# THE MORPHOLOGY AND SYNTAX OF ERGATIVITY:

# A TYPOLOGICAL APPROACH

by

Justin Rill

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Linguistics

Winter 2017

# THE MORPHOLOGY AND SYNTAX OF ERGATIVITY:

# A TYPOLOGICAL APPROACH

by

Justin Rill

Approved: _____
Benjamin Bruening, Ph.D.
Chair of the Department of Linguistics and Cognitive Science

Approved: _____
George H. Watson, Ph.D.
Dean of the College of Arts and Sciences

Approved: _____
Ann L. Ardis, Ph.D.
Senior Vice Provost for Graduate and Professional Education

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Gabriella Hermon, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Benjamin Bruening, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Peter Cole, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Jessica Coon, Ph.D.
Member of dissertation committee

# ACKNOWLEDGEMENTS

A big thank you to my family, whose unwavering support and encouragement was available when I needed it most. Finally, thank you to Maria Babsly, the love of my life, for her affection, support and perspective in good times and bad.

All errors are my own.

# TABLE OF CONTENTS

**Chapter**

# ABSTRACT

This dissertation seeks to answer three loosely-related questions regarding ergativity. First, what is the nature of the morphological phenomena known as ERG=GEN and ERG=POSS? After providing a thorough description with ample data from a wide range of sources, a typological survey of 40 ergative languages concludes that together, they are a fairly common phenomenon, occurring in over half the languages in the sample and across all types of ergative languages cross-linguistically. Furthermore, these ergative patterns of syncretism occur at a higher rate than their nominative and accusative counterparts do, according to an equivalent sample of 40 accusative languages. The next question is theoretical in nature: what is the best way to model these syncretic patterns in the grammar? ERG=GEN and ERG=POSS are shown to follow straightforwardly from existing theories of case and agreement; however, for those rarer, non-ergative patterns, it is argued that a fundamentally different process of 'recycling' takes place in a somewhat random fashion. Finally, what is the relationship between morphological ergativity and syntactic ergativity? In a marked departure from existing theories in the literature, the last major chapter in the dissertation argues that cross-linguistically, the primary function of ergative extraction asymmetries is actually disambiguation in Ā-extraction, thus weakening the close connection to morphological ergativity which has been previously claimed. Indeed, new data is presented showing that the alignment of extraction asymmetries is not always consistent with morphological alignment, increasing our current understanding of the relationship between these two components of grammar.

# LIST OF ABBREVIATIONS

| | | | | |
|---|---|---|---|---|
| 1 | first person | | INCL | inclusive |
| 2 | second person | | IND | indicative |
| 3 | third person | | INDEF | indefinite |
| A | transitive subject | | INDEP | independent clause |
| ABL | ablative | | INF | infinitive |
| ABS | absolutive | | INSTR | instrumental |
| ACC | accusative | | INT | interrogative |
| ADJ | adjective | | IRR | irrealis |
| AF | agent focus | | FIN | finite |
| ANTI | antipassive | | LOC | locative |
| AOR | aorist | | M | masculine |
| APPL | applicative | | N | noun |
| ASP | aspect | | NEG | negation |
| AUX | auxiliary verb | | NMLZR | nominalizer |
| AV | agent voice | | NOM | nominative |
| BEN | benefactive | | N.FUT | non-future |
| C | complementizer | | N.PAST | non-past |
| CAUS | causative | | O | object |
| COM | comitative | | OBJ | object verb agreement |
| CONJ | conjunction | | OBL | oblique |
| COP | copula | | P | preposition |
| CMPL | completive | | PAST | past |
| CNN | connective | | PART | participle |
| DAT | dative | | PERF | perfective |
| DECL | declarative | | PF | patient focus |
| DEF | definite | | PL | plural |
| DEIC | deictic | | PRED | predicate |
| DEM | demonstrative | | PRES | present |
| DEP | dependent clause | | PRET | preterite |
| DET | determiner | | PROG | progressive |
| DETRANS | detransitivizer | | PV | patient voice |
| DS | different subject | | REAL | realis |
| DU | dual | | RED | reduplication |
| DUR | durative | | REL | relativizer |
| ERG | ergative | | SG | singular |
| EV | evidential | | S | intransitive subject |
| EXCL | exclusive | | SS | same subject |
| F | feminine | | STAT | stative |
| FOC | focus | | SUB | subjunctive |
| FUT | future | | SUBJ | subject verb agreement |
| GEN | genitive | | TAM | tense-aspect-mood |
| IMP | imperative | | TOP | topic |
| IMPF | imperfective | | VOC | vocative |
| INC | incompletive | | WH | Wh-element |

# FOREWORD

The topic of this dissertation is ergativity. As befits such a broad and venerable topic, the dissertation consists of several academic explorations within ergativity. Some are more interrelated than others, but all still contain that common thread of ergativity binding them together. Furthermore, the tools of inquiry used in exploring this phenomenon belong in equal parts to morphology, syntax and typology. It is my hope that this dissertation serves as a successful model for joining these subfields together in the fruitful pursuit of linguistic knowledge.

The dissertation is divided into six chapters. In **Chapter 1**, I begin by defining several key concepts, such as ergative and accusative alignment. It is also useful at this point to clearly distinguish between morphological ergativity and syntactic ergativity, given that the bulk of the dissertation focuses on the former, but the last substantial chapter in the dissertation, Chapter 5, addresses the relationship between the two (more on that below). It is also in Chapter 1 that different types of morphological marking, such as head-marking and dependent-marking, are defined for future reference.

**Chapter 2** attempts to answer the first major question of the dissertation: what are ERG=GEN and ERG=POSS? A thorough description of these syncretic morphological phenomena is provided, with ample data from a wide range of languages. Then, a sample of 40 ergative languages is constructed, with the aim of providing some quantitative estimate of their prevalence cross-linguistically. The main findings from the ergative sample are as follows: together, ERG=GEN and

ERG=POSS are quite common among ergative languages, occurring in some shape or form in more than half of the languages in the sample (57.5%); furthermore, the data suggests that these patterns transcend major typological parameters since they are not confined to ergative languages of a particular type; finally, the data also suggests that more 'robustly' ergative languages are more likely to exhibit either ERG=GEN or ERG=POSS.

**Chapter 3** serves as the companion chapter to that which precedes it. An equivalent sample of 40 morphologically accusative languages is constructed to provide a basis for comparison with the results of Chapter 2. The results here are straightforward: after a descriptive portion that illustrates the phenomena in question, it is shown that the logically-possible counterparts to ERG=GEN and ERG= POSS in accusative languages actually occur at much lower rates in the sample (27.5% and 12.5%, respectively).

At this point, we take a brief pause from typology to answer the next question of central importance to the dissertation: what is the best way to model these patterns of morphological syncretism in the grammar? To that end, **Chapter 4** adopts a combined model of case and agreement, based on previous work by Marantz (1991), Bobaljik (2008) and Baker (2013), and I show that the more frequent ERG=GEN and ERG=GEN patterns can be straightforwardly accounted for within that framework. However, when it comes to the four other syncretic patterns attested in the data (ACC=GEN, ACC=POSS, NOM=POSS and ABS=POSS), the combined model makes poor predictions. As an alternative, it is proposed that these rarer patterns are generated not by extending rules of case and agreement from the clausal domain to the nominal one, but rather by an altogether different process of 'recycling', where existing case and agreement paradigms are re-used in the nominal domain for reasons of economy.

**Chapter 5** addresses the last major question in the dissertation: what is the relationship between morphological ergativity and syntactic ergativity? Answering this question from a typological perspective, I argue that the primary function of ergative extraction asymmetries is to disambiguate in potentially ambiguous Ā-extraction scenarios, illustrating with Wh-question and relative clause data from a wide range of languages. Pushing the theory even further, I present arguments that, in actuality, extraction asymmetries can be dissociated from morphological alignment, as ergative extraction asymmetries are found in morphologically accusative languages, and accusative extraction asymmetries are found in morphologically ergative languages in turn. I conclude the chapter by presenting quantitative data suggesting that the disambiguation hypothesis makes very good – but not perfect – predictions across a large sample of languages.

**Chapter 6** concludes with a summary of the findings, and suggestions for future research.

## Chapter 1

## INTRODUCING ERGATIVITY (AND OTHER USEFUL DEFINITIONS)

### 1.1  Introduction

One of the main goals of this dissertation is to explore and analyze two related morphological phenomena, ERG=GEN and ERG=POSS, whereby the morphology that encodes transitive subjects in the verbal domain overlaps with the morphology that encodes possession in the nominal domain. The first order of business is therefore to provide clear definitions of several key concepts that will be of central importance in the chapters to come. To that end, this chapter seeks to define ergative alignment, and how it differs from accusative alignment (Section 1.2). Furthermore, we will distinguish among different types of ergativity, namely morphological, syntactic and 'other' (Sections 1.3-1.5), before returning to morphological ergativity for an in-depth look at ergative dependent-marking (1.6.1), ergative head-marking (1.6.2), ergativity and 'mixed'-marking languages (1.6.3) and ergative pronouns/clitics (1.6.4). The chapter concludes with a discussion of marking types in the nominal domain (Section 1.7) and a definition of syncretism (Section 1.8).

### 1.2  What is ergativity?

Ergativity refers to an alignment pattern where the transitive object and the intransitive subject pattern together to the exclusion of the transitive subject. Ergative alignment patterns are often contrasted with accusative ones, where instead

the transitive subject and intransitive subject pattern together to the exclusion of the transitive object.[1] Borrowing the conventions from Dixon (1994), we can refer to the transitive subject, transitive object and intransitive subject as **A**, **O** and **S**, respectively. The ergative and accusative alignment patterns are schematized using this terminology in (1) below.

(1) *Schema of ergative* (L) *and accusative* (R) *alignment*



Note that the definition of ergativity above is broad enough to apply to a wide variety of linguistic phenomena. The rest of this chapter is devoted to exploring the various flavors of ergativity – **morphological**, **syntactic** and **otherwise** – and how they will be relevant in the chapters to come.

### 1.3 Morphological ergativity (a first look)

One of the most common places to find an ergative alignment pattern is in the morphology of a language. **Morphological ergativity** may therefore refer to case patterns such as in Tongan, where the morphology that encodes a transitive object is the same as that which encodes an intransitive subject. In (2a), the transitive object

---

[1] For discussion of ergative, accusative and other logically-possible alignment patterns, see Comrie (1978). See also Johns (2000), Aldridge (2008), Deal (2013), Coon & Adar (2013) and Coon, Massam & Travis (to appear) for more recent overviews of ergativity.

*e mango* 'the mango' is marked by the absolutive morpheme *'a*; in (2b), the same absolutive morpheme also marks the intransitive subject *Sione* 'John'. In contrast, a different marker is used for the transitive subject in (2a), the ergative morpheme *'e*.

Tongan (Otsuka 2000; p.50, p.50)

(2) a. *ka'e  kai  'e    Sione  'a    e    mango*
       PAST eat  ERG  John   ABS  DEF  mango
       'John ate the mango'

   b. *na'e  'alu  'a    Sione*
       PAST go    ABS  John
       'John went'

More commonly, absolutive case marking may be null, while ergative case marking is overt, as in Basque. In (3), the transitive object and the intransitive subject are realized with no obviously analyzable absolutive morpheme; however, the transitive subject is realized with an ergative suffix *-k*. This, too, is an example of morphological ergativity because the transitive object and the intransitive subject pattern together to the exclusion of the transitive subject.

Basque (Laka 1996; p.61, p.61)

(3) a. *ehiztari-a-k      otso-a    harrapatu  du*
       hunter-DEF-ERG  wolf-DEF  caught     AUX
       'the hunter has caught the wolf'

   b. *otso-a     etorri   da*
       wolf-DET  arrived  AUX
       'the wolf has arrived'

Instances of morphological ergativity such as these can be contrasted with morphologically accusative patterns, such as in Japanese. In (4), the transitive subject and the intransitive subject are realized with the same nominative marker *ga*; the transitive object, on the other hand, must be realized with the accusative marker *o*.

Japanese                                            (Kuno 1973; p.3, p.72)

(4) a. *John ga    Mary o    but-ta*
       John NOM Mary ACC hit-PAST
       'John hit Mary'

    b. *John ga ki-ta*
       John NOM come-PAST
       'John came'

In actuality, the Tongan and Basque examples in (2-3) are merely one type of morphological ergativity. As we will see, morphological ergativity can be instantiated in many other ways, which we will return to in Section 1.6. But for now let us define and discuss other types of ergativity, non-morphological ones, in order to to be clear about what kind of phenomena we will *not* talking about when we invoke the label 'ergativity' in Chapters 2, 3 and 4, which are primarily about morphological ergativity.

## 1.4 Syntactic ergativity

Unlike morphological ergativity, **syntactic ergativity** refers to syntactic phenomena in which the transitive object and intransitive subject pattern together to the exclusion of the transitive subject. Dixon (1972) first reported on such a phenomenon

4

in Dyirbal. In coordinate structures such as (5), only the absolutive-marked argument can be elided, producing an interpretation where the verb *banagan$^y$u* 'returned' can only have *nguma* 'father' as its subject.

Dyirbal                                                                (Dixon 1994; p.155)

(5) *nguma*      *yabu-nggu*    *buran banagan$^y$u*
     father.ABS mother-ERG saw     returned

     'mother saw father and (father) returned'

This is an ergative pattern because the transitive object of the first clause (O) is also the intransitive subject of the second clause (S). In other words, in this scenario, the O and S arguments are co-referential, but the A and S arguments cannot be. Furthermore, the facts in Dyirbal are in stark contrast with a language like English, where a similar sentence *mother saw father and returned* produces the opposite interpretation, i.e. one in which *mother* is the argument of *returned*.[2]

More commonly, the label 'syntactic ergativity' refers to **asymmetries in Ā-extraction** in which the transitive object and the intransitive subject pattern together to the exclusion of the transitive subject. For example, in Q'anjob'al, Wh-extracting a transitive object (6b) or an intransitive subject (6c) proceeds straightforwardly: the appropriate argument is simply converted into a Wh-form and then fronted. However, in cases of transitive subject Wh-extraction like in (6a), the verb

---

[2] For a discussion of morphological ergativity and how it relates to the notion of grammatical subject, see Anderson (1976) and Comrie (1978). They each conclude that in the majority of ergative languages, the properties that are associated with grammatical subjecthood do <u>not</u> correspond with either ERG or ABS case in a 1:1 manner. In fact, it is only in Dyirbal that these properties are consistently associated with one of these morphological cases (ABS).

<u>must</u> feature an additional morpheme *-on*, which is also known as the Agent Focus morpheme (Coon et al 2014).[3]

Q'anjob'al                                    (Coon et al 2014; p.213, p.192, p.192)

(6) a. *maktxel max-ach      il-on-i*
       who      ASP-2SG.ABS see-AF-INTR

       'who saw you?'

   b. *maktxel max y-il[-a']        naq winaq*
      who      ASP 3SG.ERG-see-TR CL   man

      'who did the man see?'

   c. *maktxel max way-i*
      who      ASP sleep-INTR

      'who slept?'

In this way, Q'anjob'al Wh-questions exhibit a syntactic type of ergativity because only in the case of transitive subject extraction (A) – and not in the case of transitive objects (O) or intransitive subjects (S) – must verbs be realized with accompanying Agent Focus morphology.[4]

Given that relativization is commonly considered a type of Ā-extraction (Chomsky 2000), ergative extraction asymmetries can also be manifested in the realm of relative clauses. For example, in West Greenlandic, the following pattern occurs: relativization of a transitive object (7b) or an intransitive subject (7c) is straightforward, with no corresponding additional morphology in the relative clause. However,

---

[3] For an overview of ergative extraction asymmetries, see Deal (2016) or Polinsky (to appear).

[4] Note in passing that Q'anjob'al also exhibits morphological ergativity, albeit of a type that we haven't discussed yet. Morphological ergativity in the system of head-marking is discussed in Section 1.6, but for now it suffices to point out that morphological and syntactic ergativity are by no means mutually exclusive.

only in the case of transitive subject relativization, the verb must first be converted to an antipassive, as indicated by the morpheme *-si*, in order to be grammatical (7a).

West Greenlandic                                    (Bok-Bennema 1991; p.69, p.69, p.239)

(7) a. <u>*piniartuq*</u>　　[　*nannu-mik*　　*tuqut-si-suq*　　]
　　　hunter.ABS　　　polar.bear-MOD kill-ANTI-PART
　　　'the hunter who killed the polar bear'

　　b. <u>*nanuq*</u>　　　[　*Piita-p*　*tuqu-ta-a*　　]
　　　polar.bear.ABS　　　Piita-ERG kill-PART-3SG
　　　'the polar bear that Piita killed'

　　c.　[　*ipis-suq*　　　]
　　　　　be.sharp-PART
　　　'the thing that is sharp'

Another example of syntactic ergativity in relative clauses occurs in Shipibo. In this language, one relativization strategy involves internally-headed relative clauses, where, as the name states, the head of the relative clause occurs within the relative clause itself.[5] An intransitive, internally-headed relative clause occurs in (8); it is internally-headed because *ainbo* 'woman' occurs within the relative clause itself (indicated with square brackets), but it is still interpreted as the head of the relative clause.

Shipibo                                                          (Valenzuela 2002; p.67)

(8) [ *mi-bé*　　<u>*ainbo*</u>　　*jo-a*　　] *-ra no-n*　　*onan-yama-ke*
　　2SG-COM woman.ABS come-PART　EV 1PL-ERG know-NEG-CMPL
　　'we don't know the woman who came with you'

---

[5] According to Valenzuela (2002, 2003), Shipibo makes use of several relativization strategies, including externally-headed, pre-nominal RCs, externally-headed post-nominal RCs, anaphoric pronoun RCs, and internally-headed RCs.

In a transitive, internally-headed relative clause, something very interesting occurs: only the transitive object can be interpreted as the head. For example, in (9), the only possible interpretation is one in which *bake* 'child' is the head of the internally-headed relative clause, and not *pitso* 'parakeet'.[6]

Shipibo                                                                                          (Valenzuela 2002; p.66)

(9) [ *pitso-n*        <u>*bake*</u>      *natex-a*    ] *-tonin-ra joshin*        *pike*
       parakeet-ERG child.ABS bite-PART   ERG-EV   banana.ABS eat-CMPL
       'the child the parakeet bit ate he banana'
     * 'the parakeet that bit the child ate the banana'

To summarize these facts, internally-headed relative clauses in Shipibo constitute another example of syntactic ergativity because only transitive objects and intransitive subjects can be interpreted as their heads; transitive subjects simply cannot act as heads in that specific configuration.[7]

One last area in which ergative extraction asymmetries are reported to occur is topicalization. For example, in Sm'algyax, basic word order is VSO, as in (10). However, arguments can be topicalized by moving them to a pre-verbal position. This syntactic operation falls under the rubric of ergative extraction asymmetries because only in the case of transitive subject Ā-extraction, a special topic marker *in* must occur (11a). In the case of transitive object or intransitive subject topicalization, that topic marker does not occur (11b-11c).

Sm'algyax                                                      (Mulder 1994, p.32, p.135, p.135, p.135)

---

[6] Throughout this dissertation, the heads of relative clauses will be <u>underlined</u> whenever possible.

[7] Note however, that they <u>can</u> act as heads in the other relative clause strategies available in Shipibo; again, see Valenzuela (2002, 2003) for more details.

(10) *yagwa-t huum =da   duus =a   hoon*
PRES-3 smell =ERG cat   =ABS fish

'the cat is sniffing the fish'

(11) a. *'yagay 'wii gyisiyaasg-at in-t   deentg =asga łgu   alasg-m*
instead great northwind-3 TOP-3 avenge ABS   little weak-CNN.ADJ
*yetsisk*
land-animal

'instead the great northwind avenged the little weak animal'

b. *waab-a        awaan nah   dzab-u*
house-CNN.NP DET    PAST make-1SG

'that's the house that I built'

c. *awta      uks-haytg-it    gi-sga      lax małiitg-m      kyoox*
porcupine toward-stand-3 DEM-CNN.P top green-CNN.ADJ grass

'it was porcupine who stood at the edge on the green grass'

In the examples (6-11) above, we have illustrated that ergative extraction asymmetries can occur in three different domains of Ā-extraction. Furthermore, it is not at all uncommon for a single language to exhibit such an asymmetry in all three areas. Gitksan is one such language. The triplets in (12-14) all illustrate the same pattern: extraction of the transitive subject in the (a) examples involves a special relativizer *ʔan*, while extraction of the transitive object (b) and the intransitive subject (c) does not.

Gitksan                                                      (Rigsby 1986; p.303, p.303, p.303)

(12) a. *naa ʔan=t               qay   łəmoo=s  Bruce*
who A.REL=3SG.ERG.DEP which help=CNN Bruce

'who helped Bruce?'

9

b. *naa=ł    ł<sup>w</sup>əmoo-yə-n*
   who=CNN help-TR-2SG.ERG

   'who did you help?'

c. *naa=ł    lim-ət*
   who=CNN sing-S.REL

   'who sang?'

Gitksan                                                  (Rigsby 1986; p.404, p.407, p.405)

(13) a. *kaʔ-ə-'y=ł              <u>kat</u>  [ ʔan-cak<sup>w</sup>-ə=ł           naks-t              ]*
        see-TR-1SG.ERG=CNN man   A.REL-kill-INCR=CNN spouse-3SG.POSS

        'I saw the man who killed his wife'

   b. *mał-tə-'y          lo:-tit    tim  k<sup>w</sup>ix<sup>w</sup>-ə=s       John=ł    <u>smax</u>  [*
      tell-TR-1SG.ERG OBL-3PL FUT shoot-TR=CNN John-CNN bear
      *łə-kaʔ-n           ]*
      DEF-see-2SG.ERG

      'I told them that John would shoot the bear that you saw'

   c. *tim  t'is-ə=ł       x̣aʔ=ł     <u>kat</u>=ł    [ pax̣-ət=kə           ]*
      FUT hit-TR=CNN slave=CNN man=CNN   run-S.REL=DIST

      'the slave will hit the man who ran'

Gitksan                                                  (Rigsby 1986; p.288, p.285, p.283)

(14) a. *niin       tim  ʔan=kit-əx̣-tiit*
        2SG.ABS FUT A.REL=ask-3PL.ERG.DEP

        'you're the one who will ask them'

   b. *ʔənaax=ł    k<sup>w</sup>ip-ə-y*
      bread=CNN eat-TR-1SG.ERG

      'it was bread that I ate'

10

c. *Mark=ł*    *pax̣-ət-əst*
   Mark=CNN run-S.REL-INTERACT

   'it's Mark who's running'

The fact that ergative extraction asymmetries can appear in the same form across all three domains isn't surprising, given that they are frequently modeled as part of the same underlying syntactic operation. What *is* surprising, however, is when asymmetries of this kind only occur in a subset of Ā-extraction domains. For example, in Inuktitut, only relative clauses exhibit the kind an ergative extraction asymmetry ergativity as described in this section (Yuan 2013). In Chapter 5, a new analysis is presented which attempts to explain these cross-linguistic differences; until then, ergative extraction asymmetries are shelved, as Chapters 2, 3 and 4 deal directly with morphological ergativity instead.

## 1.5 'Other' types of ergativity

Before returning to the topic of morphological ergativity, it may be useful to cover types of ergativity which are not typically considered morphological nor syntactic. The definition provided in Section 1.2 is broad enough to encompass a wide range of phenomena, just as long as, in some way, shape or form, the transitive object and the intransitive subject pattern together to the exclusion of the transitive subject. The purpose of this section is to identify patterns which are technically in ergative alignment, but which will not be explored in any depth in the course of this dissertation.

For example, in Nandi, only transitive objects (15b) or an intransitive subject (15c) may occur as genitive arguments of a nominalized verb. Genitive transitive subject in a nominalized verb clause are simply ungrammatical (15a).

Nandi (Creider & Creider 1989; p.54, p.54, p.55)

(15) a. * *ka-nyít-â:*           *ei:sét*    *â:p kipe:t a:p má:yyêk*
      PAST-annoy-1SG.OBJ drink.N of   Kibet of   beer

        'Kibet's drinking of beer annoyed me'

    b. *ka-nyít-â:*           *ei:sét*    *â:p má:yyê:k*
      PAST-annoy-1SG.OBJ drink.N of   beer

        'the drinking of beer annoyed me'

    c. *ka-nyít-â:*           *ei:sét*    *â:p kipe:t*
      PAST-annoy-1SG.OBJ drink.N of   Kibet

        'Kibet's drinking annoyed me'

Under the definition in Section 1.2, this too qualifies as an instance of 'ergativity', simply by virtue of the fact that the transitive object and the intransitive subject are patterning together to the exclusion of the transitive subject. It also seems fair to say that this should not count as an example of morphological ergativity, in that it does not hinge on case-marking, verb-agreement or any of the other instantiations of morphological ergativity described in Section 1.6 below; nor should this count as an example of an ergative extraction asymmetry (even though argument structure and possession are clearly regulated by the syntax). For these reasons, perhaps it is best to lump instances of ergativity such as these into a default **'other'** ergativity category.

Given that the definition of ergative alignment is so broad, it should come as no surprise that ergative patterns do actually 'pop up' like this in languages not traditionally thought of as ergative. For example, in French causative complement constructions, only transitive objects and intransitive subjects may occur as bare arguments; transitive subjects must occur with dative prepositional morphology (16).

It is also worth pointing out that unaccusative verbs have sometimes been labeled "ergative" (Cinque 1995); this is likely due to the fact that this type of verb features both a transitive (17a) and an intransitive (17b) frame, and the transitive object in the former corresponds to the intransitive subject in the latter.

French                                                    (Bobaljik & Branigan 2006; p.58, p.58)

(16) a. *Luc a    fait    acheter un livre aux    étudiants*
        Luc has made buy.INF a   book to.the students
        'Luc made the students buy a book'

   b. *Luc a    fait    travailler les  étudiants*
        Luc has made work.INF  the students
        'Luc has made the students work'

Italian                                                    (Cinque 1995; p.208, p.208)

(17) a. *il   capitano affondò la   nave*
        the captain   sank    the boat
        'the captain sank the boat'

   b. *la   nave affondò*
        the boat sank
        'the boat sank'

It is worthwhile to identify that examples such as those in (15-17) are technically examples of 'ergative alignment'. At the same time, we must point out that this type of ergativity is not at all the focus of this dissertation.[8] Chapters 2, 3 and 4

---

[8] Patterns that fall under the default rubric of 'other' ergativity are simply bound to occur rather frequently in natural language. Grammars are vast, vast systems of rules and patterns; and given three primitive elements A, S and O, simply by chance patterns that lump S and O together to the exclusion of A are bound to occur. Ergativity's definition hinges on "patterning together", which

focus exclusively on ergative alignment in morphology, while Chapter 5 explores the relationship between morphological ergativity and syntactic ergativity (specifically, ergative extraction asymmetries).

## 1.6  Morphological ergativity revisited

### 1.6.1  Ergative dependent-marking

Since the bulk of this dissertation focuses on morphological ergativity, we now return to explore this phenomenon in more detail. Recall from Section 1.3 that morphological ergativity may entail a pattern in a language's system of case marking – also known as **dependent-marking** – like in Tongan (18).

Tongan                                                              (Otsuka 2000; p.50, p.50)

(18) a. *ka'e  kai  'e    Sione  'a    e      mango*
PAST eat ERG John  ABS DEF mango

'John ate the mango'

　　b. *na'e  'alu  'a    Sione*
PAST go    ABS John

'John went'

### 1.6.2  Ergative head-marking

In addition, morphological ergativity may also manifest in a language's system of verbal **head-marking**, which is the morphology that encodes the grammatical

---

is itself very vague. Within any given sentence, there are bound to be numerous ways in which different elements "pattern together". In fact, in the French and Italian examples (16-17) above, both of which have been identified as containing instances of 'ergativity', it is easy to find equivalent accusative patterns as well. For example, in (16), both A and S "pattern together" in the sense that they occur sentence-finally. In (17), both A and S "pattern together" in the sense that they occur pre-verbally... and so on.

role of arguments on lexical verb heads (and occasionally auxiliary verb heads as well).[9] For example, in Kaqchikel, a single morphological paradigm – also known as Set B in Mayan linguistics – consistently encodes both transitive objects in (19a-19b) and intransitive subjects in (19c-19d). A different set of morphemes, known as Set A, encodes only transitive subjects, as in (19a-19b). Note that in Set B, 3SG arguments are morphologically null, but they do condition allomorphy of the preceding imperfective prefix *y-*, which is then realized as *n-* (Brown et al 2006).

Kaqchikel                                      (Brown et al 2006; p.49, p.85, p.29, p.29)

(19) a. *n-in-tz'ët*
    IMPF-1SG.ERG-see
    'I see her/him'

   b. *y-i-ru-tz'ët*
    IMPF-1SG.ABS-3SG.ERG-see
    'she/he sees me'

   c. *n-wär*
    IMPF-sleep
    'she/he sleeps'

   d. *y-i-wär*
    IMPF-1SG.ABS-sleep
    'I sleep'

Similar to the dependent-marking system of Basque in (3), ergative alignment in head-marking may be characterized by a null absolutive paradigm. Indeed, in Päri, verbs only feature ergative head-marking. In (20a), the morpheme *-ɛ̀* encodes

---

[9] Nichols (1986) originally introduced the grammatical parameters of dependent-marking and head-marking.

the 3SG transitive subject; however, no verb morphology correspondingly encodes transitive objects in (20a) or intransitive subjects in (20b).[10]

Päri                                                         (Andersen 1988; p.293, p.292)

(20) a. *ùbúr dháagɔ̀ á-yáaŋ-ɛ̀*
       Ubur woman CMPL-insult-3SG.ERG
       'Ubur insulted the woman'

   b. *ùbúr á-túuk*
      Ubur CMPL-play
      'Ubur played'

Thus a nice parallel emerges between ergative dependent-marking and ergative head-marking. On the one hand, both ergative and absolutive morphology may be overt; more commonly, however, absolutive morphology is null, and only the ergative morphology is overt. To my knowledge, the logically-possible inverse is not attested in the realm of dependent-marking: there does not appear to be a language with overt absolutive case morphology, but no corresponding overt ergative case morphology. However, there is at least one language that features only overt absolutive morphology it its system of head-marking: Hindi. According to Mahajan (1990), in Hindi perfective clauses, verbs and auxiliaries agree with the transitive object and the intransitive subject in gender, but not with the transitive subject, as in (21).

Hindi                                                         (Mahajan 1990; p.73, p.73)

(21) a. *Raam   baazaar gayaa*
       Ram.M market  go.PAST.M.SG
       'Ram went to the market'

---

[10] The above observations only apply to SOV constructions in Päri. In OVS constructions, ergative dependent-marking is observed instead. See Andersen (1988) for more details.

16

b. *Raam-ne    roTii    khaayii    thii*
   Ram.M-ERG bread.F eat.PERF.F be.PAST.F

'Ram had eaten bread'

### 1.6.3   Ergativity and 'mixed'-marking

In addition to languages which are morphologically ergative, but which feature *either* dependent-marking *or* head-marking, there are also languages which feature both. Indeed, the two are not mutually-exclusive. Hindi is actually one such language. Notice that in (21), transitive objects and intransitive subjects do not feature any kind of overt dependent-marking; however, the transitive subject *Raam* in (21a) features the ergative morpheme *-ne*.

For lack of a previously-existing term, I have labeled these languages **'mixed'-marking** languages. To be clear, these are languages that simultaneously feature dependent-marking and head-marking. Another example comes from Sm'algyax, which was already introduced in the discussion of syntactic ergativity in Section 1.4. The data in (22) illustrates the ergative system of dependent-marking in Sm'algyax, where =*da* encodes the transitive subject, and =*a* encodes the transitive object and intransitive subject. According to Mulder (1994), =*da* and =*a* are clitics that 'lean' on the phonological word to the left, while encoding the argument to their right.

The data in (23) illustrates the ergative system of head-marking in Sm'algyax: the verbal suffix *-a'nu* encodes both 1SG transitive objects and 1SG intransitive subjects, while a different suffix *-t*, in a different syntactic position, encodes 3SG transitive subjects. Example (22a) in particular illustrates that both systems are perfectly compatible: both the case clitics =*da* and =*a* and the 3SG ergative head-marking *-t* on the TAM marker occur in the same sentence.[11]

---

[11] Note, however, that the transitive subject *ol* 'bear' in (23a) is not marked with the =*da* ergative

Sm'algyax                                            (Mulder 1994; p.32, p.32, p.70, p.57)

(22) a. *yagwa-t*      *huum =da*    *duus =a*     *hoon*
       PRES-3.ERG smell =ERG cat    =ABS fish

       'the cat is sniffing the fish'

    b. *yagwa hadiks =a üüla*

       PRES swim =ABS seal

       'the seal is swimming'

(23) a. *na-t*        *'niidz-a'nu*    *ol*
       PAST-3.ERG see-1SG.ABS bear

       'the bear saw me'

    b. *dm k'yeexg-a'nu*
       FUT run.away-1SG.ABS
       'I'm going to run away'

Another 'mixed'-marking language is Paumarí, with coexisting systems of ergative dependent-marking and ergative head-marking. The sentences in (24a-24b) are transitive clauses, where the transitive subject NP is suffixed with the ergative dependent-marking *-a*; in addition, the verbal prefix *bi-* encodes this transitive subject on the verb. The post-verbal transitive object is unmarked by case, but a suffix on the verb must agree with it in gender. In the intransitive clauses (24c-24d), the sole argument is also unmarked by case, and it too must agree in gender with that

clitic as expected. According to Mulder (1994), the use of these clitics is conditioned by a wide variety of factors, such as tense and aspect, indicative *vs* subjunctive mood, common noun *vs* proper noun, and their apparent optionality in casual speech. For more details, see Mulder (1994), which is entirely devoted to discussing all facets of ergativity in Sm'algyax.

same verbal suffix, making it a form of absolute head-marking similar to Hindi above.

Paumarí                           (Chapman & Derbyshire 1986; p.164, p.250, p.195, p.186)

(24)  a. *Dono-a*  *bi-ko'diraha-'a-ha*        *ada*  *isai hoariha*
        Dono-ERG 3SG.ERG-pinch-ASP-THEME.M DEM.M child other
        'Dono pinched the other boy'

      b. *mamai-a*  *bi-n-oba-hi*              *ida*  *ava pa'itxi*
        mother-ERG 3SG.ERG-CAUS-chop-THEME.F DEM.F tree small
        'Mother chopped down a small tree'

      c. *ihamahi-'a-ha*      *ada*   *Dono*
        be.angry-ASP-THEME.M DEM.M Dono
        'Dono was angry'

      d. *noki-a-hi*          *ida*  *gora*
        see-DETRANS-THEME.F dem.f house
        'the house is visible'

To recap, Hindi, Sm'algyax and Paumarí are all examples of a 'mixed'-marking language because they happen to simultaneously feature dependent-marking and head-marking. It also happens to be the case that in these three languages, both the system of dependent-marking and the system of head-marking are in ergative alignment. Of course, there exist 'mixed'-marking languages on the nominative-accusative side, with consistently accusative dependent-marking and head-marking as well; English qualifies as one of these languages, as illustrated in (25).

English

(25)  a. *he sees you*

      b. *you see him*

c. *he runs*

d. *you run*

There are also some 'mixed'-marking languages where the system of dependent-marking is in ergative alignment, but the system of head-marking is in accusative alignment. For example, in Burushaski, the system of dependent-marking is in ergative alignment, because the transitive subject is marked with the *-e* case-marker, and the transitive object and the intransitive subject are unmarked. However, a close look at the head-marking morphology on the verb reveals that the same 'slot' – the suffix immediately following the verb root – encodes both the transitive subject *hamal* 'neighbors' in (26a), and the intransitive subject *u* 'they' in (26b). In this way, Burushaski is a 'mixed'-marking language, but it features ergative dependent-marking and accusative head-marking.

Burushaski                                                                  (Munshi 2015; p.29, p.61)

(26) a. *hamal-e         ḍarbeṣ  i-cʰarkan-uman*
        neighbour-ERG Darbes 3SG.M.OBJ-beat-3PL.SUBJ
        'the neighbors beat Darbes'

    b. *u     gaarc-uman*
        3PL run-3PL.SUBJ
        'they ran'

Curiously, the inverse configuration from that in Burushaski does not occur: there does not appear to be any language with accusative dependent-marking, but ergative head-marking.[12] To summarize our findings so far, there are three possibilities for a language's system of dependent-marking: it may be in ergative alignment, it

---

[12] The original observation is due to Anderson (1985). For a syntactic account of this typological fact, see Bobaljik (2008). See also the extended discussion of Bobaljik (2008) in Chapter 4.

may be in accusative alignment, or the language may simply lack dependent-marking altogether. The exact same three possibilities apply to head-marking as well. In total, this yields 9 logically-possible languages in terms of marking strategies, eight of which are attested in the literature. These are represented in (27).

(27)  *The nine logically-possible languages by marking type*

| dependent-marking | head-marking | example |
|---|---|---|
| ACC | ACC | English |
| ACC | ERG | unattested |
| ACC | - | Maori |
| | | |
| ERG | ACC | Burushaski |
| ERG | ERG | Sm'algyax |
| ERG | - | Tongan |
| | | |
| - | ACC | Swahili |
| - | ERG | Kaqchikel |
| - | - | Mandarin |

Importantly, for the purposes of the survey of 40 ergative languages in Chapter 2, a language will be considered ergative if either its system of dependent-marking or its system of head-marking (or both) are in ergative alignment. This means that 'mixed'-marking languages like Burushaski are considered ergative languages for the purposes of that survey, despite having accusative head-marking. It also means that its inverse, the only unattested language type in (27), *would* count should one be discovered in the future.

### 1.6.4   Pronouns and clitics

In the preceding section, we have defined ergativity in terms of head-marking and dependent-marking, and therefore defined what will 'count' as an ergative language in the typological survey of Chapter 2. However, the topic of pronouns and clitics requires special discussion before moving forward, because *a priori* it is not always clear how to categorize them in terms of dependent-marking and head-marking.

In some languages, pronouns obviously have the same morphological and syntactic distribution as lexical NPs. For example, in Burushaski, the grammar straightforwardly concatenates the ergative suffix *-e* to pronouns as well as nouns; in addition, this free pronoun occurs in the same syntactic position as lexical nouns (compare 26a and 28a). Therefore, in this particular instance, we can conclude that Burushaski pronouns participate in the system of dependent-marking just like ordinary lexical NPs.

Burushaski                                                            (Munshi 2015; p.24, p.15)

(28)  a.  *in-e*         *a-yaar-um*         *pen cʰu-mi*
          3SG-ERG 1SG-under-from pen take-3SG.SUBJ
          'he took the pen from me'

      b.  *in*    *gucʰar-imi*
          3SG walk-3SG.SUBJ
          'he walked'

Alternatively, a language's pronominal system may have dedicated ergative and absolutive forms that aren't analyzable in terms of the ergative or absolutive morphology used for lexical NPs. According to Haspelmath (1993), in Lezgian, a noun's ergative form is derived by concatenating a lexically-conditioned ergative suffix – also known as the "oblique stem" – as in (29a). However, ergative forms

of pronouns are not derived by concatenation of the ergative suffix, but rather they are idiosyncratic and morphologically simple (29b). Nevertheless, because nouns and pronouns appear to have the same syntactic distribution, we can still consider the latter to be an idiosyncratic part of Lezgian's system of dependent-marking for the purposes of classification in the survey.

Lezgian                                                                 (Haspelmath 1993; p.294, p.140)

(29) a. *Alfija-di    maǧala kĵe-na*
        Alfija-ERG article   write-AOR
        'Alfija wrote an article'

    b. *Farida, Farida, wuna    ana   wuč iji-zwa*
        Farida  Farida  you.ERG there what do-IMPF
        'Farida, Farida, what are you doing there?

So the pronominal systems in languages like Burushaski and Lezgian are still considered examples of ergativity within the realm of dependent-marking. On the other hand, sometimes a language's pronouns (or pronominal clitics) simply do not have the same morphological or syntactic distribution as lexical NPs. In Seediq, Holmer (1996) reports that pronominal clitics have a restricted distribution: they must follow the "first verbal element" in their clause, which may be a verb or a TAM marker (30).

Seediq                                                                  (Holmer 1996; p.36; p.60; p.69, p.69)

(30) a. *wada =daha     puqun damac kiya*
        PRET =3PL.ERG eat.PF food    that
        'they ate up that food'

23

b. *egu   snlhayan       =mu         han   menaq  =ku          Hori*
   much learn.LF.PRET =1SG.ERG when stay.AF =1SG.ABS Puli

   'I have learned a lot while I have been staying in Puli'

c. *wada  =ku           =na          qtaun*
   PRET =1SG.ABS =3SG.ERG see.PF

   'he saw me'

d. *prayun          =mu          bunga          nii*
   cook.PF.PRES =1SG.ERG sweet.potato this

   'I shall cook these sweet potatoes'

In situations such as these, where the pronoun's distribution is extremely restricted to verbs and verb-like elements, it is best to analyze these pronouns as instances of head-marking (at least for the purposes of our survey in Chapter 2). As one last example, let's look at Sorani Kurdish. McCarus (2009) reports a system of clitics that is in ergative alignment, as in (31a-31b). By default, if there is nothing else in the clause, these pronouns will cliticize to the verb root. However, McCarus also reports that it is possible for these pronouns to cliticize to other, non-verbal elements in the clause. In (31c), the ergative 3PL clitic has attached to the adjective *pān* 'wide', and not the verb. Due to data such as these, which illustrate a wider range of syntactic distribution, Sorani Kurdish pronominal clitics are instead considered part of the system of dependent-marking in the survey of Chapter 2.

Sorani Kurdish                                              (McCarus 2009; p.617, p.609, p.616)

(31) a. *bīnī-yān-in*
        see-3PL.ERG-2PL.ABS

        'they saw you (PL)'

b. *hāt-in*
   come-2PL.ABS

   'you (PL) came'

c. *kay   pān-yān       a-ka-n*
   when wide-3PL.ERG IMPF-do-3PL.ABS

   'when will they widen them?'

To summarize this section, it is not always clear how to analyze pronouns and clitics, since their exact grammatical properties can vary greatly from language to language, as can their description form author to author.[13] However, for the purposes of this dissertation, we will consider pronouns and clitics with a restricted distribution – restricted to verbs, auxiliaries and TAM markers – as part of a language's system of head-marking. On the other hand, if some pronouns and/or clitics show a freer distribution, beyond these verbs/auxiliaries/TAM markers, then we shall consider them part of that language's system of dependent-marking.

## 1.7   Marking types in the nominal domain

Chapter 2 consists of an in-depth examination of a particular morphological pattern, one in which the morphology that encodes transitive subjects in the verbal domain (ERG) is identical to that which encodes possession in the nominal domain (either GEN or POSS). Section 1.6 provided a detailed exploration of morphological ergativity in the verbal domain; now it behooves us to do introduce and define the different facets of morphological possession in the nominal domain.

Fortunately, the concepts of dependent-marking and head-marking defined in the verbal domain can easily be transposed to the noun phrase. For example, in

---

[13] For more discussion on the morphosyntactic differences between clitics and head-marking, see Kramer (2014).

the same way that ergative case morphology on the transitive subject constitutes an example of dependent-marking, so does genitive case morphology on a lexical NP possessor. In Trumai, the transitive subject is marked with ergative dependent-marker *-k*, as in (32a). In the same way, in the nominal domain, the possessor of a noun can be marked with genitive dependent-marker *-kate*, as in (32b).

Trumai                                                          (Guirardello 1999; p.257, p.76)

(32) a. *ine-k      atlat mapa*
        3SG-ERG pan   break
        'he broke the pan'

     b. *hakew-kate   tahu*
        Raquel-GEN knife
        'Raquel's knife'

The same parallelism holds in Ku Waru. In a transitive clause, the subject is marked with the ergative suffix *-ni* (33a); in a noun phrase, the possessor is marked with the genitive suffix *-nga* (33b).[14]

Ku Waru                                                      (Merlan & Rumsey 1991; p.340, p.338)

(33) a. *na-ni     kera laim-yl      tud*
        1SG-ERG bird cassowary-DEF hit.PERF.1SG.SUBJ
        'I killed the cassowary'

     b. *Don-nga lku    si        nyim*
        Don-GEN house crowdedness be.PERF.3SG.SUBJ
        'Don's house was crowded'

---

[14] Note also that in both Trumai and Ku Waru, the parallelism between the verbal and nominal domains extends beyond marking type to word order itself. In other words, in both languages, the verbal arguments precede the verb in the clause, and the nominal 'argument' (i.e. the possessor) precedes the possessed noun in the noun phrase. This phenomenon is referred to as headedness, and it is discussed in more detail in Chapter 2.

Thus, dependent-marking can just as easily be used to describe morphological phenomena in the nominal domain. As expected, so can head-marking. For instance, in Tenetehára, agreement morphology on the verb encodes arguments in the verbal domain (34a); in the same way, in the nominal domain, possessive morphology encodes possessors directly onto the possessed noun itself (34b). In Tiriyó, too, the same parallelism holds: head-marking encodes arguments in the verbal domain (35a), just as possessive head-marking encodes possessors in the nominal domain (35b).

Tenetehára                                                     (Bendor-Samuel 1978; p.159, p.106)

(34) a. *u-zuka     omo    taʔyraʔi kwehe we*
        3.ERG-kill certain son       long    ago
        'he killed a certain son long ago'

     b. *Mane    re-miʔu*
        Manoel 3.POSS-food
        'Manoel's food'

Tiriyó                                                         (Meira 1999; p.144, p.144)

(35) a. *ji-pawana     n-ene              wëri*
        1.POSS-friend 3>3-see.PRES.PERF woman
        'my friend has seen the woman'

     b. *ji-pawana     i-pï*
        1.POSS-friend 3.POSS-wife
        'my friend's wife'

In this way, many of the same concepts used to describe morphological phenomena in the verbal domain can be recycled to also accurately describe morphological phenomena in the nominal domain as well.

## 1.8 Syncretism

One last thing remains to be defined in this introductory chapter. As per Baerman et al (2005), **syncretism** refers to when "a single form serves two or more morphosyntactic functions". For example, in Latin nominal declensions, syncretism is extremely common. As per the paradigm for 'rose' in (36), the same form *rosa* can refer to either NOM.SG and VOC.SG; *rosae* can refer to GEN.SG, DAT.SG, NOM.PL or VOC.PL; and so on.

(36) First declension (fem.) *rosa* 'rose'　　　　　　(Oniga & Schifano 2014; p.64)

|  | SG | PL |
|---|---|---|
| NOM | *rosa* | *rosae* |
| GEN | *rosae* | *rosārum* |
| DAT | *rosae* | *rosīs* |
| ACC | *rosam* | *rosās* |
| VOC | *rosa* | *rosae* |
| ABL | *rosā* | *rosīs* |

Therefore, in the following chapters, when we refer to syncretism, we are referring to such a state of affairs in the grammar of a language. Note while that the Latin example in (36) is an example of syncretism in dependent-marking, it can also extend to head-marking as well.

## 1.9 Conclusion

The purpose of this introductory chapter was to clearly define several key concepts that are of central importance to the dissertation. To begin with, a definition of ergative and accusative alignment was provided. Then, a distinction was made

between different types of ergativity (morphological, syntactic and otherwise), noting that the dissertation primarily focuses on the first type. For this reason, morphological ergativity was explored in more depth, describing its different instantiations in the guise of dependent-marking, head-marking and 'mixed'-marking (a combination of the two). Stepping outside the realm of ergativity and alignment, it was shown how these marking types are useful for describing morphological phenomena in the nominal domain as well. Finally, a brief definition of morphological syncretism was provided.

Equipped with these clearly-defined concepts, we can now proceed to a typological survey of ERG=GEN and ERG=POSS in Chapter 2, and of ACC=GEN, ACC=POSS, NOM=GEN and NOM=POSS in Chapter 3.

## Chapter 2

## THE TYPOLOGY OF ERG=GEN AND ERG=POSS

### 2.1 Introduction

The outline of this chapter is as follows. In Section 2.2, I will provide a definition of both ERG=GEN and ERG=POSS, along with some basic examples. In Section 2.3, I will provide an in-depth description of these phenomena, illustrating with data their various instantiations across languages and language types. The point of Section 2.3 is to show to what extent these patterns of syncretism are a heterogeneous phenomenon, and yet they occur quite commonly across the spectrum of ergative languages. In the second half of the chapter, the focus changes from presentation of the data to analysis, in order to better understand just *how* common these syncretic patterns are cross-linguistically. A representative sample of 40 ergative languages is carefully selected in Section 2.4, and the findings are presented in Section 2.5. The main findings are as follows: that, together, ERG=GEN and ERG=POSS are indeed very common within the sample (2.5.1), that they are more common than comparable patterns in an equivalent sample of accusative languages (2.5.2), and that they transcend major typological parameters such as headedness, basic word order and marking type (2.5.3). Furthermore, the data suggests that less 'robustly' ergative languages are less likely to exhibit either ERG=GEN or ERG=POSS (2.5.4). Section 2.6 concludes.

## 2.2 What is ERG=GEN? What is ERG=POSS?

I define ERG=GEN as a morphological pattern that shows a high degree of overlap in the expression of case in verbal and nominal forms; specifically, in the dependent-marking morphology that encodes **transitive subjects** and **possessors of nouns**.[1]

For example, in Shipibo, the transitive subject *jiwi* 'tree' is marked with the ergative dependent-marker *-n*, as in (1a). This is an ergative pattern because the intransitive subject (such as *isá* 'bird' in 1b) receives no such marking. Crucially, in the nominal domain, possessors are marked with the same dependent-marking suffix *-n*, as in (1c). Since the genitive morpheme *-n* is identical to the ergative morpheme *-n*, this is an instance of ERG=GEN.

Shipibo                                                     (Valenzuela 2003; p.325, p.203, p.132)

(1)  a.  *jiwi-n-ra*      *Sani rishki-ke*
         tree-ERG-EV Sani hit-CMPL
         'the tree hit Sani'

     b.  *isá-ra*    *noya-i*
         bird-EV fly-INC
         'the/a bird is flying'

     c.  *nokon*    *tita-n*      *chomo*
         1SG.GEN mother-GEN jar
         'my mother's jar'

Syncretic patterns across the clausal and nominal domains are not limited to dependent-marking, however. ERG=POSS is the head-marking counterpart to

---

[1] See Section 2.3.9 for previous discussion of this phenomenon by Blake (1994), Lander (2008), Palancar (2008) and Baker (2015).

