in Tagalog.* Stanford, CA: CSLI.

LaPolla, Randy J. 2003. Overview of Sino-Tibetan morphosyntax. In Graham Thurgood & Randy LaPolla (eds.), *The Sino-Tibetan languages*, 22–42. London: Routledge.

Macaulay, Monica. 2005. On the 2>1 prominence hierarchy of Algonquian. In Rebecca Shields
(ed.), *LSO Working Papers in Linguistics 5: Proceedings of WIGL 2005*, 1–24. Madison, WI:
University of Wisconsin-Madison.

Macaulay, Monica. 2009. On prominence hierarchies: Evidence from Algonquian. *Linguistic
Typology* 13(3). 357–389.

Mallinson, Graham & Barry Blake. 1981. *Language typology*. Amsterdam: North-Holland.

Michailovsky, Boyd. 1988. *La langue hayu* (Collection sciences du langage). Paris: Editions du
Centre National de la Recherche Scientifique.

Mithun, Marianne. 1999. *The languages of native North America*. Cambridge: Cambridge
University Press.

Mithun, Marianne. 2012. Core argument patterns and deep genetic relations. In Pirkko Suihkonen,
Bernard Comrie & Valery Solovyev (eds.), *Argument structure and grammatical relations: A
crosslinguistic typology*, 257–294. Amsterdam & Philadelphia: John Benjamins.

Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago, IL: The University of
Chicago Press.

Overall, Simon E. 2007. *A grammar of Aguaruna*: Melbourne: La Trobe University dissertation.

Overall, Simon. 2009. The semantics of clause linking in Aguaruna. In R. M. W. Dixon &
Alexandra Y. Aikhenvald (eds.), *The semantics of clause linking*, 167–192. Oxford: Oxford
University Press.

Oxford, Will. 2015. Patterns of contrast in phonological change: Evidence from Algonquian
vowel systems. *Language* 91(2). 308–357.

Phillipps, Maxwell. 2013. Ergative case attrition in Central Indo-Aryan. *Studies in Language* 37.
196–216.

Rose, Françoise. 2009. A hierarchical indexation system: The example of Emerillon (Teko). In
Patience L. Epps & Alexandre Arkhipov (eds.), *New challenges in typology: Transcending
the borders and refining the distinctions*, 63–83. Berlin & New York: Mouton de Gruyter.

Rose, Françoise. 2011. *Grammaire de l'émérillon teko, une langue tupi-guarani de Guyane
française* (Langues et Sociétés d'Amérique traditionnelle 10). Leuven: Peeters.

Schachter, Paul & Fe T. Otanes. 1972. *Tagalog reference grammar*. Berkeley: University of
California Press.

Sharma, Narayan P. 2014. *Morphosyntax of Puma, a Tibeto-Burman language of Nepal*. London:
School of Oriental and African Studies dissertation.

Siewierska, Anna. 1998. On nominal and verbal person marking. *Linguistic Typology* 2. 1–55.

Siewierska, Anna. 2005. Alignment of verbal person marking. In Martin Haspelmath, Matthew S.
Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*, 406–409.
Oxford: Oxford University Press.

Witzlack-Makarevich, Alena. 2011. *Typological variations in grammatical relations*: Leipzig:
University of Leipzig dissertation.

Witzlack-Makarevich, Alena, Lennart Bierkandt, Taras Zakharko & Balthasar Bickel. 2011.
AUTOTYP database on grammatical relations. Electronic database, University of Zürich
http://www.autotyp.uzh.ch/ (accessed 16 March 2016).

Zúñiga, Fernando. 2006. *Deixis and alignment: Inverse systems in indigenous languages of the
Americas*. Amsterdam & Philadelphia: John Benjamins.

Zúñiga, Fernando. 2007. From the typology of inversion to the typology of alignment. In Matti
Miestamo & Bernhard Wälchli (eds.), *New challenges in typology: Broadening the horizons
and redefining the foundations*, 199–221. Berlin & New York: Mouton de Gruyter.