ERG=GEN, and it is defined in a similar way: a morphological pattern that shows a high degree of overlap in the expression of agreement in verbal and nominal forms; specifically, in the head-marking morphology that encodes **transitive subjects** and **possessors of nouns**. Thus, in Kaqchikel, verbs have head-marking that agrees with arguments in an ergative alignment. The morpheme that encodes the transitive object in (2a) also encodes the intransitive subject in (2b). Crucially, the same morpheme *ru-* that encodes the transitive subject in (2a) also encodes the possessor in (2c), which qualifies as an instance of ERG=POSS. Under the assumption that the same semantic relationship of possession can be encoded in either dependent-marking or head-marking, ERG=GEN and ERG=POSS are thus two closely-related variants of the same overall phenomenon.

Kaqchikel[2]

(2) a. x-      i-        ru-        tijo   -j
       PERF 1SG.ABS 3SG.ERG teach TR
       'she taught me'

    b. x-      i-        wär
       PERF 1SG.ABS sleep
       'I slept'

    c. x-      Ø-        wär
       PERF 3SG.ABS sleep
       'she slept'

    d. ri    ru-        wakx
       DET 3SG.POSS cow
       'her cow'

    In (3), the full paradigm for ERG=POSS in Kaqchikel is illustrated. With one exception, the ERG=POSS pattern systematically extends across the verbal

[2] Unless otherwise indicated, Kaqchikel data in this dissertation are from my own field work in Guatemala. See Brown *et al* (2006) for an introduction to Kaqchikel grammar.

and nominal agreement paradigms, such that the degree of similarity between them cannot be considered mere coincidence.[3]

(3)  a. *x**in**tijoj*   'I taught her'        ***nu**wakx*  'my cow'
     b. *xin**a**tijoj*   'you (SG) taught me'  ***a**wakx*   'your (SG) cow'
     c. *xi**ru**tijoj*   'she taught me'       ***ru**wakx*  'her cow'
     d. *xi**qa**tijoj*   'we taught me'        ***qa**wakx*  'our cow'
     e. *xin**i**tijoj*   'you (PL) taught me'  ***i**wakx*   'your (PL) cow'
     f. *xi**ki**tijoj*   'they taught me'      ***ki**wakx*  'their cow'

### 2.2.1   No syncretism between domains

Of course, it is not the case that all morphologically ergative languages display an ERG=GEN pattern. For example, in Yakima Sahaptin, the transitive subject *tamánwit* 'law' is marked with an ergative suffix *-nɨm*, as in (4a). This is an ergative pattern because intransitive subjects are not marked with such a suffix, as in (4b). Unlike Shipibo, however, the morpheme that encodes possessors, *-mí* in (4c), does not correspond to the ergative marker that identifies transitive subjects in the verbal domain. Thus, this is not an instance of ERG=GEN.[4]

Yakima Sahaptin                                    (Jansen 2010; p.134, p.297, p.290)

(4) a. *tamánwit-nɨm  =nash  i-nápayun-ta*
       law-ERG        1SG    3SG.SUBJ-defend-FUT
       'the law will support me'

---

[3] The Kaqchikel data in (3) is thus an example of a partial ERG=POSS pattern, which appear to be a fairly common occurrence. See Section 2.3.7 below for more discussion of partial ERG=POSS overlap and how they are analyzed in this survey.

[4] Yakima Sahaptin is also a head-marking language. Head-marking and dependent-marking are not mutually exclusive. The verbal prefix *i-* encodes 3SG subjects – both transitive and intransitive ones – and thus Yakima Sahaptin's head-marking system has an accusative alignment. For more discussion on these 'mixed'-marking systems, see Section 2.3.3.

b. *i-kwíita-na*       *Spilyáy*
   3SG.SUBJ-go.along-PST Coyote
   'Coyote was traveling along'

c. *ɨshchɨ́t kaas-mí*
   path    train-GEN
   'railroad track' (lit. 'path of the train')

Likewise, it is not the case that all morphologically ergative languages display an ERG=POSS pattern. For example, Paumarí's system of head-marking is in ergative alignment: in (5a), the 3SG transitive subject is marked by the head-marking prefix *bi-* on the verb, and the transitive object is marked by the suffix *-hi*. In (5b), the intransitive subject is indicated by the same agreement suffix as transitive objects, thus making the pattern an ergative one. Unlike Kaqchikel, however, Paumarí does not exhibit the ERG=POSS pattern because the morpheme that encodes possessors, *kidi-* 'his' in (5c), does not correspond to the ergative marker that encodes transitive subjects in the verbal domain.

Paumarí                       (Chapman & Derbyshire 1986; p.186, p.174, p.254)

(5) a. *bi-na-noki-a-hi*             *ida*    *gora*
   3SG.ERG-CAUS-see-DETRANS-THEME.F DEM.F house
   'he caused the house to become visible'

b. *adari-'i-hi*
   fall.down-ASP-THEME.F
   'she fell'

c. *voroni-'i-hi*     *da*     *kidi-hado*
   fall-ASP-THEME.F DEM.F 3SG.POSS-knife
   'his knife fell'

34

Now that we have covered what does and does not qualify as ERG=GEN and ERG=POSS, we can move on to a fuller description of these phenomena in Section 2.3.

## 2.3 A fuller description of ERG=GEN and ERG=POSS across the spectrum

### 2.3.1 More examples of ERG=GEN

Like Shipibo, many languages across the ergative spectrum feature ERG=GEN in their system of dependent-marking. In Kaluli, transitive subjects are marked with the ergative suffix *-ya:* (6a), but intransitive subjects are not (6b). In (6c), the same suffix *-ya:* also indicates the possessor of a noun.

Kaluli                                                    (Grosh & Grosh 2004; p.63, p.60, p.13)

(6) a. *Iba-ya: siabulu-wo: Hana o:lia: Sala o:lia: ma:no: sofa:*
       Iba-ERG potato-TOP Hanah ACCOMP Sarah ACCOMP food cook.PAST
       'Eva cooked sweet potatoes for Hannah and Sarah'

   b. *kabi fa: gulufo: ko:lo: a:-difa:*
       axe handle break.PAST CONJ REP-put.PAST
       'the axe handle broke so I put it on again'

   c. *tili-dabu kalu-keisale gio:, Gode-ya: nanog di-a:bi*
       pull-hear.PAST man-woman 2PL God-GEN work do-IMP
       'you believers must do God's work!'

A similar pattern occurs in Sinaugoro, Niuean and Kapampangan, three Austronesian languages. In all three cases, an ergative dependent marker modifies the transitive subject (a), but not the intransitive subject (b), thus constituting an ergative dependent-marking pattern. Again, in all three cases, the ergative marker also serves as a genitive marker as part of an ERG=GEN pattern (c).

Sinaugoro                                          (Tauberschmidt 1992; p.181, p.181, p.189)

(7) a. *tau-na     bua      e        vini-gu-to*
       man-ERG betelnut 3SG.SUBJ give-1SG.OBJ-PERF

       'the man gave me a betelnut'

    b. *mero e          ḡani-ḡani-ni*
       boy   3SG.SUBJ eat-RED-IMPF

       'the boy is eating'

    c. *au-na     motuka e        rakava-to*
       1SG-GEN car     3SG.SUBJ bad-PERF

       'my car got damaged'

Niuean                                             (Seiter 1980; p.29, p.28, p.34)

(8) a. *ne    kai he   pusi ia   e    moa*
       PAST eat ERG cat   that ABS chicken

       'that cat ate the chicken'

    b. *malona tuai   e    kapiniu ē*
       broken  PERF ABS dish      this

       'this dish is broken'

    c. *e     kapiniu he    kulī*
       ABS dish        GEN dog

       'the dog's dish'

Kapampangan                                        (Mirikitani 1972; p.77, p.77, 45)

(9) a. *isulat     (ne) ning lalaki ing  istorya*
       write.OV he-it ERG boy    ABS story

       'the boy will write the story'

36

b. *sumulat (ya) ing lalaki*

   write.AV he ABS lalaki

   'the boy will write'

c. *lapis   ning anak ini*
   pencil GEN child this

   'this is the child's pencil'

A special note is in order in the case of Kapampangan. Like many Austronesian languages, Kapampangan features voice morphology on the verb which dictates morphological case and word order. The choice of which morphosyntactic frame to use is determined by pragmatic and conversational factors such as topic, and which argument constitutes new *vs* old information (Mirikitani 1972). In any case, in order for ergative morphological case to be realized, the clause must be in the VSO Object Voice frame, as in (9a). A minimal pair constrasting VSO Object Voice and VOS Active Voice is in (10) below.

Kapampangan                                          (Mirikitani 1972; p.116, p.116)

(10) a. *isulat    ne    ning lalaki ing  poesia king mestra*
       write.OV he-it ERG  boy    ABS poem  DAT teacher

       'the boy will write the poem to the teacher'

   b. *sumulat ya ng     poesia ing  lalaki para king babai*
      write.AV he ACC? poem  ABS boy    BEN DAT girl

      'the boy will write a poem for the girl'

This additional feature is simply a fact about the grammar of Kapampangan; there exists greater flexibility such that, depending on pragmatic factors, not all transitive clauses will be realized in the Object Voice frame with accompanying

ergative case. That being said, the fact remains that there exists overlap between the expression of transitive subject dependent-marking and genitive dependent-marking, and thus this is still an instance of ERG=GEN.

Two additional examples of ERG=GEN can be found in West Greenlandic, an Inuit language of Greenland, and Burushaski, a language isolate. Again, in each case, the transitive subject is marked by ergative dependent-marking (a), while the intransitive subject is not (b). Furthermore, the dependent-marker which encodes the genitive is identical to the ergative dependent-marker (c).

West Greenlandic                                 (Bok-Bennema 1991; p.72, p.72, p.72)

(11) a. *piniartu-p   nanuq       tukut-taa*
        hunter-ERG polar.bear kill-IND.TR.3SG>3SG
        'the hunter killed the polar bear'

    b. *Piita tikip-puq*
        Piita arrive-IND.INTR.3SG.SUBJ
        'Piita has arrived'

    c. *piniartu-p   irnir-a*
        hunter-GEN son-3SG.POSS
        'the hunter's son'

Burushaski                                       (Munshi 2015; p.29, p.12, p.17)

(12) a. *hamal-e          ḍarbeṣ  i-cʰarkan-uman*
        neighbour-ERG Darbes 3SG.M.OBJ-beat-3PL.SUBJ
        'the neighbours beat Darbes'

    b. *hiles gucʰar-imi*
        boy   walk-3SG.M.SUBJ
        'the boy walked'

38

c. *in-e*      *laqpis*
   3SG-GEN handkerchief

   'his handkerchief'

Note that in both West Greenlandic and Burushaski, in addition to dependent-marking, verbs feature morphology agreeing with the subject and object. These languages are dubbed 'mixed'-marking languages, and are discussed in further detail in Section 2.3.3.

### 2.3.2   More examples of ERG=POSS

Like Kaqchikel, Itzaj is another language from the Mayan family that features ERG=POSS. Unlike Kaqchikel, however, the Itzaj ERG=POSS paradigm is complete. As per Hofling (2000), ergative person markers reference the agents of transitive verbs and nominal possessors in complete overlap (14). The glides *w-* and *y-* are morphophonologically conditioned, occuring in prevocalic contexts.

Itzaj                                      (Hofling 2000; p.37, p.37, p.37, p.25)

(13) a. *k-*    *uy-*      *il*   *-aj* *-o'on*
        CMPL 3SG.ERG see TR  1PL.ABS

      'he saw us'

   b. *tal*    *-o'on*                  c. *tal*    *-ij*
      come 1PL.ABS                        come 3SG.ABS

      'we came'                            'he came'

   d. *uy-*       *äj-* *kax*
      3SG.POSS- M   chicken

      'his chicken'

(14) *'Set A' person markers in Itzaj*                                      (Hofling 2000; p.35)

|              | **Singular** |         | **Plural** |       |          |
|--------------|--------------|---------|------------|-------|----------|
| **1st person** | *in(w)-*   | **excl.** | *ki(w)-*   |       |          |
|              |              | **incl.** | *ki(w)-*   | *...* | *-e'ex*  |
| **2nd person** | *a(w)-*    |         | *a(w)-*    | *...* | *-e'ex*  |
| **3rd person** | *u(y)-*    |         | *u(y)-*    | *...* | *-oo'*   |

Moving outside the Mayan language family, many instances of ERG=POSS can be found as well. Like Kapampangan, Seediq is an Austronesian language which features two different morphosyntactic frames which are conditioned by pragmatic factors. Holmer (1996) labels these Patient Focus (PF) and Actor Focus (AF), and notes that the pragmatic factor conditioning their use is definiteness, where a definite argument "generally" appears in clause-final position, as in (15).

Seediq                                                      (Holmer 1996; p.66, p.66)

(15) a. *mekan wawa ka    huling*
       eat.AF  meat  DEF  dog
       'the dog is eating meat'

   b. *puqun huling ka    wawa*
       eat.PF  dog     DEF  meat
       'a dog will eat the meat'

In addition, Seediq features pronominal clitics with a restricted distribution. Unlike their non-clitic counterparts, they must occur after the "first verbal element" in the clause, which may be a verb or an auxiliary (Holmer 1996).[5] Some of these

---

[5] Discussion of free pronouns in the context of ERG=GEN and ERG=POSS is found in Section 2.3.4. That these Seediq pronominal clitics are presented in the head-marking section is specifically due to their restricted distribution on "the first verbal element" in the clause.

pronominal clitics, such as 2SG *su*, can act as transitive subject, transitive object, or intransitive subject, as in (16).

Seediq                                                    (Holmer 1996; p.86, p.65, p.85)

(16) a. *snalu*        *=su*   *snuwan sapah nii*
      make.PF.PRET =2SG when    house this

      'when did you build this house?'

   b. *wada =su   qtaun =na       Awi*
      PRET =2SG see.PF =3SG.ERG Awi

      'Awi saw you'

   c. *mnekan*     *=su   di*    *(rising intonation)*
      eat.AF.PRET =2SG PERF

      'have you eaten?'

However, a subset of persons in the pronominal clitic paradigm (1SG, 3SG, 3PL) <u>can only be used</u> <u>as transitive subjects</u>, which in turn requires the PF frame. This is an ergative alignment; if these persons occur either as a transitive object or an intransitive object, an alternative strategy must be used.[6] In any case, these strictly ergative clitics also function as genitive pronouns, as in (17c). Thus, this is another instance of ERG=POSS (in fact, Holmer labels them "ErgGen" clitics).

Seediq                                                    (Holmer 1996; p.69, p.57, p.51, p.32)

(17) a. *wada =ku*        *=na*      *qtaun*
      PRET =1SG.NOM =3SG.ERG see.PF

      'he saw me'

---

[6] These alternative strategies include a dedicated absolutive clitic *ku* for 1SG, as in (17a), or "long form" free pronouns that don't have a clitic-like distribution, like *heya* 3SG in (17b).

b. *wada dehuk heya*
   PRET arrive 3SG

   'he arrived'


c. *qalux ka    uban =na*
   black DEF hair =3SG.GEN

   'its hair is black'


(18)

|          | free pronouns | clitic pronouns | |
|----------|:-------------:|:---------------:|:---:|
|          |               | **"ErgGen"**    | **Abs** |
| **1sg**  | *yaku*        | *mu*            | *ku* |
| **2sg**  | *isu*         | *su*            | |
| **3sg**  | *heya*        | *na*            | - |
| **1pl.incl** | *ita*     | *ta*            | |
| **1pl.excl** | *yami*    | *nami/mian*     | |
| **2pl**  | *yamu*        | *namu*          | |
| **3pl**  | *dheya*       | *daha*          | - |


Similar to Seediq, the grammar of Gitksan features restricted pronominal distributions that result in ergative alignment. Gitksan features free pronouns, which, in independent clauses, either encode transitive objects (19a-19b) or intransitive subjects (19c). However, in these independent clauses, transitive subjects only are encoded via suffix on the verb (see again 19a-19b). This is an ergative alignment. Crucially, these same verbal suffixes also encode possessors, as in (19d); thus, this is another instance of ERG=POSS.[7] Furthermore, the ERG=POSS in Gitksan is consistent across all persons; the facts are summarized in Table 20 below.

---

[7] For a discussion of alignment in Tsimshianic languages, see Peterson (to appear).

Gitksan                                    (Rigsby 1986; p.261, p.261, p.257, p.323)

(19)  a.  *łəmo:-yə-t*        *'nu:'m*
          help-TR-**3SG.ERG** 1PL.ABS

          'he helped us'

      b.  *łəmo:-yə-'m*       *'nit*
          help-TR-1PL.ERG 3SG.ABS

          'we helped him'

      c.  *sa:   paẋ 'nit*
          away ran 3SG.ABS

          'he ran away'

      d.  *kimxti-t*
          sister-**3SG.POSS**

          'his sister'

(20)  *Gitksan free pronouns & verb agreement in independent clauses*

                                          (adapted from Peterson to appear)

| | genitive | transitive subject | transitive object | intransitive subject |
|---|---|---|---|---|
| **free pronouns** | | | ✔ | ✔ |
| **agreement suffixes** | ✔ | ✔ | | |

Abkhaz presents an interesting scenario.  Verbs are heavily inflected with agreement prefixes, and according to Hewitt (1989), their template is as follows: ABS-DAT-ERG-V. An example of the second person singular masculine agreement marker in all three of these positions is provided in (21).

43

Abkhaz (Hewitt 1989; p.8, p.66, p.73)

(21) a. *wə-cò*        *ma*
     2SG.M.ABS-go INT
     'are you going?'

    b. *yə-wə̀-y-to-yt'*
     3SG.NH.ABS-2SG.M.DAT-3SG.M.ERG-give-FIN
     'he will give it to you'

    c. *Axra də-w-bò-w*
     Axra 3SG.M.ABS-2SG.M.ERG-see-INT
     'did you see Axra?'

Note that for the 1st and 2nd person prefixes, their form is identical regardless of grammatical function. They are only distinguished by virtue of their relative position in the template. The 1st and 2nd person prefixes are provided in (22) below.

*Abkhaz agreement prefixes (1st and 2nd person)*      (adapted from Hewitt 1989)

(22)

| | | Set I      Set II      Set III | |
|---|---|---|---|
| | | **(ABS, DAT and ERG identical in form)** | |
| | | **sg** | **pl** |
| **1st** | | *s(ə)-* | *ħ(a)-* |
| **2nd** | **male** | *w(ə)-* | *šʷ(ə)-* |
| | **female** | *b(ə)-* | |

However, if we look at 3rd person agreement markers, we can tease apart some distinctions. Looking at (23), in the 3rd person, ABS agreement markers are distinguished from DAT and ERG agreement markers, which still pattern together.

*Abkhaz agreement prefixes (3rd person)*          (adapted from Hewitt 1989)

(23)

|  |  | **Set I** (ABS) | | **Set II** (DAT and ERG identical in form) | **Set III** |
|---|---|---|---|---|---|
|  |  | **sg** | **pl** | **sg** | **pl** |
| **3rd** | **male** | *d(ə)-* | *yə-* | *y(ə)-* | *r(ə)-* |
|  | **female** |  |  | *l(ə)-* |  |
|  | **non-human** | *yə-* |  | *a-* |  |

Thus, it is clear that in order to determine whether Abkhaz agreement prefixes exhibit ERG= POSS, we must compare 3rd person transitive subjects to 3rd person possessors. Looking at (24), it turns out that ERG and POSS agreement markers do indeed overlap. Thus, Abkhaz counts as another instance of ERG=POSS, with the caveat that in this case, DAT=POSS is compatible with the data as well.

Abkhaz          (Hewitt 1989; p.150, p.140, p.129)

(24) a. *a-gəgšʷə̀g sə-r-bè-yt'*
     DET-beast 1SG.ABS-3PL.ERG-see-FIN

     'the beasts saw me'

   b. *a-jə́yas    nərcʷ       yə-n-xò-yt'*
     DET-river on.that.side.of 3PL.ABS-PREV-live-FIN

     'they live on that side of the river'

   c. *r-an         a-yʷnə̀    də-yʷnà-n*
     3PL.POSS-mother DET-house 3SG.ABS-house-STAT.PAST

     'their mother was in the house'

As a final example in this section, Canela-Kraho presents an interesting instantiation of ERG= POSS. In transitive clauses, the agreement marker that corresponds to the transitive object occurs as a prefix on the TAM morpheme, while the marker that corresponds to the transitive subject prefixes directly to the verb, as in (25a). This is an ergative pattern (albeit an unusual one), because in intransitive clauses, the marker that corresponds to the intransitive subject prefixes directly onto the verb as well, as in (25b). Much like in Abkhaz, the forms of 'ergative' and 'absolutive' agreement markers are the same, and it is merely morphosyntactic position that determines an ergative-absolutive alignment.

Canela-Kraho                    (Popjes & Popjes 1986; p.147, p.132, p.169, p.169)

(25) a. *i-te*      *a-pupun*
        1SG-PAST 2SG-see

        'I saw you'

     b. *a-jōt*
        2SG-sleep

        'you slept'

     c. *a-quêtti*              d. *i-jõ*      *wapo*
        2SG-uncle                  1SG-POSS knife

        'your uncle'               'my knife'

In addition, Canela-Kraho morphologically encodes a distinction between alienable and inalienable possession. Inalienable possession is encoded by direct prefixing onto the noun, as in (25c). However, alienable possession looks isomorphic to transitive clauses. A 'possessive' head morpheme *jõ* must occur before the possessed noun, and now the agreement morphology prefixes directly onto this morpheme. This is another instance of ERG=POSS, but again with a caveat - only

for alienable possession. See Section 2.5.5 for more discussion of ergative syncretic patterns and alienable *vs* inalienable possession.

To summarize this section, head-marking systems can vary greatly in their details from language to language. And yet, despite the heterogeneous nature of (ergative) head-marking, instances of ERG=POSS can be found across these different head-marking systems.

### 2.3.3   'Mixed'-marking languages

Sections 2.3.1 and 2.3.2 began the typological exploration of ERG=GEN and ERG=POSS by looking at instances in dependent-marking and head-marking, respectively. This section is devoted to ergative languages which simultaneously feature dependent- and head-marking, which we will label 'mixed'-marking languages.

In Section 2.3.1, we noted that West Greenlandic features ERG=GEN in its system of dependent-marking. In actuality, West Greenlandic is a 'mixed'-marking language, meaning that it also features a system of head-marking. Indeed, dependent-marking and head-marking are not mutually exclusive. Thus, in (26 – expanded from 11 above), verbal agreement suffixes follow the mood marker in West Greenlandic. In (26a), the transitive subject is marked by -∅ and the transitive object by -*a*, while in (26b), the intransitive subject is marked by -∅.

West Greenlandic                                        (Bok-Bennema 1991; p.72, p.72)

(26)   a.   *piniartu-p   nanuq     tukut-ta-a-∅*
             hunter-ERG polar.bear kill-IND.TR-3SG.OBJ-3SG.SUBJ
             'the hunter killed the polar bear'

        b.   *Piita  tikip-puq-∅*
             Piita arrive-IND.INTR-3SG.SUBJ
             'Piita has arrived'

What is the alignment of West Greenlandic's head-marking system? According to Bok-Bennema's (1991) analysis, it varies as per the participants involved. If all participants are 3rd person, as in (26), the alignment appears to be accusative. See the table in (27) for head-marking in 3 > 3 scenarios.

(27) *West Greenlandic head-marking* (3>3)      (adapted from Bok-Bennema 1991; pp.194-195)

|  | transitive subject | transitive object | intransitive subject |
|---|---|---|---|
| **3sg** | -∅ | -a | -∅ |
| **3pl** | -t | -i | -t |

On the other hand, in events featuring strictly 1st and 2nd person participants – also known as Speech Act Participants (SAP) – West Greenlandic's head-marking system is in an ergative alignment. For 1SG and 2SG arguments, the same morpheme encodes transitive objects and intransitive subjects, to the exclusion of transitive subjects (see Table 28). The alignment for 1PL and 2PL can neither be said to be ergative nor accusative.

(28) *West Greenlandic head-marking* (SAP>SAP)      (adapted from Bok-Bennema 1991; pp.196-198)

|  | genitive |  | transitive subject | transitive object | intransitive subject |
|---|---|---|---|---|---|
| **1sg** | *-ga* | | *-ga* | *-nga* | |
| **2sg** | *-t* | | *-t* | *-tit* | |
| **1pl** | *-put* | | *-vut* | *-tigut* | *-gut* |
| **2pl** | *-si* | | *-si* | | |

48

Crucially, for those grammatical persons whose alignments are unambiguously ergative, the grammar of West Greenlandic reuses the transitive subject agreement marker to mark the possessor of a noun, as in (29). Thus, this is another example of (partial) ERG=POSS.

West Greenlandic                                    (Bok-Bennema 1991; p.197, p.197)

(29) a. *pani-ga*
     daughter-1SG.POSS
     'my daughter'

   b. *panii-t*
     daughter-2SG.POSS
     'your (SG) daughter'

To be clear, saying that West Greenlandic head-marking only exhibits ERG= POSS is not accurate. When the possessor is 3SG, as in (30), the pattern is now ACC=POSS. The 3SG transitive object suffix -*a* is also used to mark a 3SG possessor.

West Greenlandic                                             (Bok-Bennema 1991; p.72)

(30) *piniartu-p   irnir-a*
     hunter-GEN son-3SG.POSS
     'the hunter's son'

The rest of the facts regarding West Greenlandic head-marking are not straightforward.[8] But the main point here is that, in addition to being 100% ERG=GEN

---

[8] In the scenario 3 > SAP, verbal head agreement is encoded by the 3 'NOM' suffix and and the SAP 'ABS' suffix, as expected. But in the inverse scenario, SAP > 3, the subject is encoded by the SAP 'ERG' suffix, but counter-intuitively, the object is encoded by the 3 'NOM' suffix, and not 'ACC'. The full set of scenarios is summarized in the table below. See Bok-Bennema (1991) for more details.

i)

|  | **3 markers** | | **SAP markers** | |
|---|---|---|---|---|
| **SAP>SAP** | | | ERG | ABS |
| **SAP>3** | | NOM | ERG | |
| **3>SAP** | | NOM | | ABS |
| **3>3** | ACC | NOM | | |

in its dependent-marking, West Greenlandic also shows traces of ERG=POSS in its complicated head-marking system.[9]

Sm'algyax is another 'mixed'-marking language, but in terms of syncretism it is the inverse of West Greenlandic: ERG=POSS occurs in the system of head-marking, and there is no ERG=GEN pattern to speak of in its system of dependent-marking. The examples in (31) illustrate the ergative alignment of Sm'algyax's case system. According to Mulder (1994), these markers occur before the word they mark in linear word order, but suffix to the word immediately preceding it. Thus, in (31a), *-sga* marks the transitive subject *awta* 'porcupine', and *-ga* marks the transitive object, an embedded clause. In (31b), the same marker *-ga* also marks the intransitive subject *sm'ooygit* 'chief', thus making Sm'algyax's dependent marking ergative in alignment. Interestingly, Sm'algyax's dependent marking does <u>not</u> display ERG=GEN; looking at (31c), it is actually ABS=GEN: the absolutive marking *-ga* also functions as a genitive in the nominal domain.[10]

Sm'algyax                                                    (Mulder 1994; p.36, p.34 p.76)

---

[9] Note also that the West Greenlandic head-marking facts lend themselves well to a line of analysis where verbs and nouns are conflated. For instance, it is possible to reanalyze (30) as a verbal construction, as in (ii) below.

ii) *piniartu-p    irnir-a-∅*
   hunter-ERG son-3SG.OBJ-3SG.SUBJ

   'the hunter's son' (or 'the son is the hunter's possessed one')

In this type of analysis, the case and agreement morphology in 'verbal' and 'nominal' domains is exactly isomorphic. In fact, there is just a single domain; they are one and the same. Among others, see Johns (1992) and Yuan (2013) for analyses in this tradition.

[10] For more discussion on ABS=GEN, see Section 2.5.6.

(31) a. *ha'ligoot -sga awta    -ga* [ *dza ła    al    dzak -sga    sts'ool -ga* ]
thought ERG porcupine ABS   C   PAST EMPH dead ABS.SUB beaver DEM

'porcupine thought that beaver was dead'

b. *t'aa    -ga  sm'ooygit-ga*
be.SG ABS chief-DEM

'there was a chief'

c. *naks    -ga   na   -wayk   -t        -ga*
spouse GEN POSS brother 3SG.POSS DEM

'their brother's wife' (lit. 'spouse of their brother')

On the other hand, Sm'algyax has an ergative system of head-marking, and it <u>does</u> feature ERG=POSS. In (32a), both transitive subject and transitive object are marked by a suffix on the verb. This is an ergative alignment because the same agreement suffix that encodes the transitive object also encodes the intransitive subject, as in (32b).[11] Crucially, the same morpheme that encodes the transitive subject in (32a) also encodes a possessor in the nominal domain, as in (32c). Thus, Sm'algyax's system of head-marking displays ERG=POSS.

Sm'algyax                                    (Mulder 1994; p.58, p.101, p.62)

(32) a. *ap'ax     -d     -u*
remember 3.ABS 1SG.ERG

'I remembered him'

b. *ada k'a        -t'aa-t*
and for.a.while sit-3.ABS

'and he sat for a while'

---

[11] In (32a), the ABS suffix has voiced due to the following vowel. See Mulder (1994) for details on the morphophonology of Sm'algyax.

c. *hak'o* *-(y)u*
   back 1SG.POSS
   'my back'

Both West Greenlandic and Sm'algyax feature ergative dependent-marking
and (at least partially) ergative head-marking. In addition, it is common to find
languages with ergative dependent-marking, but accusative head-marking, and these
too can exhibit ergative syncretic patterns.

Burushaski is one such language. As noted in Section 2.3.1, Burushaski's
dependent-marking system is in ergative alignment, and furthermore it exhibits
ERG=GEN (see examples in 12). However, Burushaski's system of head-marking is
in accusative alignment. In (33a) and (33b), the same agreement marker *-uman* en-
codes the transitive and intransitive subject, respectively. A separate prefix encodes
transitive objects in (33a). Because Burushaski's head-marking system is accusative,
it is ineligible for any kind ERG=POSS overlap in the first place.

Burushaski                                                          (Munshi 2015; p.29, p.61)

(33) a. *hamal-e*        *ḍarbeṣ*  *i-cʰarkan-uman*
        neighbour-ERG  Darbes 3SG.M.OBJ-beat-3PL.SUBJ
        'the neighbours beat Darbes'

    b. *u*    *gaarc-uman*
       3PL run-3PL.SUBJ
       'they ran'

Also in Section 2.3.1, Kaluli and Sinaugoro were both shown to exhibit ERG=
GEN in their dependent marking (see relevant examples in 6-7). However, in both
cases their system of head-marking in is accusative alignment. In (34-35) below, the

transitive subject agreement morpheme in (a) corresponds to the intransitive subject agreement morpheme in (b).

Kaluli                                                    (Grosh & Grosh 2004; p.32, p.17)

(34) a. *e    mamo wanalo: amada asul-ab*
        3SG flower yellow   very   like-3SG.SUBJ.PRES
        'he likes the yellow flowers very much'

     b. *sowa tuolon-o:  alifol-ab*
        child baby-TOP sleep-3SG.SUBJ.PRES
        'the baby is sleeping'

Sinaugoro                                                 (Tauberschmidt 1992; p.180, p.181)

(35) a. *tau-na    bua    e        vini-gu-to*
        man-ERG betelnut 3SG.SUBJ give-1SG.OBJ-PERF
        'the man gave me a betelnut'

     b. *mero e        ḡani-ḡani-ni*
        boy   3SG.SUBJ eat-RED-IMPF
        'the boy is eating'

The logically-possible inverse of languages such as Burushaski, Kaluli and Sinaugoro – accusative dependent-marking and ergative head-marking – might be expected to also display ERG=POSS, given how prevalent the phenomenon appears to be. However, this language configuration is unattested; see Bobaljik (2008) for a syntactic account of why this asymmetry exists. The table below summarizes the findings in this section regarding 'mixed'-marking languages.

(36) *Mixed-marking languages and ERG=GEN*

| 'mixed'-marking type | | ERG=GEN locus | | example |
|---|---|---|---|---|
| **dep** | **head** | **dep** | **head** | |
| ERG | ERG | ✔ | | West Greenlandic |
| ERG | ERG | | ✔ | Sm'algyax |
| ERG | ACC | ✔ | | Burushaski, Kaluli, Sinaugoro |
| ACC | ERG | language type unattested | | |

### 2.3.4   Pronominal systems

In addition to dependent-marking and head-marking, languages can also display an ergative alignment in their pronominal system. Certainly there are cases such as in Burushaski, where the ergative marker for lexical NPs is straightforwardly concatenated with pronominal forms, such as in (37-38).

Burushaski                               (Munshi 2015; p.29, p.12, p.21, p.24, p.15, p.17)

(37) a. *hamal-e      ḍarbeṣ  i-cʰarkan-uman*
     neighbour-ERG Darbes 3SG.M.OBJ-beat-3PL.SUBJ

     'the neighbours beat Darbes'

  b. *hiles gucʰar-imi*
     boy   walk-3SG.M.SUBJ

     'the boy walked'

  c. *hurmaṭ-e     oṣcum*
     Hurmat-GEN waistband

     'Hurmat's waistband'

(38) a. *in-e      a-yaar-um       pen cʰu-mi*
     3SG-ERG 1SG-under-from pen take-3SG.M.SUBJ

     'he took the pen from me'

b. *in gucʰar-imi*
   3SG walk-3SG.M.SUBJ

   'he walked'

c. *in-e laqpis*
   3SG-GEN handkerchief

   'his handkerchief'

This data indicates that the grammar of Burushaski does not make a strong distinction between pronouns and lexical NPs for the purposes of dependent-case assignment. A similar situation obtains in Shipibo. According to Valenzuela (2003), the ergative marker /-n/ has several allomorphs [-n, -an, -en, -in, -kan, -ten, -tan, -man, -nin], some examples of which are in (39) below.

Shipibo                                                   (Valenzuela 2003; p.323, p.324, p.324)

(39) a. *Inka-n shinaman manan-xawe-n pei-ki seke-nan-a iki*
        Inka-ERG think.MNS hill-turtle-GEN wing-HSY2 break-MAL-PART AUX

        'the Inka, with his mental power, broke the tortoise's wings'

   b. *jisis-in-ra Sankeman bake natex-ke*
      ischimi-ERG-EV Sanken.GEN child bite-CMPL

      'the *ischimi* ant bit Sanken's child'

   c. *Cesar-nin-ra Maria-nin wai rera-ke machito-nin*
      'Cesar-ERG-EV Maria-GEN field fell.tree-CMPL machete-INSTR

      'Cesar cleared Maria's field with a machete'

When it comes to pronouns, however, the grammar of Shipibo straightforwardly concatenates the [-n] allormorph to derive ergative (and genitive) pronouns, as in (40).

(40) a. *mi-n-ra    e-a    paran-a    iki*
2-ERG-EV 1-ABS deceive-PART AUX

'you lied to me'

b. *jawerano-ki mi-a    ka-[a]i?*
where-INT    2-ABS go-PART

'where are you going?'

c. *mi-n    ibo-n    kaya*
2-GEN owner-GEN soul

'your owner's soul'

Like Burushaski, the grammar of Shipibo does not seem to distinguish between pronouns and lexical NPs for the purposes of dependent-case assignment. That being said, pronominal forms only constitute a case of partial ERG=GEN in Shipibo. In Table 41, only cases of pronominal ERG=GEN overlap are highlighted.

(41) *Shipibo pronouns*                          (adapted from Valenzuela 2003; p.185)

|            | 1SG   | 2SG  | 3SG   | 1PL   | 2PL    | 3PL                            |
|------------|-------|------|-------|-------|--------|--------------------------------|
| Absolutive | *e-a*   | *mi-a* | *ja*    | *no-a*  | *mato*   | *jato/jabo*                      |
| Ergative   | *e-n*   | *mi-n* | *ja-n*  | *no-n*  | *mato-n* | *jato-n/jabaon/jaboan/jaboon*    |
| Genitive   | *nokon* | *mi-n* | *jawen* | *no-n*  | *mato-n* | *jato-n/jabaon/jaboan/jaboon*    |

In addition to cases such as these, where a dependent-marker concatenates with lexical NPs and pronominal forms equally, it is possible to identify languages which feature ERG=GEN in the pronominal system proper. In these cases, it not possible to tease apart a separate dependent marker in the pronominal system; the ERG=GEN forms appear to be morphologically simple.

Kapampangan is one such language. In addition to exhibiting case-marking ERG=GEN on lexical NPs, as in (42), we can also observe pronominal ERG=GEN, as in (43).

Kapampangan                  (Mirikitani 1972; p.77, p.77, p.45, p.134, p.37, p.31)

(42) a. *isulat*    *ne*      *ning* *lalaki* *ing* *istorya*
      write.OV 3SG>3SG ERG boy    ABS story
      'the boy will write the story'[12]

     b. *sumulat* *ya*      *ing* *lalaki*
      write.AV 3SG.ABS ABS lalaki
      'the boy will write'

     c. *lapis*    *ning* *anak* *ini*
      pencil GEN child this
      'this is the child's pencil'

(43) a. *kapuputut* *na*      *ning* *babai*    *ng*    *manuk*
      just.cut.OV 3SG.ERG ERG woman ABS chicken
      'the woman just cut a chicken'

     b. *maglutu* *ya*
      cook.AV 3SG.ABS
      'she will cook'

     c. *sali*    *ya*     *ng*     *libru* *ing* *babai* *king*    *pera*    *na*
      buy.AV 3SG.ABS ACC? book ABS girl    INSTR money 3SG.GEN
      'the girl will buy a book with her money'

---

[12] Kapampangan also features portmanteau pronouns, such as *ne* '3SG.3SG', which simultaneously encode several arguments at once. See Forman (1971) for more details.

Note also that, as observed in Section 2.3.1, a Kapampangan transitive clause must be realized in the Object Voice frame in order for ergative morphology to be realized. In any case, the table of Kapampangan pronouns in (44) shows the extent of ERG=GEN across the pronominal paradigm.

(44) *Kapampangan pronouns*                (adapted from Forman 1971; pp.60-62)

|  | 1SG | 2SG | 3SG | 1DU | 1PL.INCL | 1PL.EXCL | 2PL | 3PL |
|---|---|---|---|---|---|---|---|---|
| Absolutive | *ku* | *ka* | *ya* | *kata* | *támu* | *kamí, ké* | *kayú, kó* | *la* |
| Ergative | *ku* | *mu* | *na* | *ta* | *mi* | *tá* | *yu* | *da, ra* |
| Genitive | *ku* | *mu* | *na* | *ta* | *mi* | *tá* | *yu* | *da, ra* |

Another language whose pronouns fall under the rubrick of ergative syncretism is Sorani Kurdish. At first blush, Sorani Kurdish looks like it simply has an ergative head-marking system. In (45a), the clitic morpheme *-yān* encodes the transitive 3PL subject, while *-in* encodes the transitive 2PL object. In (45b), the same clitic *-in* encodes the intransitive 2PL subject, making the pattern an ergative one. In (45c), the possessor is encoded by the same ergative morpheme, which appears to qualify Sorani Kurdish as an ERG=POSS language.[13]

Sorani Kurdish                (McCarus 2009; p.617, p.609 / Blau 1980; p.63)

(45)  a.  *bīnī-yān-in*
          see-3PL.ERG-2PL.ABS
          'they saw you (PL)'

---

[13] Sorani Kurdish is actually in a split-ergative alignment. In this case, it means that the ergative alignment is only observed in the past tense; present tense is in accusative alignment. See Section 2.5.4.1 for more on split-ergativity and syncretic patterns, and McCarus (2009) and Blau (1980) for more on split-ergativity in Sorani Kurdish in particular.

b. *hāt-in*

come-2PL.ABS

'you (PL) came'

c. *qamik-yān*

thumb-3PL.POSS

'their thumb'

However, looking at more Sorani Kurdish data reveals that the distribution of these clitics is not strictly limited to verbs and/or auxiliaries. If there are no other words in the clause, agreement clitics will indeed attach to the verb as above. According to McCarus (2009), should there be any overt nominals or adjectives in the clause, they prefer to cliticize onto these elements, as in (46). Thus, agreement clitics in Sorani Kurdish have quite a free distribution, and are probably not best analyzed as being instances of head-marking, but rather clitic pronouns. This means that the accurate characterization of ergative syncretism in Sorani Kurdish is ERG=GEN instead.[14]

Sorani Kurdish (McCarus 2009; p.616)

(46) *kay   pān-yān      a-ka-n*
    when wide-3PL.ERG IMPF-do-3PL.ABS

'when will they widen them?'

---

[14] In addition to the possessive construction in (45c), Sorani Kurdish also possesses an alternate possessive strategy called the *izāfa* construction. It consists of a head noun followed by the suffix *-ī* and the possessor in independent pronoun form, as in (i). The *izāfa* construction can be used for a broader range of semantic relationships than just possession, as illustrated in (ii-iii).

Sorani Kurdish (McCarus 2009; p.613, p.613, p.613)

i) *nāw-ī    min*          ii) *hēlka-ī   tāza*        iii) *māŋ-ī    galāwēž*
   name-POSS 1SG                egg-POSS fresh              month-POSS Gelawezh
   'my name'                   'fresh eggs'                'the month of Gelawezh'

### 2.3.5 Headedness, ERG=GEN and ERG=POSS

Greenberg (1963) famously made the observation that certain syntactic properties tend to correlate with one another cross-linguistically. For instance, if a language has VO basic word order, it is also likely have prepositions; however, if a language has OV basic word order, it is likely to have postpositions. Chomsky (1981) expanded on this idea, postulating a headedness parameter whereby all languages are either head-initial or head-final.

While exceptions to these generalizations are known to exist, it is nevertheless useful to classify languages as being head-initial or head-final for typological purposes. In (47), some of Greenberg's universals are listed which seem like seem like relevant diagnostics for syntactic headedness.

(47) *Headedness and Greenbergian universals*

| | relevant Greenbergian universal(s) | | | | | |
|---|---|---|---|---|---|---|
| | **2** | **2,3,4** | **3,4,13,15,16** | **13** | **15** | **16** |
| **head-initial** | P NP | N $NP_{poss}$ | VO | V [sub. clause] | Aux V | V [rel. clause] |
| **head-final** | NP P | $NP_{poss}$ N | OV | [sub. clause] V | V Aux | [rel. clause] V |

Returning now to ergative patterns of syncretism, it is apparent in the data that neither ERG=GEN nor ERG=POSS are restricted to languages of a specific headedness configuration. Indeed, both head-initial and head-final ergative languages have ERG=GEN. For example, Kaqchikel (48) and Kapampangan (49) both exhibit consistently head-initial behavior, while Sinaugoro (50) and Shipibo (51) both exhibit consistently head-final behavior, and yet all are valid instances of ERG=GEN and ERG=POSS, respectively. In the examples below, the head in question is underlined for ease of exposition.

Kaqchikel (head-initial)

(48) a. <u>*pa*</u> *toq'a*
P    night

'at night'

b. <u>*ru-wakx*</u>      *ri*    *a Xuan*
3SG.POSS-cow DET Juan

'Juan's cow'

c. <u>*n-u-tz'ët*</u>      *ri*    *öj*      *ri*    *tz'i*
IMPF-3SG.ERG-see DET avocado DET dog

'the dog sees the avocado'

d. <u>*n-inw-ajo*</u>      [ *n-in-b'än*      *ri*    *q'utu'n* ]
IMPF-1SG.ERG-want    IMPF-1SG.ERG-do DET food

'I want to make the food'

e. (no free Auxiliary in Kaqchikel)

f. *ri*    <u>*wakx*</u> [ *ri*    *x-u-loq'*      *rija* ]
DET cow      DET PERF-3SG.ERG-buy 3SG

'the cow that he bought'

Kapampangan (head-initial)    (Mirikitani 1972; p.72, p.45, p.77, p.221, p.41, p.190)

(49) a. *ati*    *ing*   *lalaki* <u>*king*</u> *eskwela ngeni*
COP ABS boy   P     school    now

'the boy is at school now'

b. <u>*lapis*</u>    *ning anak ini*
pencil GEN child this

'this is the child's pencil'

c. *isulat    ne       ning lalaki ing  istorya*
   write.OV 3SG>3SG ERG  boy    ABS story

   'the boy will write the story'

d. *susubukan* [ *keng        bubuklat ing  awang*  ]
   try          1SG>3SG.C open     ABS window

   'I'm trying to open the window'

e. *pota    munta ing  mestro  king Japan*
   might go     ABS teacher P    Japan

   'the teacher might go to Japan'

f. *ing babaing* [ *tuturung Inglis    king anak ku*        ]
   ABS woman     teach    English P    child 1SG.POSS

   'the woman who is teaching English to my child'

Sinaugoro                          (Tauberschmidt 1992; p.189, p.181 /

                                   Tauberschmidt 1999; p.55, p.80, p.74)

(50) a. *ḡoraḡani  ḡai        Kemaia ḡana ḡa         iago-to*
     yesterday 1PL.EXCL Kemaia P     1PL.EXCL go-P

     'yesterday we went to Kemaia'

   b. *au-na     motuka e        rakava-to*
      1SG-GEN car      3SG.SUBJ bad-PERF

      'my car got damaged'

   c. *tau-na     bua       e        vini-gu-to*
      man-ERG betelnut 3SG.SUBJ give-1SG.OBJ-PERF

      'the man gave me a betelnut'

   d. [ *mutu gabu ḡa        ] iaḡo-ni*
        grass burn 1PL.EXCL   go.IMPF

      'we are going to burn the grass'

62

e. (no free Auxiliary in Sinaugoro)

f. [ *ḡarivata e        lema-ia-to             ] <u>tau-na</u>,       e*
   banana   3SG.SUBJ steal-3SG.OBJ-PERF    man-3SG.POSS 3SG.SUBJ
*raga-rekwa-rekwa-to*
run-loose-RED-PERF

'the man who stole the bananas ran away'

Shipibo                        (Valenzuela 2003; p.175, p.132, p.325, p.418, p.313, p.446)

(51) a. *moa     icha   baritia <u>pekáo</u>*
       already many time    after

'after many years'

b. *nokon     tita-n        <u>chomo</u>*
       1SG.GEN mother-GEN jar

'my mother's jar'

c. *jiwi-n-ra     Sani <u>rishki-ke</u>*
       tree-ERG-EV Sani  hit-CMPL

'the tree hit Sani'

d. [ *ja         jo-ai       ] <u>oin-taanan-ki</u> joni sotameet-a     ki*
       3SG.ABS come-PART   see-PSS-EV2    man spy.MID-PART AUX

'seeing him coming, the man hid to spy'

e. *rama-ra   nee-ti     <u>atipan-ke</u>*
       Rama-EV climb-INF can-CMPL

'Rama can climb'

f. [ *papa-n      rete-ibat-a     ] <u>jono-ra</u>            moa     no-n*
       father-ERG kill-PAST-PART   collared.peccary-EV already 1PL-ERG
*keyo-ke*
finish-CMPL

'we already finished the collared peccary father killed yesterday'

This is important because, *a priori*, this did not have to be the case. Ergative patterns of syncretism could have been restricted according to this particular parameter. However, this does not appear to be the case.

### 2.3.6   Basic word order, ERG=GEN and ERG=POSS

It also appears that ERG=GEN and ERG=POSS are attested across a wide range of languages in terms of basic word order. In this chapter, we have already observed extensive ergative syncretism in many V-initial languages from the Austronesian family (Niuean, Kapampangan, Seediq), the Mayan family (Kaqchikel, Itzaj), and Tsimshianic family (Gitksan, Sm'algyax).

In addition, it is clear that many ergative languages outside of these families have an SOV basic word order, and furthermore than many of *them* feature either ERG=GEN or ERG=POSS (Abkhaz, Burushaski, Canela-Kraho, Kaluli, Shipibo, Sinaugoro, Sorani Kurdish, West Greenlandic). Indeed, according to the *World Atlas of Language Structures* (Dryer 2013), it appears that SOV is the most common basic word order for languages with an ergative alignment in either the dependent-marking or head-marking system. The *WALS* data is presented in (52-53) below, with accusative alignments also included for comparison.

(52) **dependent-marking alignment X basic word order**

(*WALS* features 81A X 98A)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **ergative** | VSO 1 | VOS 2 | SVO 1 | <span style="color:red">SOV 16</span> | OVS 0 | OSV 0 | *none* 7 |
| **accusative** | VSO 3 | VOS 1 | SVO 9 | SOV 29 | OVS 1 | OSV 0 | *none* 7 |

(53) **head-marking alignment X basic word order**

(*WALS* features 81A X 100A)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **ergative** | VSO 2 | VOS 0 | SVO 0 | SOV 4 | OVS 1 | OSV 1 | *none* 5 |
| **accusative** | VSO 14 | VOS 8 | SVO 65 | SOV 71 | OVS 3 | OSV 0 | *none* 29 |

Mahajan (1997) observed that the vast majority of ergative languages are 'verb-peripheral', meaning that the verb is either clause-initial or clause-final in that language's basic word order. However (as noted by Mahajan himself), there are exceptions to this generalization. And there exists at least one 'verb-medial' ergative language which exhibits ERG=POSS, Päri.

According to Andersen (1988), the unmarked word order in Päri is OVS, as in (54a) below. The alignment of dependent-marking is ergative because the transitive object and the intransitive subject NPs are unmarked, but the post-verbal transitive subject is marked with the ergative suffix *-i*.

Päri                                                                           (Andersen 1988; p.292, p.292)

(54) a. *dháagɔ̀ á-yàaɲ        ùbúrr-ì*
        woman CMPL-insult Ubur-ERG
        'Ubur insulted the woman'

    b. *dháágɔ̀ á-ŋὲɛ-ɔ̀*
        woman CMPL-laugh-SUF
        'the woman laughed'

In addition, Päri features an alternate SOV word order, as in (55a). The post-verbal subject has moved to the front of the clause, and now the clause-final verb features a suffix marker which agrees with the transitive subject, *-ὲ*. This pattern, too, is ergative because that suffix does not coindex either intransitive subjects or transitive objects, regardless of configuration.

(55) a. *ùbúr  dháagɔ̀  á-yáaɲ'-ɛ̀*
     Ubur woman CMPL-insult-3SG.ERG

     'Ubur insulted the woman'

   b. *tyend-e*
      foot-3SG.POSS

      'his feet'

As it turns out, this marker is reused for possessive morphology, as in (55b). To recap, in Päri, the morphological marker that is used to indicate transitive subjects in the SOV construction is reused for possessors of NPs, which is a clear case of ERG=POSS. The full paradigm for these ergative and genitive suffixes is provided in (56) below.[15,16]

---

[15] Suffixed vowels in Päri are subject to Advanced Tongue Root (ATR) harmony: they must always agree with the [± ATR] value of the stem vowel.

[16] According to Andersen (1988), simple stems undergo a consonant alternation in the verb root for a subset of persons in the paradigm, while in derived stems that person is represented by an additional suffix on the verb, preceding the ergative suffix. A simple vs derived stem minimal pair is provided in (i-iv) below.

Päri                                                (Andersen 1988; p.295, p.295, p.296, p.296)

i. *dháagɔ̀  á-cúɔl'-à*
   woman CMPL-call-1SG.ERG
   'I called the woman'

iii. *dháagɔ̀  á-cừɔnd'-á*
   woman CMPL-call.MULT-1SG.ERG
   'I called the woman (repeatedly)'

ii. *dháagɔ̀  á-cúɔnd'-à*
   woman CMPL-call-1PL.EXCL.ERG
   'we called the woman'

iv. *dháagɔ̀  á-cúɔnd'-ừ-wà*
   woman CMPL-call.MULT-SUF-1PL.EXCL.ERG
   'we called the woman (repeatedly)'

(56) *Päri ERG=GEN head-marking suffixes*

(adapted from Andersen 1988; p.297 / Simeoni 1978; p.35)

| | 1SG | 2SG | 3SG | 1PL.INCL | 1PL.EXCL | 2PL | 3PL |
|---|---|---|---|---|---|---|---|
| possessive | *-a* | *-i/-ɪ* | *-e/-ɛ* | *-ó/-ɔ́* | *-wa* | *-u/-ʋ* | *-gi* |
| ergative  simple | *-a* | *-i/-ɪ* | *-e/-ɛ* | *-ó/-ɔ́* | *-a* | *-u/-ʋ* | *-e/-ɛ* |
| derived | | | | | *-ù -wà/-ʋ̀ -wà* | | *-ì -gì/-ì -gì* |

### 2.3.7  Partially syncretic paradigms

In this survey of the ERG=GEN and ERG=POSS phenomena, we have encountered cases of complete paradigmatic overlap, as well as cases of partial paradigmatic overlap. Those cases which have already been discussed are in (57) below.

Given (57), it seems as if partially syncretic patterns are quite common, and thus an important component of the typology of these phenomena. In this light, we will now provide three more examples to illustrate further. As noted in Sections 2.3.1 and 2.3.3, Burushaski features ERG=GEN in its system of dependent-marking. However, there is a little quirk to the paradigm of ergative and genitive pronouns in Burushaski.

(57) *Complete and partial paradigms of ERG=GEN*

| complete overlap | | | partial overlap | | |
|---|---|---|---|---|---|
| **language** | **type** | **examples** | **language** | **type** | **examples** |
| Abkhaz | head-marking | (22-23) | Kapampangan | pronoun | (44) |
| Gitksan | head-marking | (20) | Kaqchikel | head-marking | (3) |
| Itzaj | head-marking | (14) | Päri | head-marking | (56) |
| Sm'algyax | head-marking | (31-32) | Seediq | pronoun | (17) |
| S. Kurdish | pronoun | (45-46) | Shipibo | pronoun | (41) |
| | | | W. Greenlandic | head-marking | (28-27) |

67

As per Munshi (2015), the genitive pronominal form is derived straightforwardly by concatenating the suffix -*e* to the free pronoun. Sometimes some allomorphy takes place in the shape of vowel assimilation, as in $/mi+e/ \rightarrow [mii]$ 'our'. Curiously, however, for 3SG.FEM possessives <u>only</u>, a clearly different genitive suffix -*mo* must be used. The full paradigm is in (58); as noted by Munshi, this one outlier in the paradigm is an exception, and so it serves as another good example of actual language data being 'messy'.

(58) *Burushaski pronouns*                                   (adapted from Munshi 2015; p.50)

|          | 1SG | 2SG | 3SGM | 3SGF | 1PL | 2PL | 3PL |
|----------|-----|-----|------|------|-----|-----|-----|
| free     | *ǰe* | *un* | *in* | *in* | *mi* | *ma* | *u* |
| ergative | *ǰaa* | *une* | *ine* | *ine* | *mii* | *maa* | *uwe* |
| genitive | *ǰaa* | *une* | *ine* | *inmo* | *mii* | *maa* | *uwe* |

Ayutla Mixe is another ergative language which features a partial ERG=POSS paradigm in its system of head-marking, but with many nuances. In (59a), an agreement prefix *y-* agrees with the transitive subject *Carlos* 'Carlos'. The 3rd person transitive object is not marked on the transitive verb. In (59b), the 3rd person intransitive subject is not marked on the verb either, thus making the pattern an ergative one.[17]

---

[17] Ayutla Mixe's system of head-marking is hierarchical, meaning that at most one argument is marked on the verb via agreement. The argument in question is the highest one in the hierarchy $1 > 2 > 3$. In (59a), the highest available argument is the 3rd person transitive subject, therefore the verb is realized with the *y-* 3ERG prefix. However, in a different scenario – one in which a third person subject is acting upon a first person object – it is the 1st person object which must be indexed on the verb, as in (i).

i) *yë'ë    letsy x-jojt-pëk-p*
   DEM.M milk  1.ABS-stomach-hurt-INDEP
   'the milk gave me a stomach ache'

See Romero-Méndez (2009) for more details.

(59) a. *Carlos tu'uk uk   y-ex-yp*
        Carlos one    dog **3.ERG**-see-INDEP.TR

        'Carlos saw a dog'

    b. *tsu'uts-p    yë'ë    uk*
        bite-INDEP  DEM.M  dog

        'this dog bites'

    c. *y-wet*
        **3.POSS**-cloth

        'his cloths'

Ayutla Mixe does not display ergative head-marking across the board, how-ever. According to Romero-Mendez (2009), in independent clauses, only 3rd person arguments display an ergative alignment. In (60), we see that only for the 3rd person do transitive objects and intransitive subjects pattern together to the exclusion of the transitive subject. The 2nd person displays neutral alignment in the sense that the same morpheme is used for all three grammatical roles, while the 1st person displays tripartite alignment in the sense that a different morpheme is used for each grammatical role.

(60) *Ayutla Mixe independent clause and possessive morphology*

(adapted from Romero-Méndez 2009; p.295, p.258)

|              | possessor | trans. subject | trans. object | intr. subject |
|--------------|-----------|----------------|---------------|---------------|
| 1st person   | *n-*      | *n-*           | *x-*          | $\emptyset$-  |
| 2nd person   | *m-*      | *m-*           | *m-*          | *m-*          |
| 3rd person   | *y-*      | *y-*           | $\emptyset$-  | $\emptyset$-  |

Comparing the independent clause verb agreement morphology and possessor morphology in (60), it is indeed the case that the morpheme that encodes the transitive subject always corresponds to the possessor. The problem is, outside of the 3rd person, the transitive subject morpheme is not an *ergative* one. Therefore, Ayutla Mixe head-marking does constitute an example of ERG=POSS, but only in the 3rd person.

As a final example of partially syncretic paradigms, let's consider Ngiyaamba. In this language, lexical NPs show an ergative alignment. In (43a), the transitive subject *mayiŋ* 'person' is marked with an ergative suffix *-gu*. However, in (43b) and (43c), the same noun occurs in its unmarked absolutive form when acting as transitive object and intransitive subject, respectively.

Ngiyambaa                                    (Donaldson 1980; p.90, p.219, p.117, p.230)

(61) a. *ga:-nhi*      =*naŋ-gal mayiŋ-gu*   *gana:-ga*
        carry-PAST =ABS-PL person-ERG shoulder-LOC

    'the men carried them on (their) shoulders'

   b. *mu:n-miyi*    =*lu*        *dhi:rba-nhi mayi*
      all-TR.PAST 3SG.ERG know-PAST person

      'he knew everybody'

   c. *dhiŋga:-ŋa:n*              *mayi*   *ga-ɽa*
      meat-skilled.at.catching person be-PRES

      '(she) is a person who is always catching meat'

   d. *mayiŋ-gu*    *ŋiya*
      person-DAT law

      'blacks' law'

Looking at (43d) suggests that Ngiyambaa lexical NPs exhibit ERG=GEN.[18] After all, the same suffix that encodes transitive subjects also encodes possessors. In actuality, only a subset of lexical NPs exhibit ERG=GEN. The table in (62), taken from Donaldson (1980), illustrates the range of nominal subclasses in the Ngiyambaa lexicon. If it is representative of the whole, it suggests that ERG=GEN applies to approximately half of all Ngiyambaa lexical NPs.

(62)  *Ngiyambaa nominal case forms*          (adapted from Donaldson 1980; p.82)

| gloss | absolutive | ergative | dative |
|---|---|---|---|
| 'spear' | *mura* | *muragu* | *muragu* |
| 'dog' | *miri* | *mirigu* | *mirigu* |
| 'snake' | *dhuru* | *dhuruŋgu* | *dhuruŋgu* |
| 'sand goanna' | *dhuli:* | *dhuli:ŋgu* | *dhuli:ŋgu* |
| 'child' | *bura:y* | *bura:dhu* | *bura:ygu* |
| 'emu' | *ŋuruy* | *ŋurundhu* | *ŋuruyŋgu* |
| 'mosquito' | *gamugin* | *gamugindu* | *gamugingu* |
| 'stone' | *gaɽul* | *gaɽulu* | *gaɽulgu* |
| 'prickle' | *mugar* | *mugaru* | *mugargu* |
| 'rabbit' | *yura:bad* | *yura:badu* | *yura:badgu* |
| 'fence' | *badig* | *badigu* | *badigu* |
| 'trap' | *dhara:b* | *dhara:bgu* | *dhara:bgu* |

Lexical NPs are not the only grammatical domain of Ngiyambaa which shows hints of ERG=GEN, either. In Ngiyambaa, 3rd person pronominal clitics show an

---

[18]  Like many Pama-Nyungan languages, in Ngiyambaa the genitive case is syncretic with the dative, and thus often labeled as such. See Donaldson (1980) for details.

ergative alignment. In (63a), the 3SG transitive subject is encoded by the clitic =*lu*, however in (63b) and (63c), it is the clitic =*na* which corresponds to the 3SG transitive object and 3SG intransitive subject, respectively.

Ngiyambaa                                                         (Donaldson 1980; p.43, p.94, p.89, p.90)

(63)  a.  *ya:la:-bu*        =*lu*        =*na*        *má-l-ágá*
          thus.EST-UNIV 3SG.ERG 3SG.ABS make-CM-IRR

          ''that's just how he would make it!'

      b.  *bulagar-a* =*dhu*        =*na*        *ŋa:-nhi*
          two-LOC    1SG.NOM 3SG.ABS see-PAST

          'I saw him twice'

      c.  *wanhdha-la-wa:*              =*na*        *ga-ṛa*
          which.place-LOC-EXCLAM 3SG.ABS be-PRES

          'whereat is it?'

      d.  *ŋa:-y-aga*    =*ndu*      *dhina-lugu*    *miŋga-ga*
          see-CM-IRR 2SG.NOM foot-3SG.POSS burrow-LOC

          'you will see his foot(mark) by the burrow'

The shape of the genitive 3rd person marker is intriguing. On the surface, it appears as =*lugu*, as in (63d). But, as noted by Donaldson, it can straightforwardly be broken down into two components, the ergative 3SG clitic =*lu*, and the lexical NP ergative/dative/genitive clitic =*gu*. If his analysis is correct, this is an extremely curious case of ERG=GEN: a little unusual, but still revealing a close relationship between ergativity and possession in Ngiyambaa.[19]

---

[19]  Note: while Ngiyambaa lexical NPs and 3rd person pronouns show an ergative alignment, 1st and 2nd persons show an accusative alignment, meaning that Ngiyambaa exhibits split ergativity based on person, which in turn, makes 1st and 2nd persons ineligible for ERG=GEN. For comparison, some 1st/2nd person free pronoun examples are in (i-iii) below, while 1st/2nd person clitic pronoun examples are in (iv-vi).

## 2.3.8  Borderline cases?

On occasion, some languages exhibit a pattern which exists right on the cusp of being either ERG=GEN or ERG=POSS, and thus they merit further discussion. Take, for example, Lhasa Tibetan. According to Denwood (1999), the ergative marker consists of lengthening and changing the vowel quality of pronouns and lexical NPs alike. In (64a), the ergative marker has concatenated with the transitive subject *nga* '1SG' to produce *ŋɛ:* after morphophonological processes have taken place. The pattern is an ergative one because neither transitive objects nor intransitive subjects receive this marking, as in (64b) and (64c), respectively.[20]

Lhasa Tibetan                                          (Denwood 1999; p.197, p.196, p.123, p.101)

(64) a. ⌟*ŋɛ:*      ⌟*jɪgɪ*   ⌝*taŋdʑɔ*
        1SG.ERG letter send-VOL

        'I can easily send a letter'

---

i) *ŋindu*      *ŋadhi:*    *ŋa:-nhi*
  2SG.NOM 1SG.ACC see-PAST

  'you saw me'

ii) *ŋinu:*      *ŋadhu*    *ŋa:-nhi*
  1SG.NOM 2SG.ACC see-PAST

  'I saw you'

iii) *ya:nhdhu*    *ŋindu*      *badhiyi*
  at.this.time 2SG.NOM arrive.PAST

  'at last you've come'

iv) *minjaŋ-gu-wa:*      *=ndu*    *manabi-giri ŋiya-ɽa*
  what-DAT-DISCOV 2SG.NOM hunt-PURP   say-PRES

  'why do you say "hunt"?'

v) *ŋa:-y-a:li-y-aga*      *=dhu*    *=nu:*
  see-CM-AGAIN-CM-IRR 1SG.NOM 2SG.ACC

  'I'll see you again!'

vi) *waŋa:y* *=ndu*    *badha-l-buna-y-aga-la*
  NEG    2SG.NOM come-CM-BACK-CM-IRR-THEN

  'then you won't come back'

There are no 3rd person free pronouns in Ngiyambaa, demonstratives are used instead. See Donaldson (1980) for more details.

[20] According to Denwood (1999), the ergative marker is optional in (64b) when honorific concord between special forms of the 3SG pronoun *kʰɔ̃:* and the verb *tã:* make identification of the subject straightforward. Denwood also notes that the same sentence, with the ergative marker, is accepted by informants.

b. ⌐$k^h\tilde{ɔ}$:    ⌐$ŋa$ ⌐$t\tilde{a}$: ⌐$naŋbəre$:
   3SG.HON 1SG send do-LINK-AUX

‘he sent me’

c. ⌐$ŋa$ ⌐$naŋlə$   ⌐$jø$:
   1SG home-LOC exist

‘I was at home’

d. <span style="color:red">⌐$ŋɛ$:</span>   ⌐$lagbə$
   <span style="color:red">1SG.GEN</span> hand

‘my hand’

An extremely similar genitive particle exists, which is spelled differently in the orthographical system of Tibetan but which produces the same morphophonological effect, as seen in (64d). Denwood writes "[the ergative marker] can be difficult to distinguish from the genitive particle [...] in rapid speech, but differs in that its vowel is often longer and closer and may bear falling pitch, and in that it is more likely to be followed by a pause" (p.193).

What to make of this? It appears to be a match for ERG=GEN as per (64). Combing through the data in Denwood (1999) further, it is possible to find more examples where the ergative and genitive forms do overlap (3SG $k^hɔ$ in 65), but also some where they don't overlap (3SG.HON $k^h\tilde{ɔ}$: in 66), and some where they sometimes overlap, and they sometimes don't (2SG.HON $kjer\tilde{a}$: in 67).

Lhasa Tibetan     (Denwood 1999; p.145, p.240, p.198, p.149, p.156, p.221, p.205)

(65) a. <span style="color:red">⌐$k^hø$:</span>   ⌐$jɪgɪ$ ⌐$taŋmas\tilde{ʊ}$
    <span style="color:red">3SG.ERG</span> letter send-NEG-AUX

‘he didn't send a letter’

b. ⌐*kʰø:* ⌐*lʌbdã:lə* ⌐*tɛ:nə* ⌐*kʰambə* ⌐*jɪmbədra*
   **3SG.GEN** speak-NOM-LOC look-SUB Khamba be-LINK-AUX

   'to judge by his accent, he is probably a Khamba'


(66) a. ⌐*kʰɔŋgi:* ⌐*dɪmi:* ⌐*nam* ⌐*pʰe:sũ*
   3.HON.ERG key bring-AUX

   'it was he who brought the key'

   b. ⌐*kʰɔŋgɪ* ⌐*tsʰɛnlə* ⌐*ɕrestʰa* ⌐*ɕʊdzã:* ⌐*pɛ:bʊ* ⌐*jɪmbəre:*
   3SG.HON.GEN name-LOC Shrestha say-SUB Nepalese be-LINK-AUX

   'as his name is Shrestha, he will be Nepalese'


(67) a. ⌐*kʰɛ:sə* ⌐*kjeraŋgɪ* ⌐*ŋa:* ⌐*lɤpnə* ⌐*ŋɛ:* ⌐*kje:* ⌐*yɔŋbəjø:*
   yesterday **2SG.HON.ERG** 1SG.LOC tell-SUB 1SG.ERG carry come-LINK.AUX

   'if you had told me yesterday, I would have brought it'

   b. ⌐*kjeraŋgi:* ⌐*lɛ:gə* ⌐*tɕɛ:* ⌐*tsʰa:nɪ* ⌐*ŋa* ⌐*jʊŋgyjĩ:*
   2SG.HON.ERG work did finish-SUB 1SG come-LINK-AUX

   'I'll come after you've finished work'

   c. ⌐*kjeraŋgɪ* ⌐*tønda:lə* ⌐*re:*
   **2SG.HON.GEN** benefit-LOC be

   'it is for your benefit'


Given the data in (64-67), we can consider Lhasa Tibetan to exhibit ERG=GEN, but with the caveat that the facts are not always 100% clear.

A similar scenario is found in Lezgian, which shows an ergative alignment in its system of dependent-marking. Absolutive forms are either bare, as in (68a), or marked by the suffix *-m*, as in (69a). Ergative forms are derived by suffixation of the 'oblique stem', which is conditioned by lexical class – the oblique stem is *-di* and *-da* in (68b) and (69b), respectively. Finally, genitive forms consist of first deriving

the 'oblique' (i.e. ergative) form, and then concatenating the genitive suffix *-n*, as in (68c) and (69c).

Lezgian                                    (Haspelmath 1993; p.170, p.104, p.84, p.85, p.97, p.86)

(68) a. *kac      stol-di-n  k'anikaf    xkec'-na*
        cat.ABS table-GEN from.under go.out-AOR
        'the cat came out from under the table'

    b. *gada-di  utanmišwil-äj wiči-n    wil-er  čünüx-iz alaqʰ-na*
        boy-ERG shame-INEL    self-GEN eye-PL hide-INF strive-AOR
        'the boy tried to hide his eyes out of shame'

    c. *koridor-di-n cl-a*
        hall-GEN      wall-INESS
        'on the wall of the hall'

(69) a. *a-m       ask'an buj-di-n,      q̄umral jak'-ar-i-n    žehil   tir*
        3SG-ABS low     stature-GEN dark     flesh-PL-GEN youth COP.PAST
        'he was a young man of low stature and dark complexion'

    b. *a-da       Šafiga-di-n amal-r-i-kaj    fikir-zawa-j*
        3SG-ERG Šafiga-GEN prank-PL-SBEL think-IMPF-PAST
        'he thought about Šafiga's pranks'

    c. *a-da-n     rik'-e       mühübat-di-n c'elxem kük'ün-zawa-j*
        3SG-GEN heart-INESS love-GEN    spark    burn-IMPF-PAST
        'the spark of love was burning in her heart'

According to Haspelmath (1993), in exactly this environment genitive forms are likely to be neutralized 'back' into to the oblique stem: "in the sequence /Vn/ [...] the [n] is often deleted with accompanying nasalization of the vowel', e.g. *zun*

'I' [zun]/[zũ]" (p.35). This is extremely interesting given that there is experimental evidence that non-phonemic vowel nasalization is not perceived differently from oral vowels (Beddor & Krakow 1999). If this is correct, it indicates that Lezgian forms such as *ada* '3SG-ERG' and *adan* '3SG-GEN' may have converged, providing yet another example of ERG=GEN.

Sometimes, there is an ERG=GEN pattern lurking within the grammar of a language, even though it is not common or immediately obvious. Cavineña is one such language. As a general rule, the use of the ergative is constrained in Cavineña – it only arises in a subset of transitive clauses. According to Camp (1985), in transitive clauses where a pronoun is acting upon a lexical NP, the ergative suffix is simply not used; in the inverse scenario, however, the ergative *-ra* suffix does occur. These two types of clauses are illustrated in (70a) and (70b), respectively.

Cavineña                                                           (Camp 1985; p.44, p.44)

(70) a. *tudʸa ya-ce waka kʷana ba-ca-čine           umada*
      so   1-DU cow  PL    see-arriving.object-PAST many
      'so we two saw many cows coming'

    b. *a-y     ya-ce iyakʷa waka-ra*
      do-PRES 1-DU now   cow-ERG
      'the cows will get us now'

In addition, in transitive scenarios where one pronoun is acting upon another, the use of the ergative suffix is constrained by a strict hierarchy $1 > 2 > 3$. Only if a lower ranked pronoun is acting upon a higher-ranked one does the ergative suffix arise as in (71a); in the inverse scenario, ergative morphology is not licensed (71b-71c).[21]

---

[21] In addition to not showing any ergative morphology, in sentences like (71b-71c) where a higher-ranked pronoun is acting upon a lower ranked one, an additional constraint requires that OSV word order must be observed, instead of the more basic SOV. See Camp (1985) for more details.

Cavineña (Camp 1985; p.45, p.45, p.49)

(71) a. *tudʸa tu-na-ra   e-kʷana aahe-kʷare*
    so   3-PL-ERG 1-PL   chase-REM.PAST
    'so they chased us' (3>1)

   b. *yusurupai mi-ke   e-kʷana a-ya*
    thank   2-FORM 1-PL   do-PRES
    'we thank you' (1>2)

   c. *hadʸa tu-ke-mi   e-a-ti-u*
    thus   3-FORM-2 might-do-go-might
    'thus you might get it' (2>3)

However, in the event that two equally-ranked (i.e. third person) arguments are involved in a transitive scenario, something interesting occurs: exceptionally, the genitive pronominal form may be used to encode a 3rd person transitive subject acting upon a 3rd person object, as in (72a-72c). Again, it seems as if the grammar of Cavineña has a system of constraints in place to avoid using ergative morphology except in a specific set of scenarios (ones in which lower-ranked arguments act upon higher-ranked arguments). An example of 'regular' use of the genitive is provided in (73), thus providing evidence for yet another instance of ERG=GEN.

Cavineña (Camp 1985; p.52, p.52, p.52, p.39)

(72) a. *ehe  buča ara-na-e       re-ha-ri           tuhuri       a-ma*
    ig(?) like   eat-come-manner 3.PROX-GEN-3.PROX mosquito.net do-NEG
    'it would never eat this mosquito net'

   b. *tueke-dʸa tu-ha-tu medu-kʷare     e-kʷe   babi-či*
    next-EMP 3-GEN-3 claw-REM.PAST 1-GEN grandfather.DIM
    'next it clawed my grandfather'

78

c. *ta-ce kas-tere-wa-hu*           *kʷita baa-cu-pa*     *tu-ha-tu*
3-DU strength-finish-REC.PAST-when$_{ds}$ very see-when$_{ss}$-REP 3-GEN-3
*amena kučiru cewe e-ka-ce*    *amena iye-kʷare*
finally knife with FORM-DU finally kill-REM.PAST

'after he saw that they were very tired, he killed the two with a knife'

(73) *tume tawi-ya-ke-dʸa-tu*     *Munišu-ha*     sapatu *tiru-kʷare*
then sleep-PRES-when-EMP-3 Minishu-GEN shoe burn-REM.PAST

'then while he was sleeping, Munishu's shoe burned'

As additional evidence that Cavineña generally avoids ergative morphology when possible, Camp identifies several other strategies. For example, "if the state of affairs rather than the action is in focus", transitive clauses can be de-transitivized morphologically as in (74a) (p.54). Another strategy is to use reflexive morphology in a transitive clause that clearly does not have reflexive semantics, as in (74b). In any case, the realization of ergative morphology is complicated by many factors in Cavineña. Its use is limited to a strict subset of pragmatically-conditioned scenarios. But, in just the right scenario, we see that Cavineña, too, exhibits ERG=GEN.

Cavineña                                                           (Camp 1985; p.54, p.56)

(74) a. *tume-ke*     *tee-hu-dʸa-tu*   *dukʷeri kʷawe*   *ara hu-kʷare*
that-which field-in-EMP-3 deer     manioc eat do-REM.PAST

'in that field the deer was eating manioc'

b. *hadʸa hua-cu-tu*    *kučiru ka-haka-ti-kʷare*         *ekʷita*
thus do-when$_{ss}$-3 knife SELF-sharpen-SELF-REM.PAST man

'afterward the man sharpened his knife'

### 2.3.9 Previous typological work on ERG=GEN

Several authors have written on ERG=GEN patterns previously. The largest typological survey is by Palancar (2008). Reporting on a survey of 140 languages, he cites instrumental case as the most common grammatical category in syncretic

relationship with ergative case. Palancar also reports possession as "another semantic category that ergative markers [...] express cross-linguistically, but in overall frequency, the pattern is much less common", citing the Inuit and Tibeto-Burman languages families, as well as "sporadic" cases in some Mixe-Zoquean and Caucasian languages.

As we will see, the results in Section 2.5 indicate that ERG=GEN (and its head-marking counterpart ERG=POSS) are a fairly common occurrence, occurring in just over half the languages in the sample. One obvious reason for this apparent difference in results is that Palancar focuses exclusively on clear-cut cases of ergative dependent-marking. In this survey, instances of dependent-marking ERG=GEN are considered, but also head-marking ERG=POSS (Section 2.3.2). This is particularly significant given that ERG=POSS seems to be the most common sub-type of the pattern, occurring in over two-thirds of head-marking languages in the sample. In addition, it is not clear how Palancar treats cases of partial paradigmatic overlap, which appear frequently in the sample as well (Section 2.3.7), or cases that we have identified as 'borderline' (Section 2.3.8). Thus, in this sample, only 8 languages would qualify for Palancar's stricter definition of ergative syncretism, namely Niuean, Kapampangan, Burushaski, Lhasa Tibetan, West Greenlandic, Sinaugoro, Kaluli and Shipibo.[22]

Other researchers have also commented on ERG=GEN patterns in passing. In the same anthology as Palancar, Lander (2008) remarks in his chapter on genitive case that "languages often code the possessor in a similar way to the marked participant in a transitive construction", citing Eskimo, Niuean, Ladakhi, Lak, Circassian

---

[22] Palancar (2008) also reports DAT, LOC and CAUSE/SOURCE as other attested syncretisms with ergative dependent-marking.

languages, Philippine languages and Burushaski.[23] Baker (2015) writes that "a fair percentage of ergative languages use ergative case for possessors inside nominals".[24] Finally, Blake (1994) mentions Zoque, some Caucasian languages and Eskimo languages as having a single case "common to A function and possessor function".

What all these previous sources have in common is a narrow focus on morphological ergativity realized as dependent-marking. Only Blake (1994) mentions an instance of syncretism instantiated in head-marking (Abaza, p.152). Thus, in comparison, the contribution of the survey to be conducted in this chapter can best be viewed as properly expanding the scope of morphological ergativity and ERG=GEN to include broader instances such as head-marking and pronominal systems, for example. The idea is to get a sense of how often ergative morphology – globally-speaking – overlaps with genitive morphology cross-linguistically. Additionally, the other major contribution of this survey is to freely provide the relevant data; see Appendices B and C for language-by-language data that the findings in Chapters 2 and 3 are based on, respectively.

### 2.3.10   Interim conclusion

In Section 2.3, we explored many of the possible dimensions of the ERG=GEN phenomenon, from dependent-marking to head-marking to pronominal systems. We examined major typological factors (such as headedness and basic word order), as well as the common occurrence of partial paradigms. We also discussed specific borderline cases, and whether or not they conform to ERG=GEN as defined at the outset.

---

[23] Lander also mentions Martuthunira and Karachai-Balkar as examples of languages where genitive and accusative morphology overlap, but doesn't elaborate further.

[24] See Section 4.2.1 for discussion of Baker's (2015) analysis of ERG=GEN.

The data seems clear: despite its heterogeneous nature, ERG=GEN is quite common among the world's ergative languages. The next step is to try to determine just *how* common it is as a phenomenon. To answer this question, in Section 2.4 we will construct a diverse, representative sample of the world's ergative languages, and thus calculate an estimate of the prevalence of ERG=GEN cross-linguistically.

## 2.4 The sample of ergative languages

### 2.4.1 Selecting the sample

To get a sense of how common ERG=GEN and ERG=POSS are among ergative languages, a representative sample must be constructed which accurately reflects the set of all ergative languages cross-linguistically. To quote Bakker (2011) and his article on language sampling in linguistics, "as in other cases where a population is too large to be studied in full, one has to restrict oneself to a small but representative subset by taking a sample". But building the sample isn't simply a matter of choosing a set of languages at random, and several issues must be considered.

#### 2.4.1.1 How many languages is enough?

In principle, a bigger sample size is always better than a smaller one. But for the purposes of this study, how many is enough? In answering this question, Bakker identifies two kinds of samples, *variety samples* and *probability samples*. The former involves exploring the distribution of a single grammatical feature cross-linguistically; for example "what is the distribution of lexical gender categories?" Following Bakker's guidelines, answering such a question requires a larger sample size which often numbers in the hundreds. Variety samples typically involve coding for several categories across the selected languages; in the lexical gender example above, categories could possibly be 'no grammatical gender', '2 grammatical genders', '3

grammatical genders' and '4 or more grammatical genders'. The wider range of possible categories thus warrants a larger sample size.

On the other hand, *probability samples* involve measuring the presence or absence of a single feature $X$ within languages of type $Y$. The study of ERG=GEN/ERG=POSS phenomena in question thus falls under this type of sample, as the research question is essentially, "how common is it to feature a morphological pattern of ergative syncretism (the $X$) within languages that are ergative (the $Y$)". The number of categories involved in coding is much smaller, typically either 'yes' or 'no'. Given that the scope of this kind of study is restricted from the outset, Bakker writes that these samples can be relatively small and yet still be meaningful.[25]

Bakker cites "between 50 and 200 languages" as being an appropriate size for a *probability sample*. Unfortunately, the proposed ERG=GEN study runs into a problem. While, according to some estimates, up to one-quarter of the world's languages may be ergative (Dixon 1994), there still remains of relative paucity of thorough, descriptive grammars for ergative languages, especially compared to accusative ones.[26] Thus we encounter the problem that Bakker refers to as "running into the bibliographic wall".

In the end, a sample size of **40 languages** was decided on for the purposes of practicality. This sample, while small, is good enough to provide an accurate representation of 'ergativity' broadly-speaking. The following sections detail the

---

[25] In Chapter 3, the results from the ergative sample are compared to a parallel calculation of syncretic patterns in accusative languages; see also Section 2.5.2 in this chapter which briefly anticipates the results from Chapter 3.

[26] *The World Atlas of Language Structures (WALS)* measures alignment patterns of dependent-marking in Chapter 98. Of the 190 languages sampled, 52 feature an accusative alignment and 32 feature an ergative alignment. Chapter 100 measures alignment patterns of head-marking, where the disparity is even more striking: of the 380 languages sampled, 212 feature an accusative alignment and 19 feature an ergative alignment.

precautionary measures and concerns that were considered to avoid bias in the sample.

### 2.4.1.2 Avoiding common biases in the sample

The 40 languages that were ultimately selected are listed in (75). They were selected based on availability of grammars in the literature, partially guided by the online database *WALS*, but also with a consideration for avoiding common typological sampling biases, as detailed below.

(75) *the 40 languages in the ergative sample*

**North America**

| 1. | Kaqchikel | Mayan |
| 2. | Itzaj | Mayan |
| 3. | Ayutla Mixe | Mixe-Zoquean |
| 4. | Halkomelem | Salishan |
| 5. | Sm'algyax | Tsimshianic |
| 6. | Gitksan | Tsimshianic |
| 7. | Yakima Sahaptin | Penutian |
| 8. | West Greenlandic | Eskimo-Aleut |

**South America**

| 9. | Trumai | isolate |
| 10. | Shipibo | Panoan |
| 11. | Canela-Kraho | Gê |
| 12. | Tiriyó | Carib |
| 13. | Paumarí | Arauan |
| 14. | Cavineña | Tacanan |
| 15. | Epena Pedee | Embera |
| 16. | Ika | Chibchan |
| 17. | Tenetehára | Tupian |

**Europe & Africa**

| 18. | Basque | isolate |
| 19. | Päri | Nilotic |

**Asia**

| 20. | Burushaski | isolate |
| 21. | Sumerian | isolate |
| 22. | Niuean | Austronesian |
| 23. | Kapampangan | Austronesian |
| 24. | Seediq | Austronesian |
| 25. | Tibetan | Sino-Tibetan |
| 26. | Lezgian | NE Caucasian |
| 27. | Abkhaz | NW Caucasian |
| 28. | Georgian | Kartvelian |
| 29. | Hindi | Indo-European |
| 30. | Sorani Kurdish | Indo-European |
| 31. | Hurrian | Hurro-Urartian |

**Australia & New Guinea**

| 32. | Ngiyambaa | Pama-Nyungan |
| 33. | Yingkarta | Pama-Nyungan |
| 34. | Warlpiri | Pama-Nyungan |
| 35. | Yukulta | Pama-Nyungan |
| 36. | Sinaugoro | Austronesian |
| 37. | Kaluli | TNG |
| 38. | Suena | TNG |
| 39. | Ku Waru | TNG |
| 40. | Hua | TNG |

84

Deferring to Bakker (2011) and his guidelines, the sample was carefully selected so as to avoid what he identifies as the five kinds of possible bias in a language sample. *Bibliographic bias* occurs when a sample relies too heavily on languages that have been well-documented. Thus, if a sample includes exclusively well-studied languages (such as English, Arabic, Mandarin Chinese, Hebrew, Japanese and other mainstays of the generative linguistics literature), it does not constitute a random sample that accurately captures the full range of typological variation.

Anticipating this source of sample bias, the languages in the present sample do not disproportionately belong to 'well-studied' languages or language families, as in (75). The risk of *bibliographic bias* is inherently lower anyway, given that ergative languages tend to be understudied compared to accusative ones. But just to be safe, special care was taken not to overly rely on some of the more well-known ergative languages and language families, such as Mayan or Pama-Nyungan.

The next type of sample bias to avoid is *genetic bias*. The risk is straightforward: a sample is a bad representation of cross-linguistic variation if it relies exclusively on members of one or a handful of language families. For example, a sample of ten languages that is made up of Serbo-Croatian, English, Irish, Portuguese, Czech, Russian, Hungarian, Hebrew, Warlpiri and Japanese is not a good one because Indo-European languages are over-represented. According to *Ethnologue*, only 437 of the 7,102 currently known living languages are Indo-European, which corresponds to ~6% of the world's known living languages. But in the above scenario, Indo-European languages are overrepresented because they comprise 60% of the sample.

*Areal bias* is also a risk because genetically-unrelated languages may share typological features simply due to areal diffusion. Therefore a reliable sample shouldn't feature an over-concentration of languages from a given geographical area. Note that

these biases aren't mutually exclusive; on the contrary, the dummy example discussed above is contaminated not only with *genetic bias*, but also *areal bias* and *bibliographic bias* as well.

Turning back to (75), special care was taken to ensure the sample was reasonably free of these first three types of biases. *Bibliographic bias* is inherently less of a concern when dealing with ergative languages, but still I was mindful not to rely too much on some of the well-known ergative languages families, such as Mayan and Pama-Nyungan. *Genetic bias* should not be a concern either, given that the 40 languages selected still represent 28 separate language families. Finally, *areal bias* is avoided by drawing on languages from geographically disparate regions, but keeping in mind that ergativity does cluster in several well-known 'hot spots', such as the Caucasus, the Amazon, Australia and Papua New-Guinea.[27]

Still following Bakker's guidelines, *typological bias* is also to be avoided in a language sample. *Typological bias* may occur if a random-looking sample nevertheless features a disproportionate amount of languages with one or more salient typological features, thus potentially skewing the data. For example, among languages with an S-initial dominant word order, they are approximately split half and half between SVO and SOV (Dryer 2013). But a survey of S-initial languages would be fatally flawed if it consisted of SVO languages exclusively.

With this in mind, the languages in the sample feature a balanced distribution in terms of three typological macro-parameters: basic word order, headedness and marking strategy. In (76), the 40 languages are listed by basic word order. At first glance, the sample does seem excessively weighted toward SOV basic word order, but

---

[27] To a certain extent, these ergative 'hot spots' are simply due to extremely high language diversity. In other words, regions such as Papua New-Guinea and the Amazon feature a high number of ergative languages, but they also feature a high number of languages overall.

this is justified. In (77), languages are cross-listed in terms of dependent-marking alignment and basic word order using *WALS* data. The relevant point here is that, as far as we can tell, languages with ergative dependent-marking tend to be SOV (accusative language data is also provided for reference). In (78), languages are similarly cross-listed by head-marking alignment and basic word order. Although less clear due to a low sample size, the same point stands: languages with ergative head-marking tend to be SOV (accusative data is again provided for reference). Thus, the fact that the 40 language sample is heavily weighted towards SOV basic word orders reflects a simple typological fact: ergative languages, by and large, tend to be SOV.[28]

(76) *the languages of the ergative sample by basic word order*

| b.w.o. | # | languages |
|---|---|---|
| SOV | 24 | Abkhaz, Ayutla Mixe, Basque, Burushaski, Canela-Kraho, Epena Pedee, Georgian, Hindi, Hua, Hurrian, Ika, Kaluli, Ku Waru, Lezgian, Lhasa Tibetan, Ngiyambaa, Shipibo, Sinaugoro, Sorani Kurdish, Suena, Sumerian, Trumai, Warlpiri, West Greenlandic |
| SVO | 4 | Paumarí, Yakima Sahaptin, Yingkarta, Yukulta |
| VSO | 5 | Gitksan, Halkomelem, Niuean, Sm'algyax, Tenetehára |
| VOS | 4 | Itzaj, Kapampangan, Kaqchikel, Seediq |
| OVS | 2 | Päri, Tiriyó |
| OSV | 0 | - |
| none | 1 | Cavineña |

---

[28] See Mahajan (1997) for more discussion of generalizations regarding ergativity and basic word order.

(77) **dependent-marking alignment X basic word order**

<div align="right">(<i>WALS</i> features 81A X 98A)</div>

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **ergative** | VSO 1 | VOS 2 | SVO 1 | <span style="color:red">SOV 16</span> | OVS 0 | OSV 0 | *none* 7 |
| **accusative** | VSO 3 | VOS 1 | SVO 9 | SOV 29 | OVS 1 | OSV 0 | *none* 7 |

(78) **head-marking alignment X basic word order**

<div align="right">(<i>WALS</i> features 81A X 100A)</div>

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **ergative** | VSO 2 | VOS 0 | SVO 0 | <span style="color:red">SOV 4</span> | OVS 1 | OSV 1 | *none* 5 |
| **accusative** | VSO 14 | VOS 8 | SVO 65 | SOV 71 | OVS 3 | OSV 0 | *none* 29 |

Let's now look at another major typological parameter, headedness. By and large, languages can be classified as either head-initial or head-final, according to the criteria in Greenberg (1963).[29] The table in (79) lists the languages in the ergative sample by headedness value. As might be expected, the majority of the languages in the sample are head-final, but again, the justification is the same: the majority of the world's ergative languages tend to be head-final.[30]

Finally, let's look at the languages in the sample in terms of their marking strategy. As mentioned previously, some languages have ergative head-marking, and some have ergative dependent-marking. As the two are not mutually exclusive, some

---

[29] See also Section 2.3.5 previously for discussion of headedness as per Greenbergian universals.

[30] Despite headedness criteria based on classic works such as Greenberg (1963), classifying the languages of the world into discrete categories such as 'head-initial' and 'head-final' is not always straightforward. In (79), languages in *italics* were judged to present less clear evidence for headedness than their peers, but still sorted into discrete categories due to partial fulfillment of the criteria. In the case of Yingkarta and Warlpiri, data was altogether insufficient in Dench (1998) and Keen (1972), respectively, to make a clear judgment one way or another.

languages have both - these have been labeled 'mixed'-marking languages. As the table in (80) shows, there is a nice distribution between the three major types of marking strategies.

(79) *the languages of the ergative sample by headedness*

| headedness | # | languages |
| --- | --- | --- |
| head-initial | 12 | Gitksan, Halkomelem, Itzaj, Kapampangan, Kaqchikel, Niuean, Päri, Paumarí, Seediq, Sm'algyax, Sorani Kurdish, Yakima Sahaptin |
| head-final | 26 | Abkhaz, *Ayutla Mixe*, Basque, Burushaski, Canela-Kraho, *Cavineña*, Epena Pedee, Georgian, Hindi, Hua, Hurrian, Ika, Kaluli, Ku Waru, Lezgian, Lhasa Tibetan, *Ngiyambaa*, Shipibo, Sinaugoro, Suena, Sumerian, Tenetehára, Tiriyó, Trumai, Warlpiri, West Greenlandic |
| unclear | 2 | Yingkarta, Yukulta |

languages have both - these have been labeled 'mixed'-marking languages. As the table in (80) shows, there is a nice distribution between the three major types of marking strategies.

The data and tables in (76-80) are meant illustrate that the sample of 40 ergative languages is reasonably free from *typological bias* as defined by Bakker (2011). Finally, the last type of bias discussed by Bakker is what he calls *cultural bias*. This arises when a language sample contains a disproportionate amount of languages of one cultural 'type', which in this case essentially refers to number of speakers. Here Bakker is referring to the observation that certain grammatical features appear to

only exist in "languages with only a few hundred or a few thousand speakers", citing O-initial basic word orders as a very uncommon grammatical feature that is only realized in tiny speech communities. Thus, a language sample that is too homogeneous in terms of the number speakers of each language (either homogeneously 'big' or 'small') is to be avoided. To illustrate that this is not the case in the sample of 40 ergative languages here, the table in (81) lists each language and its estimated number of speakers.[31]

(80) *the languages of the ergative sample by marking strategy*

| marking strategy | # | languages |
|---|---|---|
| head | 9 | Abkhaz, Ayutla Mixe, Canela-Kraho, Gitksan, Halkomelem, Itzaj, Kaqchikel, Sorani Kurdish, Tenetehára |
| 'mixed' | 19 | Basque, Burushaski, Epena Pedee, Georgian, Hindi, Hua, Hurrian, Ika, Kaluli, Ku Waru, Päri, Paumarí, Sinaugoro, Sm'algyax, Suena, Sumerian, Tiriyó, West Greenlandic, Yakima Sahaptin |
| dependent | 12 | Cavineña, Kapampangan, Lezgian, Lhasa Tibetan, Niuean, Ngiyambaa, Seediq, Shipibo, Trumai, Warlpiri, Yingkarta, Yukulta |

---

[31] All numbers in (81) are from *Ethnologue* (www.ethnologue.com).

(81) *the languages in the ergative sample by number of speakers*

| North America | | |
|---|---|---|
| 1. | Kaqchikel | 451,000 |
| 2. | Itzaj | 12 |
| 3. | Ayutla Mixe | 11,500 |
| 4. | Halkomelem | 285 |
| 5. | Sm'algyax | 180 |
| 6. | Gitksan | 350 |
| 7. | Yakima Sahaptin | 25 |
| 8. | West Greenlandic | 57,000 |

| South America | | |
|---|---|---|
| 9. | Trumai | 51 |
| 10. | Shipibo | 26,000 |
| 11. | Canela-Kraho | 4,400 |
| 12. | Tiriyó | 1,400 |
| 13. | Paumarí | 290 |
| 14. | Cavineña | 1,680 |
| 15. | Epena Pedee | 3,500 |
| 16. | Ika | 7,000 |
| 17. | Tenetehára | 13,100 |

| Europe & Africa | | |
|---|---|---|
| 18. | Basque | 545,000 |
| 19. | Päri | 28,000 |

| Asia | | |
|---|---|---|
| 20. | Burushaski | 87,000 |
| 21. | Sumerian | - |
| 22. | Niuean | 7,767 |
| 23. | Kapampangan | 1,900,000 |
| 24. | Seediq | 20,000 |
| 25. | Tibetan | 1,173,000 |
| 26. | Lezgian | 789,000 |
| 27. | Abkhaz | 112,000 |
| 28. | Georgian | 4,327,000 |
| 29. | Hindi | 260,330,000 |
| 30. | Sorani Kurdish | 5,000,000 |
| 31. | Hurrian | - |

| Australia & New Guinea | | |
|---|---|---|
| 32. | Ngiyambaa | 12 |
| 33. | Yingkarta | 5 |
| 34. | Warlpiri | 2,510 |
| 35. | Yukulta | 1 |
| 36. | Sinaugoro | 18,000 |
| 37. | Kaluli | 2,500 |
| 38. | Suena | 3,600 |
| 39. | Ku Waru | 40,900 |
| 40. | Hua | 21,000 |

## 2.4.2   Coding the sample

Now that I have explained the methodology for how the ergative sample was selected, it's time to explain how each individual language was analyzed. Each of the 40 languages was analyzed in the same way. Typically, one reliable source was obtained as a main reference for each language, although sometimes I used two or three sources for a given language as needed. These sources were usually published

grammars or field linguistics dissertations. In some cases, grammatical sketches were sufficient for the purposes of this sample. The one outlier is Kaqchikel, and the research for that language comes from my own field work in Guatemala from 2012-2014. That field work is largely consistent with existing accounts of Kaqchikel grammar; see Brown et al (2006) for comparison. Each language's primary source is listed in (82).

(82) *the languages in the ergative sample by primary source*

| | **North America** | | | **Asia** | |
|---|---|---|---|---|---|
| 1. | Kaqchikel | field research | 20. | Burushaski | Munshi (2015) |
| 2. | Itzaj | Hofling (2000) | 21. | Sumerian | Edzard (2003) |
| 3. | Ayutla Mixe | R.-M. (2009) | 22. | Niuean | Seiter (1980) |
| 4. | Halkomelem | Galloway (1993) | 23. | Kapampangan | Mirikitani (1972) |
| 5. | Sm'algyax | Mulder (1994) | 24. | Seediq | Holmer (1996) |
| 6. | Gitksan | Rigsby (1986) | 25. | Tibetan | Denwood (1999) |
| 7. | Yakima Sahaptin | Jansen (2010) | 26. | Lezgian | Haspelmath (1993) |
| 8. | West Greenlandic | B.-B. (1991) | 27. | Abkhaz | Hewitt (1989) |
| | | | 28. | Georgian | Harris (1981) |
| | **South America** | | 29. | Hindi | Montaut (2004) |
| 9. | Trumai | Guirardello (1999) | 30. | Sorani Kurdish | McCarus (2009) |
| 10. | Shipibo | Valenzuela (2003) | 31. | Hurrian | Bush (1964) |
| 11. | Canela-Kraho | P. & P. (1986) | | | |
| 12. | Tiriyó | Meira (1999) | | **Australia & New Guinea** | |
| 13. | Paumarí | C. & D. (1986) | 32. | Ngiyambaa | Donaldson (1980) |
| 14. | Cavineña | Guillaume (2008) | 33. | Yingkarta | Dench (1998) |
| 15. | Epena Pedee | Harms (1994) | 34. | Warlpiri | Legate (2002) |
| 16. | Ika | Frank (1985) | 35. | Yukulta | Keen (1972) |
| 17. | Tenetehára | B.-S. (1972) | 36. | Sinaugoro | Tauberschmidt (1998) |
| | | | 37. | Kaluli | G. & G. (2004) |
| | **Europe & Africa** | | 38. | Suena | Wilson (1974) |
| 18. | Basque | H. & O. (2003) | 39. | Ku Waru | M. & R. (1991) |
| 19. | Päri | Simeoni (1978) | 40. | Hua | Haiman (1980) |

Using these reference sources, each of the 40 languages was combed through for the same set of morphosyntactic features. First and foremost, each language was coded for the presence or absence of ERG=GEN (see Section 2.2 for a definition of the phenomenon). In principle, this measure will provide a simple baseline for its prevalence cross-linguistically.

In addition, all 40 languages were coded for additional morphosyntactic properties that seemed relevant as per the discussion in Section 2.3. The goal was to be able to tease apart additional cross-linguistic properties of ERG=GEN and ERG=POSS, should the data bear them out.

With this in mind, the languages in the sample were coded any of the attested variants in ergative syncretism discussed thus far: dependent-marking (ERG=GEN) head-marking (ERG=POSS) and pronominal systems (could be either). In addition, they were coded for their marking type (Sections 2.3.1-2.3.4), taking special note of 'mixed'-marking languages and the two subtypes therein (ERG-ABS vs NOM-ACC verb agreement).

The languages in the sample were also coded for their headedness (Section 2.3.5) and basic word order (Section 2.3.6). Also, as per the discussion in Section 2.3.7, languages which did exhibit a syncretic pattern were coded for whether they exhibit full or partial paradigms matches. In fact, languages exhibiting some kind of syncretic pattern were ultimately sorted into three categories, which hopefully capture the degree of paradigmatic overlap on an intuitive level.

A subset of languages exhibit **'full'** paradigmatic overlap, and thus were coded as such. The idea is as follows: regardless of marking type, these languages exhibit full overlap between ergative forms broadly-speaking and possessive forms in their grammar. For example, Sinaugoro belongs in this category because the only instance of ergativity in the grammar – the dependent-marking suffix *-na* in (7) – overlaps

with the possessive suffix. Itzaj Maya also belongs in this category because <u>all</u> of its instances of ergativity – the head-marking paradigm in (14) – overlaps completely with the possessive-marking paradigm. Finally, a language such as Kapampangan also belongs in this category, because both the ergative dependent marker (9) and the entire ergative pronominal paradigm (42) overlap completely with the corresponding possessive forms.

Another subset of languages in the sample exhibit paradigmatic overlap in most instances (i.e. more than half, but not all possible instances). This category is labeled **'partial'** overlap. As an example, Kaqchikel belongs in this category because five out of six cells of its ergative head-marking and possessive head-marking paradigms overlap (3). In the same way, Shipibo pronominal forms (41) and Burushaski (58) pronominal forms overlap in most – but not all – cells of their ergative and possessive paradigms.

Finally, a subset of languages only exhibit **'minor'** syncretic overlap, which is defined as follows: less than half of the possible instances of ergativity in the grammar exhibit either ERG=GEN or ERG=POSS, but crucially 3rd person and/or lexical NP forms are among them. For example, it was shown in (8) that Niuean lexical NPs use the same morphology for ergative and possessive forms. However, this is not the case for pronominal forms, which concatenate *ha-* to derive pronominal possessives, but *e* for ergative pronouns, as in (83).[32]

---

[32] Note: in Niuean, the ergative marker for pronouns is *e*, but the absolutive marker for lexical NPs is also *e*. The full ergative-absolutive paradigm for all forms is in (i) below. See Seiter (1980) for more details.

i)

|            | ERG | ABS |
|------------|-----|-----|
| lexical NPs | *he* | *e* |
| pronouns    | *e*  | *a* |
| proper nouns | *e* | *a* |

Niuean

(83) a. *kua  nākai kitia e    au  e    pusi*
       PERF NEG   see   ERG 1SG ABS cat

     'I haven't seen the cat'

  b. *ne  fia   taute e    ia   e    motokā ha-aku*
      PST want fix   ERG 3SG ABS car      POSS-1SG

     'he was willing to fix my car'

Another language which belongs in the 'minor' category is Ngiyambaa. Recall from the table in (62) that a large subset of Ngiyambaa lexical NPs exhibit ERG=GEN, as per Donaldson (1980). Because this is an important subset of the lexicon, but still not more than half, Ngiyambaa is sorted into this category.

Finally, recall from (6) that Kaluli lexical NPs exhibit ERG=GEN. It turns out, however, that another area of the Kaluli grammar, demonstrative forms, are in an ergative-absolutive alignment. Furthermore, these ergative demonstrative forms do not overlap with possessive demonstratives, as in (84). Therefore, because a subset of ergative forms, which is less than half but crucially includes at least lexical NPs, exhibits ERG=GEN, it is eligible for being sorted into the 'minor' category.[33]

(84) *Kaluli dependent marking and demonstratives*

---

[33] Note that by this metric, if a language exhibits partial paradigmatic overlap, but crucially lexical NPs/3rd person forms are not among them, it is not qualified for membership in this category, and for the purposes of this typological survey it is not considered to exhibit either ERG=GEN or ERG=POSS at all.

|  | ergative form | possessor | overlap? |
|---|---|---|---|
| lexical NPs | *-ya:/-wa:/-a:* | *-ya:/-wa:/-a:* | ✔ |
| 'this' | *wema:* | *ene* | ✘ |
| 'that' | *a:ma:* | *ene* | ✘ |
| 'other' | *noma:* | *ene* | ✘ |

Although it wasn't discussed specifically in Section 2.3, split ergativity is a well-known phenomenon within the ergativity literature (Comrie 1978, Dixon 1994, Coon 2013). Given the importance of split ergativity, the presence or absence of ergative splits was also coded for in the sample of 40 ergative languages.

Split ergativity occurs when only a subset of clauses exhibit an ergative-absolutive alignment, and they tend to occur in several well-known varieties. For example, Georgian exhibits an **aspectual split**, meaning that only in perfective aspect/past tense (85a) is an ergative-absolutive alignment observed. In non-perfective/non-past tenses, Georgian dependent marking is in the accusative alignment, as in (85b).

Georgian                                                                   (Harris 1981; p.1, p.1)

(85) a. *glex-ma      datesa        simind-i*
      peasant-ERG he-sowed-it-II-I corn-NOM

      'the peasant sowed corn'

   b. *glex-i       tesavs        simind-s*
      peasant-NOM he-sows-it-I-I corn-DAT

      'the peasant is sowing corn'

Another common ergative split among ergative languages is based on grammatical person, also known as a **person split**. Typically, events involving third persons follow an ergative-absolutive alignment, while a nominative-accusative alignment is observed elsewhere. In Yukulta, dependent-marking is in ergative alignment

when the event involves third persons; compare the transitive and intransitive constructions in (86a-86b). However, in a construction involving non-third persons, no ergative marking is observed on the transitive subject (86c).[34]

Yukulta                                                    (Keen 1972; p.121, p.244, p.115)

(86) a. *tir-iya      kanta    pa:tja mantuwara*
        snake-ERG TR.PAST bite    boy

    'the snake bit the boy'

   b. *wartja   yinka tula:tula:tja mantuwara*
       quickly PAST descend       boy

    'the boy got down quickly'

   c. *kuritja-ŋanpu-ŋari ŋata*
       see-1SG>3PL-PRES 1SG

    'I'm looking at them'

Finally, while aspectual splits and person splits are the most commonly attested in the literature (Silverstein 1976), there do exist instances of split ergativity which are not conditioned by either of these factors; for lack of a better term, we will label these **'other' ergative splits**. Being a sort of 'elsewhere case', the exact conditioning factor is bound to be more heteregeneous; however, an example is provided below for Niuean.

According to Massam et al (2006), specificity of arguments can condition whether a transitive clause will exhibit ergative alignment in Niuean. Thus, in (87a), the transitive subject *ekekafo* 'doctor' is marked with the ergative prefix as expected.

---

[34] Like many Pama-Nyungan languages, Yukulta also features a system of agreement clitics. These clitics are in a nominative-accusative alignment, and have a freer distribution than simple head-agreement, meaning that they don't necessarily cliticize onto verbal forms. Thus in (86c), the portmanteau *ŋanpu* encodes a 1SG subject and 3PL object. For the purposes of the ergative split, however, what matters is that the free 1SG pronoun *ŋata* does <u>not</u> feature an ergative suffix.

However, in (87b), because the object *tama fifine* 'a girl' is <u>non-specific</u> (marked by non-specific morpheme *taha*), *mautolu* '1PL.EX' is realized with an absolutive prefix, despite being the subject of a transitive clause. Note however that this does not constitute a canonical example of an ergative split, since the alignment in (87b) does not revert to nominative-accusative; in actuality, non-specific direct objects merely 'don't count' for the case computing algorithm, and thus the transitive subject in (87b) receives absolutive case like in ordinary intransitive clauses.

Niuean                                                                  (Seiter 1980; p.29, p.40)

(87) a. *to    lagomatai he    ekekafo a    ia*
        FUT help        ERG doctor   ABS 3SG

        'the doctor will help him'

    b. *ne    kumi   a    mautolu   he taha tama fifine   ne    taute pasikala afi*
        PAST search ABS 1PL.EXCL at NSP  child female N.FUT fix    bicycle  fire

        'we're looking for a girl who can fix motorbikes'

To summarize, all the morphosyntactic features that were coded for in the sample are listed in table (88) below, along with the possible values for each category.

(88) *Summary of relevant morphosyntactic features in the survey of ergative languages*

| feature | value |
|---|---|
| - ERG=GEN | - Y/N |
| - marking type | - dependent-marking |
| | - head-marking |
| | - 'mixed'-marking (ERG-ABS verb agreement) |
| | - 'mixed'-marking (NOM-ACC verb agreement) |
| | - pronominal |
| - headedness | - head-initial |
| | - head-final |

| | |
|---|---|
| - basic word order | - SVO |
| | - SOV |
| | - *etc...* |
| - degree of overlap | - 'full' |
| | - 'most' |
| | - 'some' |
| - split ergativity | - aspect split |
| | - person split |
| | - 'other' |

## 2.5 Results

### 2.5.1 ERG=GEN and ERG=POSS are generally common

Applying the methodology from Section 2.4.2, the main result is that **a majority of ergative languages in the sample display either ERG=GEN or ERG=POSS to some degree**. In other words, 23 out of 40 (57.5%) displayed one of the two patterns of syncretism to some degree ('full', 'partial' or 'minor').

(89) *Prevalence of syncretic patterns in the sample of ergative languages*

| **ERG=GEN or ERG=POSS** (either 'full', 'partial' or 'minor') | **<u>no</u> syncretic pattern whatsoever** |
|---|---|
| Abkhaz, Ayutla Mixe, Burushaski, Canela-Kraho, Cavineña, Gitksan, Halkomelem, Itzaj, Kaluli, Kapampangan, Kaqchikel, Lezgian, Lhasa Tibetan, Ngiyambaa, Niuean, Päri, Seediq, Shipibo, Sinaugoro, Sm'algyax, Sorani Kurdish, Sumerian, West Greenlandic | Basque, Epena Pedee, Georgian, Hindi, Hua, Hurrian, Ika, Ku Waru, Paumarí, Suena, Tenetehára, Tiriyó, Trumai, Warlpiri, Yakima Sahaptin, Yingkarta, Yukulta |
| *23/40 = <u>57.5%</u>* | *17/40 = <u>42.5%</u>* |

99

Breaking things down even further, we see that there is a fairly even distribution between the different 'degrees' of paradigmatic overlap as per Section 2.4.2. Eleven languages fall into the 'full' category, five into 'partial', and seven into 'minor'. This is an interesting result because it indicates that within these languages, full overlap is the largest subgroup, making the phenomenon quite a robust one.

(90) *Degrees of paradigmatic overlap: 'full', 'partial' and 'minor'*

| *full overlap* | *partial overlap* | *minor overlap* |
|---|---|---|
| Abkhaz, Ayutla Mixe, Canela-Kraho, Gitksan, Itzaj, Kapampangan, Lhasa Tibetan, Päri, Seediq, Sinaugoro, Sorani Kurdish | Burushaski, Kaqchikel, Shipibo, Sm'aglyax, West Greenlandic | Cavineña, Halkomelem, Kaluli, Lezgian, Ngiyambaa, Niuean, Sumerian |

(91) *Breakdown of the sample by degrees of paradigmatic overlap*
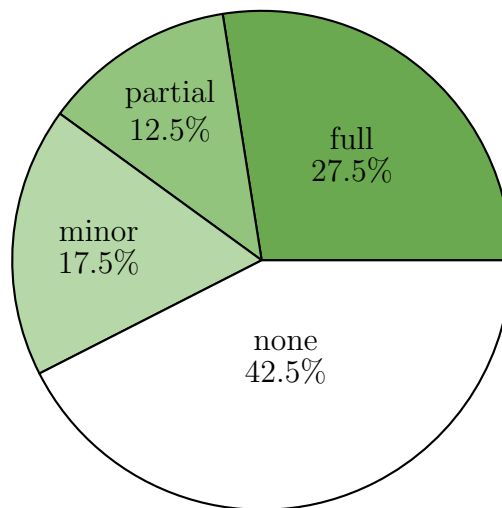
### 2.5.2  More common than what, exactly?

While Section 2.5.1 illustrates that more than half the languages in the sample exhibit either ERG=GEN or ERG=POSS, providing evidence that the phenomenon is quite common cross-linguistically, a reasonable question to ask as a follow-up is: "more common than what, exactly?" There are at least two possible ways to answer this question.

One thing to consider is the prevalence of comparable syncretic patterns in languages of nominative-accusative alignment. Anticipating the discussion and results of Chapter 3, which measures rates of NOM=GEN, NOM=POSS, ACC=GEN and ACC=POSS in a comparable sample of 40 accusative languages, we find that ergative patterns are indeed more common. In that sample, nominative and accusative syncretic patterns occur at a rate of 12.5% (n=5) and 27.5% (n=11), respectively. This difference between the ergative and accusative samples is statistically significant ($\chi^2$=7.598, $p$-value=0.022).

(92) *Nominative* (l) *and accusative* (r) *syncretism in the accusative sample*
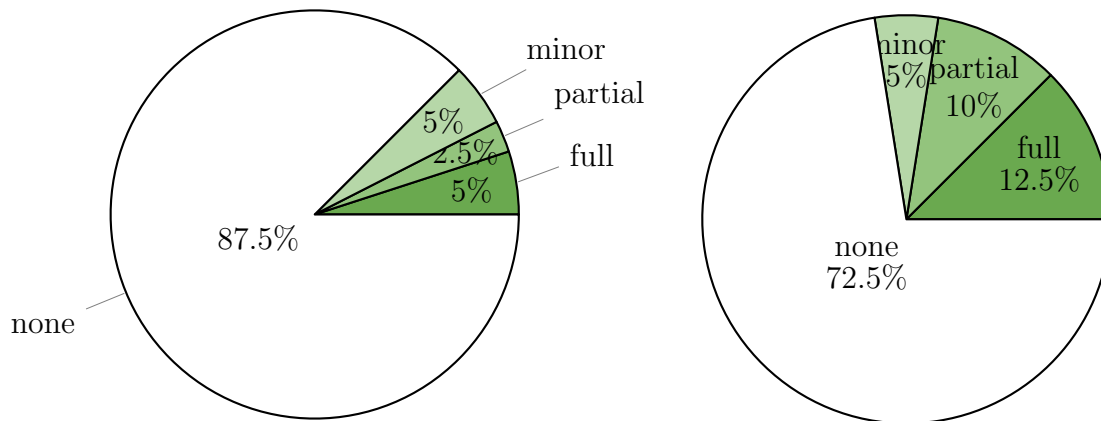


Another way to potentially answer that question is by calculating the baseline likelihood of <u>any</u> morpholgically-syncretic pattern. Given that languages sometimes

re-use functional morphology in different areas of the grammar (Ben Bruening, p.c.), we want to know whether ERG=GEN and ERG=POSS are occurring at a higher rate than chance given the range of possible syncretic options in a language.

For example, in Itzaj, the ERG=POSS pattern is realized because the ergative marker in (93a) corresponds to the possessive marker in (93d). However, there is another head-marking 'slot' on the verb in Itzaj – an absolutive marker – and therefore, another entire paradigm which could (in principle) have been recycled to encode possession in the nominal domain. Therefore, if a language like Itzaj is going to realize any syncretic pattern at all between the verbal and nominal domain, *a priori* there is a 50% chance that it will actually be ERG=POSS.

Itzaj                                                    (Hofling 2000; p.37, p.37, p.37, p.25)

(93) a. *k-      uy-      il   -aj -o'on*
        CMPL 3SG.ERG see  TR  1PL.ABS

    'he saw us'


    b. *tal    -o'on*                    c. *tal    -ij*
       come 1PL.ABS                         come 3SG.ABS

    'we came'                            'he came'


    d. *uy-        äj- kax*              e. * *äj- kax      -ij*
       3SG.POSS- M   chicken                 M   chicken

    'his chicken'                        'his chicken'

In Burushaski, the ERG=GEN pattern is instantiated in the dependent-marking morphology: the ergative suffix in (94a) corresponds to the genitive suffix in (94b). However, according to Munshi (2015), there are four other morphological

cases that can suffix to nouns in the verbal domain (94c-94f). By that same reasoning, the base likelihood of instantiating an ERG=GEN pattern (as opposed to another syncretic pattern) is 20% in a language like Burushaski.

Burushaski                                    (Munshi 2015; p.29, p.17, p.46, p.46, p.46, p.46)

(94) a. *hamal-e*       *ḍarbeṣ*   *i-cʰarkan-uman*
        neighbour-**ERG** Darbes 3SG.M.OBJ-beat-3PL.SUBJ
        'the neighbours beat Darbes'

   b. *in-e*       *laqpis*
        3SG-**GEN** handkerchief
        'his handkerchief'

   c. *haal-ar*
        home-DAT
        'to home'

   d. *haal-ulo*
        home-LOC
        'in the home'

   e. *cʰur-ate*
        knife-INSTR
        'with the knife'[35]

   f. *in-cum*
        3SG-ABL
        'from him'

Proceeding in this manner, we can calculate a language's *potential* number of syncretic patterns and baseline likelihood of any of them occurring should a syncretic

---

[35] Munshi (2015) also lists a second locative suffix in forms such as *teebal-ate* 'on the table'. However, since this does not constitute a different <u>form</u> from the instrumental *-ate* suffix, for the purposes of calculating this 'baseline likelihood' of ERG=GEN it is not counted twice.

pattern be realized. To round out this analysis, we need to include *all* instances of verbal syncretism with genitive morphology in the sample, and not just ergative ones. There are four additional non-ergative instances of syncretism in the sample: two cases of DAT=GEN (Yingkarta and Warlpiri), and two cases of ABS=POSS (Tiriyó and Tenetehára).[36,37]

(95) *Verbal-nominal syncretism in the sample of ergative languages*

| potential patterns | instances of ERG=GEN | expected rate | actual rate |
|:---:|:---|:---:|:---:|
| 1 | Päri (n=1) | 100% | 100% (1/1) |
| 2 | Ayutla Mixe, Canela-Kraho, Gitksan, Halkomelem, Itzaj, Kaqchikel, Sm'algyax$_h$, Sorani Kurdish, Sumerian, Tenetehára, Tiriyó, West Greenlandic$_h$ (n=12) | 50% | 83% (10/12) |
| 3 | Abkhaz, Kapampangan, Kaluli, Seediq, Sm'algyax$_d$ (n=5) | 33% | 100% (5/5) |
| 4 | Ngiyambaa, Sinaugoro (n=2) | 25% | 100% (2/2) |
| 5 | Burushaski, Cavineña, Lhasa Tibetan, Niuean, Warlpiri (n=5) | 20% | 80% (4/5) |
| 7 | West Greenlandic$_d$ (n=1) | 14.2% | 100% (1/1) |
| 8 | Lezgian, Shipibo, Yingkarta (n=3) | 12.5% | 66.6% (2/3) |

In (95) above, the bottom line is this: if an ergative language has a syncretic pattern between possessive morphology in the nominal domain, and something else from the verbal domain, it overwhelmingly happens to be with ergative morphology, at a rate far higher than one might expect if syncretic patterns were just randomly distributed between all the potential patterns available *a priori*.

---

[36] See Section 2.5.6 for more discussion on ABS=POSS, the much rarer counterpart to ERG=POSS.

[37] Sm'algyax and West Greenlandic occur twice in (95) because they are 'mixed'-marking languages that actually instantiate two independent ERG=GEN patterns (one in the <u>d</u>ependent-marking, one in the <u>h</u>ead-marking).

### 2.5.3 ERG=GEN and ERG=POSS transcend macroparameters

The next main finding is that ERG=GEN and ERG=POSS are fairly well-represented across major typological divisions in the sample: headedness, marking type and basic word order.[38]

### 2.5.3.1 Ergative syncretism and headedness

The data indicates that ergative patterns of syncertism are well-represented both among head-initial and head-final languages in the sample. The table in (96) lists the languages of the sample in terms of whether they exhibit ergative syncretism and by headedness following the guidelines in Section 2.3.5. Ten total head-initial languages exhibit a syncretic pattern, as do thirteen head-final languages.

(96)

| head-initial languages | | head-final languages | |
|---|---|---|---|
| ✔ | ✘ | ✔ | ✘ |
| - Gitksan | - Paumarí | - Abkhaz | - Basque |
| - Halkomelem | - Yakima Sahaptin | - Ayutla Mixe | - Epena Pedee |
| - Itzaj | | - Burushaski | - Georgian |
| - Kapampangan | | - Canela-Kraho | - Hindi |
| - Kaqchikel | | - Cavineña | - Hua |
| - Niuean | | - Kaluli | - Hurrian |
| - Päri | | - Lezgian | - Ika |
| - Seediq | | - Lhasa Tibetan | - Ku Waru |
| - Sm'algyax | | - Ngiyambaa | - Suena |
| - Sorani Kurdish | | - Shipibo | - Tenetehára |
| | | - Sinaugoro | - Tiriyó |
| | | - Sumerian | - Trumai |
| | | - West Greenlandic | - Warlpiri |

---

[38] In this section's charts and tables, the symbol '✔' refers to instances of ERG=GEN and/or ERG=POSS in the sample, while '✘' refers to languages with no observed syncretism.

(97)
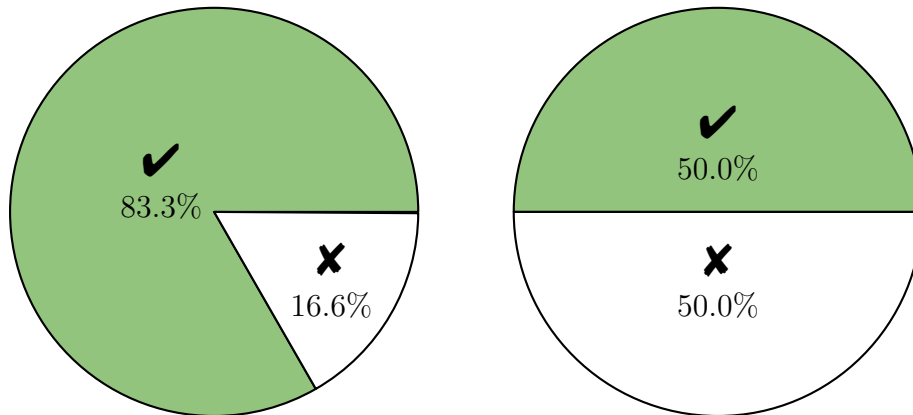


However, an observation is in order: while it is true that there is a healthy number of both head-initial and head-final languages exhibiting syncretism, the proportions aren't quite the same. It appears as if head-initial languages have a greater likelihood of exhibiting a syncretic pattern (83%), compared to the head-final ones (50%). This correlation is just outside of the range of being statistically significant ($\chi^2$=3.818, $p$-value=0.051).

In any case, this data is interesting because 'it didn't have to be that way'. In other words, it could have been the case that only head-initial languages are ERG=GEN. Or, that all head-final languages are ERG=POSS. But that is not borne out by the data, which suggests that **there is something inherent about ergativity itself**, transcending its exact morphosyntactic instantatiation, that makes ergative syncretic patterns as a whole likely to occur.

#### 2.5.3.2 Ergative syncretism and marking type

In the same vein, the data also indicates that ERG=GEN and ERG=POSS are well-represented among all three types of languages discussed in Sections 2.3.1-2.3.3: dependent-marking, 'mixed'-marking and head-marking. The table in (98) lists the languages of the sample by marking strategy.

(98)

| +DEPENDENT MARKING | | +HEAD MARKING |
|---|---|---|
| Cavineña, Kapampangan, Lezgian, Lhasa Tibetan, Ngiyambaa, Niuean, Shipibo (n=7) | Burushaski, Kaluli, Päri, Sinaugoro, Sm'algyax, Sumerian, West Greenlandic (n=7) | Abkhaz, Ayutla Mixe, Canela-Kraho, Gitksan, Halkomelem, Itzaj, Kaqchikel, Seediq, Sorani Kurdish (n=9) |
| Trumai, Warlpiri, Yingakarta, Yukulta (n=4) | Basque, Epena Pedee, Georgian, Hindi, Hua, Hurrian, Ika, Ku Waru, Paumarí, Suena, Tiriyó, Yakima Sahaptin (n=12) | Tenehetára (n=1) |

(99)
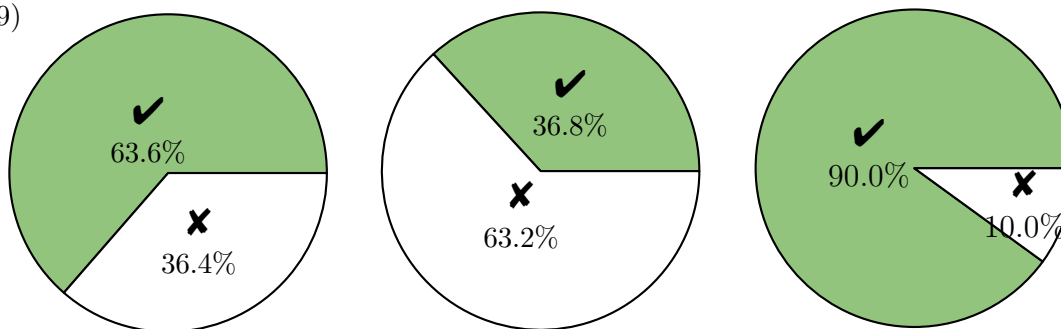


Again, it is perhaps surprising that languages across this spectrum exhibit some form of ergative syncretism. These languages belong in comparable numbers to the dependent-marking group, the 'mixed'-marking group, and the head-marking group. It needn't have turned out that way; it could have been the case that 'mixed'-marking languages <u>never</u> show ERG=GEN or ERG=POSS, for example. Therefore this data, too, suggests that there is something about ergativity itself which is strongly compatible with ergative patterns of syncretism.

107

That being said, something conspicuously jumps out of the table in (98): all types of marking are fairly evenly represented among languages that have a syncretic pattern, but the same cannot be said about languages that do <u>not</u>. Among languages that lack ERG=GEN or ERG=POSS, 'mixed'-marking are by far the most numerous (n=12), followed by dependent-marking (n=4) and finally head-marking (n=1). This correlation is statistically significant ($\chi^2$=7.810, $p$-value=0.020).

The data in (98) suggests that of all marking types, the head-marking variety of ergative syncretism, ERG=POSS, is most likely to occur. Further evidence for this hypothesis is provided if we unpack the 'mixed'-marking category back into binary head-marking and dependent-marking categories (for example, if we count a language like West Greenlandic 'twice', both as an instance of head-marking ERG=POSS and as an instance dependent-marking ERG=GEN). Remember that while all 'mixed'-marking languages feature ergative dependent-marking, only a subset of them also feature ergative head-marking (many of them instead featuring NOM-ACC head-marking).

(100) *ERG=GEN* (l) *vs ERG=POSS* (r)
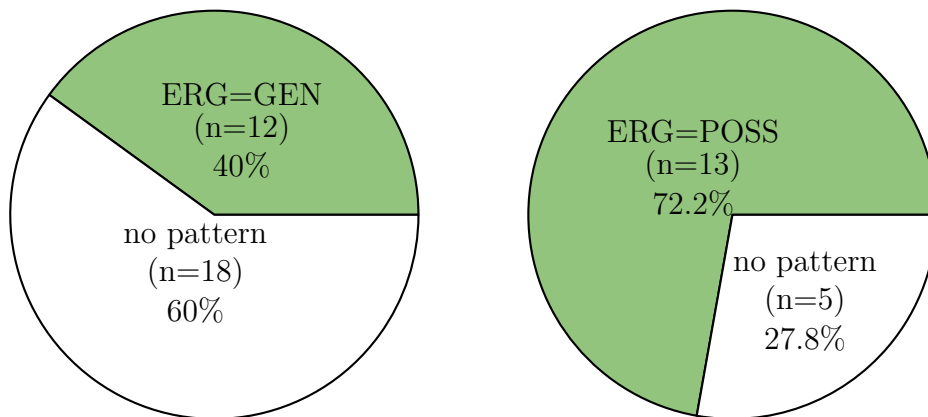
Thus the chart in (100) illustrates the data if we reapportion instances of ergative dependent-marking and ergative head-marking from 'mixed'-marking languages back into those two separate categories. In other words, (100) tracks the **total instances** of ergative dependent-marking (n=30) and head-marking (n=18) in the sample. Again, the same observations hold: generally speaking, ergativity seems to lend itself well to both ERG=GEN and ERG=POSS. That being said, however, ERG=POSS does occur proportionally more often than ERG=GEN in the sample. This difference between marking types is statistically significant ($\chi^2$=4.680, $p$-value=0.031).

### 2.5.3.3 Ergative syncretism and basic word order

As discussed in Section 2.4.1.2, the languages in the survey feature a range of basic word orders. The data reveals that globally, ERG=GEN and ERG=POSS are represented in all three major groups (V-initial, V-medial and V-final), but not equally ($\chi^2$=7.764, $p$-value=0.021). This is consistent with the hypothesis that there is something inherent to ergativity, transcending major typological parameters, that lends itself to ergative syncretic patterns, and yet that some typological settings are more likely than others to feature them.

(101) *ERG=GEN and basic word order*
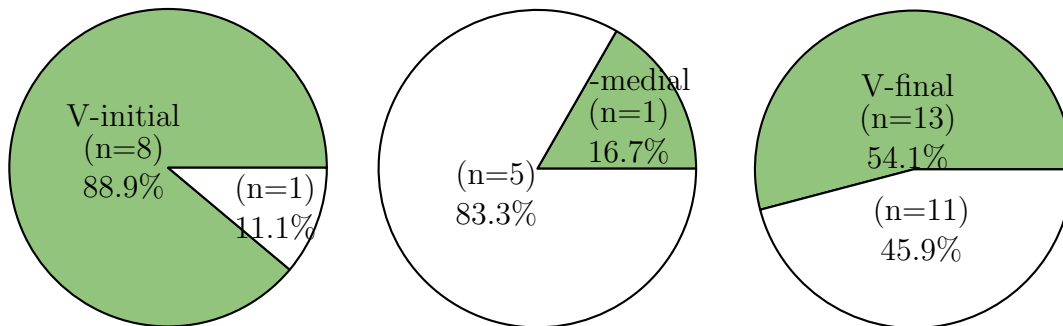
To summarize the findings reported in Section 2.5.3, ergative patterns of syncretism seem to occur across the spectrum of morphologically ergative languages, and yet the numbers in the sample indicate that certain kinds of languages are more likely to instantiate such a pattern. Head-initial, V-initial languages are more likely to instantiate a syncretic pattern that head-final, non-V-initial ones. Furthermore, ERG=POSS occurs at a proportionally higher rate than ERG=GEN. Why should a cluster of grammatical properties (head-initial, V-initial, head-marking) favor higher rates of syncretism? The data in the sample seems to hint at an important difference between these languages types cross-linguistically. Unfortunately, as of right now I do not have any explanation for these differences as this point in time, and I leave the subject open to further research.

### 2.5.4 'Ergative robustness'

Section 2.5.4 develops the argument that languages that are less 'robustly ergative' are more likely to lack the any kind of ergative syncretic pattern.

### 2.5.4.1 Ergative splits

As discussed in Section 2.4.2, many languages in the survey feature ergative splits of some kind. The data reveals that languages lacking both ERG=GEN and ERG=POSS almost all feature an ergative split of some kind (88.2% of cases), while languages possessing either ERG=GEN or ERG=POSS do sometimes feature ergative splits, but much less often (34.7%). The diagrams in (102) illustrate this result, breaking down the data by type of ergative split as per Section 2.4.2 ($\chi^2$=11.429, $p$-value<0.001).

110

(102)
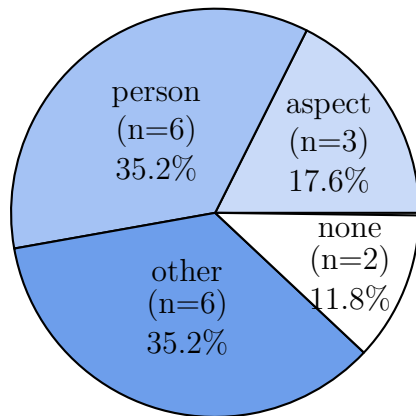


*splits in languages with neither*

*ERG=GEN nor ERG=POSS*
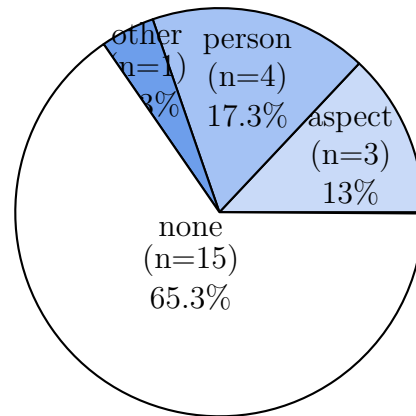
*splits in languages with either*

*ERG=GEN or ERG=POSS*

### 2.5.4.2 Optional ergative marking

Another fact that emerges from the survey is that it is fairly common for ergative dependent-marking to be optional. In other words, there are many languages which feature an ergative case marker that marks transitive subjects, but the usage of this marker is optional. Hua is one such language. In (103a), the noun *fu* 'pig' is marked with a topic suffix, and due to the language's flexible word order the sentence can either mean 'Buro saw the pig' or 'the pig saw Buro', depending on context. However, in (103b), *fu* 'pig' is marked with an ergative suffix; now disambiguated, the sentence can only mean 'the pig saw Buro'.

Hua                                                                  (Haiman 1980; p.361, p.361)

(103) a. *Buro' fu-mo    ke*
        Buro   pig-TOP see.3SG
      'Buro saw the pig' / 'the pig saw Buro'

   b. *Buro' fu-mu    ke*
        Buro   pig-ERG see.3SG
      *'Buro saw the pig' / 'the pig saw Buro'

111

In this way, ergative dependent marking can be optional. Haiman writes: "this, essentially, is the extent of ergativity in Hua. Not only is the ergative inflection optional, however, there are other factors which weaken ergativity in Hua to virtual insignificance." Therefore it seems safe to say that languages which display this kind of optional ergative marking are 'less robustly ergative' than their counterparts which don't. As it turns out, a high proportion of languages lacking ERG=GEN feature dependent marking which is optional (62.5%); on the other hand, among languages that feature ERG=GEN, the proportion is much lower (35.7%) ($\chi^2$=2.143, $p$-value>0.05).

(104)
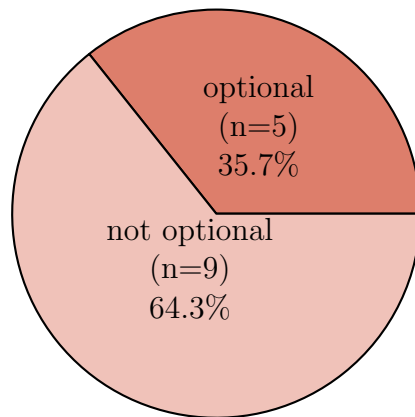


*optional ERG in*
*languages with ERG=GEN*
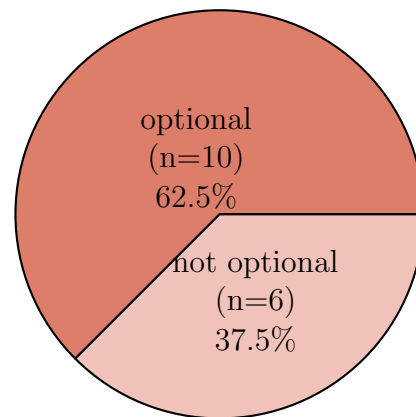
*optional ERG in*
*languages without ERG=GEN*

### 2.5.4.3 'Mixed'-marking, again

Another way to tangibly distinguish between 'more' and 'less' ergative languages is to look at 'mixed'-marking languages again. Remember that there are 19 'mixed'-marking languages in the survey. Looking more closely, we distinguish a subset that is consistently ergative (ERG/ABS dependent-marking **and** ERG/ABS head-marking) and a subset that is not (ERG/ABS dependent-marking but NOM/ACC

head-marking). While the sample sizes are small, here too the trend seems consistent: among the consistently ergative 'mixed'-marking languages, the total rate of ERG=GEN and ERG=POSS is fairly high (50%); but in the other subset, the rate is substantially lower (27.2%) ($\chi^2$=1.916, $p$-value>0.05).

(105)
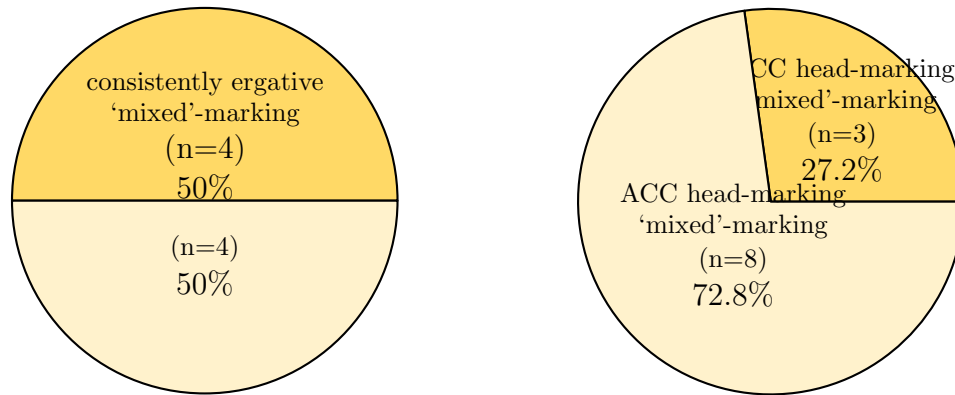


#### 2.5.4.4 ERG-ABS pronouns

Finally, another way to distinguish between 'more' and 'less' ergative languages is to look at whether these ergative languages possess ergative and absolutive pronominal forms, as per Section 2.3.4. Again, it turns out that languages that have these ergative and absolutive pronominal forms are more likely to exhibit a pattern of ergative syncretism (30.4%) than those who don't (17.6%) ($\chi^2$=0.85, $p$-value>0.05).

(106)
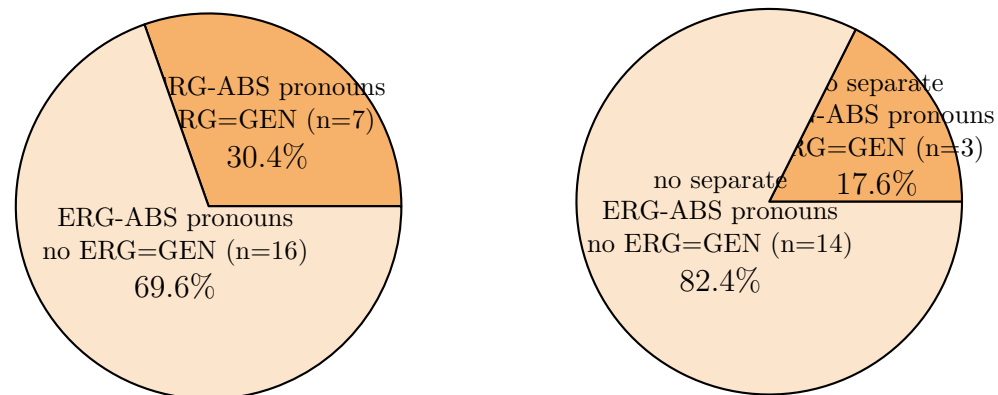
To summarize, Section 2.5.4 has argued that a reliable predictor for not having either ERG=GEN or ERG=POSS is whether a language is 'less robustly' ergative. This was measured by looking at four separate morphosyntactic areas (ergative splits, optional ergative marking, 'mixed'-marking, and the availability of ERG-ABS pronouns), and in all four cases the trend is consistent: languages that are 'less' ergative by any of these metrics are less likely to exhibit syncretism.[39]

### 2.5.5 Discussion: alienable *vs* inalienable possession

A total of 14 languages in the survey show a morphological distinction between alienable and inalienable possession. Sometimes, this is simply manifested in the presence of an arbitrary extra morpheme on either alienably or inalienably possessed nouns. For example, in Itzaj Maya, alienably possessed nouns are simply marked by possessive head-marking which agrees with the possessor, as in (107a); however, inalienably possessed nouns (such as body parts) must in addition feature an extra suffix *-el* after the nominal root, as in (107b).

Itzaj                                                                 (Hofling 2000; p.36, p.112)

(107)  a.  *u-naj*                                    b.  *u-tzo'otz-el*
          3SG.POSS-house                           3SG.POSS-hair-IA

        'his house'                               'his hair'

However, for a small subset of languages, the exact instantiation of the alienable-inalienable distinction does not seem so arbitrary. Indeed, in five of these languages,

---

[39] Why should more robustly ergative languages be more likely to exhibit syncretism across domains? Again, it is not immediately clear, but see Section 4.2 for syntactic theories which seek to capitalize on the structural similarities between ergative and genitive arguments in their respective domains.

it is specifically **alienable** possession that shows a syncretic pattern, while inalienable possession does not. For instance, in Sinaugoro, transitive subjects are marked with the ergative suffix -na, as in (108).

Sinaugoro                                          (Tauberschmidt 1992; p.181, p.188, p.189)

(108)  *tau-na      bua      e       vini-gu-to*
       man-ERG betelnut 3SG.SUBJ give-1SG.OBJ-PERF

       'the man gave me a betelnut'

(109)  a.  (*au*) *kwaku-gu*              b.  *au-na      motuka e      rakava-to*
           1SG  leg-1SG.OBJ                  1SG-GEN car    3SG.SUBJ bad-PERF

           'my leg'                          'my car got damaged'

However, in nominal constructions, only possessors of alienable possessions are marked with the same -na suffix, as in (109b). Inalienable possession is formed by a wholly different strategy: the possessor, if overt, features no additional morphology, and instead the possessed noun is suffixed with an agreement morpheme that normally encodes transitive objects on verbs![40]

Burushaski features a very similar distinction in how alienable and inalienable possession is encoded. Alienable possessors (111b), like transitive subjects (110), are marked with the suffix -e; however, inalienable possession is accomplished by an altogether different strategy. Instead, the possessed noun is suffixed with an agreement morpheme that normally encodes transitive objects on verbs (111a).

Burushaski                                         (Munshi 2015; p.27, p.13/ Holst 2014; p.43)

(110)  *hamal-e        ḍarbeṣ  i-cʰarkan-uman*
       neighbour-ERG Darbes 3SG.M.OBJ-beat-3PL.SUBJ

       'the neighbours beat Darbes'

---

[40]  This is an ACC=POSS pattern of syncretism, which is discussed in detail in Chapter 3.

(111) a. *gu-rén*
       2SG.OBJ-hand

    b. *hurmaṭ-e    oṣcum*
       Hurmat-GEN waistband

    'your hand'

    'Hurmat's waistband'

Kaluli is another language that features a similar distinction. Transitive subjects (112) and alienable possessors (113) are both marked with the suffix *-ya:*. However, kin terms are formed by a very different strategy, one that does not seem consistently morphologically decomposable (see the table in 114).

Kaluli                                                    (Grosh & Grosh 2004; p.63, p.13, p.15)

(112) *Iba-ya:   siabulu-wo: Hana  o:lia:   Sala  o:lia:    ma:no: sofa:*
      Iba-ERG potato-TOP Hanah ACCOMP Sarah ACCOMP food     cook.PAST

    'Eva cooked sweet potatoes for Hannah and Sarah'

(113) *tili-dabu      kalu-keisale gio:, Gode-ya: nanog di-a:bi*
      pull-hear.PAST man-woman 2PL  God-GEN work   do-IMP

    'you believers must do God's work!'

(114)

| 1st person | 2nd person | 3rd person | kin |
|------------|------------|------------|-----|
| *do* | *gol* | *iya* | father |
| *no:* | *go:* | *ano* | mother |
| *na:la:* | *ga:la:* | *ida:* | daughter |
| *no:l* | *go:l* | *inso* | son |
| *nao* | *gao* | *ao* | brother |
| *nado* | *gado* | *ado* | sister |

Ngiyambaa also exhibits a similar distinction. It is ERG=GEN in the sense that transitive subjects (115) and alienable possessors (116b) are both encoded by

the suffix *-gu*. However, inalienable possession follows a simpler strategy: simple apposition (116a).

Ngiyambaa                                            (Donaldson 1980; p.90, p.231, p.231)

(115)  *ga:-nhi=naŋ-gal*         *mayiŋ-gu*   *gana:-ga*
        carry-PAST=3ABS-PL person-ERG should-LOC

      'the men carried them on (their) shoulders'

(116) a. *gurugun ŋamu*                    b. *gurugun-gu ŋamu*
         cow        breast/milk              cow-GEN    breast/milk

     'cow's udder'                          'milk (in a glass)' or 'butchered udder'

     Finally, Canela-Kraho is the last language in the sample to feature this kind of sub-pattern. In (117), the transitive subject is encoded by an agreement morpheme *i-* '1SG' on the auxiliary; the transitive object is encoded by an agreement morpheme *a-* '2SG' on the verb root itself. Constructions of alienable possession feature a dummy possession morpheme *jõ* in exactly the same slot as the auxiliary in transitive clauses, to which the possessor attaches (118b); however, in inalienable constructions, there is no dummy possession morpheme and the possessor attaches directly to the noun itself, similarly to objects in transitive clauses (118a).

Canela-Kraho                                        (Popjes & Popjes 1986; p.147, p.169, p.169)

(117)  *i-te*        *a-pupun*
        1SG-AUX 2SG-see

     'I saw you'

(118) a. *a-quêtti*                    b. *i-jõ*        *wapo*
         2SG-uncle                        1SG-POSS knife

     'your uncle'                          'my knife'

To summarize, it seems important that in five unrelated languages from vastly different parts of the world, a consistent pattern emerges: ergative syncretism is reflected only in alienable possession. In no other language in the survey does the other logically-possible pattern emerge: a language which has, for instance, ERG=GEN but only in the case of inalienable possession. These facts require further study, but for now they do hint at a possible syntactic or semantic relationship between transitive subjects and alienable possessors of nouns, specifically.

### 2.5.6  Discussion: where are ABS=GEN and ABS=POSS?

One final topic remains to be addressed in this chapter: the existence of the other logically-possible counterpart to ergative syncretism: absolutive syncretism. Is it attested? How common is it?

As alluded to briefly in Section 2.5.2, only two languages in the survey consistently exhibit absolutive syncretism in plain, transitive clauses. In Tenetehára, a Tupian language, the same morpheme encodes the subject of unaccusative[41] intransitive verbs (119a) and the object of transitive verbs (119b). In addition, this same morpheme encodes the possessor of nouns, as in (119c). Thus Tenetehára qualifies as ABS=POSS.

Tenetehára                    (Harrison 1986; p.421/Bendor-Samuel 1972; p.91, p.106)

(119)  a. *he-rurywete*
           1SG-be.happy

       'I am happy'

---

[41] Tenetehára actually features a split system whereby a different set of morphemes encodes the sole argument of unergative intransitives (ex.: *a-zə* 'I run'). There is also a person hierarchy 1>2>3 such that at most one person is encoded on transitive verbs, which can be the transitive subject if it outranks the object (ex.: *a-petek* 'I beat him'. See Harrison (1986) and Bendor-Samuel (1972) for more details.

b. *he-petek*
   1SG-beat

   'he beats me'

c. *he-mukaw*
   1SG-gun

   'my gun'

In addition, Tiriyó, a Cariban language, reflects a similar situation. The same morpheme encodes the subject of intransitive verbs (120a), the object of transitive verbs (120b), and possessors (120c).Thus, this is also an ABS=POSS pattern.[42]

Tiriyó                                              (Meira 1999; p.283, p.290, p.201)

(120) a. *ë-emamina*
         2SG-play

         'you have played'

      b. *ë-eta*
         2SG-hear

         'it has heard you'

      c. *ë-pata*
         2SG-village

         'your village'

Thus Tenetehára and Tiriyó display ABS=POSS in basic transitive clauses. However, it is also possible to discern some glimmers of absolutive syncretism in isolated grammatical areas of other languages. For instance, Sm'algyax is classified as an

---

[42] Like Tenetehára, Tiriyó features a person hierarchy such that, in transitive verbs, the subject will be encoded if it outranks the object (ex.: *m-eta* 'you have heard it'). Perhaps it comes as no surprise that the Tupian, Cariban and Macro-Gê language families have been proposed to all have a common origin. See Rodrigues (2000) for more details.

ERG=POSS language in the survey because its system of head-marking was shown to exhibit the pattern (see the example in Section 2.3.3). However, for Sm'algyax dependent-marking, the results are more complicated: most dependent-marking patterns *do* exhibit ERG=GEN, but crucially common nouns in the indicative mood – presumably a very common occurence – exhibit ABS=GEN instead.[43]

(121) *Instances of absolutive syncretism in the survey of ergative languages*

| Language | Extent of absolutive syncretism |
|---|---|
| *Tenetehára* | ABS=POSS in basic transitive clauses |
| *Tiriyó* | ABS=POSS in basic transitive clauses |
| *Sm'algyax* | common noun ABS=GEN in indicative mood **only** |
| *Gitksan* | ABS=POSS in subordinate clauses **only** |
| *Akbhaz* | ABS=POSS in non-finite verbs **only** |
| *Canela-Kraho* | ABS=POSS in inalienable possession **only** |

In Gitksan, another Tsimshianic language of the Pacific Northwest, subordinate clauses switch to an ABS=POSS pattern (Rigsby 1986). In Abkhaz, finite transitive clause head-marking exhibits the ERG=POSS pattern. However, in non-finite clauses, a reversal occurs: the slot that encodes both ERG and POSS in finite clauses now encodes ABS and POSS (Hewitt 1989). Finally, as observed in the Section 2.5.5, Canela-Kraho inalienable possession is essentially ABS=POSS.

---

[43] See the dependent-marker *-ga*, which marks the transitive object in (31a), the intransitive object in (31b) and the possessor in (31c). However, dependent-marking of proper nouns in the indicative mood, as well as common nouns in the subjunctive mood, both exhibit ERG=GEN. See also discussion in Mulder (1994), particularly Table 3 p.33, for a fuller discussion of case in Sm'algyax.

These tiny bubbles of absolutive syncretism, in addition to the wider paradigmatic coverage of Tenetehára and Tiriyó, are the extent of the pattern in the survey. It seems clear that, by and large, absolutive patterns of syncretism are not nearly as common as ergative ones.

## 2.6 Conclusion

In this chapter, the twin phenomena of ERG=GEN and ERG=POSS were approached from two different angles. In the first half of the chapter, an abundance of data was presented to illustrate their many shapes and permutations across the spectrum of ergative languages. Following this descriptive portion, a quantitative analysis was undertaken to impart a modicum of statistical significance to the descriptive work in Section 3. A sample of 40 ergative languages was carefully built, as free of common typological biases as can reasonably be expected, to represent the sum of ergative languages cross-linguistically.

If the methodology is sound, and the sample does indeed represent ergative languages well as a whole, then the results suggest that ergative patterns of syncretism are indeed very common cross-linguistically, perhaps even occurring in roughly ~50% of ergative languages in some shape or form. Furthermore, the results of the quantitative analysis suggest that they are indeed more common than their accusative counterparts (NOM=GEN, NOM=POSS, ACC=GEN and ACC=POSS), that they don't seem to be bound to any major typological properties (such as marking type, headedness and basic word order) and that they are more likely to occur in 'robustly' ergative languages. While these results should be taken as a preliminary attempt to understand ERG=GEN and ERG=POSS, and more work absolutely remains to be done, the sum of all these findings strongly suggest that there is something about ergativity which inherently makes it compatible with an ERG=GEN or

ERG=POSS grammatical configuration.

In the next chapter, we turn to the sample of 40 accusative languages mentioned previously. The structure of the that chapter is modeled after this one: first, a thorough description of the phenomenon is presented, followed by quantitative analysis.

## Chapter 3

## THE SAMPLE OF ACCUSATIVE LANGUAGES

### 3.1 Introduction

The outline of this chapter is as follows. In Sections 3.2-3.3, I will provide a basic definition of four morphological patterns of syncretism found in accusative languages: ACC=GEN, ACC=POSS, NOM=GEN and NOM=POSS. Section 3.5 consists of in-depth description of these phenomena, illustrating with ample data their various instantiations across languages and language types. In Section 3.6, a representative sample of 40 accusative languages is constructed to serve as the counterpart to the ergative sample in Chapter 2. The findings are presented in Section 3.7; they are mainly that accusative and nominative syncretic patterns occur at lower rates than ergative ones in the data. In short, this brief chapter is the companion chapter to that which precedes it, and it's structure and methodology are deliberately intended to be identical

### 3.2 What are ACC=GEN and ACC=POSS?

ACC=GEN and ACC=POSS are the two logically-possible, accusative counterparts to the ERG=GEN and ERG=POSS patterns from Chapter. They are defined in the same way: a morphological pattern that shows a high degree of overlap in the expression of case in verbal and nominal forms. Specifically, ACC=GEN refers to a high degree of overlap in dependent-marking morphology that encodes **transitive**

**objects** and **possessors of nouns**; likewise, ACC=POSS refers to a high degree of overlap in the head-marking morphology that encodes those same two entities.

An example of ACC=GEN is found in the dependent-marking system of Comanche. In (1a-1b), both the intransitive subject and the transitive subject are unmarked, while the transitive object is indicated with the suffix -*ʔa*.[1] This is an accusative alignment. In the nominal domain, the same suffix -*ʔa* encodes the possessor of a noun, and therefore this is an example of ACC=GEN.

Comanche                                                          (Charney 1993; p.202, p.94, p.56)

(1) a. *tena-pɨ-tsa      waʔi-pɨ-ʔa      puni*
   man-ABS-TOP woman-ABS-ACC see
   'the man sees the woman'

 b. *ɨ-tsaa-tɨi-tsa      miʔa-tɨ*
   your-good-friend-TOP go-GEN.ASP
   'your good friend is leaving'

 c. *Mia-ʔa-tsa      satiʔi*
   Mia-GEN-TOP dog
   'Mia's dog'

In Indonesian, argument NPs are not marked for overt morphological case. However, according to Sneddon (1996), if a transitive verb features the *meN-* verbal prefix, transitive objects can be encoded by a bound suffix on the verb, as in (2a). This is an accusative alignment, because this option is only available to transitive

---

[1] Both N heads in (1a) are marked with what Charney calls the absolutive suffix. This is not a true absolutive suffix in the ergative-absolutive sense, but rather an unfortunate coincidental convention from Uto-Aztecan linguistics. In Comanche, this 'absolutive' suffix tends to be affixed to a human referent N head when that N head is not the member of a compound, postpositional phrase or incorporated construction. See Charney (1993) Section 3.1.2 for more details.

objects.[2] Furthermore, in the nominal domain, that same bound suffix also encodes the possessor of a noun (2c), and therefore this constitutes an example ACC=POSS in Indonesian's system of head-marking.

Indonesian (Sneddon 1996; p.170, p.134, p.171)

(2) a. *Narti me-nunggu-ku*
Narti *meN*-wait.for.AV-1SG.OBJ

'Narti is waiting for me'

b. *saya tinggal di rumah ini*
1SG live.OV P house DET

'I live in this house'

c. *rumah-ku*
house-1SG.POSS

'my house'

## 3.3 What are NOM=GEN and NOM=POSS?

Looking at nominative patterns of syncretism now, an example is found in the pronominal clitics of Cocama. In (3a-3b), both the transitive 3SG subject and the intransitive 3SG subject, respectively, are encoded by the clitic *ya=*. However, transitive 3SG objects are encoded by a separate form, the clitic *=ay*, as in (3c),

---

[2] Alternatively, the transitive object can also be expressed as a full independent pronoun, as in (i). See Sneddon (1996) Section 2.71 for more details.

Indonesian (Sneddon 1996; p.171)

i) *Narti me-nunggu aku*
Narti *meN*-wait.for.AV 1SG

'Narti is waiting for me'

making the alignment of these clitics accusative. In the nominal domain, the same clitic *ya=* encodes 3SG possessors of nouns (3d), and thus this is an example of NOM=GEN.

Cocama                                    (Vallejos Yopán 2010; p.169, p.227, p.166, p.275)

(3) a. *yaepe, este= pura ya=      mutsana-ka   =tsuri*
       there   1SG   FOC  3SG.SUBJ medicine-REI PAST3

   'then, he cured/bewitched me'

   b. *ya-puka        ya=      pura aykua   =tsuriay*
      like.that-when 3SG.SUBJ FOC  be.sick PAST3

   'at that moment she became sick'

   c. *inu= ray tseta muna =ay      ukua    =tsuriay*
      3PL   SPE want steal  3SG.OBJ used.to PAST3

   'they seem to have always wanted to steal it'

   d. *ya=      mena    uka   uwari*
      3SG.POSS husband house fall

   'her husband's house falls down'

In Hungarian, the verbal suffix *-ja* in (4a) agrees with the subject *a lány* 'the girl' in person and number. In (4d), the same morpheme is re-used in the nominal domain to encode the possessor *Pál* 'Paul', and it also suffixes to the head of its phrase, the possessed N *hibá* 'mistake'.[3] Thus, this is now an instance of NOM=ACC in the head-marking system of Hungarian.

---

[3] Hungarian head-marking is accusative, but a different set of subject agreement morphemes is used for intransitive verbs (4b) and transitive verbs with indefinite objects (4c). See Kenesei et al (1998) Section 2.1.3.2, and also this chapter's Section 3.5.1 for more discussion of tense marking in Hungarian.

Hungarian                                    (Kenesei et al 1998; p.195, p.231, p.195, p.209)

(4) a. *a    lány ír-ja          a   level-et*
    the girl  write-3SG.DEF the letter-ACC

    'the girl is writing the letter'

   b. *a   lány áll-∅*
    the girl  stand-3SG.INDEF

    'the girl is standing'

   c. *ez  az óra   tíz dollár-t   ér-∅*
    this the watch ten dollar-ACC be.worth-3SG.INDEF

    'this watch is worth ten dollars'

   d. *Pál  hibá-ja          tudatlanság-ból fakad-t*
    Paul mistake-3SG.POSS ignorance-ELA   stem-PAST.3SG.INDEF

    'Paul's mistake stemmed from his ignorance'

### 3.4   No syncretism between domains

It is not the case that all nominative-accusative languages display one of these four syncretic patterns. To illustrate, I will provide examples from languages which lack this pattern as well. In Maori, neither transitive nor intransitive subjects are overtly marked for case, as in (5a-5b). On the other hand, objects are marked for case with the accusative marker *i*, as in (5b). But because a different marker *a* is used for the genitive in the nominal domain in (5c), Maori does not qualify as an instance of dependent-marking ACC=GEN.

Maori                                           (Bauer 1993; p.266, p.268, p.108)

(5) a. *kua tae   mai  ngaa manuhiri*
    TAM arrive hither the.PL visitor

    'the visitors have arrived'

b. *ka    kapo    au    i    te    puu*
   TAM snatch 1SG ACC the gun

   'I snatched the gun'

c. *te    kurii a    te    tamaiti*
   the dog    GEN the child

   'the child's dog'

In Fongbe, subject agreement clitics immediately precede the verb, such as *ùn* '1SG.NOM in (6a-6b). Object clitics must follow the verb, such as *mì* '1SG.ACC' in (6c). The fact that the transitive and intransitive subject clitics pattern together to the exclusion of the object clitic make this an accusative alignment. But since this object clitic does not overlap with the corresponding genitive clitic *cè* '1SG.GEN' in (6d), Fongbe cannot be said exhibit ACC=POSS either.

Fongbe                           (Lefebvre & Brousseau 2002; p.63, p.139, p.63 p.68)

(6) a. *ùn        mɔ̀ Kɔ̀kú*
       1SG.NOM see Koku

       'I see Koku'

   b. *ùn        wá,    é        kó        tɔ́*
      1SG.NOM come 3SG.NOM already go.out

      'I came but he had already left'

   c. *Kɔ̀kú mɔ̀ mì*
      Koku see 1SG.ACC

      'Koku saw me'

   d. *xwé    cè*
      house 1SG.GEN

      'my house'

In Manipuri, the dependent-marking system does not reflect any NOM=GEN pattern. In (7a-7b), subjects are marked with the nominative case marker *-nə*. This nominative marker does not overlap with the genitive marker *-gi* in nominal constructions, however, so this is not a case of dependent-marker NOM=GEN.

Manipuri                                       (Bhat & Ningomba 1997; p.63, p.103, p.78)

(7) a. *pulis-nə*      *huranbə-bu pha-re*
       police-NOM thief-ACC    catch-PERF
       'the police have caught the thief'

    b. *əŋaŋ-nə*      *kəppi*
       child-NOM cried
       'the child cried'

    c. *oja-gi*       *məca-du siŋŋi*
       teacher-GEN son-that clever
       'the teacher's son is clever'

Finally, in Pipil, subject prefixes appear at the beginning of the verbal complex, such as *ti-* '2SG.SUBJ' in (8a-8b). Object prefixes appear next in the linear order, such as *nech-* '1SG.OBJ' in (8a). However, in nominal constructions, the prefix that encodes possession on the possessed N in (8c) does not overlap with the subject agreement marker, therefore, this is not an instance of NOM=POSS in head-marking.

Pipil                                          (Campbell 1985; p.56, p.54, p.43)

(8) a. *ti-nech-ita-k*
       2SG.SUBJ-1SG.OBJ-see-PAST
       'you saw me'

b. *ti-kuch-ki*
   2SG.SUBJ-sleep-PAST
   'you slept'

c. *mu-chi:l*
   2SG.GEN-chili
   'your chili pepper'

Now that we have provided a definition of these four nominative-accusative syncretic patterns, and established what does and does not constitute a valid instance of them, we can move on a fuller description of these phenomena as they occur across the spectrum of accusative languages. Furthermore, in Sections 3.6-3.7, an equivalent sample of 40 languages and a quantative analysis will follow, modeled on the structure of Chapter 2.

## 3.5 A description of ACC=GEN, ACC=POSS, NOM=GEN and NOM= POSS across the spectrum

At first glance, it seems fairly common for accusative languages to exhibit ACC= POSS in their system of head-marking. For example, in Lower Grand Valley Dani, a nearly identical agreement paradigm encodes transitive objects on verbs and possessors on noun heads. Not only are the forms the same in all cases except 3SG, but the agreement morpheme itself is linearized in the same relative position in both domains. The possessive prefix occurs word-initially and preceding the noun head and the possession prefix *a-*, while the object prefix occurs word-initially immediately preceding the verb head.

Lower Grand Valley Dani                    (adapted from Bromley 1981; p.190, p.191)

(9)  a.  *n-a-su*    'my net'                    *n-ath-e*    'he hit me'

     b.  *h-a-su*    'your (SG) net'             *h-ath-e*    'he hit you (SG)'

     c.  ∅-*a-su*    'his, her net'              *w-ath-e*    'he hit him, her'

     d.  *nin-a-su*  'our net'                   *nin-ath-e*  'he hit us'

     e.  *hin-a-su*  'your (PL) net'             *hin-ath-e*  'he hit you (PL)'

     f.  *in-a-su*   'their net'                 *in-ath-e*   'he hit them'

In Africa, languages from several unrelated language families exhibit the same kind of ACC=POSS pattern as well. In (10-12) are examples from Lango (Nilotic), Jola Bandial (Niger-Congo A), and Lele (Afro-Asiatic); and although the particularities may differ, the generalization is the same: the morpheme that encodes transitive objects in the (a) examples is reused in the nominal domain to encode possessors in the (c) examples.

Lango                                              (Noonan 1992; p.35, p.31, p.78)

(10) a.  *ò-*         *nɛ̀n -á*
         3SG.SUBJ see  1SG.OBJ
         'he saw me'

     b.  *ò-*         *ɔɲ*
         3SG.SUBJ spill
         'it spilled'

     c.  *tyɛ̌n -á*
         leg   1SG.POSS
         'my leg'

131

Jola Bandial                                   (Bassène 2007; p.92, p.49, p.54)

(11) a. *Atejo na-sen-om          si-rálam*
        Atejo 3SG.SUBJ-give-1SG.OBJ 4-money
        'Atejo gave me the money'

     b. *a-vví aku    fíllim    na-cel-e*
        1-king 1.DEM last.year 3SG.SUBJ-die-TAM
        'the king died last year'

     c. *ga-aɲen-om*
        9-POST-hand-1SG.POSS
        'my hand'

Lele                                           (Frajzyngier 2001; p.187, p.93, p.16)

(12) a. *tamá-ŋ    jè  tè-gè          hire*
        woman-DEF IMPF chase.away-3PL.OBJ often
        'the woman often chased them away'

     b. *tamá   go  lay-wa     è  jè*
        woman REF pretty-REF go VENT
        'a prettier woman came'

     c. *dìngàw dí-gè*
        ferocity GEN.PL-3PL.POSS
        'their ferocity'

Of course, unrelated languages such as Lango, Jola Bandial and Lele are bound to differ in the details. For instance, the degree of ACC=POSS paradigmatic overlap differs between full overlap (Jola Bandial) and partial (Lango, Lele), as illustrated in Table (13).

(13) *Paradigmatic ACC=GEN head-marking in Lango, Jola Bandial and Lele*

**Lango**
(Noonan 1992; p.78, p.96)

|  | obj. | poss. A | poss. B |
|---|---|---|---|
| 1sg | -á | -á | -ná |
| 2sg | -í | -í | -ní |
| 3sg | -ɛ́ | -ɛ̀/-ɛ́ | -mɛ́rɛ̂ |
| 1pl | -wá | -wá | |
| 2pl | -wú/-wùnú/-ú | -wú | |
| 3pl | -gí | -gí | |

**Jola Bandial**
(Bassène 2007; p.92, p.54)

|  | obj. | poss. |
|---|---|---|
| 1sg | -om | -om |
| 2sg | -i | -i |
| 3sg | -ol | -ol |
| 1pl.incl | -olal | -olal |
| 1pl.excl | -óli | -óli |
| 2pl | -ul | -ul |

**Lele**
(Frajzyngier 2001; p.109, p.61)

|  | obj. | poss. |
|---|---|---|
| 1sg | -iŋ | -(i)ŋ |
| 2sg.m | -gi | -m |
| 2sg.f | -me | -te |
| 3sg.m | -iy | -(d)i/-iy |
| 3sg.f | -du | -do/-ro |
| 1du.incl | -ŋga | -dí-ŋgà |
| 1pl.incl | - | -dí-ŋgà-ŋgù |
| 1pl.excl | -ni | -dí-nì |
| 2pl | -ŋgu | -dú-ŋgù |
| 3pl | -ge | -dí-gè |

Some more of these differences are as follows: according to Noonan (1992), Lango distinguishes between two different possessive paradigms: Set A for consonant-final inalienable nouns, and Set B for vowel-final inalienable and all alienable nouns.

In Jola Bandial, the possessive suffixes in (13) only apply to human possessors; according to Bassène (2007), non-human possessors are encoded via the appropriate lexical class suffix, as in (14). And finally, all plural possessives must additionally feature the plural genitive morpheme *-di/-du* in Lele, according to Frayzyngier (2001).

Jola Bandial                                                    (Bassène 2007; p.55, p.55)

(14)  a. *fi-léj  e-ssiko*                         b. *fi-léj-yo*
          7-tail 3-cat                                 7-tail-3.POSS

         'the cat's tail'                              'its tail'


      c. *mi-tiɲ   bu-nunuk*                     d. *mi-tiɲ-bo*
          10-fruit 5-tree                             10-fruit-5.POSS

         'the tree's fruit'                            'its fruit'

But aside from these language-specific particularities, the overall generalization remains valid. In these languages, there is a high degree of overlap between the head-marking morphology that encodes objects in the transitive clause, and possessors in the noun phrase.

Sometimes an ACC=POSS head-marking pattern exists in the grammar, but it is obscured by other factors, such as a person hierarchy which simultaneously operates in the system of verb agreement. Hixkaryana is one such language. At most one verbal argument can be marked on the verb in Hixkaryana. If the transitive subject outranks the transitive object as per the hierarchy I > II > III, the transitive subject is marked on the verb. Thus the first person subject is marked in (15a) in the transitive scenario I → II, and the second person argument is marked in (16a) in the transitive scenario II → III. Although not identified as such by Derbyshire (1985), we can consider these agreement markers nominative, because these same

134

markers also occur when that person is the sole argument of an intransitive verb, as in (15b) and (16b).

Hixkaryana                                    (Derbyshire 1985; p.181, p.4, p.181, p.70)

(15) a. *kɨ-hananɨhɨ-no*
      1SG.SUBJ-teach-IMM.PAST
      'I taught you'

   b. *kɨ-omokɨ-no*
      1SG.SUBJ-teach-IMM.PAST
      'I have come'

(16) a. *mɨ-hananɨhɨ-no*
      2SG.SUBJ-teach-IMM.PAST
      'you taught him'

   b. *huhyaye    mɨ-to-no*
      downriver 2SG.SUBJ-go-IMM.PAST
      'did you go downriver?'

However, if we reverse the transitive scenarios in (15a) and (15b), producing II → I and III → II, respectively, the same grammatical persons must be encoded on the verb as per the person hierarchy. This is represented in (17) and (18) below; and we can consider these agreement morphemes accusative in the sense that they only ever encode transitive objects (and only if the right hierarchical conditions are met).[4]

---

[4] Forms like (17) are the exception to the generalization that at most one argument is encoded on the verb in Hixkaryana. Only in this transitive scenario, II → I, is this required. See discussion in Derbyshire (1985) Section 1.2.2 for more details.

(17)  *uro-mɨ-onytxa-no*
      1SG.OBJ-2SG.SUBJ-hear-IMM.PAST
      'you heard me"

(18)  *o-momokɨ-yaha*
      2SG.OBJ-wait.for-N.PAST
      'he is waiting for you'

Even in III → III transitive scenarios, it is possible to tease apart nominative and accusative agreement morphology. According to Derbyshire (1985), If the 3rd person transitive object is not overt, but the verb is clearly transitive, the 3rd person subject is indicated by the prefix *n-*, as in (19a). Note that this is the same prefix that marks sole 3rd person arguments of intransitive verbs, as in (19b). However, if the transitive 3rd person object is overtly realized in its canonical pre-verbal position, this triggers accusative verb agreement, which must now be realized with the prefix *y-* as in (19c).

Hixkaryana                                    (Derbyshire 1985; p.32, p.31, p.32)

(19)  a.  *n-ahosɨ-ye*                    *kamara*
          3.SUBJ-grab-DIST.PAST.CMPL jaguar
          'the jaguar grabbed him'

      b.  *n-eweh-yatxhe*                 *worɨskomo komo*
          3.SUBJ-take.a.bath-COLL.N.PAST women      COLL
          'the women are taking a bath'

      c.  *toto  y-ahosɨ-ye*              *kamara*
          man 3.OBJ-grab-DIST.PAST.CMPL jaguar
          'the jaguar grabbed the man'

In this manner, we can consider *n-* and *y-* to be the 3rd person nominative and 3rd person accusative verbal agreement forms, respectively. The discussion of Hixkrayana verb agreement morphology thus far is summarized in (20).[5]

(20) *Summary of Hixkaryana verbal agreement*

(adapted from Derbyshire 1985; p.188)

|        | NOM  | ACC  |
|--------|------|------|
| 1SG    | *kɨ-* | *uro-* |
| 2SG    | *mɨ-* | *o-*  |
| 3SG/3PL | *n-* | *y-*  |

Bringing the discussion back to syncretism, when the grammar Hixkaryana encodes a possessor within the noun phrase, it reuses the morphology from the <u>accusative</u> verbal agreement set, as in (21). Thus, this too is an example of head-agreement ACC=POSS, even if it was not as immediately apparent as in the previous examples from Lower Grand Valley Dani, Lango, Jola Bandial and Lele.

Hixkaryana                                   (Derbyshire 1985; p.199, p.61, p.110)

(21) a. *ro-kanawa-rɨ*
        1SG.POSS-canoe-POSS

    'my canoe'

    b. *henta  o-he-txe*
        where 2SG.POSS-wife-POSS

    'where is your wife?'

---

[5] Hixkaryana verb agreement actually features two additional grammatical persons in the paradigm, 1PL.INCL and 1PL.EXCL. The generalizations stated in this section still apply. See Derbyshire (1985) Appendix B for more details.

c. *toto y-owa-nɨ*
   man 3.POSS-chest-POSS

   'the man's chest'

Sometimes, a pattern of ACC=POSS in a language's head-marking system is clearly related to its free pronouns. In Warao, transitive and intransitive subjects are encoded as suffixes on the verb (22a-22b), while transitive objects are encoded as prefixes (22c). Possessors are also encoded with those same prefixes, making it a straightforward ACC=POSS pattern (22d).

Warao                                    (Romero-Figeroa 1997; p.78, p.65, p.23, p.5)

(22) a. *raukaba     sinar-ine*
        slashed.field measure-1SG.SUBJ

        'I measured my slashed field'

    b. *Wauta omi   naru-ki-tia-ine*
       Wauta PRIV go-INTENT-HAB-1SG.SUBJ

       'I am going to see Wauta as usual'

    c. *sina ma-yehe-bu-te-ra*
       who 1SG.OBJ-call-ITER-N.PAST-INT

       'who calls me repeatedly?'

    d. *ma-rahe haya-te*

       1SG.POSS-brother run-N.PAST

       'my brother runs'

The interesting thing is that the relationship between free pronouns and object/possession markers is especially transparent in Warao. Compare the paradigms in (23): the object/possession markers are clearly related to the free accusative pronouns. This is consistent with diachronic analyses going back to Givón (1976) in

138

which bound agreement morphology is descended from free forms. In Givón's theory, free pronouns which are immediately adjacent to the verb can be reanalyzed as bound morphology over time. If this is indeed the case, presumably the process of reanalysis was ongoing at the time of Romero-Figeroa's writing in Warao, given that both free forms and bound forms were reported to co-exist synchronically.[6]

(23) *Free pronouns and bound pronominal markers in Warao*

(adapted from Romero-Figeroa 1997; p.44, p.66, p.64)

| | free ACC pronoun | object prefix | possessive prefix | free NOM pronoun | subject suffix |
|---|---|---|---|---|---|
| 1sg | *ma* | *m(a)-* | *m(a)-* | *ine* | *-ine* |
| 2sg | *hi* | *h(i)-* | *h(i)-* | *ihi* | *-ihi* |
| 3sg | *tai* | - | - | *tai* | - |
| 1pl | *ka* | *k(a)-* | *k(a)-* | *oko* | *-oko* |
| 2pl | *yatu* | - | - | *yatu* | - |
| 3pl | *tatuma* | - | - | *tatuma* | - |

---

[6] Note also that nominative-accusative bound morphology in Warao only extends to three grammatical persons: 1SG, 2SG, 1PL. For the others, Warao relies on its OSV basic word order to disambiguate. The examples in (i-iii) below illustrate for *tatuma* '3PL'.

Warao                                                    (Romero-Figeroa 1997; p.65, p.22, p.66)

i) *naku     tatuma mi-a-e*                                            (transitive O<u>S</u>V)
   monkey 3PL     see-PUNC-PAST
   'they saw a monkey'

ii) *naba muhoko tatuma waba-komoni t-a-e-ra*                          (intransitive <u>S</u>V)
    river side.of   3PL     die-NEG.POT AUX-PUNC-PAST-INT
   'couldn't they die on that side of the river?'

iii) *tatuma Wauta mi-n-a-e*                                           (transitive <u>O</u>SV)
    3PL     Wauta see-SG-PUNC-PAST
   'Wauta saw them'

It is also possible for an ACC=GEN pattern to manifest itself strictly in the pronominal system, but this time with no obvious connections to bound agreement morphology. In Aguaruna, nominative pronominal forms are bare (24a-24b), but accusative ones feature the suffix *-na* (24c). According to Overall (2007), in singular pronouns only, the same forms are re-used in the nominal domain to encode possessors, as in (24d).

Aguaruna                                    (Overall 2007; p.303, p.506, p.260, p.130)

(24) a. *amɨ mi-na    duʃi-mitika-ha-mɨ-i*
        2SG 1SG-ACC laugh-CAUS-1SG.OBJ.IMPF-2SG.SUBJ-DECL
        'you are making me laugh'

   b. *amɨ wɨ-tʃau-aita-ku-mɨ-ĩ-ka*                      *ʃiiha*
      2SG go.PERF-NEG.REL-COP-SIM-2SG.SUBJ-DS-COND well
      *anɨa-sa-nu*              *puhu-mai-inu-aita-ha-i*
      be.happy-SBD-1.SUBJ.SS live-POT-NR-COP-1SG.SUBJ-DECL
      'if you had not gone, I would be happy'

   c. *ami-na apahui tuki    puhu-wa=nu       yaĩ-pa-ka-ti*
      2SG-ACC god    always live-3SG.SUBJ=ANA.REL help-2SG.OBJ-INTS-JUSS
      'may God, who lives forever, help you'

   d. *ami-na   apa*
      2SG-GEN father
      'your father'

Overall (2007) speculates that genitive marking in Aguaruna arose out of accusative marking in the following manner diachronically. Accusative case marks not just transitive objects in Aguaruna, but other "core non-subject NPs" such as benefactives (p.218). Therefore, accusative case may have been used to mark possessors

in the sense of being benefactives (ex.: *the hat for John*), before being reanalyzed as possessors in their own right (ex.: *the hat of John*). Overall also claims that there exists close syntactic relationship between what he considers "core'118 ' and "non-core" NPs in either domain (possessor$_{\text{NON-CORE}}$-possessum$_{\text{CORE}}$, object$_{\text{NON-CORE}}$-verb$_{\text{CORE}}$), which may have speeded this reanalysis.[7] In any case, singular pronouns of Aguaruna are identical in their accusative and genitive forms, as per (25).

(25) *Singular pronoun forms in Aguaruna*     (adapted from Overall 2007; p.218)

|  | NOM | ACC | GEN |
| --- | --- | --- | --- |
| 1sg | *wíi* | *mína* | *mína* |
| 2sg | *ámɨ* | *ámina* | *ámina* |
| 3sg | *níĩ* | *nína* | *nína* |

In a similar vein, in Northern Saami, there is ACC=GEN paradigmatic overlap, but in singular pronouns only. Northern Saami pronouns are in accusative alignment: transitive subjects (26a) and intransitive subjects (26b) pattern together to the exclusion of transitive objects (26c). In addition, in singular forms only, there is ACC=GEN overlap, as in (26d). The entire pronominal paradigm for NOM, ACC and GEN forms in Northern Saami is in (27).

Northern Saami                 (Wilbur 2014; p.120, p.244; p.214; p.89)

(26) a. *ma-v*      *dån*     *sida*
        what-ACC.SG 2SG.NOM want.2SG.PRES
        'what do you want?'

---

[7] However, see discussion of Baker (2015) in Section 4.2, who argues the exact opposite: that the structural relationship between possessor and possessee is fundamentally different from that between verb and internal argument.

b. *suovade         dån?*
   smoke.2SG.PRES 2SG.NOM

   'do you smoke?'

c. *båtsoj    máhtta       duv      nala  båhte-t*
   reindeer can.3SG.PRES **2SG.GEN** upon come-INF

   'the reinder can attack you' (lit. 'come upon you')[8]

d. *men ådtjo        sáme           giela-v            ságasti-t  duv*
   but  may.2SG.PAST Saami.GEN.SG language-ACC.SG speak-INF **2SG.GEN**
   *åbeni-j*
   sibling-CMPL.PL

   'but were you allowed to speak the Saami language with your siblings?'

(27) *Northern Saami pronouns*                    (adapted from Wilbur 2014; p.114)

|   | SG | | | | DU | | |
|---|------|------|------|---|--------|--------|--------|
|   | NOM | ACC | GEN | | NOM | ACC | GEN |
| 1 | *mån/månna* | *muv* | *muv* | | *måj/måjå* | *månov* | *munuo* |
| 2 | *dån/dånna* | *duv* | *duv* | | *dåj/dåjå* | *dånov* | *dunuo* |
| 3 | *sån/sånna* | *suv* | *suv* | | *såj/såjå* | *sånov* | *sunuo* |

|   | PL | | |
|---|------|------|------|
|   | NOM | ACC | GEN |
| 1 | *mij/mija* | *mijáv* | *mijá* |
| 2 | *dij/dija* | *dijáv* | *dijá* |
| 3 | *sij/sija* | *sijáv* | *sijá* |

To conclude this wide-ranging discussion of accusative syncretic patterns cross-linguistically, we now return to Comanche. In Section 3.2, Comanche was briefly

---

[8] To my knowledge, this is the best example provided by Wilbur (2014) of a singular pronoun being used as a transitive object. While it is not clear that the example in (26c) is indeed a transitive clause (Wilbur has glossed *duv* as genitive), his Table 6.1 p.114 makes it clear that the ACC and GEN forms of singular pronouns do indeed overlap in Northern Saami.

demonstrated to exhibit ACC=GEN in its dependent-marking. Going into more detail now, Comanche actually features quite widespread accusative syncretism throughout the grammar. For example, in (1), dependent-marking ACC=GEN was illustrated with the re-using of the -*ʔa* suffix. In actuality, there are several lexical noun classes in Comanche, and they each condition their own accusative/genitive suffix. According to Charney (1993), the overlap between these forms is extremely consistent across lexical classes, as in (28).[9]

(28) *Dependent-marking overlap across lexical classes in Comanche*

(adapted from Charney 1993; p.56)

| ACC | GEN |
|---|---|
| -∅ | -∅ |
| -ʔa | -ʔa |
| -ha | -ha |
| -i | -i |
| -e | -e |
| -ta/-Hta | -ta/-Hta |
| -tɨi | -tɨna |

---

[9] According to Charney (1993), the one potential difference between the accusative and genitive suffix is that former contains an abstract "lenis final feature" that may condition lenition in an obstruent immediately following that suffix. For example, the root for the verb 'see' is underlyingly /puni/, but it is realized with a word-initial [v] in (i). However, Charney also notes that sometimes this final feature "seems to be suspended", resulting in no lenition in sentences such as (ii). Charney thus speculates that this final feature "may be in the process of being lost"; see his Section 3.1.4 for more details.

Comanche                                                    (Charney 1993; p.54, p.54)

i) *uhka    nɨɨ  posa-rena-pɨ-ʔa     vuni-n*
   that.ACC 1SG crazy-man-ABS-ACC  see-CMPL
   'I saw the crazy man'

ii) *uhka    nɨɨ  tsuku-pɨ-ʔa     puni-n*
    that.ACC 1SG old.man-ABS-ACC see-CMPL
    'I saw the old man'

Furthermore, accusative alignment is observed not only in Comanche's dependent-marking, but in its head-marking as well. For non-NP arguments, transitive and intransitive subjects occur in their full pronominal form (usually in second position in the sentence), as in (29a-29b). However, transitive objects are marked by a prefix on the verb complex, as in (29c). This an accusative alignment because transitive and intransitive subjects pattern together to the exclusion of transitive objects.

Comanche                                     (Charney 1993; p.144, p.81, p.127, p.87)

(29) a. *hakaniti ɨnnɨ nápukuwáá-ʔa tsahani-ka*
        how.ACC 2SG car-ACC        drive-STAT
        'what kind of car do you drive?'

   b. *hakah-nai ɨnnɨ kɨma-yu*
       where-DIR 2SG come-PROG
       'where are you coming from?'

   c. *ɨ-pɨa-tai-tuʔi            nɨɨ*
       2SG.OBJ-leave-go.out-IRR 1SG
       'I'm going to divorce you'

   d. *ɨ-ámawóó-mati     nɨɨ kɨh-kaʔa-tuʔi*
       2SG.POSS-apple-PRT 1SG teeth-break-IRR
       'I'm going to bite off a piece of your apple'

In the nominal domain, the same set of agreement prefixes is re-used to encode singular pronominal possessors, as in (29d), thus producing a partial ACC=POSS pattern. The full paradigms for subject, object and possessor are illustrated in (30).

(30) *Free and bound pronominal forms in Comanche*

(adapted from Charney 1993; p.98, p.99, p.99, p.92)

| | | SUBJECT | OBJECT | POSSESSOR |
|---|---|---|---|---|
| SG | 1 | *nɨɨ* | *nɨ-/nɨe-/nɨeti-* | *nɨ-* |
| | 2 | *ɨnnɨ/nɨɨ* | *ɨ-/ɨmmɨ-/immɨ-* | *ɨ-* |
| | $3_{near}$ | *i* | *ihka-* | *ihka-* |
| | $3_{mid}$ | *o* | *ohka-* | *ohka-* |
| | $3_{far}$ | *u* | *uhka-* | *uhka-* |
| | $3_{unknown}$ | *ma* | *mahka-* | *mahka-* |
| DU | 1INCL | *tah/takwɨh* | *tahi-* | *tahi-* |
| | 1EXCL | *nɨkwɨh* | *nɨhi-* | *nɨhi-* |
| | 2 | *mɨh/mɨkwɨh/nɨɨ* | *mɨhi-* | *mɨhi-* |
| | $3_{near}$ | *itɨkwɨh* | *itɨhii-* | *itɨhɨ-* |
| | $3_{mid}$ | *otɨkwɨh* | *ohti-* | *ohtɨ-* |
| | $3_{far}$ | *utɨkwɨh* | *uhti-* | *uhtɨ-* |
| | $3_{unknown}$ | *matɨkwɨh* | *mahti-* | *mahtɨ-* |
| PL | 1INCL | *tannɨ* | *tai-/tammi-* | *taa-/tammi-* |
| | 1EXCL | *ninnɨ* | *nimmi-* | *nimmi-* |
| | 2 | *minnɨ/mɨmmɨ* | *mɨii-/mɨmmi* | *mɨii-/mɨmmɨ-* |
| | $3_{near}$ | *itɨɨ* | *itii-* | *itɨɨ-* |
| | $3_{mid}$ | *otɨɨ* | *otii-* | *otɨɨ-* |
| | $3_{far}$ | *utɨɨ* | *utii-* | *utɨɨ-* |
| | $3_{unknown}$ | *matɨɨ* | *matii-* | *matɨɨ-* |

To summarize, Comanche exhibits widespread accusative syncretism throughout the grammar, both in dependent-marking on lexical NPs, but also in head-marking as well. In this sense, we can consider Comanche the accusative counterpart to languages such as West Greenlandic and Sm'algyax from Chapter 2, which themselves exhibited widespread ergative syncretism in distinct areas of the grammar.

145

To a lesser extent, there are more attested examples of NOM=POSS as defined in Section 3.3. A straightforward example can be found in the head-marking system of Maybrat. In this language, verbs agree with transitive and intransitive subjects via prefix, as in (31a-31b).[10] On the other hand, transitive objects are not encoded on the verb, and so if they are overt, they can only be realized as lexical NPs or free pronouns (31c). This is an accusative alignment because the transitive and intransitive subject pattern together to the exclusion of the transitive object.

Maybrat                                                  (Dol 2007; p.77, p.51, p.75, p.63, p.66)

(31)  a. *y-po*          *ku*   *kiniah*
         3M.SG.SBJ-hold child small
         'he holds the small child'

     b. *y-ros*
        3M.SG.SBJ-stand
        'he stands'

     c. *m-ape*          *ait*
        3U-carry.on.back 3M.SG
        'she carries him on her back'

     d. *y-ana*
        3M.SG-POSS-head
        'his head'

---

[10] According to Dol (2007), not all verbs exhibit subject agreement in Maybrat. Verbs with underlying roots of the shape CVCV, CCV or CCVC have 'covert agreement' instead, as in (i) below. See her Section 3.1.2 for more details.

Maybrat                                                                    (Dol 2007; p.116)

i) *peroh*  *ait*    *Ø-kpat  Kocu Ata*
   wrong 3M.SG Ø-leave Kocu Aata
   'no, he leaves Kocu Ata'

e. *amah r-ait*
   house POSS-3M.SG
   'his house'

In the nominal domain, inalienable possessions re-use that subject agreement prefix to encode the possessor, as in (31d), thus instantiating a NOM=POSS pattern. For alienable possession, a different strategy is used: the possessive morpheme *r-* prefixes to the free pronoun, as in (31e).

Maricopa's system of head-marking also exhibits NOM=POSS. Comparing the forms in (32-33), the same agreement prefix encodes the subject of intransitive verbs, and the subject of transitive verbs when the object is in the third person.

Maricopa                              (Gordon 1986; p.19, p.19, p.19, p.17, p.18, p.16)

(32) a. *'-ashvar-k*
        1.SUBJ-sing-REAL
        'I sang'

     b. *m-ashvar-k*
        2.SUBJ-sing-REAL
        'you sang'

     c. *∅-ashvar-k*
        3.SUBJ-sing-REAL
        'he sang'

(33) a. *'-wik-k*
        1.SUBJ-help-REAL
        'I helped him'

     b. *m-wik-k*
        2.SUBJ-help-REAL
        'you helped him'

c. ∅-*wik-k*
   3.SUBJ-help-REAL
   'he helped him'

Questions of alignment are obscured in transitive clauses where the object is a Speech Act Participant (i.e. either 1st or 2nd person). The paradigm in (34) hints at a person hierarchy vaguely reminiscent of Hixkaryana, but the analysis is not as clear; Gordon (1986) discusses their historical origin as segmentable subject and object prefixes, but ultimately considers them unanalyzable fusional morphemes synchronically.

(34) *Pronominal prefixes in Maricopa* (adapted from Gordon 1986; p.21)

|   | intransitive | transitive | | | possessive |
|---|---|---|---|---|---|
|   |   | 1<sub>OBJ</sub> | 2<sub>OBJ</sub> | 3<sub>OBJ</sub> |   |
| 1 | '- | - | *ny-* | '- | '- |
| 2 | *m-* | '*nym-* | - | *m-* | *m-* |
| 3 | ∅- | *ny-* | *m-* | ∅- | ∅- |

In any case, under the assumption that Maricopa's system of verbal head-marking is essentially in accusative alignment (with some instances of fusional suppletion), we can observe a NOM=POSS pattern in the nominal domain. The same morphemes that encode intransitive and transitive (3rd person object only) subjects are re-used to encode possessors. In (35), inalienable possession is marked by direct prefixation onto the noun root; in (36), alienable possession is achieved by prefixation to an additional morpheme *ny-*. Crucially, in all these examples, it is never the person prefixes of the shape *ny-/'nym-* that encode possession – it is the 'nominative' ones.

c. ∅-*wik-k*
   3.SUBJ-help-REAL
   'he helped him'

Questions of alignment are obscured in transitive clauses where the object is a Speech Act Participant (i.e. either 1st or 2nd person). The paradigm in (34) hints at a person hierarchy vaguely reminiscent of Hixkaryana, but the analysis is not as clear; Gordon (1986) discusses their historical origin as segmentable subject and object prefixes, but ultimately considers them unanalyzable fusional morphemes synchronically.

(34) *Pronominal prefixes in Maricopa* (adapted from Gordon 1986; p.21)

|   | intransitive | transitive | | | possessive |
|---|---|---|---|---|---|
|   |   | $1_{\text{OBJ}}$ | $2_{\text{OBJ}}$ | $3_{\text{OBJ}}$ |   |
| 1 | '- | - | *ny-* | '- | '- |
| 2 | *m-* | '*nym-* | - | *m-* | *m-* |
| 3 | ∅- | *ny-* | *m-* | ∅- | ∅- |

In any case, under the assumption that Maricopa's system of verbal head-marking is essentially in accusative alignment (with some instances of fusional suppletion), we can observe a NOM=POSS pattern in the nominal domain. The same morphemes that encode intransitive and transitive (3rd person object only) subjects are re-used to encode possessors. In (35), inalienable possession is marked by direct prefixation onto the noun root; in (36), alienable possession is achieved by prefixation to an additional morpheme *ny-*. Crucially, in all these examples, it is never the person prefixes of the shape *ny-/'nym-* that encode possession – it is the 'nominative' ones.

Maricopa                                      (Gordon 1986; p.31, p.31, p.31, p.32, p.32, p.32)

(35) a.  *'-haav*
         1.POSS-shirt
         'my shirt'

     b.  *m-mpur*
         2.POSS-hat
         'your hat'

     c.  *Bonnie ∅-avhay*
         Bonnie 3.POSS-dress
         'Bonnie's dress'

(36) a.  *'-ny-va*
         1.POSS-POSS-house
         'my house'

     b.  *m-ny-kwr'ak*
         2.POSS-POSS-old.man
         'your husband'

     c.  *'iipaa ∅-ny-hat*
         man   3.POSS-POSS-dog
         'a/the man's pet'

Examples of nominative syncretism outside of head-marking appear to be quite rare; to find some, we need to look further than plain, declarative main clauses. For example, in Japanese, main clauses do not show any kind of syncretism between NOM, ACC and GEN, as in (37).

149

Japanese <span>(Storm 2003; p.62, p.79, p.121)</span>

(37) a. *Yamada-san ga   pan   o   tabema-shita*
Mr. Yamada NOM bread ACC eat-PERF
'Mr. Yamada ate the bread'

b. *Ima Mita-san ga   ofisu ni   imasu*
Ms. Mita      NOM office LOC now
'Ms. Mita is in the office now'

c. *watashi no   heya*
1SG      GEN room
'my room'

However, in relative clauses, Japanese allows subjects to be marked with either nominative *ga* or genitive *no*, which is known as Nominative-Genitive Conversion in the Japanese syntax literature (Hiraiwa 2000). This phenomenon is illustrated in (38).

Japanese <span>(Hiraiwa 2000; p.68, p.68)</span>

(38) a. *kinoo    John ga  katta       hon*
yesterday John NOM buy.PAST.ADN book
'the book which John bought yesterday'

b. *kinoo    John no  katta       hon*
yesterday John GEN buy.PAST.ADN book
'the book which John bought yesterday'

A similar situation exists in Nuuchahnulth. In main declarative clauses, Nuuchahnulth subject clitics are in accusative alignment because they can only encode transitive and intransitive subjects, as in (39a-39b). Possession is encoded by a different set of clitics altogether, as in (39c).

150

(39) a. *ńaacsaa='aλ=maa=aḥ     suẃa*
     see=TEMP=IND=1SG.SUBJ 2SG

'I see you'

b. *ʔi:ḥʷ='aλ=maa=aḥ        haanaʔa='aλ*
   very=TEMP=IND=1SG.SUBJ play.lehal=TEMP

'I was playing *lehal* (a gambling game) in a big way'

c. *ʔaλ-qimł=uk=qaa=s                    kuunaa*
   two-many.round-objects=POSS=DEF=1SG.POSS schooner

'my two schooners"

However, in a large set of non-indicative clauses – interrogative, conditional, subordinate, relative and inferential – the grammar of Nuuchahnulth re-uses the possessive clitic pronouns to encode subjects, as in (40). Again, this is technically a NOM=GEN pattern, but it does not occur naturally in basic, declarative clauses. The fact that we must dig into 'deeper' areas of the grammar to find such tokens of NOM=GEN suggests that it is a fairly rare phenomenon overall.

(40) *ʔaḥkuu ʔaya-aqλ=ʔii     haʔum yaqʷ-'iis='aλ=qaa=s*
     DEM    much-inside=ART food  that.which-consume=TEMP=DEF=1SG.SUBJ
     'this expensive food I am now eating'

### 3.5.1   Partially syncretic patterns

Partial paradigmatic overlap seems common when discussing syncretic patterns such as ACC=POSS and NOM=POSS as well. Recall in Section 3.5 that

Lower Grand Valley Dani, Lango, Lele and Comanche all exhibited a partial pattern of accusative syncretism between verbal and nominal domains.

On the NOM=POSS side, this state of affairs is also attested. In Section 3.2, Hungarian was cited as a language with head-marking NOM=POSS, illustrating with data such as (4), reiterated in (41) below.

Hungarian                                (Kenesei et al 1998; p.195, p.231, p.195, p.209)

(41) a. *a   lány ír-ja            a   level-et*
      the girl   write-**3SG.SUBJ.DEF** the letter-ACC
      'the girl is writing the letter'

  b. *a   lány áll-∅*
      the girl   stand-3SG.SUBJ.INDEF
      'the girl is standing'

  c. *ez   az  óra   tíz dollár-t    ér-∅*
      this the watch ten dollar-ACC be.worth-3SG.SUBJ.INDEF
      'this watch is worth ten dollars'

  d. *Pál   hibá-ja          tudatlanság-ból fakad-t*
      Paul mistake-**3SG.POSS** ignorance-ELA   stem-PAST.3SG.SUBJ.INDEF
      'Paul's mistake stemmed from his ignorance'

In actuality, the situation is a little more nuanced. There are two 'nominative' paradigms in Hungarian. One coindexes transitive subjects in clauses with definite objects (41a), while the other coindexes intransitive subjects (41b) as well as transitive subjects in clauses with indefinite objects (41c). Turning back to NOM=POSS, it appears that the possessive paradigm borrows extensively from both nominative sets. Therefore it is the case that nearly all possessive suffixes in Hungarian overlap

with a nominative one, but in the end both nominative paradigms are pilfered only partially. See (42) below for all three paradigms listed in full.

(42) *Nominative* (x2) *and genitive paradigms in Hungarian*

(adapted from Kenesei et al 1988; p.21)

|  | definite | indefinite |
|---|---|---|
| 1sg | *-om/-em/-öm* | *-ok/-ek/-ök/-om/-em/-öm* |
| 2sg | *-od/-ed/-öd* | *-ol/-el/-sz* |
| 3sg | *-ja/-i* | *-∅/-ik* |
| 1pl | *-juk/-jük* | *-unk/-ünk* |
| 2pl | *-játok/-itek* | *-tok/-tek/-tök* |
| 3pl | *-ják/-ik* | *-nak/-nek* |

|  | possessive |
|---|---|
| 1sg | *-m/-om/-em/-öm* |
| 2sg | *-d/-od/-ed/-öd* |
| 3sg | *-a/-e/-ja/-je* |
| 1pl | *-nk/-unk/-ünk* |
| 2pl | *-tok/-tek/-tök/-otok/-etek/-ötök* |
| 3pl | *-uk/-ük/-juk/-jük* |

Also on the NOM=POSS side, Turkish exhibits a partial pattern. Turkish verbs agree with transitive subjects (43a) and intransitive subjects (43b). Objects are not encoded on the verb; if realized overtly, they take the shape of lexical NPs or free pronouns (43c). This is a NOM=POSS pattern because the same subject agreement suffix may be reused to encode the possessor of nouns (43d).

Turkish                              (Göksel and Kerslake 2005; p.141, p.48, p.143, p.24)

(43) a. *her    gün  çikolata    yi-yebil-ir-im*
        every day chocolate eat-PSB-AOR-1SG.SUBJ
        'I could eat chocolate every day'

153

b. *gid-ebil-ir-im*           *de*
   go-PSB-AOR-1SG.SUBJ also

   'and I can indeed go'

c. *beni  anlama-mak-ta        inat        ed-iyor*
   1SG  understand-SUB-LOC obstinate AUX-IMPF

   'she obstinately refuses to understand me'

d. *hal-im*
   condition-1SG.POSS

   'my condition'

However, close examination of verbal and nominal paradigms reveals that the amount of NOM= POSS overlap is partial at best; these paradigms are in (44) below, and approximately half the subject agreement markers are re-used as possession markers in the nominal domain.

(44) *Turkish 'Group 1' verbal suffix and possessive suffixes*

(adapted from Göksel and Kerslake 2005; p.82, p.66)

|  | subject agreement | possessive suffix |
| --- | --- | --- |
| 1sg | *-m* | *-m/-Im* |
| 2sg.familiar | *-n* | *-n/-In* |
| 2sg.formal | *-nIz* | *-nIz/-InIz* |
| 3sg | *-∅* | *-sIn/-sI/-I/In* |
| 1pl | *-k* | *-mIz/-ImIz* |
| 2pl | *-nIz* | *-nIz/-InIz* |
| 3pl | *-lAr* | *-lArI/-lArIn* |

To summarize, it seems that cross-linguistically, it is fairly common for ACC= POSS and NOM=POSS patterns to only occur in partial paradigmatic overlap. Of

154

course, it is also common for these patterns to occur in full paradigmatic overlap as well. The descriptive survey in Section 3.5 suggests that both occur in equal measure in the data; furthermore, these findings are consistent with that of ergative patterns in Chapter 2. The table in (45) indicates the set of nominative and accusative syncretic patterns that have been covered thus far.

(45) *Full and partial paradigmatic overlap in nominative-accusative languages*

| full overlap | | | partial overlap | | |
|---|---|---|---|---|---|
| language | type | examples | language | type | examples |
| Comanche | dependent-marking | (1), (28) | L. G. V. Dani | head-marking | (9) |
| Indonesian | head-marking | (2) | Lango | head-marking | (10), (13) |
| Cocama | pronoun | (3) | Lele | head-marking | (12), (13) |
| J. Bandial | head-marking | (11), (13) | Aguaruna | pronoun | (24-25) |
| Hixkaryana | head-marking | (15-21) | N. Sami | pronoun | (26-27) |
| Warao | head-marking | (22-23) | Comanche | head-marking | (29-30) |
| Maybrat | head-marking | (31) | Hungarian | head-marking | (41-42) |
| Maricopa | head-marking | (32-36) | Turkish | head-marking | (43-44) |

### 3.5.2 Borderline cases?

Much like in Chapter 2 for ergative languages, sometimes some languages exhibit a pattern which exists right on the cusp of being ACC=POSS or NOM=POSS, and they thus merits further discussion. For example, in Modern Hebrew, there are sets of verb agreement suffixes that correspond either to (transitive and intransitive) subjects (46a-46b) or transitive objects (46c), which is an accusative alignment. That same transitive object suffix is re-used in the nominal domain, like in (46d), producing a classic head-marking ACC=POSS pattern.

Modern Hebrew                                    (Glinert 1989; p.275, p.125, p.52 , p.52)

(46) a. *sha'-alt      éfo    hayi-ti*
ask-2SG.M.SUBJ where be-1SG.SUBJ
'you asked where I was'

155

b. *kshe-tsilts-alt,*      *ani bidiyuk difdáf-ti*      *bo*
when-ring-2SG.M.SUBJ 1SG just   leaf.through-1SG.SUBJ P+it
'when you rang, I was just leafing through it'

c. *lehazir-Ha*
warn.INF-2SG.M.OBJ
'to warn you'

d. *iyum-Ha*
threat-2SG.M.POSS
'your threat'

The only problem, as Glinert observes, is that "[this usage] is formal and rather uncommon" (p.52). This intuition still applies for current speakers of Modern Hebrew (Sandy Abu Adas, p.c.). Instead, Modern Hebrew speakers are far more likely to opt for a construction as in (47), where the transitive object marker is affixed to a free accusative 'head', *ot*. This breaks down the linear order parallelism between the verbal and nominal domains which has characterized such patterns thus far.

Modern Hebrew                                  (Glinert 1989; p.256)

(47) a. *halo*      *ani makir ot-Ha*
after.all 1SG know ACC-2SG.M.OBJ
'I know you after all'

Therefore, since constructions such as (46c) are still technically grammatical, Modern Hebrew does count as an ACC=POSS language, but with a caveat.

In Section 3.5, Aguaruna was demonstrated to unambiguously have ACC=GEN in its singular pronouns (see Table 25). However, there is reason to believe the pattern may extend to regular lexical NPs as well. At first glance, there doesn't seem

to be a match. Underlyingly, accusative NPs are marked with with the suffix *-na*
(48a), while possessor NPs are unmarked (48b).

Aguaruna                                                              (Overall 2007; p.507, p.217)

(48) a. *nu-na      waʃi-na              duʃiki-a-tʃa-ku-n-u-ka*
        ANA-ACC spider.monkey-ACC laugh.at-IMPF-NEG-SIM-1SG.SS-COND
        *tuku-mai-nu    awakɨ-ka-ha-i*
        shoot-POR-NR overcome-1SG-DECL
        'if I hadn't laughed at that monkey, I would have been able to shoot it'

    b. *waʃi-∅      yaka-hĩ*
        monkey-GEN arm-PERT.1PL/3
        'the monkey's arm'

However, as Overall (2007) discusses in his Section 4.6.2.1, two successive
phonological processes take place in accusative forms. First, an extremely common
process of apocope deletes vowels in a word-final light syllable (49).

Aguaruna                                                              (Overall 2007; p.60)

(49) /*nahana-ta*/    →   [na.há.nat]
        create-ACTNR
        'to create'

In addition, this process of apocope can feed a more specific phonological
rule where word-final nasals are dropped, "particularly from accusative forms that
directly precede their governing element" (p.217). The end result is that underlying
forms such as *waʃi-na* 'monkey (ACC)' in (48a) are often realized as /waʃi-na/ →
[wa.ʃi-n] → [wa.ʃi] on the surface, thus making accusative and genitive forms entirely
homophonous.

For Overall, these phonological rules provide additional evidence for the common diachronic origins of accusative and genitive case in Aguaruna. For our purposes, the facts are that singular pronominal forms do clearly overlap in Aguaruna (as covered in Section 3.5), but in addition, they are highly likely to overlap in lexical NPs as well due to the aforementioned phonological processes.[11]

The last borderline case to discuss involves Cocama, which was introduced in Section 3.2 as an example of pronominal NOM=POSS. Recall that these clitic pronouns are in accusative alignment because (transitive and intransitive) subject clitics (50a-50b) pattern differently from object clitics (50c). These subject clitics also serve as possessive clitics, thus resulting in a syncretic pattern (50d).

Cocama                                       (Vallejos Yopán 2010; p.169, p.227, p.166, p.275)

(50) a. *yaepe, este= pura ya=        mutsana-ka    =tsuri*
       there   1SG    FOC   3SG.SUBJ medicine-REI PAST3

       'then, he cured/bewitched me'

   b. *ya-puka        ya=        pura aykua  =tsuriay*
       like.that-when 3SG.SUBJ FOC  be.sick PAST3

       'at that moment she became sick'

   c. *inu= ray tseta muna =ay     ukua    =tsuriay*
       3PL   SPE want steal  3SG.OBJ used.to PAST3

       'they seem to have always wanted to steal it'

   d. *ya=      mena     uka    uwari*
       3SG.POSS husband house fall

       'her husband's house falls down'

---

[11] This process of word-final nasal deletion, resulting homophony of differently case-marked NPs in Aguaruna, is strongly reminiscent of another 'borderline case' concerning Lezgian NPs in Section 3.8 of Chapter 2.

The only problem is that 3rd person clitic pronouns are the total extent of accusativity in the language. No other persons in the pronominal paradigm of Cocama exhibit an accusative alignment in this fashion, nor is there any dependent-marking on NPs or true head-marking on verbs. Therefore, yes, strictly-speaking, this is a NOM=POSS pattern because in that one instance of 'nominativity' in the grammar, that same clitic is reused to encode possessors of nouns. However, it is worthy of a caveat to mention that Cocama isn't a very 'accusative' language in the first place.[12]

### 3.5.3 Interim conclusion

In Section 3.5, we explored many of the possible dimensions of accusative and nominative syncretic patterns cross-linguistically, from dependent-marking to head-marking to pronominal systems. The discussion touched on patterns that seem quite common (e.g. head-marking ACC=POSS) and others which appear rarer (e.g. NOM=GEN in general).

In the next section, we go a step further, constructing a sample of 40 accusative languages deliberately modeled as the accusative counterpart to the ergative sample in Chapter 2. The goal is simple: if the sample is well-chosen and representative of accusative languages as a whole, we can calculate an estimate of the prevalence of these mirror nominative-accusative patterns cross-linguistically, and thus compare them with the results of the ergative sample in Chapter 2.

---

[12] An interesting property of Cocama is that its pronoun system (clitic and non-clitic alike) is gendered by speaker. The 3rd person pronouns in (50a-50d) belong to the register for female speakers. The corresponding forms for male speakers are *ra=* 3SG.SUBJ/POSS and *=ura* 3SG.OBJ. See Vallejos Yopán (2010) Section 5.3 for more details.

## 3.6 The sample of accusative languages

### 3.6.1 Selecting the sample

To get a sense of how common accusative and nominative syncretic patterns are among accusative languages, we must proceed exactly like we did for ergative languages in Chapter 2. A representative sample must be studied which accurately reflects the set of all accusative languages cross-linguistically. Therefore, his sample of accusative languages is deliberately designed to serve as the counterpart to the ergative sample in Chapter 2. As before, the guidelines in Bakker (2011) were followed to produce an appropriate sample numbering 40. In light this deliberate parallelism, the reader can refer back to Chapter 2 for a more thorough discussion of language sampling techniques; Sections 3.6.1 and 3.6.2 address the same issues of methodology, but more succinctly.

#### 3.6.1.1 Avoiding common biases in the sample

The 40 languages that were ultimately selected are listed in (51). They were selected semi-randomly, based on availability of grammars in the literature and partially guided by the online database *WALS*, but also with a consideration for avoiding common typological sampling biases, as detailed below.

To avoid *bibliographic bias*, care was taken to select languages from a wide range, from 'better' studied (ex.: Hebrew, Swahili, Hungarian, Turkish, Japanese) to understudied (ex.: Pipil, Cocama, Northern Sami, Grand Valley Dani, Aguaruna, Awa Pit, *etc*).

*Genetic bias* occurs when drawing excessively from one or more language families. In this sample, five language families are sampled from twice (Austronesian, Niger-Congo A, Nilotic, Uralic, Uto-Aztecan) and two language families are sampled

from thrice (Trans-New Guinea, Afro-Asiatic). In all these instances, however, the related languages are distinct enough typologically to serve adequately in the sample; for example, Hebrew and Somali may both be Afro-Asiatic, but they feature SVO and SOV basic word orders, respectively. The end result is that 31 language families are represented in total, which is roughly equal to the 28 language families represented in the ergative sample of Chapter 2.

(51) *the 40 languages in the accusative sample*

### North America

1. Koasati — Muskogean
2. Southern Pomo — Pomoan
3. Misantla Totonac — Totonacan
4. Sierra Miwok — Utian
5. Comanche — Uto-Aztecan
6. Pipil — Uto-Aztecan
7. Nuuchahnulth — Wakashan
8. Maricopa — Yuman

### South America

9. Aymara — Aymaran
10. Awa Pit — Barbacoan
11. Hixkaryana — Carib
12. Warao — isolate
13. Aguaruna — Jivaroan
14. Barasano — Tucanoan
15. Cocama — Tupian

### Australia & New Guinea

16. Maori — Austronesian
17. Maybrat — Papuan
18. Kayardild — Pama-Nyungan
19. Daga — Trans-New Guinea
20. Grand Valley Dani — Trans-New Guinea
21. Mian — Trans-New Guinea

### Europe

22. Latin — Indo-European
23. Hungarian — Uralic
24. Northern Sami — Uralic

### Africa

25. Lele — Afro-Asiatic
26. Somali — Afro-Asiatic
27. Swahili — Bantu
28. Sandawe — Khoisan
29. Fongbe — Niger-Congo A
30. Jola Bandial — Niger-Congo A
31. Nandi — Nilotic
32. Lango — Nilotic

### Asia

33. Hebrew — Afro-Asiatic
34. Indonesian — Austronesian
35. Kannada — Dravidian
36. Japanese — Japonic
37. Manipuri — Sino-Tibetan
38. Manchu — Tungusic
39. Turkish — Turkic
40. Kolyma Yukaghir — Yukaghir

*Areal bias* is managed by drawing on different regions of the world in equal parts. With the exception of Europe, all regions are roughly equally represented by between six and eight languages. The justification for Europe's low number comes from the fact that Europe is actually a linguistically homogeneous area of the world comparatively – according to *Ethnologue*, only 6% of the world's 7,102 living languages are in Europe. Finally, like for the ergative sample, the accusative sample is supposed to reflect the world's accusativity 'hotstpots'. As such, the sample features a high number of languages from sub-Saharan Africa, New Guinea and the Amazon.

*Typological bias* is managed by ensuring that no major typological feature is disproportionately represented in the sample. As per the *World Atlas of Language Structures* (Dryer 2013), we know that the distribution of basic word orders is not even across accusative languages. The tables in (52-53) illustrate that S-initial word orders, in particular SOV, dominate among accusative languages cross-linguistically (ergative data is provided for comparison).

(52) **dependent-marking alignment X basic word order**

(*WALS* features 81A X 98A)

| **ergative** | VSO 1 | VOS 2 | SVO 1 | SOV 16 | OVS 0 | OSV 0 | *none* 7 |
|---|---|---|---|---|---|---|---|
| **accusative** | VSO 3 | VOS 1 | SVO 9 | SOV 29 | OVS 1 | OSV 0 | *none* 7 |

(53) **head-marking alignment X basic word order**

(*WALS* features 81A X 100A)

| **ergative** | VSO 2 | VOS 0 | SVO 0 | SOV 4 | OVS 1 | OSV 1 | *none* 5 |
|---|---|---|---|---|---|---|---|
| **accusative** | VSO 14 | VOS 8 | SVO 65 | SOV 71 | OVS 3 | OSV 0 | *none* 29 |

162

With this in mind, the sample of 40 accusative languages aims to reflect these typological facts as closely as possible. In (54), the languages in the accusative sample are listed by basic word order. Like in the *WALS* data, S-initial word orders dominate (especially SOV), while non-S-initial word orders are underrepresented. Nuuchahnulth and Misantla Totonac are listed in the category 'VSO/VOS' because, according to Davidson (2002) and MacKay (1999), respectively, the basic word order is simply 'V-initial'. In the case of Kayardild and Sierra Miwok, insufficient information was available in Evans (1995) and Broadbent (1964), respectively, to assign to them a basic word order.

(54) *the languages of the accusative sample by basic word order*

| b.w.o. | # | languages |
|---|---|---|
| SOV | 19 | Aguaruna, Awa Pit, Aymara, Comanche, Daga, Japanese, Kannada, Koasati, Kolyma Yukaghir, Latin, Lower Grand Valley Dani, Manchu, Manipuri, Maricopa, Mian, Sandawe, Somali, Southern Pomo Turkish |
| SVO | 11 | Cocama, Fongbe, Hungarian, Indonesian, Jola Bandial, Lango, Lele, Maybrat, Modern Hebrew, Northern Sami, Swahili |
| VSO | 2 | Maori, Nandi |
| VSO/VOS | 2 | Misantla Totonac, Nuuchahnulth |
| VOS | 1 | Pipil |
| OVS | 2 | Barasano, Hixkaryana |
| OSV | 1 | Warao |
| no dominant order | 2 | Kayardild, Sierra Miwok |

Another major typological feature to consider when avoiding *typological bias* is headedness. As per the table in (55), the languages in the sample are fairly evenly split between head-initial and head-final. Like in Chapter 2, the criteria for headedness are drawn from Greenberg (1963); languages that only partially fulfilled these criteria are listed in *italics*, and for two languages (Kayardild and Sierra Miwok), not enough information was provided to even tentatively sort them into a category.

(55) *the languages of the accusative sample by headedness*

| headedness | # | languages |
|---|---|---|
| head-initial | 17 | *Cocama*, *Fongbe*, *Hungarian,* Indonesian, Jola Bandial, Lango, *Latin*, Lele, Maori, Maybrat, Misantla Totonac, Modern Hebrew, Nandi, Northern Sami, Nuuchahnulth, Pipil, Swahili |
| head-final | 21 | Aguaruna, Awa Pit, *Aymara*, Barasano, Comanche, Daga, Hixkaryana, Japanese, Kannada, Koasati, Kolyma Yukaghir, Lower Grand Valley Dani, Manchu, Manipuri, Maricopa, Mian, Sandawe, Somali, Southern Pomo, Turkish, Warao |
| unclear | 2 | Kayardild, Sierra Miwok |

Finally, the table in (56) lists the languages in the sample as per their marking strategy. While the distribution does initially seem weighted toward head-marking ('mixed'-marking refers to the presence of both head-marking and dependent-marking in the same language), in actuality the proportions in (56) accurately reflect the fact that head-marking is a disproportionately popular strategy among accusative languages worldwide (see again 52-53). In sum, the data and tables in (52-56) are

meant illustrate that reasonable precautions were taken to ensure that the sample of accusative languages is not contaminated by *typological bias*.

(56) *the languages of the accusative sample by marking strategy*

| marking strategy | # | languages |
|---|---|---|
| head-marking | 15 | Cocama, Daga, Fongbe, Hixkaryana, Indonesian, Jola Bandial, Lango, Lele, Lower Grand Valley Dani, Maybrat, Mian, Misantla Totonac, Nandi, Pipil, Swahili |
| 'mixed'-marking | 17 | Aguaruna, Awa Pit, Aymara, Barasano, Comanche, Hungarian, Kannada, Kolyma Yukaghir, Latin, Maricopa, Modern Hebrew, Northern Sami, Sandawe, Sierra Miwok, Somali, Turkish, Warao |
| dependent-marking | 8 | Japanese, Kayardild, Koasati, Manchu, Manipuri, Maori, Nuuchahnulth, Southern Pomo |

The last type of bias to avoid is *cultural bias*, which occurs when languages in a sample disproportionately belong to either very large speech communities, or very small ones. Addressing this concern, the table in (57) lists the languages in the accusative sample by number of speakers.[13]

---

[13] All numbers in (57) from *Ethnologue* (www.ethnologue.com).

(57) *the languages in the accusative sample by number of speakers*

### North America

| | | |
|---|---|---|
| 1. | Koasati | 350 |
| 2. | Southern Pomo | unknown |
| 3. | Misantla Totonac | 500 |
| 4. | Sierra Miwok | 7 |
| 5. | Comanche | 100 |
| 6. | Pipil | 20 |
| 7. | Nuuchahnulth | 130 |
| 8. | Maricopa | 100 |

### South America

| | | |
|---|---|---|
| 9. | Aymara | 2,800,000 |
| 10. | Awa Pit | 13,000 |
| 11. | Hixkaryana | 600 |
| 12. | Warao | 28,100 |
| 13. | Aguaruna | 38,300 |
| 14. | Barasano | 1,890 |
| 15. | Cocama | 250 |

### Australia & New Guinea

| | | |
|---|---|---|
| 16. | Maori | 148,000 |
| 17. | Maybrat | 25,000 |
| 18. | Kayardild | 23 |
| 19. | Daga | 9,000 |
| 20. | Grand Valley Dani | 20,000 |
| 21. | Mian | 1,400 |

### Europe

| | | |
|---|---|---|
| 22. | Latin | - |
| 23. | Hungarian | 12,600,000 |
| 24. | Northern Sami | 25,700 |

### Africa

| | | |
|---|---|---|
| 25. | Lele | 26,000 |
| 26. | Somali | 6,460,000 |
| 27. | Swahili | 15,400,000 |
| 28. | Sandawe | 60,000 |
| 29. | Fongbe | 1,400,000 |
| 30. | Jola Bandial | 13,000 |
| 31. | Nandi | 949,000 |
| 32. | Lango | 1,490,000 |

### Asia

| | | |
|---|---|---|
| 33. | Hebrew | 5,302,000 |
| 34. | Indonesian | 43,000,000 |
| 35. | Kannada | 37,700,000 |
| 36. | Japanese | 127,000,000 |
| 37. | Manipuri | 1,485,000 |
| 38. | Manchu | 10 |
| 39. | Turkish | 66,500,000 |
| 40. | Kolyma Yukaghir | 370 |

### 3.6.2 Coding the sample

Now that I have explained the methodology for how the accusative sample was selected, it's time to explain how each individual language was analyzed. Each

of the 40 languages was analyzed in the same way. Typically, one reliable source was obtained as a main reference for each language, although sometimes I used two or three sources for a given language as needed. These sources were usually published grammars or field linguistics dissertations; each language's primary source is listed in (58).

(58) *languages in the accusative sample by primary source*

### North America

1. Koasati — Kimball (1991)
2. Southern Pomo — Walker (2013)
3. Misantla Totonac — MacKay (1999)
4. Sierra Miwok — Broadbent (1964)
5. Comanche — Charney (1993)
6. Pipil — Campbell (1985)
7. Nuuchahnulth — Davidson (2002)
8. Maricopa — Gordon (1986)

### South America

9. Aymara — Coler (2014)
10. Awa Pit — Curnow (1997)
11. Hixkaryana — Derbyshire (1985)
12. Warao — R.-F. (1997)
13. Aguaruna — Overall (2007)
14. Barasano — J. & J. (1991)
15. Cocama — V. Y. (2010)

### Australia & New Guinea

16. Maori — Bauer (1993)
17. Maybrat — Dol (2007)
18. Kayardild — Evans (1995)
19. Daga — Murane (1974)
20. L. G. V. Dani — Bromley (1981)
21. Mian — Fedden (2007)

### Europe

22. Latin — O. & S. (2014)
23. Hungarian — K. et al (1998)
24. Northern Sami — Wilbur (2014)

### Africa

25. Lele — Frajzyngier (2001)
26. Somali — Saeed (1999)
27. Swahili — Loogman (1965)
28. Sandawe — Steeman (2011)
29. Fongbe — L. & B. (2002)
30. Jola Bandial — Bassène (2007)
31. Nandi — C. & C. (1989)
32. Lango — Noonan (1992)

### Asia

33. Hebrew — Glinert (1989)
34. Indonesian — Sneddon (1996)
35. Kannada — Sridhar (1990)
36. Japanese — Storm (2003)
37. Manipuri — B. & N. (1997)
38. Manchu — Gorelova (2002)
39. Turkish — G. & K. (2005)
40. Kolyma Yukaghir — Maslova (2003)

Using these reference sources, each of the 40 languages was combed through for the same set of morphosyntactic features. Furthermore, these features are by design supposed to closely mirror those in the ergative sample, to provide a baseline for comparison and analysis.

First and foremost, each language was coded for the presence or absence of the accusative and nominative syncretic patterns, as defined in Sections 3.2-3.3. In principle, this simple measure provides an estimate of their prevalence cross-linguistically.

In addition, all 40 languages were coded for additional morphosyntactic properties that seemed relevant in the discussion of Section 3.5, as well as in the description of ergative patterns in Chapter 2. The goal was to be able to tease apart additional cross-linguistic properties of these different syncretic patterns, and as always to compare these findings back to the ergative sample.

With this in mind, the languages in the sample were coded for their marking type (head-marking, dependent-marking, and 'mixed'-marking), headedness and basic word order. In addition, following the criteria set in Chapter 2, instances of paradigmatic overlap are further sorted into three categories: 'full', 'partial' and 'minor'.

One morphosyntactic feature from Chapter 2 that no longer seemed relevant is ergative splits. At no point in the data did I encounter an 'accusative' split in the classic sense of Silverstein (1976). In Chapter 2, several languages were shown to switch from a default ergative-absolutive alignment to nominative-accusative when conditioned by either tense or person. In the accusative sample, at best some languages went from nominative-accusative alignment to simply neutral alignment in certain circumstances. For example, in Aguaruna, transitive clauses with 1PL, 2SG

or 2PL transitive subjects do not trigger ACC marking in the transitive object (compare 59a-59b). Furthermore, it is not the case that the transitive subject 1PL, 2SG or 2PL feature ergative morphology either (59c).

Aguaruna                                              (Overall 2007; p.216, p.444, p.444)

(59) a. *ataʃu    yu-a-tata-hi-i*
        chicken eat-HIAF-FUT-1PL.SUBJ-DECL
        'we will eat a chicken'

   b. *ataʃu-na    yu-a-tata-ha-i*
      chicken-ACC eat-HIAF-FUT-1SG.SUBJ-DECL
      'I will eat a chicken'

   c. *hutii ainau-ti ama-sa-tata-himɨ-i        tsamau atumɨ ainau*
      1PL  PL-SAP  give-ATT-FUT-1>2.PL-DECL banana 2PL   PL
      'we gave you (PL) bananas'

In Awa Pit, the verbal head-marking system is nominative-accusative in the past, but it switches to a hierarchical system with neutral alignment in non-past clauses (Curnow 1997). Similarly, in Aymara, head-marking is nominative-accusative when the transitive event involves 3rd persons only; when it involves Speech Act Participants, head-marking is fusional and also shows no alignment (Coler 2014). These facts of Awa Pit and Aymara are interesting empirically, but they do not constitute an accusative 'equivalent' to ergative splits in the sense of Georgian or Yukulta (i.e. fully switching from one morphological alignment to another).

At best, the greatest extent of ergative alignment in nominative-accusative languages occurs in the possession of nominalized clauses. For example, in Hungarian, intransitive subjects (60a) and transitive objects (60b) can serve as bare possessors in nominalizations. Transitive subjects, on the other hand, cannot be bare in

169

these constructions; they must be marked by the postposition *által* 'by' (60c). That being said, this too does not constitute a true example of an 'accusative split', and therefore the issue will not be pursed further.[14],[15]

Hungarian                                            (Kenesei et al 1998; p.206, p.206, p.206)

(60) a. *Ilona  meg-érkez-és-e*
Helen PFX-arrive-NMLZR-3SG.POSS
'Helen's arrival'

b. *a    könyv el-olvas-ás-a*
the book  PFX-read-NMLZR-3SG.POSS
'the reading of the book'

c. *a    könyv-nek János által-i   el-olvas-ás-a*
the book-DAT John   by-ADJ PFX-read-NMLZR-3SG.POSS
'John's reading of the book'

To summarize, all the morphosyntactic features that were coded for in the accusative sample are listed in table (61) below, along with the possible values for each category.

(61) *Summary of relevant morphosyntactic features in the accusative sample*

---

[14] For more on this topic, see Alexiadou (2001). Her analysis posits that a subtype of verb nominalizations, event nominals, are essentially an ergative construction. Furthermore, she predicts that these event nominals are the diachronic origin of fully ergative alignment in finite main clauses.

[15] In the accusative sample, several other languages besides Hungarian exhibited an ergative pattern of possession in nominalized clauses: Nandi (Creider & Creider 1989), Maori (Bauer 1993), Comanche (Charney 1993), Hixkaryana (Derbyshire 1985) and Kayardild (Evans 1995).

| feature | value |
|---|---|
| - ACC=GEN | - Y/N |
| - NOM=GEN | - Y/N |
| - marking type | - head-marking |
| | - 'mixed'-marking |
| | - dependent-marking |
| - headedness | - head-initial |
| | - head-final |
| - basic word order | - SVO |
| | - SOV |
| | - *etc...* |
| - degree of overlap | - 'full' |
| | - 'partial' |
| | - 'minor' |

## 3.7   Results

### 3.7.1   Accusative and nominative syncretic patterns are not very common

The results of the accusative survey reveal that across the board, ACC=GEN, ACC=POSS, NOM=GEN and NOM=POSS occur at lower rates than similar patterns in the ergative sample. **Only 11 out of 40** (27.5%) of languages displayed an accusative pattern of syncretism, and only **5 out of 40 (12.5%)** languages displayed a nominative pattern, respectively.

(62) *Prevalence of accusative and nominative syncretism in the accusative sample*

| **accusative syncretism** | **nominative syncretism** | **neither** |
|---|---|---|
| Aguaruna, Comanche, Hixkaryana, Indonesian, Jola Bandial, Lango, Lele, Lower Grand Valley Dani, Modern Hebrew, Northern Sami, Warao | Cocama, Hungarian, Maybrat, Maricopa, Turkish | Awa Pit, Aymara, Barasano, Daga, Fongbe, Hixkaryana, Japanese, Kannada, Kayardild, Koasati, Kolyma Yukaghir, Latin, Manchu, Manipuri, Maori, Mian, Misantla Totonac, Nandi, Nuuchahnulth, Pipil, Sandawe, Sierra Miwok, Somali, Southern Pomo, Swahili |

$$11/40 = \underline{\mathit{27.5\%}} \qquad 5/40 = \underline{\mathit{12.5\%}} \qquad 24/40 = \underline{\mathit{60\%}}$$

For comparison, (63) shows the distribution of <u>all</u> attested syncretic patterns, by alignment and across their respective samples. As reported in Chapter 2, this difference between the ergative and accusative samples is statistically significant ($\chi^2$=7.598, $p$-value=0.022).

(63) *ACC=GEN* (l), *NOM=GEN* (m) *and ERG=GEN* (r) *across both samples*



172
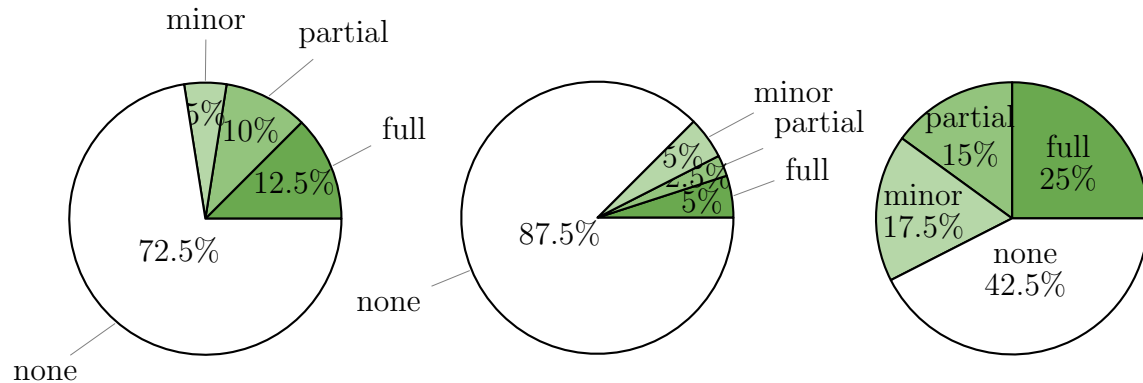
### 3.7.2 Typological macroparameters

The data reveals that for both accusative and nominative syncretic patterns, headedness is not a predictor. As the table in (64) shows, rates of syncretism occurred roughly equally in head-initial and head-final languages, and therefore the results are not statistically significant ($\chi^2$=1.523, $p$-value>0.05).

(64) *Headedness and nominative-accusative syncretisms*

|  | ACC | NOM | no pattern |
|---|---|---|---|
| head-initial (n=17) | 6 | 3 | 8 |
| head-final (n=21) | 5 | 2 | 14 |

However, the data comparing rates of nominative-accusative syncretism to basic word order is more revealing. As in (65), verb-initial languages never exhibited any of the four patterns in question. In contrast, verb-medial and verb-final languages did exhibit ACC=GEN, ACC=POSS and NOM=POSS. And while the sample is small for head-initial languages (n=5), statistical analysis indicates that these results are significant ($\chi^2$=11.326, $p$-value=0.023), and which in turn suggests that basic word order may indeed be a predictor after all.[16]

(65) *Basic word order and nominative-accusative syncretisms*

|  | ACC | NOM | no pattern |
|---|---|---|---|
| V-initial (n=5) | 0 | 0 | 5 |
| V-medial (n=13) | 7 | 3 | 3 |
| V-final (n=20) | 4 | 2 | 14 |

---

[16] Note that this is the inverse distribution from the ergative sample in Chapter 2: in that sample, V-initial languages were <u>more</u> likely to instantiate an ergative pattern of syncretism.

Finally, looking at rates of nominative-accusative syncretisms vis-à-vis marking type is perhaps most revealing. Sorting languages of the survey into three categories (head-marking, 'mixed'-marking and dependent-marking), the data reveals that the first two types exhibit both types of patterns, but the latter type does not (66).

(66) *Marking type and nominative-accusative syncretisms*

|  | ACC | NOM | no pattern |
|---|---|---|---|
| head-marking (n=15) | 6 | 2 | 7 |
| 'mixed'-marking (n=17) | 5 | 3 | 9 |
| dependent-marking (n=8) | 0 | 0 | 8 |

While the results in (66) are not statistically significant ($\chi^2$=7.162, $p$-value> 0.05), there is another way of 'slicing the data' which may be helpful. Remember that, by definition, 'mixed'-marking languages are languages that feature both head-marking and dependent-marking at the same time. Like in Chapter 2, we can sort instances of head-marking from 'mixed'-marking languages into the same category as instances from strictly head-marking languages. We can also sort instances of dependent-marking from 'mixed'-marking languages into the same category as instances from strictly dependent-marking languages. Redistributing the data in this way yields 32 total instances of head-marking and 25 total instances of dependent marking in the accusative sample, as in (67).

(67) *Marking type and nominative-accusative syncretisms (redux)*

|  | ACC | NOM | no pattern |
|---|---|---|---|
| <u>all</u> head-marking (n=32) | 9 | 5 | 18 |
| <u>all</u> dependent-marking (n=25) | 3 | 0 | 22 |

The data in (67) indeed suggests that both accusative patterns of syncretism and nominative patterns of syncretism are more likely to be instantiated in head-marking than dependent-marking, and this difference is statistically significant ($\chi^2$= 7.656, $p$-value=0.022). This trend observed in the accusative sample is consistent with that in the ergative sample: syncretic patterns are more likely to be instantiated in head-marking than in dependent-marking. Again, I have no immediate explanation for these facts, and leave the subject open to further research.

## 3.8 Conclusion

In this chapter, a sample of 40 accusative languages was constructed to serve as the counterpart to the ergative sample in Chapter 2. Its purpose was to study the ACC=GEN, ACC=POSS, NOM=GEN and NOM=POSS syncretic patterns, themselves the counterparts to ERG=GEN and ERG=POSS in ergative languages. The patterns were clearly defined, and ample data was presented to illustrate their many permutations.

The main finding is that globally the set of nominative-accusative syncretisms occur at significantly lower rates than the ergative ones do in their respective samples. Some minor findings to report also include the fact that unlike for the ergative sample, major typological parameters <u>do</u> seem to predict the occurrence of syncretic patterns in this sample. Both families of nominative-accusative syncretic patterns were found to occur significantly less often in V-initial languages, and in dependent-marking.

As always, more research is needed in the form of larger samples to either confirm or disconfirm these preliminary findings. However, under the assumption that the findings in Chapters 2 and 3 represent all languages, cross-linguistically, we can now proceed to a formal syntactic analysis. In Chapter 4, a indepdently-motivated model of case and agreement in the clausal domain is tested against the syncretism findings from Chapters 2 and 3. The result ultimately models the fundamental distinction between ergative syncretic patterns on the one hand – which seem to occur at rates near 50% – and non-ergative patterns on the other, which occur much more infrequently.

# Chapter 4

# THEORETICAL IMPLICATIONS OF ERG=GEN, ERG=POSS AND OTHER SYNCRETIC PATTERNS

The goal of this chapter is to evaluate whether current theories of case and agreement are compatible with the typological findings from Chapters 2 and 3. In Section 4.1, I will outline a general model of case and agreement which is built on previous work by Marantz (1991), Bobaljik (2008) and Baker (2013). A brief overview of each component of the model is provided, showing that all together they can account for a wide range of morphological patterns but also some well-known typological facts. In Section 4.2, the discussion shifts to an analysis of the syncretic patterns from Chapters 2 and 3. In Section 4.2.1, the general model of case and agreement is shown to to successfully account for the ergative syncretic patterns from Chapter 2 (ERG=GEN and ERG=POSS). However, in Sections 4.2.2-4.2.3, this model is shown to make poor predictions when it comes to rarer syncretic patterns such as ACC=GEN, NOM=POSS and ABS=POSS. As a result, Section 4.3 proposes that while ergative syncretisms are generated by extending the computation of ergative case and agreement to the nominal domain, non-ergative syncretisms are generated by a different process altogether. These rarer syncretic patterns occur when case-marking and agreement paradigms are 'recycled' in the nominal domain for reasons of economy, a process which can potentially be influenced by outside factors such as morphological markedness and linear word order.

## 4.1 A model of case and agreement

Section 4.1 serves as an outline of the general model of case and agreement which I will be assuming throughout this chapter, and against which I will ultimately be testing the range of syncretic patterns described in Chapters 2 and 3. The next few subsections will show that the model is well-motivated in that it can account for a range of empirical and typological facts. But for now it may be useful to provide a snapshot of the model's overall architecture. It is fundamentally a linear, two-step model: first, syntactic structure provides the input for the case computation algorithm, which produces morphological case as its output. Then, if necessary, the output of the case computation algorithm serves as the input to the agreement algorithm in turn, producing verbal agreement as its output. Thus, according to Bobaljik (2008), the relationship between the computation of case and agreement is best characterized as one of **feeding**, where the output of the first computation (case) feeds the next step in the derivation (agreement). This linear, two-step model is schematized in (1) below.

(1) *The Marantz-Bobaljik-Baker model of case and agreement*

syntactic structure    →    case algorithm    →    agreement algorithm

## 4.1.1 Marantz (1991) and dependent case

The first major component of this model of case and agreement is Marantz's (1991) theory of dependent case. The major intuition formalized by Marantz is that morphologically-marked case will only occur if there are two eligible (i.e. unmarked) NPs within the clausal domain of case computation. In other words, the realization of morphologically-marked case (either ERG or ACC, depending on alignment) is

**dependent** on the existence of another eligible NP within that domain. His exact formalism for dependent case is in (2).

(2) Dependent case is assigned by V+I to a position governed by V+I when a distinct position governed by V+I is:

    a. not "marked" (not part of a chain governed by a lexical case determiner)

    b. distinct from the chain being assigned dependent case

      Dependent case assigned up to subject: ergative

      Dependent case assigned down to object: accusative

Furthermore, a disjunctive hierarchy posits that within that domain governed by V+I, the case assignment algorithm assigns different types of case following a strict sequence. According to (3), the sequence goes as follows: first, lexical case is assigned by specific lexical heads idiosyncratically; then dependent case is assigned according to the formalism in (2); finally, default case is assigned last in the sequence.

(3) Case realization disjunctive hierarchy

    1. lexically-governed case

    2. dependent case (ergative and accusative)

    3. default case

In this way, Marantz's theory of dependent case can account for simple cases of ergative morphological case, such as in Yukulta. In (4a), the case computation algorithm identifies two eligible, unmarked NPs in the clausal domain, *tir* 'snake' and *mantuwara* 'boy'. Because Yukulta's morphological case is in ergative alignment, the algorithm first assigns dependent ergative case to the highest unmarked NP in the domain, *tir* 'snake'. Afterwards, default absolutive case is assigned to all remaining NPs in the domain. The same algorithm applies to an intransitive clause of Yukulta;

179

in (4b), the algorithm fails to find an appropriate candidate for ergative case, given that only one NP is eligible (*kukuwi:nta* 'crocodile'). As a result, dependent case cannot be assigned, and the intransitive subject instead receives default absolutive case much like the transitive object did in (4a). Note that in (4b) *wampal* 'bank' does not count as eligible for the purposes of the algorithm, having already been assigned lexical LOC case at a prior stage in the derivation.[1,2]

Yukulta                                                              (Keen 1972; p.121, p.243)

(4)  a.  *tir-iya*      *=kanta*    *pa:tja*  *mantuwara*
         snake-ERG  TR.PAST  bite       boy

'the snake bit the boy'



---
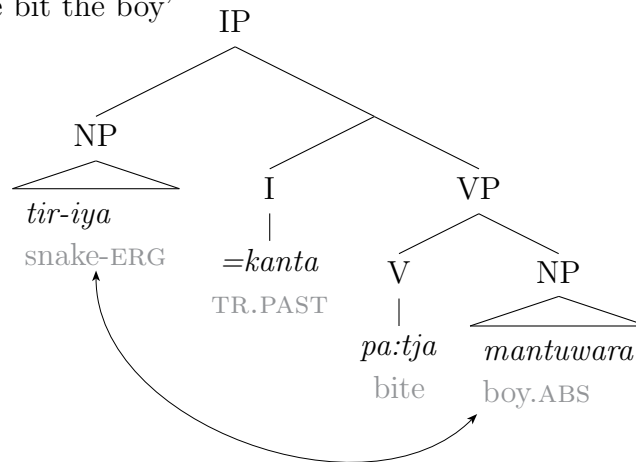
[1]  Accusative dependent case is a matter of parameterization in Marantz's account. It has the same requirements as dependent ergative case – two eligible NPs in the clausal domain governed by V+I – but instead the structurally-<u>lower</u> eligible NP receives dependent ACC case. As a result, both the transitive and the intransitive subject will end up receiving default NOM case.

[2]  In this chapter, derivations are modeled with the grammatical subject in Spec-IP; however, I remain agnostic as to whether it is base-generated in that position, or whether it is generated VP-internally (Koopman & Sportiche 1991).

b. *kukuwi:nta =yiŋka ḷaṭitija wampal-i*

crocodile PAST wait bank-LOC

'the crocodile waited on the bank'

```
                        IP
                   ╱         ╲
              NP              ╲
          ╱────────╲       ╱──────╲
         kukuwi:nta      I          VP
        crocodile.ABS    │       ╱─────╲
                       =yiŋka    V       PP
                        PAST     │    ╱─────╲
                              ḷaṭitija  wampal-i
                                wait   bank-LOC
```
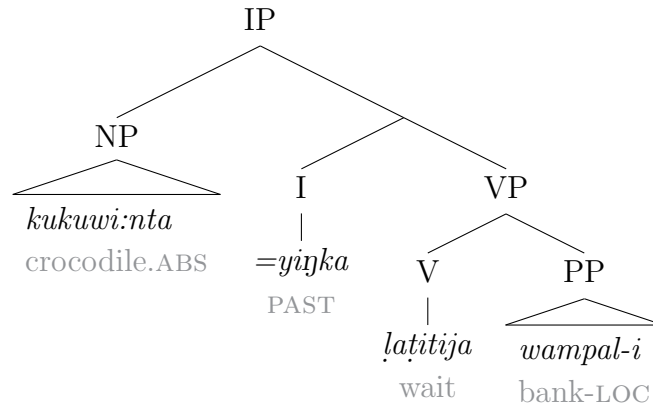
Marantz's theory of dependent case is useful because it can capture not only typical cases of ergative case such as in Yukulta, but atypical ones as well. In Shipibo, a typical transitive clause consisting of an external argument and an internal argument is computed exactly like in Yukulta. In (5a), dependent ergative case is assigned to the external argument *jiwi* 'tree' by virtue of being structurally higher than the other eligible candidate, the internal argument *Sani* 'Sani'. As before, default absolutive case is then assigned to any remaining NPs in the domain, namely the internal argument *Sani*.

Shipibo                     (Valenzuela 2003; p.325 / Baker & Bobaljik to appear; p.10)

(5) a. *jiwi-n-ra    Sani rishki-ke*
       tree-ERG-EV Sani hit-CMPL

   'the tree hit Sani'
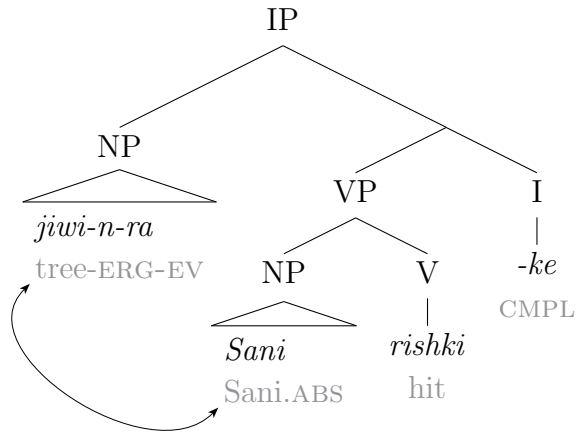
b. *bimi-n-ra    Rosa joshin-xon-ke*
   fruit-ERG-EV Rosa ripen-APPL-CMPL

'the fruit ripened for Rosa'
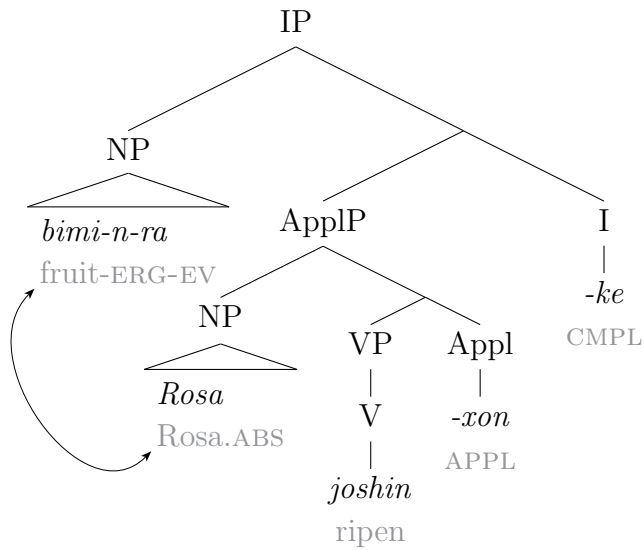


However, as Baker and Bobaljik (to appear) observed, ergative case can also be assigned in clauses which are arguably transitive, but which do not feature a canonical external argument and internal argument. In (5b), *bimi* 'fruit' is obviously an internal argument (being the sole argument of *joshin* 'ripen'); nevertheless, if the

verb is augmented via applicative morphology, and therefore a benefactive NP also occurs in this ostensibly unaccusative clause, even an internal argument such as *bimi* 'fruit' may receive ergative case. While not a typical transitive clause, this scenario is perfectly well accounted-for by Marantz's case-computing algorithm: *bimi* 'fruit' receives ergative case simply by virtue of being structurally higher in the clausal domain.

Note that in (5b) above, *bimi* 'fruit' is modeled in the Spec-IP position, despite being the internal argument of *joshin* 'ripen', an unaccusative verb. The exact details of how and why it moves to Spec-IP aren't important at this point, but the main observation is crucial: an NP does not have to be an external argument in order to receive dependent ergative case. Indeed, Baker and Bobaljik provide similar examples from West Greenlandic, Yup'ik and Chukchi, indicating that this pattern is well-attested. Furthermore, this simple typological fact can easily be captured with a dependent case approach such as Marantz (1991), providing further justification for its inclusion in the general model under discussion.[3]

### 4.1.2 Bobaljik (2008) and verb agreement

Marantz's (1991) theory of dependent case thus serves as the first major component of the general model of case and agreement in (1). Building on Marantz's work, Bobaljik's (2008) theory of agreement serves as the next major component, and like Marantz, his theory covers a wide range of case and agreement data, but it also makes good typological predictions.

---

[3] For alternative theories where ergative is modeled as inherent case, see Aldridge (2004, 2007, 2008) and Legate (2008, to appear). Those theories are not pursued here because of issues like the one raised by Shipibo in (5b), which they struggle to account for.

Bobaljik's key insight was that morphological case appears to feed the computation of agreement. Specifically, verb agreement is determined by the highest **accessible** NP in the domain governed by V+I. Whether a given NP is accessible or not depends on a language's parameterization with regard to the hierarchy in (6). In some languages, only NPs marked with default case are accessible, while in others, both dependent case and default case make an NP accessible for verb agreement. A language in which all case-marked NPs are eligible is also allowed in this theory, but crucially there can be no language in which, for example, dependent case is eligible for verb agreement but unmarked case is not. The accessibility hierarchy in (6) is implicational: if dependent case is eligible then so is unmarked case, but not the other way around.

(6) *The accessibility hierarchy in Bobaljik (2008)*[4]

unmarked case ≫ dependent case ≫ lexical/oblique case

The locus of agreement in Bobaljik's theory can be any verbal head, as long as it occurs in the clausal domain governed by V+I. According to Bobaljik, the relationship between the eligible NP and the head on which agreement is actually realized is not characterized by c-command or any other structural relationship in the sense of Chomsky (2000). Rather, "an NP need bear no relation to a verb other than satisfying morphological accessibility and locality in order to trigger agreement on that verb". Therefore in the derivations that follow, I am going to assume that agreement takes place either with the I head, or with a combined V+I after head-movement has taken place.

---

[4] Bobaljik's (2008) accessibility hierarchy is a revised version of Moravcsik's (1974, 1978).

To illustrate Bobaljik's theory, the derivation of a transitive sentence in Hindi proceeds as follows. In perfect aspect, case marking in Hindi is in ergative alignment; therefore the computation of dependent ergative case and default absolutive case will proceed exactly like in Yukulta and Shipibo. The algorithm identifies two eligible, unmarked NPs in the clausal domain, *Raam* 'Ram' and *RoTii* 'bread', and therefore it assigns dependent ergative case to the highest unmarked NP in the domain, which is the transitive subject. Afterwards, default absolutive case is assigned to all remaining NPs in the domain.

Hindi <span style="float:right">(Bobaljik 2008; p.18)</span>

(7)    *Raam-ne*     *RoTii*      *khaayi*     *thii*
       $\text{Ram}_M$.ERG   $\text{bread}_F$.ABS   eat.PERF.F   be.PAST.F
       'Ram had eaten bread'



⇓

185
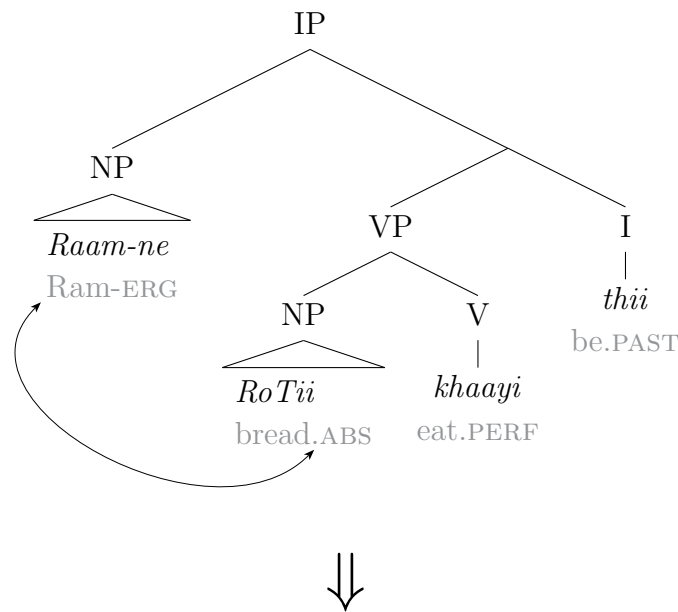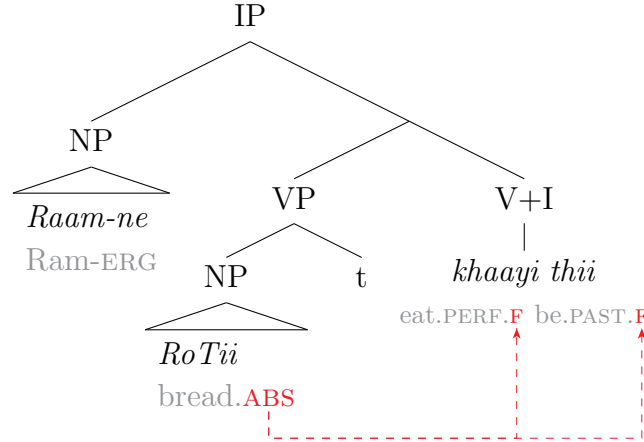
Then, the output of the case computing algorithm serves as the input to the agreement algorithm. In Hindi, dependent case is ineligible for verb agreement, therefore the agreement algorithm cannot form an agreement relationship with the highest NP in the domain, the ergative-marked transitive subject. The algorithm must keep searching for the next-highest eligible NP, which will be the absolutive-marked transitive object. The exponent of this agreement relationship is realized on both the verb and the auxiliary in Hindi, but again, the details of how exactly this occurs are not very important – the main point is that in Bobaljik's theory, morphological case (specifically, the type of morphological case) is relevant for determining which arguments will enter into an agreement relation with verbal heads, and in Hindi ergative-marked arguments are simply ineligible.

In an intransitive clause of Hindi, the derivation is extremely straightforward, as in (8). First, because there is only one unmarked NP in the domain, the requirements for dependent ergative case are not met and thus the sole argument *Raam* 'Ram' receives default absolutive case. In the next stage, the agreement algorithm searches for the highest eligible NP to agree with; this happens to be the only NP in the clause, that same absolutive-marked intransitive subject. Here, the locus of

agreement is modeled in V+I, assuming that verbal head-movement takes place in clauses with no overt auxiliaries. In this way, Bobaljik's theory can capture the fact that (in perfect aspect, at least) both case and agreement in are ergative alignment in Hindi.

Hindi                                                                    (Mohanan 1994; p.71)

(8) *Raam*      *giraa*
    Ram$_M$.ABS fall.PERF.M

'Ram fell hard'
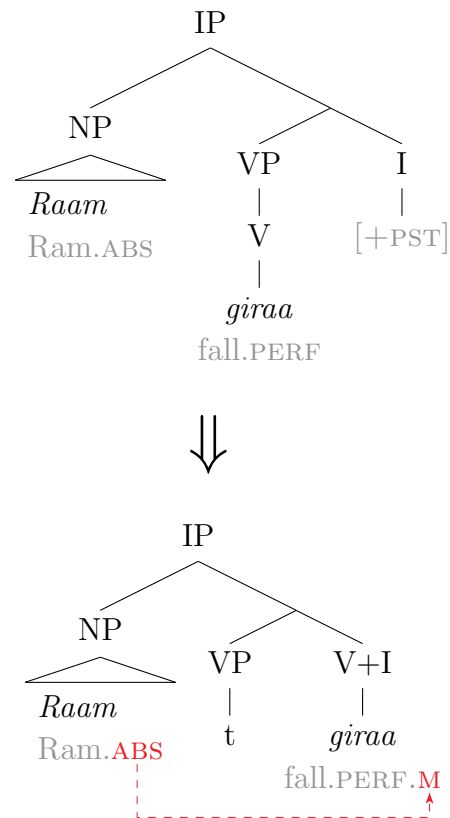
By manipulating just a single point of parameterization, Bobaljik's theory can also account for languages which exhibit ergative-absolutive case marking and nominative-accusative agreement, a configuration which appeared frequently in Chapter 2. For example, in Epena Pedee, dependent ergative case and default absolutive case are assigned as normal in (9). However, in this language, dependent-marked arguments <u>are</u> eligible for agreement; therefore, when the agreement algorithm searches among eligible NPs in the domain, the V head can enter into an agreement relationship with the structurally-highest NP, the dependent-marked transitive subject *ɨmɨkʰĩraa-rã́* 'men'.

Epena Pedee                                                                (Harms 1994; p.10)

(9) *ɨmɨkʰĩraa-rã́-pa hápa       hidiu-hi-dá*
    man-PL-ERG    canoe.ABS drag-PAST-PL
    'men dragged the canoe'



188
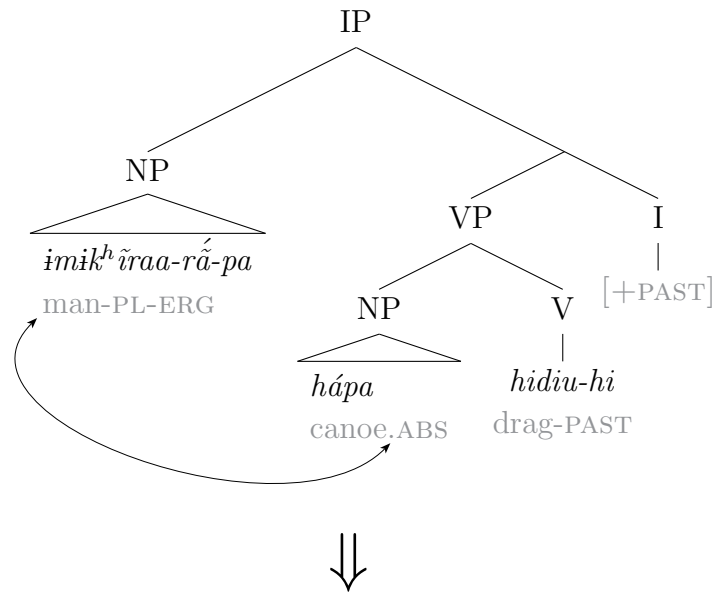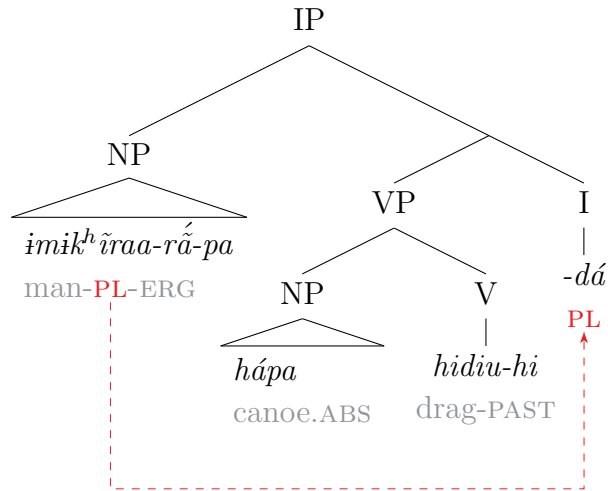
Like for Hindi, the derivation of an intransitive clause in Epena Pedee is extremely straightforward. In (10), the requirements for dependent case are not met, therefore the intransitive subject receives default absolutive case. Then, the agreement algorithm scans the domain, ultimately agreeing with the only available argument. The end result is a language with ergative-absolutive case marking (the transitive subject is marked ERG, while the transitive object and the intransitive subject are marked ABS) but nominative-accusative head-marking (the verb agrees with either the transitive or the intransitive subject).

Epena Pedee                                                                 (Harms 1994 1994; p.10)

(10)  *wẽrá-rã*       *če-hi-dá*      *barrée*
      woman-PL   come-PAST-PL downstream
      'the women came downstream'

189
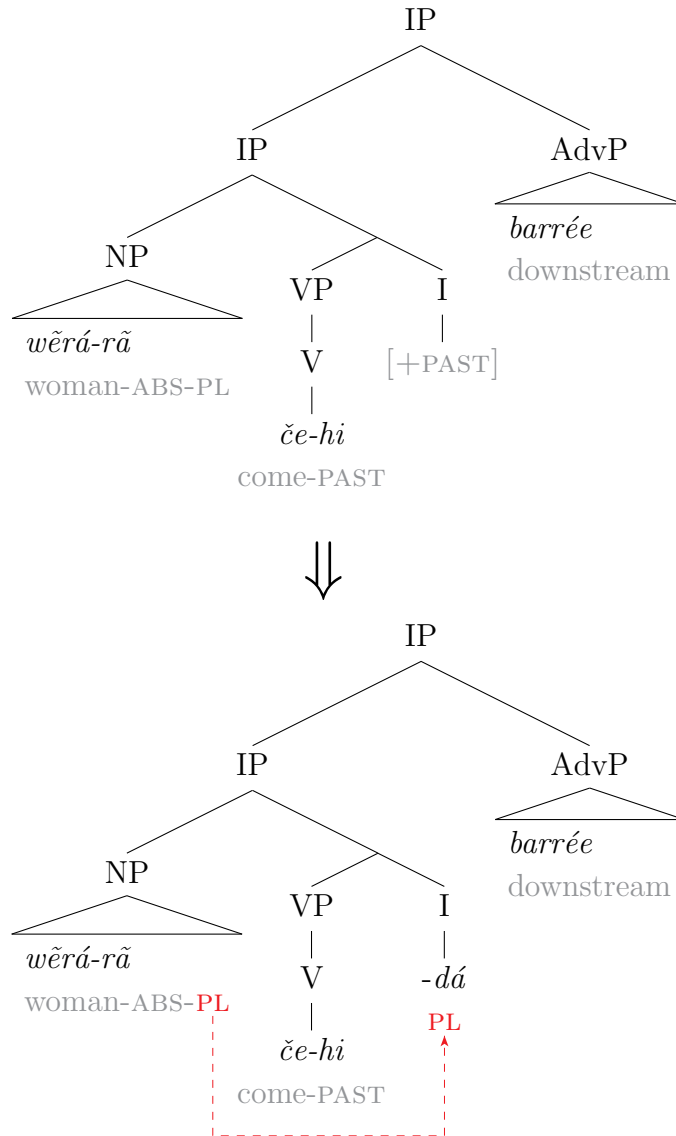
Finally, Bobaljik is also able to show that his theory can account for languages with consistently nominative-accusative alignment in case and agreement, such as Icelandic. After nominative-accusative case has been computed, in this theory nominative-accusative agreement is the only possible option: scanning the domain in such a language will always yield agreement with the structurally-higher

transitive subject, exactly like in Epena Pedee, regardless of whether dependent-marked arguments are eligible or not.

Icelandic                                                                 (Wood 2012; p.16)

(11) *Bjartur       málaði   hurðina*
     Bjartur.NOM paint.SG door.the.ACC
     'Bjartur painted the door'
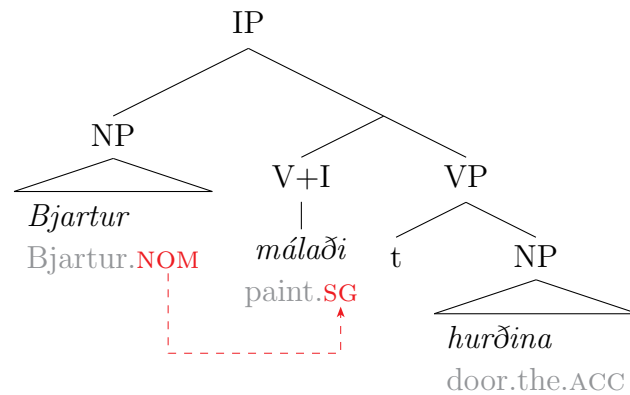


Indeed, the strength of Bobaljik's analysis is that it captures a very important typological asymmetry: ergative-absolutive case is compatible with either ergative-absolutive agreement or nominative-accusative agreement – both languages types are attested – but nominative-accusative case only ever co-occurs with nominative-accusative agreement. The asymmetry is schematized in (12) below, and the fact that Bobaljik's analysis can account for this well-known typological fact provides further motivation for its inclusion in a general model of case and agreement.

(12) *Typology of case and agreement (mis)matches*

| language type | case alignment | agreement alignment | example |
|---|---|---|---|
| Type 1 | ERG | ERG | Hindi |
| Type 2 | ERG | ACC | Epena Pedee |
| Type 3 | ACC | ERG | unattested! |
| Type 4 | ACC | ACC | Icelandic |

191

### 4.1.3   Baker (2013) and 'double' agreement

Extending the model even further, Baker (2013) proposes an analysis which aims to capture the fact that some languages exhibit agreement with both core arguments of a transitive clause. Drawing upon evidence from non-finite clauses, Baker comes to the conclusion that two configurations are possible for these languages: either a) the I head 'agrees twice', simultaneously with two different NPs or b) I agrees once with one NP, and another, lower functional projection agrees once with another NP.

Looking at data from Nez Perce, Baker notices that while in regular, finite clauses the verb agrees with both the subject and object (13a), in non-finite clauses neither subject nor object agreement is realized (13b).

Nez Perce                                                          (Baker 2013; p.15, p.15)

(13) a. *'iin wéet'u 'itúu-ne   'aa-p-sá-qa*
        1SG NEG   INDEF-ACC 1SG>3SG-eat-IMPF-PAST
        'I didn't eat anything'

   b. [ *weet'u 'ituu-ne   mic'ii-t'as* ] *qooqoX hi-wc'ee-y-e*
        not    what-ACC hear-PART   raven   3SG.SUBJ-become-P(?)-PAST
        'she become a raven, unable to hear anything'

On the other hand, in Mapudungun, the following observation holds: in regular, finite clauses the verb likewise agrees with its subject and object (14a), but in non-finite clauses only object agreement persists (14b).

Mapudungun                                                         (Baker 2013; p.15, p.16)

(14) a. *llüka-l-ka-ke-fu-fi-y-ngün*                  *ti   nge-n     ruka*
        afraid-CAUS-FAC-HAB-PAST-3SG.OBJ-IND-3PL.SUBJ the have-NMLZR house
        'they used to intimidate the house owner'

192

b. *fey ayü-w-üy*       [ *iñché ñi*      *pe-me-fi-el*      *fey* ]
    he  love-REFL-3SG.SUBJ.IND  I    1SG.POSS see-TH-3SG.OBJ-INF he

    'he was happy at my going to see him'

Thus Baker proposes to model this distinction in the following way. For languages such as Nez Perce, the locus of agreement for both arguments is in the I head, as in (15a). On the other hand, in languages such as Mapudungun, the locus of agreement is split: subjects agree with the I head, but objects agree with the transitive v head. The non-finite facts in (13-14) thus follow from this point of parameterization. In (13a), the lack of an I head in the non-finite clause means that neither subject nor object agreement are available. However, in (14b), the fact that object agreement is located in [+TRANS] v means that it alone remains available in non-finite environments.[5]

(15) *Locii of 'double' agreement*            (adapted from Baker 2013)



---
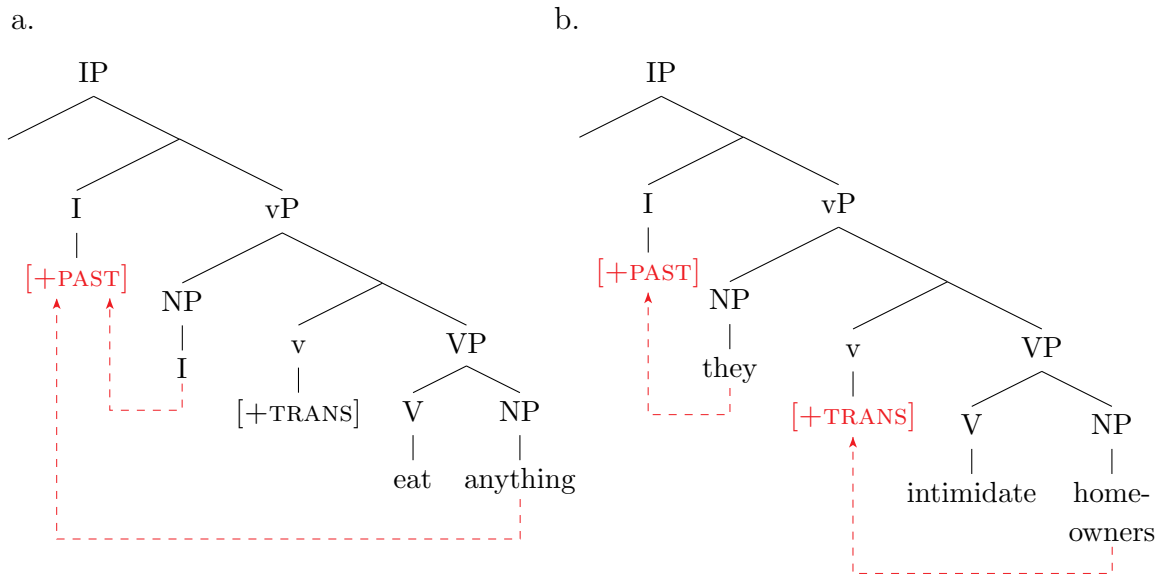
[5] See Legate (2008) for arguments that 'absolutive' case in some ergative languages is actually exactly as described for Mapudungun agreement in (15b): transitive object 'ABS' is available in non-finite clauses, but intransitive subject 'ABS' is not.

Like Marantz (1991) and Bobaljik (2008), Baker's parameterized, split-locus analysis captures some typological facts – namely that object agreement is a heterogeneous phenomenon with respect to finiteness – while still providing a framework to model double verb agreement itself. In this light, Baker's analysis can be incorporated into the general model to account for a language such as Paumarí, which features consistently ergative-absolutive alignment, both in its dependent-marking and head-marking.

In (17), after case has been computed, the agreement algorithm makes use of that case information to determine agreement as before, except now there are two agreement 'slots' to fill. First, the absolutive agreement slot is computed exactly like in Hindi: transitive subjects marked with dependent ergative case are ineligible, and so the algorithm must scan downward to agree with the next-highest eligible NP, the absolutive-marked transitive object (17a). Then, the second agreement 'slot' must be filled, and here dependent-marked arguments must obviously be eligible in order to enter into an agreement relationship with the ergative-marked transitive subject (17b). In this extension of Bobaljik's theory, a 'double agreement' language such as Paumarí is derivationally identical to a single agreement language such as Hindi, except that after first computing absolutive agreement, Paumarí must go a step further by computing a second agreement relationship with the ergative-marked transitive subject if at all possible.

Paumarí                                     (Chapman & Derbyshire 1986; p.164)

(16) *Dono-a*     *bi-ko'diraha-'a-ha*              *ada*     *isai*   *hoariha*
     Dono-ERG 3SG.ERG-pinch-ASP-THEME.M DEM.M child other
     'Dono pinched the other boy'

194

(17) a.

IP
- NP
  - *Dono-a*
    Dono-ERG
- I
  - I_ERG
  - I_ABS
    - *-ha*
      THEME.M
- VP
  - V
    - *ko'diraha-'a*
      pinch-ASP
  - NP
    - *ada isai hoariha*
      DEM.M other child.ABS

b.

IP
- NP
  - *Dono-a*
    Dono-ERG
- I
  - I_ERG
    - *bi-*
      3SG.ERG
  - I_ABS
    - *-ha*
      THEME.M
- VP
  - V
    - *ko'diraha-'a*
      pinch-ASP
  - NP
    - *ada isai hoariha*
      DEM.M other child.ABS
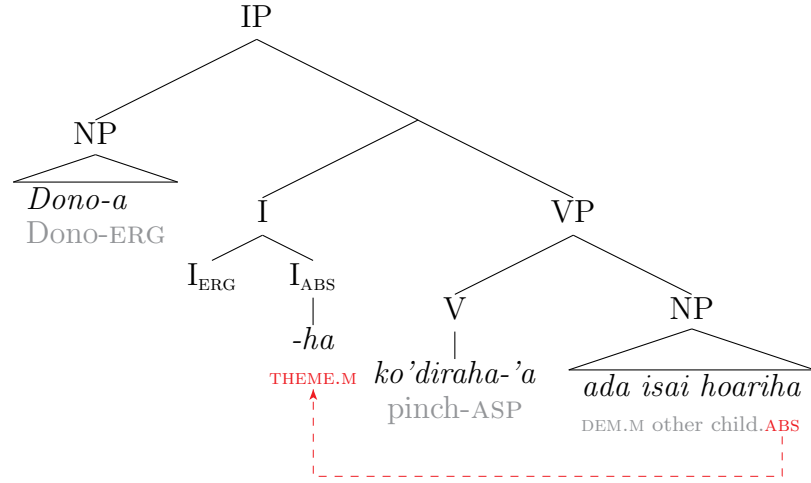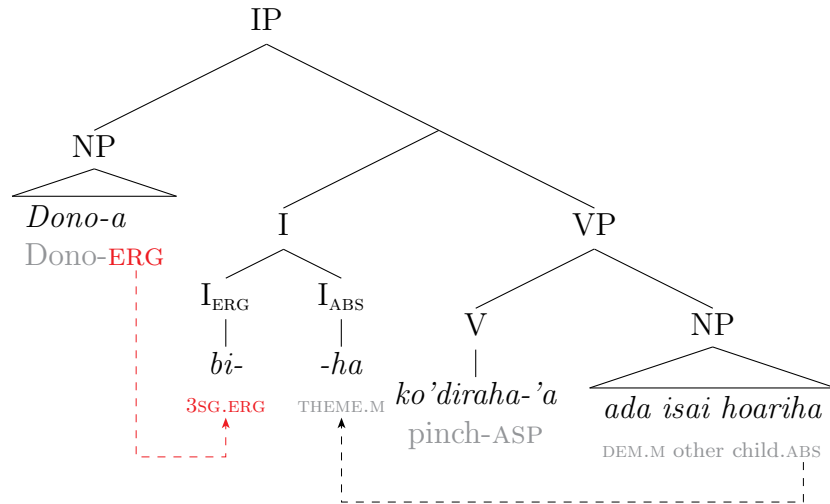
Despite the fact that Paumarí is a 'double agreement' language, its intransitive clauses will not require that extra step in the computation of agreement, since by definition they can only agree with a single argument. Therefore the derivation of an intransitive clause in Paumarí is going to look identical to one in Hindi. In (18), after case has been computed and the sole argument has been assigned unmarked absolutive case, the agreement algorithm enters into an agreement relationship with

the intransitive subject. Even though it doesn't really matter in (17b), the eligibility requirements for agreement with the intransitive subject are the same as they were for the first agreement slot in (17a): dependent-marked NPs are ineligible. In this view, this is exactly where verb agreement's alignment is parameterized in a language such as Paumarí: the grammar simply knows that the agreement relationship which is computed first in transitive clauses is also computed in intransitive clauses.

Paumarí                                              (Chapman & Derbyshire 1986; p.195)

(18) *ihamahi-'a-ha*            *ada*     *Dono*
     be.angry-ASP-THEME.M DEM.M Dono

     'Dono was angry'
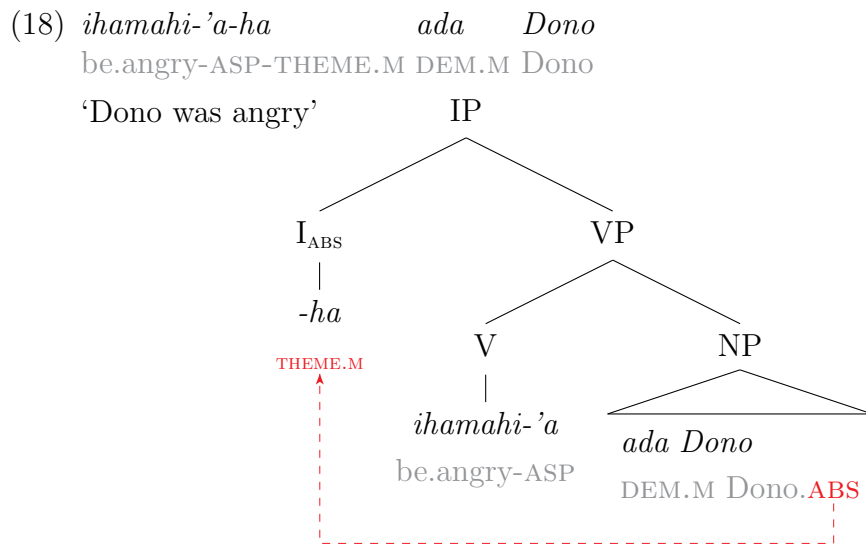


This extension of Bobaljik's agreement algorithm can also account for languages with 'asymmetric' alignment like Epena Pedee, and double verbal agreement like Paumarí. Burushaski is exactly that type of language: it features ergative-absolutive case marking, and verb agreement with both the subject and object but which is nominative-accusative in alignment. In (20a), after case has been computed, two agreement 'slots' must be filled. First, nominative agreement is computed exactly like in Epena Pedee: because NPs marked with dependent case are eligible, an agreement relationship forms with the ergative-marked transitive subject (20a).

Then, the second agreement slot must be filled, but at this point it doesn't really matter whether dependent case is eligible or not, because the only remaining NP in the domain is the transitive object marked with default absolutive case (20b).

Burushaski (Munshi 2015; p.29)

(19) *hamal-e* *ḍarbeṣ* *i-cʰarkan-uman*
  neighbour-ERG Darbes 3SG.M.OBJ-beat-3PL.SUBJ
  'the neighbours beat Darbes'

(20) a.



 b.



197
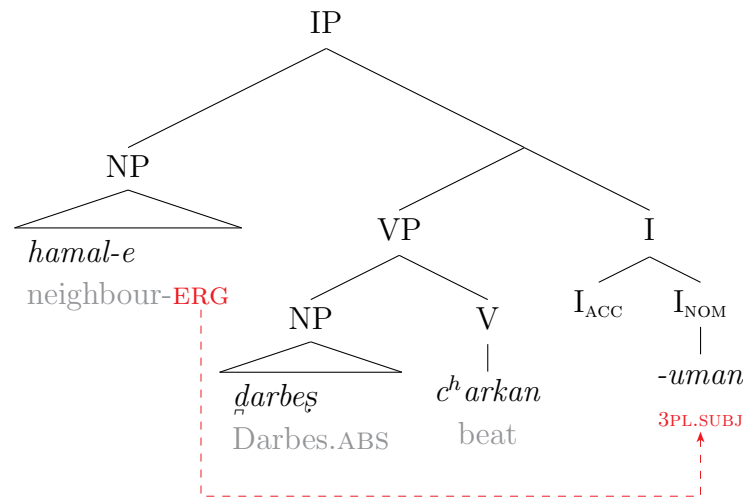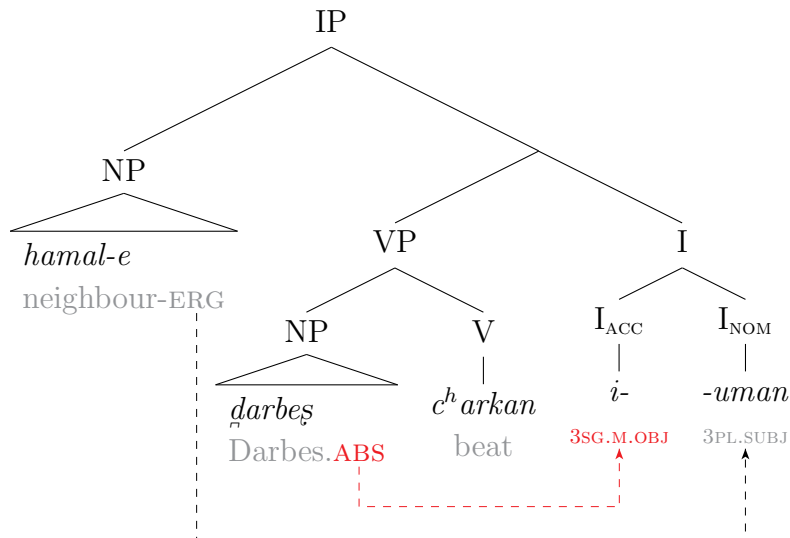
Again, deriving an intransitive clause in Burushaski is more straightforward because there is only one agreement 'slot' to fill. In (21), after the intransitive subject has been assigned default absolutive case, it will seek to enter into an agreement relationship with that same slot which was filled first in the derivation of a transitive clause ($I_{\text{NOM}}$), and according to the same eligibility requirements. Here, even though dependent-marked NPs are in principle eligible, it doesn't matter because the only available argument is marked with default absolutive case, which is of course always eligible for agreement.

Burushaski (Munshi 2015; p.12)

(21) *hiles guc$^h$ar-imi*
boy walk-3SG.M.SUBJ
'the boy walked'
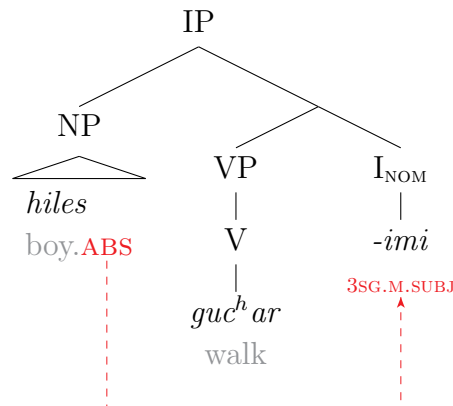


Finally, extending the model of Bobaljik-style agreement can also cover the last kind of double-agreement language in the literature, a language which is consistently nominative-accusative in both its case marking and its verbal agreement. In a transitive clause of Aguaruna, morphological case is computed exactly like in Icelandic: if there are two eligible NPs in the domain, assign dependent accusative case

198

to the structurally-lower one, and then assign default nominative case to any remaining NPs after that. Then, the output of the case assignment algorithm gets passed off to the agreement algorithm to compute two agreement relationships. First, one agreement slot looks to agree with the highest eligible NP in the domain; like in Icelandic, it doesn't matter whether dependent-marked NPs are eligible or not because the highest eligible NP will be the transitive subject marked with default nominative case (23a). Then, unlike in Icelandic, one additional agreement relationship must be computed. Here dependent-marked arguments must be eligible, because that second agreement slot will form an agreement relationship with the accusative transitive object (23b).

Aguaruna                                                          (Overall 2007; p.303)

(22) *amɨ mi-na    duʃi-mitika-ha-mɨ-i*
     2SG 1SG-ACC laugh-CAUS-1SG.OBJ+IMPF-2SG.SBJ-DECL

     'you are making me laugh'

(23) a.

b.

IP
- NP
  - *amɨ*
  - 2SG.NOM
- VP
  - NP
    - *mi-na*
    - 1SG-ACC
  - V
    - *duʃi-mitika*
    - laugh-CAUS
- I
  - I_ACC
    - *-ha*
    - 1SG.OBJ
  - I_NOM
    - *-mɨ*
    - 2SG.SUBJ
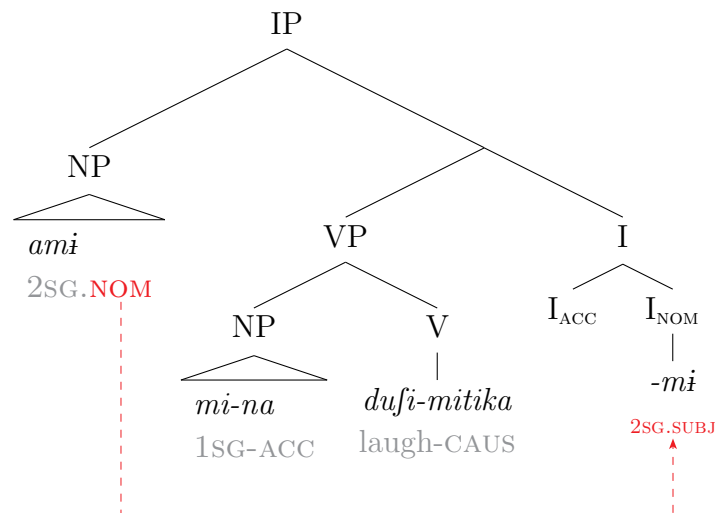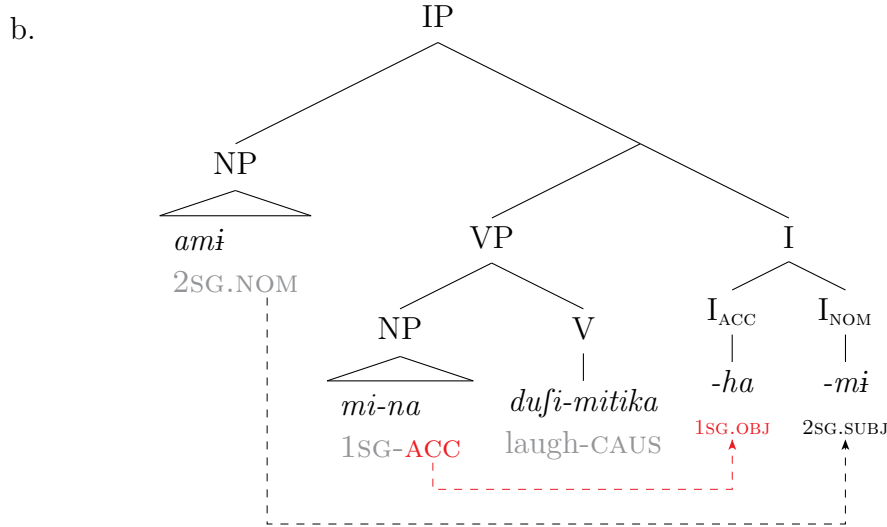
To summarize the last two sections, Bobaljik's model of verb agreement can be extended to languages which feature 'double' verb agreement. As a result, we can account for the entire range of attested permutations of case, agreement and alignment from the samples in Chapters 2 and 3, as in (24) below.[6]

(24) *Attested permutations of case, agreement & alignment in Chapters 2 & 3*

| | case | agreement | | | ex. |
|---|---|---|---|---|---|
| Hindi | ERG-ABS | slot #1: | dependent case is not eligible | ABS | (7-8) |
| Epena Pedee | ERG-ABS | slot #1: | dependent case is eligible | NOM | (9-10) |
| Icelandic | NOM-ACC | slot #1: | doesn't matter | NOM | (11) |
| Paumarí | ERG-ABS | slot #1: | dependent case is not eligible | ABS | (16-18) |
| | | slot #2: | dependent case is eligible | ERG | |
| Burushaski | ERG-ABS | slot #1: | dependent case is eligible | NOM | (19-21) |
| | | slot #2: | doesn't matter | ACC | |
| Aguaruna | NOM-ACC | slot #1: | doesn't matter | NOM | (22-23) |
| | | slot #2: | dependent case is eligible | ACC | |

[6] Note that because double agreement is modeled as an extra derivational 'step' compared to single agreement, the same languages types that were excluded in the original analysis of single agreement (NOM-ACC case, ERG-ABS single agreement) will also be systematically excluded in double agreement (NOM-ACC case, ERG-ABS double agreement). This is desirable because as far as anyone can tell the typological generalization from Bobaljik (2008) extends to languages with double agreement.

### 4.1.4   What and where is morphological case?

Now that we have a general model of case and agreement, we are almost ready to move on to a discussion of syncretic patterns. However, two issues of theoretical importance must be addressed first: what is the relationship between abstract morphological case and overt morphological case, and where in the grammar are the computations described in Sections 4.1.1-4.1.3 located?

The first question is relevant because many morphologically ergative languages are morphologically ergative by virtue of their system of head-marking, and yet they do not feature any overt dependent-marking; Karitiâna is one such language (25).

Karitiâna                            (Storto 1999; p.157, p.157, p.157, p.157, p.157, p.157)

(25) a. *yn  a-ta-oky-j          an*
      1SG 2SG.ABS-DECL-hurt-IRR 2SG
      'I will hurt you'

    d. *y-ta-opiso-t          yn*
      1SG.ABS-DECL-listen-N.FUT 1SG
      'I listened'

    b. *an  y-ta-oky-t          yn*
      2SG 1SG.ABS-DECL-hurt-N.FUT 1SG
      'you hurt me'

    e. *a-ta-opiso-t          an*
      2SG.ABS-DECL-listen-N.FUT 2SG
      'you listened'

    c. *yjxa ∅-na-ahee-t          iso*
      1PL  3SG.ABS-DECL-blow-N.FUT fire
      'we blew the fire'

    f. *∅-naka-hỹrỹja-t          taso*
      3SG.ABS-DECL-sing-N.FUT man
      'the man sang'

Earlier models of generative syntax such as Government and Binding (GB) made a distinction between morphological case proper and abstract case. For instance, in Haegeman (1991) all of the underlined arguments in (26a-26b) are considered to be licensed by abstract case, even though only those in (26b) feature visible morphological case which corresponds to it. Furthermore, in GB abstract case is considered a universal property of language, occurring in the syntactic component

of every language, while the realization of morphological case is language-specific and is computed by the morphological component of the grammar on a language-by-language basis.[7]

(26) a. [ *that* <u>*the butler*</u> *attacked* <u>*the robber*</u> ] *is surprising*

    b. [ *that* <u>*he*</u> *attacked* <u>*him*</u> ] *is surprising*

For the purposes of this dissertation, *a priori* there is no immediate reason to maintain this theoretical distinction. Moving forward, I will assume that even in languages with no overt morphological case such as Karitiâna, Marantz's algorithm assigns case to eligible arguments in exactly the same way as for all other languages. The morphological exponent of case in these languages is simply null, but otherwise it is considered the 'same' in nature as overt case marking (27).

Karitiâna                                          (Storto 1999; p.157)

(27)  *yn-∅*      *a-ta-oky-j*                *an-∅*
     1SG-ERG 2SG.ABS-DECL-hurt-IRR 2SG-ABS
     'I will hurt you'

Turning to the next theoretical question of importance in this section, where exactly does the computation of case and agreement take place? Clearly, Marantz (1991) considers the computation of morphological case a post-syntactic phenomenon, arguing for its elimination from the theory of syntax. Bobaljik (2008) also explicitly endorses this position, arguing that if the computation of case is post-syntactic, and if it feeds the computation of agreement, then logically agreement itself must also be a post-syntactic process.

---

[7] For more recent overviews and discussion on the nature of abstract and morphological case, see Butt (2006), Bobaljik & Wurmbrand (2008), Legate (2008), Markman (2010), Pesetsky & Torrego (2011) and Baker (2015).

However, some more recent work in the generative tradition indicates that case and agreement may occur in the syntax after all. Preminger (2014) argues, based on A-movement facts from a variety of languages, that case and agreement must take place in the syntax. For instance, in French, dative arguments cannot undergo A-movement to the subject position (28a), but in Icelandic, they can (29a). This implies that since A-movement is a syntactic operation, and morphological case feeds A-movement, then (in these languages at least) the computation of case must take place in the syntax too.

French                                                    (McGinnis 1998; p.90 / Bošković 2007; p.603)

(28) a. * *à  Marie$_i$  semble* [ __$_i$ *Jean avoir    du    talent* ]
          to Marie   seem          Jean have.inf some talent
          'To Marie Jean seems to have talent'

     b.   *il    semble* [ *au    général être    arrivé    deux soldats en ville* ]
          EXPL seem       to.the general be.INF arrive.PART two  soldiers in  town
          'there seem to the general to have arrived two soldiers in town'

Icelandic                                                 (Holmberg & Hróarsdóttir 2003; p.998, p.998)

(29) a. *mér$_i$    virðast* [ __$_i$ *hestarnir      vera seinir* ]
          me.DAT seem.PL        horses.the.PL.NOM be   slow
          'it seems to me that the horses are slow'

     b. *það    virðist* [ *einhverjum manni    hestarnir      vera seinir* ]
          EXPL seem      some.DAT   man.DAT horse.the.PL.NOM be   slow
          'a man finds the horses slow'

Baker & Vinokurova (2010) argue that it is possible for a single language to feature both syntactic and post-syntactic morphological processes of case assignment.

In their view, Sakha accusative and dative case are assigned post-syntactically (30a), but nominative case is assigned at an earlier derivational stage, in the syntax proper (30b).

Sakha (Baker & Vinokurova 2010; p.4, p.4)

(30) a. *Masha Misha-qa    at-y        bier-de*
        Masha Misha-DAT horse-ACC give-3SG.SUBJ.PAST
        'Masha gave Misha a horse'

    b. *min        kel-li-m*
        1SG.NOM come-PAST-1SG.SUBJ
        'I came'

Given that the exact location of morphological case and agreement in the architecture of grammar remains an open theoretical question, the combined model endorsed in this chapter is agnostic in this regard. The model is compatible with either view: either agreement and case are computed post-syntactically (31a) or in the syntax itself (31b). The important thing is the derivational feeding relationship between case and agreement which successfully captures the important typological generalizations of Sections 4.1.1-4.1.3.

(31) *The exact location of case and agreement in the grammar - two possibilities*



204
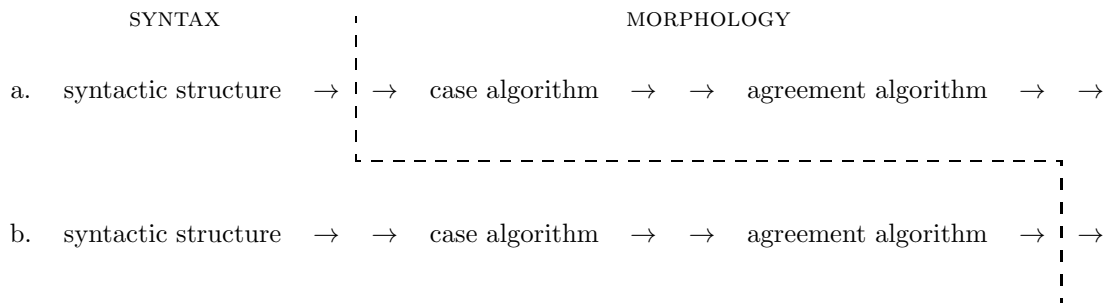
Therefore, if we indeed assume that – regardless of where in the grammar the computation happens – languages such as Karitiâna still receive case via the same algorithm described in Section 4.1.1, deriving sentences such as (27) becomes straightforward. In (32), morphological case is computed as normal. First, dependent ergative case is assigned given that the clausal requirements are met, then default absolutive case is assigned to all remaining arguments. Remember that in this language both ergative and absolutive morphology are simply null. The output of the case-assignment algorithm can then serve as the input to the agreement algorithm as usual. The fact that the NPs in the clause have been assigned morphological case, albeit null phonologically, gives the agreement algorithm something to interpret, and thus allows it to successfully complete the derivation. The highest argument in the domain is presumably ineligible for agreement due to being marked with dependent case, like in Hindi, and thus the algorithm scans down to the next eligible target. The V+I complex ultimately agrees with the only remaining argument, the transitive object (verb movement up to I not pictured below).

(32)

$$\Downarrow$$

```
                        IP
            ┌───────────┴───────────┐
           NP                       ┌──────┴──────┐
        ┌──┴──┐                     I            VP
       yxja-∅                       │        ┌────┴────┐
      1PL-ERG                       ∅-       V        NP
                                 3SG.ABS     │      ┌──┴──┐
                                          na-ahee-t  iso-∅
                                       DECL-blow-N.FUT fire-ABS
```
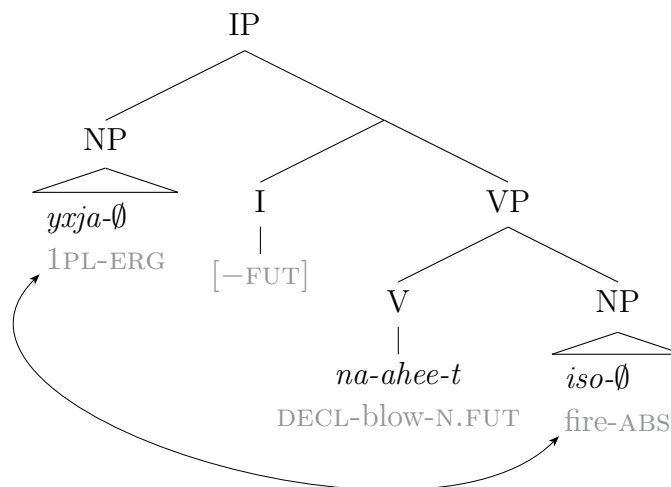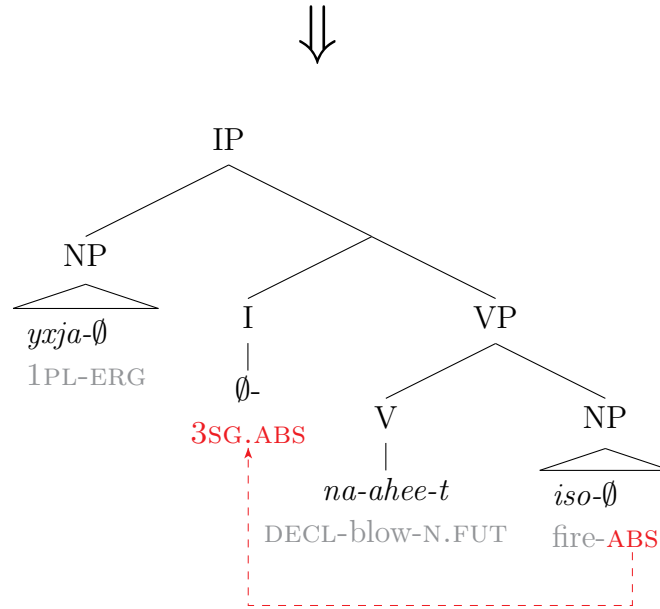
Note in passing that under our system, there is a theoretical precedent for morphologically null case serving as the input for the agreement algorithm. For example, in our analysis of Hindi, null absolutive case serves as the input for absolutive agreement, even though ergative case is overt (7-8). In Aguaruna, null nominative case serves as the input for nominative agreement, even though it is accusative case which is overt (23). In Burushaski, null absolutive case serves as input for either object (20) or subject agreement (21), even though ergative case is overt...and so on. Therefore, if the line of analysis we are pursuing here is on the right track, and morphological case really does feed the computation of agreement, then it follows that if the grammar allows configurations like Hindi where it is the null morphological case which feeds agreement, then it should also allow configurations where both instances of morphological case are null, and which, in turn, feed agreement.

(33)

```
                              CP
                     ┌────────┴────────┐
                     C                 IP
                     │          ┌───────┴───────┐
              Ø-naka-hỹrỹja-t   NP          ┌───┴───┐
              DECL-sing-N.FUT   △           I       VP
                             taso-Ø         │       │
                             man-ABS        t       V
                                                    │
                                                    t
```

$$\Downarrow$$

```
                              CP
                     ┌────────┴────────┐
                  C+I+V               IP
                     │          ┌───────┴───────┐
              Ø-naka-hỹrỹja-t   NP          ┌───┴───┐
         3SG.ABS-DECL-sing-N.FUT △          I       VP
                             taso-Ø         │       │
                             man-ABS        t       V
                                                    │
                                                    t
```
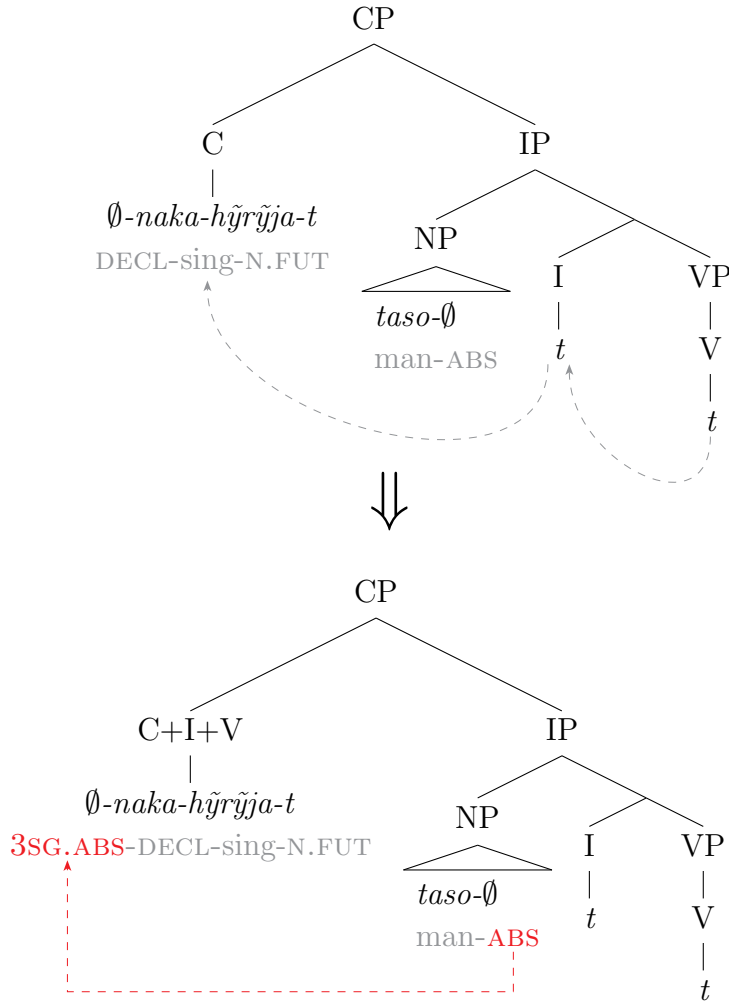
The derivation of an intransitive sentence in Karitiâna will likewise proceed as usual. In (33), the requirements for dependent case are not met, and so the intransitive subject receives (null) absolutive case. Then, the agreement algorithm can interpret this null absolutive case to form an agreement relationship with the verb. In sum, aside from one assumption – that null morphological case can serve as the input to the agreement algorithm – the derivations in (32-33) are in accordance

with the rest of the model as outlined in Sections 4.1.1-4.1.3.[8]

Fortunately, this one assumption also allows the general model to account for double agreement languages with no overt morphological case, such as Kaqchikel. In (34), morphological case in a transitive clause is computed exactly like in Karitiâna, as long the requirements are met. Then, despite being null, this morphological case can serve as an interpretable input for the computation of agreement, yielding a pattern of ergative double agreement on the surface.

Kaqchikel

(34) x-        a-          ru-         tz'ët  rat   ri     tijonel
    PERF  2SG.ABS  3SG.ERG  see    2SG  DET  teacher

'the teacher saw you'



---
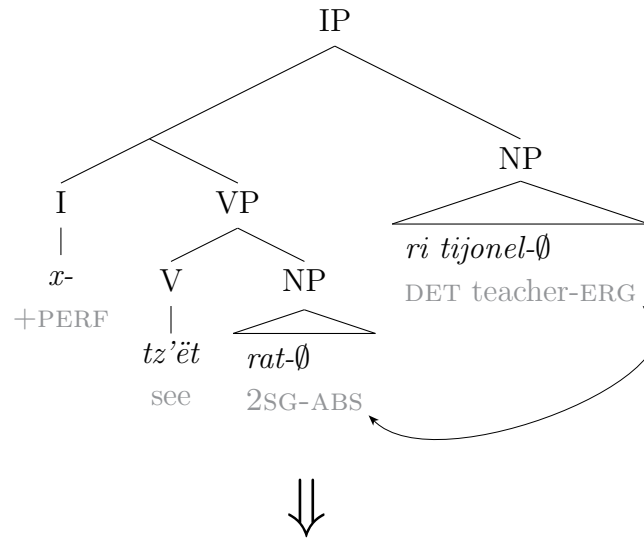
[8] In order to derive V S surface word order in intransitive clauses, the V+I complex is assumed to undergo additional head-movement to C in (33).

IP
├─ I
│  ├─ x-   a-
│  │ +PERF 2SG.ABS
│  └─ V — ru-tz'ët (3SG.ERG-see)
├─ VP
│  ├─ V — ru-tz'ët (3SG.ERG-see)
│  └─ NP — rat-∅ (2SG-ABS)
└─ NP — *ri tijonel-∅* (DET teacher-ERG)
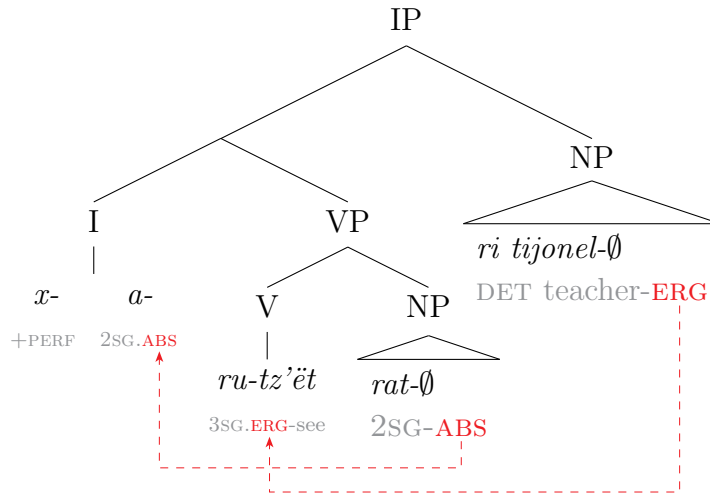
In Section 4.1.3, Baker's (2013) theory of double agreement captured the fact that in some languages, object agreement is still available in non-finite environments, but in other languages, it is altogether absent. In languages like Mapudungun, nominative agreement is located in the I head, while accusative agreement is located in the [+TRANS] v head (see 15 above). And yet, Baker's theory did not explicitly have anything to say about <u>ergative</u> agreement in non-finite clauses. Assuming that the same two locii are available in languages with ergative-absolutive agreement – I and [+TRANS] v – which is associated with which argument?

In Kaqchikel, non-finite data indicates that it is ergative agreement which is realized in [+TRANS] v, and thus still available in those contexts. In (35b), absolutive agreement cannot be realized in the embedded non-finite clause (it would otherwise be 3PL.ABS *e'-*, compare with 35a), but the ergative agreement morpheme *a-* that corresponds to the 2SG transitive subject still occurs.

Kaqchikel

(35) a. *x-e'-a-k'ul*       *ri*    *ak'wala rat*
       PERF-3PL.ABS-2SG.ERG-meet DET child.PL 2SG

       'you met the children'

   b. *x-∅-a-chäp*      [   *a-k'ul-ïk*      *kik'in ri*   *ak'wala* ]
       PERF-3SG.ABS-2SG.ERG-grab    2SG.ERG-meet-NOML REL    DET child.PL

       'you began your meeting with the children'

Generalizing from this fact of Kaqchikel, I will assume the following as an extension of Baker's (2013) theory of double agreement and split locii: in ergative double agreement systems, either a) the I head agrees twice, simultaneously with two different NPs or b) I agrees once with the absolutive NP, and v [+TRANS] agrees once with the ergative NP.[9] The proposed typology I'll be assuming, an extension from Baker (2013), is in (36).

(36) *Extended typology of split locii in double agreement*

| | type | locus of agreement | | example |
|---|---|---|---|---|
| NOM–ACC | 1 | I = NOM$_{AGR}$, ACC$_{AGR}$ | | Nez Perce |
| | 2 | I = NOM$_{AGR}$ | v$_{[+TRANS]}$ = ACC$_{AGR}$ | Mapudungun |
| ERG–ABS | 3 | I = ABS$_{AGR}$, ERG$_{AGR}$ | | Paumarí |
| | 4 | I = ABS$_{AGR}$ | v$_{[+TRANS]}$ = ERG$_{AGR}$ | Kaqchikel |

---

[9] See Coon (to appear) for a similar theory of 'low' ergative agreement in the v head for Chol, another Mayan language. The assumption that ergative agreement is localized in v is also consistent with Legate (2008), who found that in non-finite clauses it tends to be absolutive case which becomes unavailable.

Finally, note that in addition to being able to account for more cross-linguistically common patterns of agreement such as absolutive only (Hindi, Karitiâna) and ergative and absolutive agreement (Paumarí, Kaqchikel), the model can also account for rarer patterns that feature ergative agreement only.[10] For instance, in Halkomelem, third person transitive subjects are indicated with the suffix -es, but neither transitive objects or intransitive objects are indicated by head-marking (37).

Halkomelem                                                          (Wiltschko 2006; p.197, p.197)

(37) a. *q'ó:y-t-es*      *te*  *Strang te*  *sqelá:w*
         kill-TR-3SG.ERG DET Strang DET beaver
      'Strang killed the beaver'

     b. *í:mex*   *te*   *Strang*
         walking DET Strang
      'Strang is walking'

Given that in Section 4.1.3, absolutive agreement is taken to always take place prior to ergative agreement, the logical conclusion for a language like Halkomelem is to assume that in those rare cases, the absolutive agreement paradigm is uniformly null in that language. After null absolutive agreement has first been computed, then overt ergative agreement can take place, technically making Halkomelem a 'two-agreement' language as well.

Indeed, it is not at all uncommon for third person absolutive agreement to be null, as in Kaqchikel. In the derivation of a clause such as (38), the model will first produce null 3SG.ABS agreement, and then overt 2SG.ERG agreement. Therefore

---

[10]  In the sample of 40 ergative languages from Chapter 2, only two were found to exhibit consistently overt ergative agreement and consistently absent absolutive agreement: Halkomelem and Päri.

it stands to reason that if some languages allow absolutive agreement to be null for a subset of persons, perhaps some languages also feature absolutive agreement which is null across the entire paradigm. Though it involves an increased degree of abstraction, this key assumption allows the model to account for ergative-agreement-only languages such as Halkomelem while still ruling out the typological gap from Bobaljik (2008).[11]

Kaqchikel

(38) x-      ∅-        a-        tz'ët ri   tijonel rat
     PERF 3SG.ABS 2SG.ERG see   DET teacher 2SG
     'you saw the teacher'

### 4.1.5   Interim conclusion

In Sections 4.1.1-4.1.4, I have provided an outline of the general model of case and agreement which I'll be assuming moving forward. Its main strengths – and therefore the main motivations for its adoption independently of syncretic patterns – are listed in (39). In the next section, this model is evaluated in light of the syncretic data from Chapters 2 and 3, to get a sense of whether or not it can adequately account for that data, too.

---

[11] Alternatively, see Coon (2017) for a different account of 'preferential targeting' of ergative subjects. In her account, there are actually two different ways to derive verb agreement which is in ergative-absolutive alignment. Languages such as Hindi are derived exactly like in Bobaljik (2008); however, purely head-marking languages such as Halkomelem feature inherent ergative case assigned by the [+TRANS] $v$ head, which in turn enters into a local agreement relationship with the transitive subject generated in Spec-vP. In this view, languages such as Halkomelem preferentially target transitive subjects by skipping the assignment of absolutive agreement altogether.

(39) *The main strengths of the Marantz-Bobaljik-Baker model of case and agreement*

| theory | theoretical contribution | typological contribution |
|---|---|---|
| Marantz (1991) | ✔ case computing algorithm | ✔ dependent ergative case is not always an external argument |
| Bobaljik (2008) | ✔ agreement computing algorithm | ✔ predicts typological gap in (12) |
| Baker (2013) | ✔ theory of double agreement | ✔ parameterization of locii of NOM-ACC agreement |
| Section 4.1.4 | ✔ theory of agreement in languages lacking overt morphological case | ✔ parameterization of locii of ERG-ABS agreement |

## 4.2 Theories of syncretism

### 4.2.1 Baker (2015), ERG=GEN & ERG=POSS

To account for ERG=GEN data such as in Shipibo, Baker (2015) proposes to extend Marantz-style theories of dependent case to the nominal domain. In (40a), the transitive subject is marked with the ergative marker *-kan* in the clausal domain; in (40b), the possessor is marked with the genitive marker *-kan* in the nominal domain.

Shipibo                                                         (Baker 2015; p.166, p.166)

(40) a. *Jose-kan ochiti ben-ai*

  José-ERG dog seek-IMPF

  'José is looking for a/the dog'

  b. *Jose-kan ochiti*

  José-GEN dog

  'José's dog'

The formalism he proposes is in (41).

(41) Suppose that X bears index [i], no other phrase Z properly contains X and also bears [i], and there is a term Y that bears a distinct index [k] in the same spell out domain as X.

   a. Then if X c-commands Y, assign X ergative.
   b. Then if X is c-commanded by Y, assign X accusative.

In most languages with dependent case, the rule in (41) will only be specified for the clausal domain. However, as a point of parameterization, languages like Shipibo will apply the rule in both the clausal and the nominal domain. In both cases, the formalism will produce the correct ERG/GEN case on the transitive subject/possessor. In (42a), $DP_k$ is not properly contained in another phrase bearing that same index, and furthermore within that domain it c-commends $DP_i$, which bears a different index. Therefore, according to (41) it is eligible to receive ergative case. In (42b), the same requirements are met by $DP_k$ there: it is not properly contained in another phrase bearing that same index $k$, and within that domain it c-commends $NP_i$, which bears a different index. Therefore, by that same formalism it is eligible to receive 'ergative' (i.e. genitive) case.

Note that some details in (42a) are different from the original computation of dependent case in Section 4.1.1: the structure has been enriched with functional projections such as DP and vP, and the transitive subject is base-generated in Spec-vP. Crucially, however, the basic mechanics of dependent case assignment remain the same, and furthermore the single formalism in (41) applies equally-well to either the clausal or nominal domain.

(42) a.

```
                    IP
                   /  \
                  vP   I
                 /  \  |
               DP_k  \ -ai
              /\     \ IMPF
          Jose-kan    \
          José-ERG     VP   v
                      /  \  [+TRANS]
                    DP_i  V
                   /\     |
               ochiti    ben
               dog.ABS   seek
```

b.

```
                    DP_i
                   /   \
                 nP_i   D
                /   \
             DP_k   NP_i   n
            /\       |
        Jose-kan    N_i
        José-GEN     |
                   ochiti
                   dog.'ABS'
```
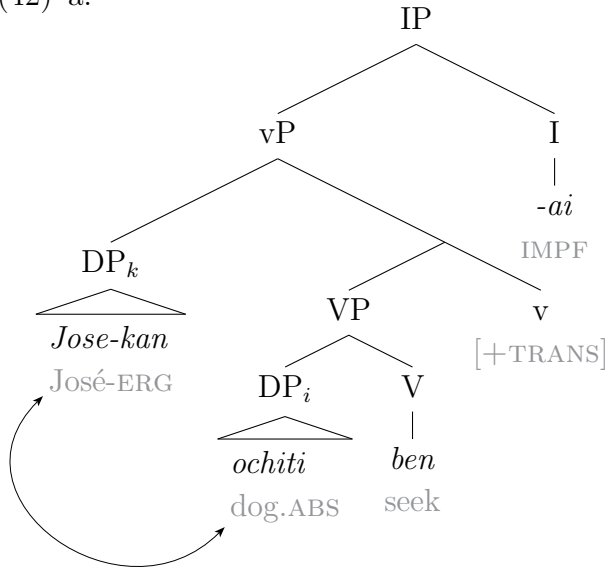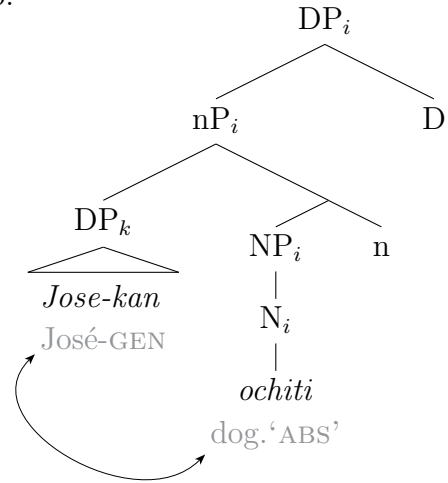
In this way, Baker's (2015) formalism can successfully account for ERG=GEN patterns observed in Chapter 2. What about ERG=POSS patterns, then? Baker doesn't explicitly have anything to say about syncretic patterns involving head-marking. But under the assumption that case feeds the computation of agreement, it follows that if dependent case can be generated in the nominal domain as in (42b), presumably in some languages this dependent case can feed the computation of agreement in the nominal domain as well. In other words, some languages will only compute agreement within the clausal domain. However, as another point of parameterization in the world's languages, some languages will likewise extend the agreement computation algorithm to both the clausal and nominal domains, producing an ERG=POSS pattern.

As a reminder, descriptively in Kaqchikel the same agreement paradigm encodes both transitive subjects (43a) and possessors (43b). In accordance with the general model of case and agreement in (1), the computation of head-marking in

the nominal domain in Kaqchikel will therefore proceed isomorphically to the clausal domain. First, dependent case is computed in the nominal domain, exactly like in Shipibo and according to Baker's (2015) formalism. Because the possessor $DP_k$ is not contained within a clause that has the same index, but within that domain it c-commands another phrase with a distinct index, $NP_i$, it is eligible to receive ergative case. Then, the output of the case algorithm serves as the input for the agreement algorithm. The agreement algorithm scans the domain, looking for the highest eligible DP. Assuming that dependent-marked arguments are accessible, an agreement relationship forms between the genitive-marked DP and the n head, the nominal-domain analog to the v head in clauses.

Kaqchikel

(43) a. x-      i-        ru-       tijo   -j  rin  rija
        PERF 1SG.ABS 3SG.ERG teach TR 1SG  3SG
        'she taught me'

     b. ri    ru-        wakx  rija
        DET 3SG.POSS cow   3SG
        'her cow'

(44)

$$\Downarrow$$

```
                        DP_i
                       /    \
                      /      \
                    D          nP_i
                    |         /    \
                    ri       /      \
                    DET     /        DP_k
                           /        /‾‾‾\
                          n   NP_i  rija-∅
                          |    |    3SG-GEN
                          ru-  N_i
                       3SG.POSS  |
                               wakx-∅
                              cow-'ABS'
```
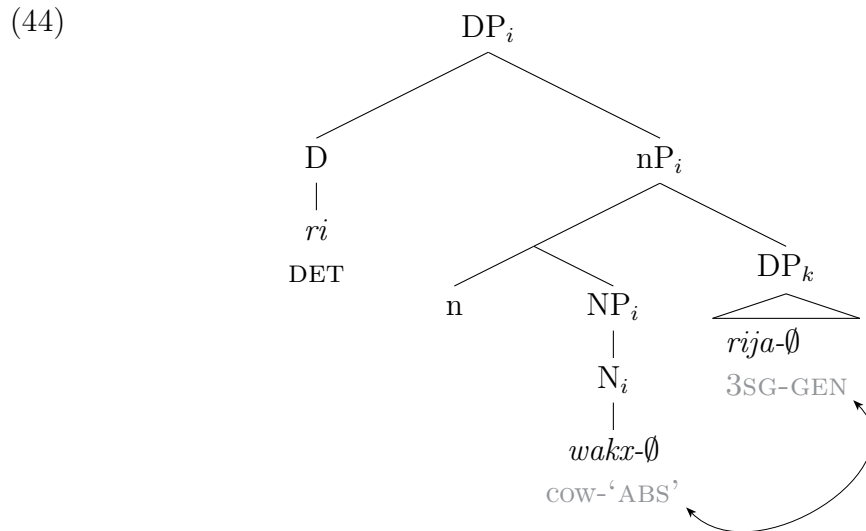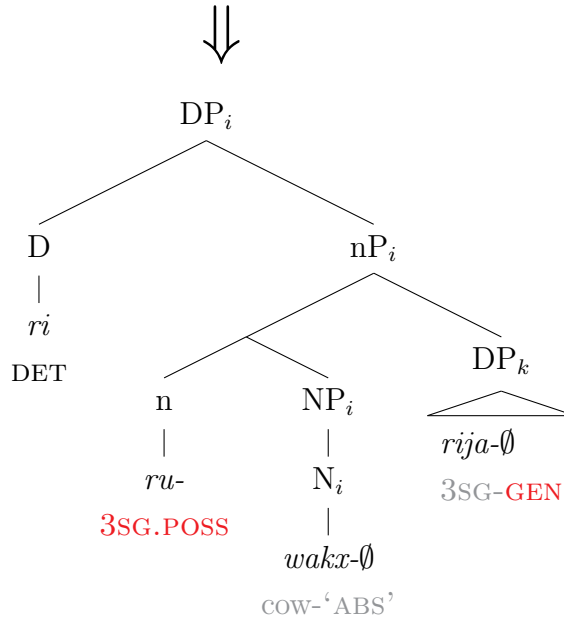
Some lingering questions remain regarding the derivation of nominal agreement in (44). For instance, is default 'absolutive' case assigned in the nominal domain, on analogy with the clausal domain? Baker (2015) doesn't address the issue, and so for now we can leave the question open to further research. In (44), the possessed noun *wakx* 'cow' is modeled as indeed receiving null, default case, but it shouldn't affect the derivation either way.

The reason it shouldn't affect the derivation either way is the following: by definition, nominals lack the IP functional projection, which occurs solely in finite clauses. This explains why, for instance, 'absolutive' agreement is not realized in (44), despite the fact that normally Kaqchikel is a double-agreement language. In this respect, Kaqchikel really is the ergative-absolutive counterpart to Mapudungun from Section 4.1.3: in non-finite environments, agreement from the I head is simply unavailable. Therefore even if *wakx* 'cow' does receive some version of 'absolutive'

case in (44), it doesn't matter because the relevant functional projection for interpreting that input and converting it to agreement morphology is unavailable in the first place.

The derivation of nominal agreement as modeled in (44) rests on another crucial assumption: that n is indeed the nominal counterpart to v from the clausal domain. Recall from Section 4.1.4 that in split-locii languages with ergative double agreement, v was assumed to be the locus of ergative agreement, based on nonfinite evidence from Kaqchikel such as (35). In (44) the assumption is naturally extended to the nominal domain: namely that the locus of possessor agreement in Kaqchikel nominals is n, the structurally-analogous functional projection that selects a lexical phrase in its complement and introduces a structurally higher argument in its specifier.

In this view, possessive agreement in the nominal domain in (44) is very similar to clausal 'single agreement', like in Hindi (8). Two morphological cases, one dependent and one default, are computed within the domain, but ultimately there is only one 'slot' for agreement. In Hindi, that one slot encodes the absolutive argument and is located in the I head; however, in Kaqchikel possessed nominals, the only available locus for agreement is n, and by hypothesis only ergative agreement can be assigned there.

To sum up this section, Baker (2015) provides a straightforward account of ERG= GEN patterns of syncretism such as in Shipibo. Building on his intuition that, in some languages, the computation dependent case is extended to the nominal domain, the general model of case and agreement from Section 4.1 can also account for patterns of ERG=POSS syncretism, such as in Kaqchikel. This requires a few additional assumptions regarding some of the finer details of the derivation, but ultimately the model works. In the next section, I will analyze whether the current

model of case and agreement can also account for rarer, but still attested patterns of syncretism involving accusative case and accusative agreement.

### 4.2.2  Can the current model account for ACC=GEN and ACC=POSS?

Baker's (2015) analysis was designed to cover more than just ERG=GEN patterns of Shipibo. Indeed, Baker deliberately crafted his analysis of dependent case in nominal constructions to allow ERG=GEN patterns, but also to disallow ACC=GEN patterns. The formalism in (41) clearly states that in order to receive accusative case, a phrase must not be contained within a larger phrase that bears the same index, but also it must be c-commanded by another phrase bearing a different index within that domain. In (45), the possessor $DP_k$ is indeed not contained within a larger phrase bearing that same index $k$, but it doesn't meet the second requirement: it is not c-commanded by another phrase in that domain bearing a different index. Since none of $DP_i$, $nP_i$, $NP_i$ or $N_i$ c-command $DP_k$, according to Baker possessors will never receive accusative dependent case within the nominal domain.

(45)
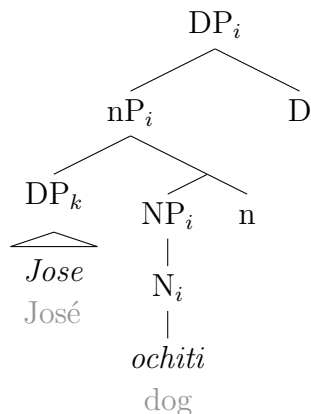


If the computation of agreement is predicated on the computation of case – as we have been assuming throughout this chapter – then ACC=POSS patterns

shouldn't occur under any circumstances either. The survey of accusative languages in Chapter 3 provides clear counter-examples to this prediction. Among those 40 languages, nine of them exhibited some form of systematic ACC=POSS syncretism, including Lango (46), Jola Bandial (47), Lower Grand Valley Dani (48), Indonesian (49) and Warao (50).

Lango                                                        (Noonan 1992; p.35, p.31, p.78)

(46) a. *ò-      nὲn -á*
       3SG.SUBJ see  1SG.OBJ
       'he saw me'

   b. *ò-      ɔɲ*
       3SG.SUBJ spill
       'it spilled'

   c. *tyἔn -á*
       leg   1SG.POSS
       'my leg'

Jola Bandial                                                (Bassène 2007; p.92, p.49, p.54)

(47) a. *Atejo na-sen-om           si-rálam*
       Atejo 3SG.SUBJ-give-1SG.OBJ 4-money
       'Atejo gave me the money'

   b. *a-vví  aku   fíllim   na-cel-e*
       1-king 1.DEM last.year 3SG.SUBJ-die-TAM
       'the king died last year'

   c. *ga-aɲen-om*
       9-POST-hand-1SG.POSS
       'my hand'

Lower Grand Valley Dani                          (Bromley 1981; p.191, p.130, p.190)

(48) a. *h-ath-e*
       2SG.OBJ-hit-3SG.SUBJ

       'he hit you'

    b. *ekken hulelh-e*
       seed    sprout-3SG.SUBJ

       'the seed sprouted'

    c. *h-oppase*
       2SG.POSS-father

       'your father'

Indonesian                                   (Sneddon 1996; p.170, p.134, p.171)

(49) a. *Narti me-nunggu-ku*
       Narti *meN*-wait.for.AV-1SG.OBJ

       'Narti is waiting for me'

    b. *saya tinggal di rumah ini*
       1SG   live.OV   P   house   DET

       'I live in this house'

    c. *rumah-ku*
       house-1SG.POSS

       'my house'

Warao                                          (Romero-Figeroa 1997; p.23, p.65, p.5)

(50) a. *sina ma-yehe-bu-te-ra*
       who 1SG.OBJ-call-ITER-N.PAST-INT

       'who calls me repeatedly?'

b. *Wauta omi naru-ki-tia-ine*
   Wauta PRIV go-INTENT-HAB-1SG.SUBJ

   'I am going to see Wauta as usual'

c. *ma-rahe haya-te*

   1SG.POSS-brother run-N.PAST

   'my brother runs'

So the fact that ACC=POSS patterns are well-attested is potentially problematic, given that accusative case is assumed to feed accusative agreement, but accusative case is predicted by Baker to be categorically unavailable on possessors. And yet an even more direct challenge to Baker's (2015) formalism in (41) is that ACC=GEN patterns themselves are actually attested as well. Looking again at the sample from Chapter 3, the pattern is rarer, but it is nevertheless still attested in three languages: Aguaruna (51), Comanche (52) and Northern Saami (53).

Aguaruna                                              (Overall 2007; p.260, p.506, p.130)

(51) a. *ami-na apahui tuki puhu-wa=nu yaĩ-pa-ka-ti*
       2SG-ACC god always live-3SG.SUBJ=ANA.REL help-2SG.OBJ-INTS-JUSS


       'may God, who lives forever, help you'

b. *amɨ wɨ-tʃau-aita-ku-mɨ-ĩ-ka* *ʃiiha*
   2SG go.PERF-NEG.REL-COP-SIM-2SG.SUBJ-DS-COND well
   *anɨa-sa-nu* *puhu-mai-inu-aita-ha-i*
   be.happy-SBD-1.SUBJ.SS live-POT-NR-COP-1SG.SUBJ-DECL

   'if you had not gone, I would be happy'

c. *ami-na apa*
   2SG-GEN father

   'your father'

Comanche                                                   (Charney 1993; p.202, p.94, p.56)

(52) a. *tena-pɨ-tsa*      *waʔi-pɨ-ʔa*       *puni*
        man-ABS-TOP woman-ABS-ACC see

        'the man sees the woman'

    b. *ɨ-tsaa-tɨi-tsa*          *miʔa-tɨ*
        your-good-friend-TOP go-GEN:ASP

        'your good friend is leaving'

    c. *Mia-ʔa-tsa*       *satiʔi*
        Mia-GEN-TOP dog

        'Mia's dog'

Northern Saami                                            (Wilbur 2014; p.214, p.244, p.89)

(53) a. *båtsoj*   *máhtta*        *duv*     *nala*  *båhte-t*
        reindeer can.3SG.PRES 2SG.GEN upon come-INF

        'the reinder can attack you' (lit. 'come upon you')[12]

    b. *suovade*          *dån?*
        smoke.2SG.PRES 2SG.NOM

        'do you smoke?'

    c. *men*  *ådtjo*       *sáme*          *giela-v*       *ságasti-t*  *duv*
        but  may.2SG.PAST Saami.GEN.SG language-ACC.SG speak-INF 2SG.GEN
        *åbeni-j*
        sibling-COM.PL

        'but were you allowed to speak the Saami language with your siblings?'

---

[12] As discussed in Chapter 3, ACC=GEN overlap in Northern Saami is limited to singular pronouns, and the datapoint in (53c) is the best example provided by Wilbur (2014) of a singular pronoun being used as a transitive object. See Table 6.1 p.114, which makes it clear that the ACC and GEN forms of singular pronouns do indeed overlap in Northern Saami.

In a footnote, Baker does address the fact that some Uto-Aztecan languages feature what on the surface looks like an ACC=GEN pattern, which is predicted never to occur. His claim is that these languages have an extremely impoverished case system, and thus it is not exactly accusative which is being generated on possessors, but a catch-all 'oblique' form for all non-nominative nouns. This certainly is true for Comanche: the same case-marking morphology *-ʔa* encodes possessors, direct objects and indirect objects in that language, and there are no other overt case-markers in the language (Charney 1983). However, the same cannot be said for Aguaruna. In addition to a syncretic ACC=GEN marker *-na*, Aguaruna also features comitative *-haĩ*, locative *-numa/-nĩ*, instrumental *-(a)i*, and ablative *-ia* (Overall 2007). Therefore, in Aguaruna at least, it cannot be the case that what looks like ACC=GEN is actually some kind of more general syncretism like OBL=GEN. Rather, it is specifically accusative which is being generated on transitive objects the verbal domain and on possessors in the nominal domain.

To summarize, the weight of the evidence in (46-53) indicates that, contrary to Baker's theory, possessors can be marked with accusative dependent case within the nominal domain, and furthermore that this dependent accusative case can in turn feed the computation of agreement. In other words, insofar as dependent case (and by extension agreement) can take place in nominal constructions, there is no 'built-in' asymmetry which allows dependent ergative case on possessors, but which definitely rules out dependent accusative case on possessors.[13]

### 4.2.3 Can the current model account for NOM=POSS and ABS=POSS?

NOM=POSS is another rare, but nevertheless attested syncretic pattern from Chapters 2 and 3. In the sample of 40 accusative languages, NOM=POSS occurs a

---

[13] See Section 4.3 for an alternative account for these ACC=GEN and ACC=POSS facts from Chapter 3.

total of five times, in languages such as Maybrat (54), Hungarian (55) and Maricopa (56).

Maybrat                                                  (Dol 2007; p.77, p.51, p.75, p.63)

(54) a. *y-po          ku    kiniah*
        3SG.M.SUBJ-hold child small
        'he holds the small child'

   b. *y-ros*
        3SG.M.SUBJ-stand
        'he stands'

   c. *m-ape              ait*
        3SG.U.SUBJ-carry.on.back 3SG.M
        'she carries him on her back'

   d. *y-ana*
        3SG.M-POSS-head
        'his head'

Hungarian                                (Kenesei et al 1998; p.195, p.231, p.195, p.209)

(55) a. *a   lány ír-ja          a   level-et*
        the girl  write-3SG.SUBJ.DEF the letter-ACC
        'the girl is writing the letter'

   b. *a   lány áll-∅*
        the girl   stand-3SG.SUBJ.INDEF
        'the girl is standing'[14]

---

[14] The nominative agreement morphology that encodes transitive subjects in Hungarian comes in two varieties, depending on whether the transitive object is definite or indefinite. Possessive agreement in Hungarian overlaps with different portions of the definite and indefinite nominative paradigms. See Chapter 3 for more details.

c. *ez   az   óra    tíz dollár-t     ér-∅*
   this the watch ten dollar-ACC be.worth-3SG.SUBJ.INDEF

   'this watch is worth ten dollars'

d. *Pál   hibá-ja          tudatlanság-ból fakad-t*
   Paul mistake-**3SG.POSS** ignorance-ELA   stem-PAST.3SG.SUBJ.INDEF

   'Paul's mistake stemmed from his ignorance'

Maricopa                                    (Gordon 1986; p.18, p.19, p.17, p.31)

(56) a. *m-wik-k*
        **2.SUBJ**-help-REAL

        'you helped him'

   b. *m-ashvar-k*
        **2.SUB**J-sing-REAL

        'you sang'

   c. *'-wik-k*
        1.SUBJ-help-REAL

        'I helped him'

   d. *m-mpur*
        **2.POSS**-hat

        'your hat'

It is not immediately clear how to model NOM=POSS in Baker (2015) or within the greater framework we have been assuming. Under the assumption that case always feeds agreement, in principle we would want to generate nominative case on the possessor, in order to feed a later computation of 'nominative' possessor agreement. The most obvious way to accomplish this is to first assign dependent accusative case to the possessee, and then default nominative case on the possessor. However, Baker (2015) is very clear that not only should possessors never be marked

with accusative case, possessees too should never be eligible either. To rule this second option out, his dependent case formalism (repeated in 57), contains a tiny but crucial asymmetry in how the requirements for ergative and accusative case are met.

(57) Suppose that X bears index [i], no other phrase Z properly contains X and also bears [i], and there is a term Y that bears a distinct index [k] in the same spell out domain as X.

    a. Then if X c-commands Y, assign X ergative.

    b. Then if X is c-commanded by Y, assign X accusative.

(58)
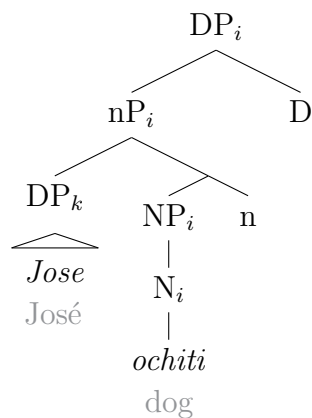


As stated, the requirements for ergative dependent case in the nominal domain are only that the possessor *Jose* in (58) c-commands <u>any</u> phrase bearing the index $i$ which corresponds to the possessee, *ochiti* 'dog'. Given that $DP_k$ c-commands $NP_i$, ergative (=genitive) dependent case may be assigned to the possessor. However, when it comes to assigning dependent accusative case to the possessee, neither $NP_i$ nor $DP_i$ meet the requirements in (57): the former because it is properly contained within another phrase which bears the same index (and therefore not "maximal" to

use Baker's terminology), and the latter because it is simply not c-commanded by the possessor.

So Baker's formalism rules out the possibility that dependent accusative case could be assigned to a possessee, after which an agreement relationship could form with a nominative-marked possessor within the nominal domain. Another possibility would be as follows: if the requirements for accusative case will never be met in a nominal construction, as per Marantz's (1991) disjunctive hierarchy in (3), default nominative case should be assigned to <u>both</u> the possessor and possessee within that domain. However, if Baker's approach is correct, and the possessee's index percolates to the very highest projection $DP_i$ in (58), the highest eligible nominative NP will always be the possessee, not the possessor that the agreement algorithm is trying to form an agreement relationship with.

One more potential problem for any theory of NOM=POSS involves the locus of agreement itself. There is near-universal agreement in the literature that nominative agreement is realized in the I head, and that IP is a functional projection which is unique to finite clauses (see among others Legate 2008, Bobaljik 2008, Baker 2013, Baker 2015). And yet even if nominative agreement is not directly fed by nominative case in a two-step, Bobaljik-style model, it is not at all clear how nominative agreement would even come about in (54-56), given that by definition nominal constructions lack the IP functional projection.

Finally, one last syncretic pattern from the end of Chapter 2 deserves special mention here due to its theoretical implications. Though rare, ABS=POSS is clearly attested in at least two languages of the world, Tenetehára (59) and Tiriyó (60).

Tenetehára                                  (Harrison 1986; p.421/Bendor-Samuel 1972; p.91, p.106)

(59) a. *he-rurywete*
         1SG.ABS-be.happy

         'I am happy'

     b. *he-petek*
         1SG.ABS-beat

         'he beats me'

     c. *he-mukaw*
         1SG.POSS-gun

         'my gun'

Tiriyó                                                (Meira 1999; p.283, p.290, p.201)

(60) a. *ë-emamina*
         2SG.ABS-play

         'you have played'

     b. *ë-eta*
         2SG.ABS-hear

         'it has heard you'

     c. *ë-pata*
         2SG.POSS-village

         'your village'

Like NOM=POSS, ABS=POSS faces the immediate problem of which functional head would serve as the locus of nominal agreement in these languages, given that the I head is also associated with absolutive agreement in the theoretical literature (Legate 2008). It is also unclear how ABS=POSS agreement would actually happen, even if we look past the fact that the I head is unavailable in nominal constructions.

Assuming that agreement is computed using case as its input, deriving ABS= POSS would first require assigning dependent ergative case to the possessor, and then default absolutive case to the possessee, exactly like in Shipibo (61). Then, this morphological case would in principle serve as the input to the agreement algorithm, but here we run into a real problem: the head seeking to form an absolutive agreement relationship (let's call it D for now) is searching for the absolutive-marked argument, but the absolutive-marked argument is the possessee, not the possessor. It wouldn't make sense for possessive agreement to encode the $\phi$-features of the possessee and not the possessor. Thus the derivation runs into a fundamental problem because the agreement algorithm ultimately wants to form an agreement relationship with the possessor, but by hypothesis possessors can only bear the wrong morphological case, ergative.

(61)

DP$_i$

D    nP$_i$

D$_{\text{ABS}}$    D$_\phi$    DP$_k$    n    NP$_i$

*???*    *???*    *Jose*-∅    N$_i$
                  José-**ERG**

*ochiti*-∅
dog-**ABS**
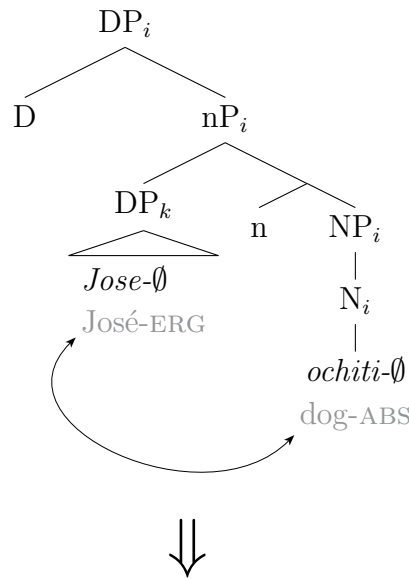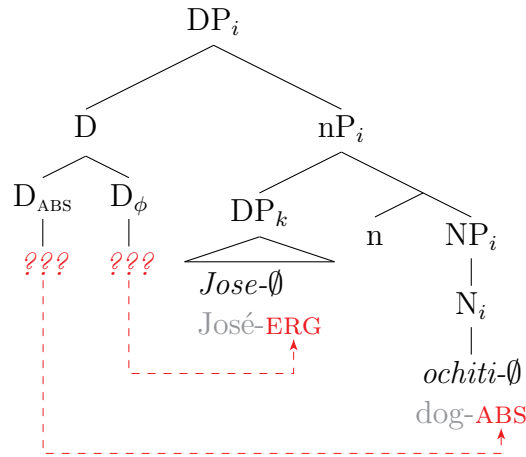
This hypothetical scenario is different in nature from 'asymmetrical' alignments in a language such as Burushaski. In that language, case alignment and agreement alignment do not line up perfectly, but at the very least they share the commonality that intransitive subjects are always 'unmarked': they bear default absolutive case, and they form an agreement relationship with the nominative paradigm. On the other hand, the derivational scenario described for ABS=POSS above requires a complete inversion: the ergative-marked possessor would be encoded by the absolutive agreement paradigm, an incongruous inversion for which there does not seem to be any precedent (62).

(62)  *'Asymmetrical' alignment in Burushaski and inverted alignment in ABS=POSS*

|  | case | agreement |  |  | case | agreement |
|---|---|---|---|---|---|---|
| **trans. subject** | ERG | NOM | | **possessor** | ERG | ABS |
| **intr. subject** | ABS | NOM | | **possessee** | ABS | ERG |
| **trans. object** | ABS | ACC | | | | |

To sum up the discussion in this section, existing theories of case and agreement do not seem compatible with all the syncretic patterns observed in Chapters 2 and 3. Baker (2015) can indeed account for ERG=GEN patterns, and furthermore the feeding model of Bobaljik (2008) can be extended to the nominal domain to produce a straightforward account of ERG=POSS as well. However, beyond this point, we run into some issues.

Baker (2015) makes the strong prediction that ACC=GEN patterns shouldn't exist, and yet they are directly observable in several languages. In fact, Baker's theory predicts that <u>only</u> dependent ergative case may be assigned in the nominal domain (and to the possessor only), which is problematic because under the assumption of the general model of case and agreement in (1), case must feed agreement, and yet ACC=POSS, NOM=POSS and ABS=POSS are all attested. NOM=POSS and ABS=POSS run into the additional problem of lacking the IP functional projection by definition, but requiring the I head for agreement according to a variety of existing theories. Finally, ABS=POSS runs into the most problems within our general model, because it logically forces us to assume an alignment 'inversion' such that an absolutive agreement relationship is formed with the ergative-marked possessor.

So we find ourselves at an impasse. Do we do away with Bobaljik-style theories which connect case and agreement altogether, and lose our account of the typological asymmetry in (12)? Do we do away with that close link, but only in the nominal domain? In the next section, I will argue for a third option. The parameterization of Baker's (2015) ERG=GEN analysis accounts for the very high number of ergative syncretisms cross-linguistically, but other, rarer syncretic patterns are generated by a different process altogether. They are not generated by extending dependent case assignment rules from one domain to another, rather inflectional forms are recycled for simple reasons of economy, a somewhat random process that nevertheless allows

for the existence of these rarer syncretic patterns.

## 4.3 Syncretism as 'recycling'

In this section, I will continue to assume that the general model of case and agreement from Section 4.1 is essentially correct. I will also assume that the analysis in Baker (2015) is at least partially correct: some ergative languages have the option of extending ergative dependent case to the nominal domain. The availability of this parametric option is due to the extremely close parallelism between the clausal and nominal domains in this regard. In (63), both possessors and external arguments are generated in the specifier of vP/nP. Furthermore, they c-command the other argument in the domain, thus meeting the requirements for dependent case according to Baker.
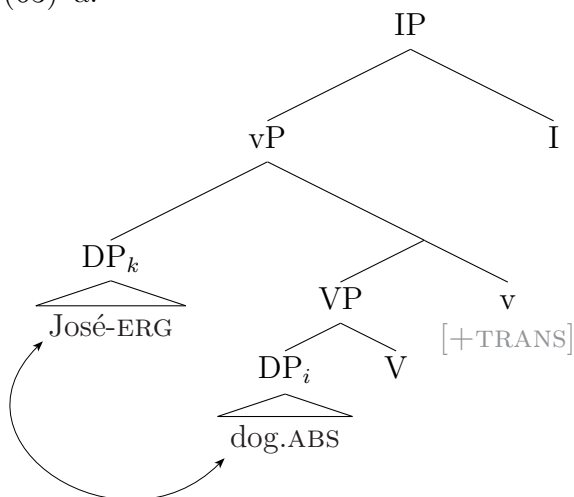
(63) a.                                        b.
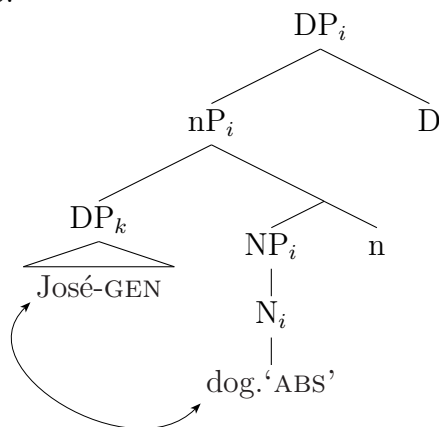


(64) The ERG=GEN parameter: ergative languages may or may not extend the rule that assigns dependent ergative case in the clausal domain to possessors in the nominal domain.

233

Formalizing Baker's ERG=GEN rule as a parameter explains the high rates of syncretism observed in the sample from Chapter 2. Recall that in that sample, approximately half the languages in the survey were found to exhibit some kind of ergative syncretism. Therefore it follows that if all ergative languages have one of two options with regards to (64), approximately half of them will extend the rule, either producing an ERG=GEN pattern or the ERG=POSS pattern that derives from it, and approximately half of them will not.

Baker (2015) makes it clear that the ERG=GEN parameter in (64) is limited to ergative languages. I am going to assume that he is correct in saying that there is no equivalent for absolutive, nominative or accusative case. However, the fact remains that other syncretic patterns such as ACC=GEN, ACC=POSS, NOM=POSS and ABS=POSS are still attested. Therefore it cannot be the case that these patterns should be categorically ruled out. These patterns still ought to be generated in the grammar, but perhaps not via the same mechanisms as ERG=GEN and ERG=POSS.

What happens if an ergative language does not go the ERG=GEN route? How does it select its possessive morphology? What happens in accusative languages, in which the ERG=GEN parameter is not available in the first place? I will assume that one of two options present themselves in these languages: 1) those languages will simply innovate separate possessive forms, or 2) those languages will choose to recycle existing forms for reasons of economy, resulting in an **accidental homophony** that is not structurally-conditioned.

In Maybrat, for example, there is only one paradigm available in the clausal domain: the nominative paradigm that encodes transitive and intransitive subjects on verbs (65). Therefore it makes sense that a language like Maybrat may opt to re-use that existing paradigm to mark possession, thus taking up less valuable space

in the lexicon.[15]

(65) *Person prefixes in Maybrat*[16] (Dol 2007; p.63)

| 1SG | t- | | |
|-----|-----|-----|-----|
| | | 1PL | p- |
| 2SG | n- | | |
| | | 2PL | n- |
| 3M.SG | y- | | |
| | | 3PL | m- |
| 3U.SG | m- | | |

However, if non-ergative syncretisms were merely a matter of random recycling of existing paradigms from other parts of the grammar, we might expect a more even distribution among them. In (66), ergative syncretisms are by far the most common cross-linguistically, a fact that has been captured by the availability of the parameter in (64). However, accusative patterns appear to be the next most common type of pattern, consistently in head-marking and in dependent-marking.

(66) *Summary of syncretic patterns across both the ergative and accusative sample*

| syncretism | head-marking (=POSS) | | dependent-marking (=GEN) | | total | |
|------------|------|---------|------|--------|-------|---------|
| ERG | 13/18 | (72.2%) | 12/30 | (40%) | 23/40 | (57.5%) |
| ACC | 9/32 | (28.1%) | 3/25 | (12%) | 11/40 | (27.5%) |
| NOM | 5/32 | (15.6%) | 0/25 | (0%) | 5/40 | (12.5%) |
| ABS | 2/18 | (11.1%) | 0/30 | (0%) | 2/40 | (5%) |

---

[15] Naturally, there is an equally-valid reason <u>not</u> to re-use the same paradigm in two domains: dedicated verbal and nominal paradigms unambiguously identify the root they are attaching to as verbs and nouns, respectively. Thus the tension between these two functional pressures can be understood as a classic tug-of-war between the needs of the speaker (lower storage requirements in the lexicon) and the needs of the hearer (less difficulty in parsing).

[16] The grammar of Maybrat marks a gender distinction in the third person singular only: 'M' is used for masculine entities, and 'U' refers to 'unmarked', and it is used for non-masculine entities.

This suggests that the distribution of these accidentally-homophonous, non-ergative syncretic patterns is not quite random, but influenced by additional factors. For instance, **morphological markedness** could be a factor that explains the higher rates of accusative syncretism cross-linguistically when compared to nominative and absolute. The schema in (67) illustrates: in ergative alignments, by definition all clauses will feature ABS case, but only transitive clauses will feature ERG case. Similarly, in accusative alignments, by definition all clauses will feature NOM case, but only transitive clauses will feature ACC case.

(67)  *Marked case in ergative alignment, accusative alignment, and nominal domains*

<div style="text-align:center">ergative alignment</div>

intransitive  [ *John*$_{\text{ABS}}$ *runs* ]

transitive  [ *John*$_{\text{ERG}}$ *throws the ball*$_{\text{ABS}}$ ]

<div style="text-align:center">nominal domain</div>

un-possessed  [ *the hat*$_?$ ]

possessed  [ *John's*$_{\text{GEN}}$ *the hat*$_?$ ]

<div style="text-align:center">accusative alignment</div>

intransitive  [ *John*$_{\text{NOM}}$ *runs* ]

transitive  [ *John*$_{\text{NOM}}$ *throws the ball*$_{\text{ACC}}$ ]

In the same way, genitive/possessive morphology is also the 'marked' form within the nominal domain: only a subset of NPs (possessed NPs) will feature genitive/possessive morphology. Therefore, under the assumption that any kind parallelism between domains may potentially contribute to the increased likelihood of recycling clausal morphology to indicate possession, it makes perfect sense that accusative forms may be selected for this purpose more often.[17]

---

[17] The line of reasoning pursued here implies that, on some level, possessed nouns are 'transitive' (in the sense that they involve two participants) exactly like transitive clauses are. Or perhaps a

To illustrate with two brief examples, consider first Aguaruna. As mentioned previously, Aguaruna features a broad range of morphological cases in the clausal domain, including accusative *-na*, comitative *-haĩ*, locative *-numa/-nĩ*, instrumental *-(a)i*, and ablative *-ia*. Under the hypothesis we are pursuing here, the grammar of Aguaruna opts to recycle a form from the clausal domain as genitive case in the nominal domain. However, none of the non-accusative cases can be said to belong to an alignment, much less correspond to the marked form within that alignment. None of the others are defined in terms of a 'transitive' relationship between two arguments. Therefore it makes sense that accusative morphology would be chosen for this purpose, since its transitive usage in the clausal domain more consistently corresponds to its 'transitive' usage in the nominal domain. To be clear, it is <u>not</u> the case that the ACC=GEN pattern in Aguaruna results from extending the application of dependent accusative case from the clausal domain to the nominal domain. That option is only available to ergative languages, and for ergative syncretisms only, according to the parameter in (64). In this view, any non-ergative syncretism reflects a coincidence of morphological spell-out, not necessarily a structural parallelism between domains.

(68) *Subject and object agreement in Lower Grand Valley Dani*

(Bromley 1981, p.191, p.192)

| subject agreement | | object agreement | |
|---|---|---|---|
| *-i* | 1SG.SUBJ | *n-* | 1SG.OBJ |
| *-in* | 2SG.SUBJ | *h-* | 2SG.OBJ |
| *-e* | 3SG.SUBJ | *w-* | 3SG.OBJ |
| *-u* | 1PL.SUBJ | *nin-* | 1PL.OBJ |
| *-ip* | 2PL.SUBJ | *hin-* | 2PL.OBJ |
| *-a* | 3PL.SUBJ | *in-* | 3PL.OBJ |

more accurate characterization is that possessed nouns are transitive much like Shipibo applicatives in (5b) are, meaning that transitivity is not a requirement, but the construction can nevertheless be augmented by an additional argument.

Consider also Lower Grand Valley Dani, which exhibits an ACC=POSS pattern (48). *A priori*, there are two clausal agreement paradigms to choose from, a nominative and an accusative one (68). It seems to follow, then, that all other things being equal, if the language will opt to recycle one of them to encode possession, it will be the accusative one because of its inherent connection with transitivity between two arguments. Indeed, the hypothesis being advanced here is that when grammars opt to recycle a non-ergative case marker or paradigm, they do so in a pseudo-random way. In principle, any form is eligible, but still some minor factors may make some specific syncretisms more likely that others.

Another factor that may explain the more frequent occurrence of ACC=POSS *vis-à-vis* NOM= POSS is **simple linear word order**. In principle, possessed nouns and possessors should almost always occur adjacently in linear word order, either as N-NP$_{POSS}$ or NP$_{POSS}$-N (depending on headedness). Cross-linguistically, verbs and objects are also very likely to occur adjacently in linear word order. V and O occur adjacently in approximately 90% of the world's languages (SVO, SOV, VOS and OVS) according to Tomlin (1986). On the other hand, verbs and subjects are much less likely to occur adjacently in linear word order cross-linguistically. This only happens in approximately 55% of the world's languages (SVO, VSO, OSV and OVS).

Therefore if a randomly-selected language has elected to reuse a verbal agreement paradigm in the nominal domain to encode possession for reasons of economy, it is more likely to have an inherent linear word order parallel between N-NP$_{POSS}$ and V-O than with V-S. To illustrate, Language X$_1$ exhibits SVO basic word order in clauses and N-NP$_{POSS}$ word order in possessed NPs (69). It follows that recycling the agreement paradigm that occurs between the V and O elements to also occur between N and NP$_{POSS}$ is more natural than extending the prefix that occurs between

238

the S and O elements.[18]

(69) *Language $X_1$*

| clause: | NP$_{\text{SUBJ}}$ | SUBJ-V-OBJ | NP$_{\text{OBJ}}$ |
|---|---|---|---|
| NP: | | N | NP$_{\text{POSS}}$ |
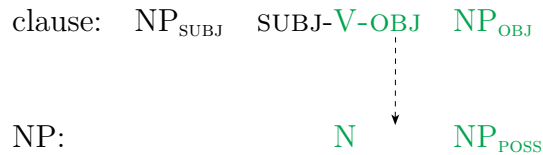
The schema in (69) corresponds exactly to a language such as Lango. Transitive objects and possessors occur as lexical NPs following their governing heads (70); as expected, the suffixal object agreement paradigm is recycled as suffixal possessor agreement as well (71). Furthermore, this exact kind of configuration occurs four more times in the survey of accusative languages (Jola Bandial, Lele, Modern Hebrew, Indonesian), indicating it is fairly common cross-linguistically among SVO languages.

Lango                                                    (Noonan 1992; p.119, p.53, p.120, p.159)

(70) a. *lócɔ̀ ò-càmò̀     dɛ̀k*
      man 3SG.SUBJ-eat stew

    'the man ate stew'

   b. *tyɛ̀n ɲɔ̂ŋ*
      leg   crocodile

    'a crocodile's leg'

---

[18] See Siewerska's (2004) survey on correlations between basic word order and location of affixes. In SVO languages, subject agreement occurs as a prefix 73% of the time, and object agreement occurs as a suffix 64% of the time, mirroring clausal basic word order. In SOV languages, object agreement mirrors basic word order and occurs as a prefix in 56% of languages, but subject agreement mirrors basic word order and occurs as a prefix only 30% of the time. See also Greenberg (1963) for correlations between order of V and O, and order of possessor and possessee.

(71) a. *lócɔ̀ ò-nèn-á*
    man 3SG.SUBJ-see-1SG.OBJ
    'the man saw me'

  b. *tyɛ̀n-á*
    leg-1SG.POSS
    'my leg'

Consider now Language $X_2$, which exhibits SOV basic word order in clauses and $\text{NP}_\text{POSS}$-N word order in possessed NPs. Again, it follows that recycling the object agreement paradigm for use as possessor agreement is extremely natural due to the inherent parallel between V heads and objects in the clausal domain, and N heads and their possessors in the nominal domain (72).

(72) *Language $X_2$*

clause: $\text{NP}_\text{SUBJ}$   $\text{NP}_\text{OBJ}$   OBJ-V-SUBJ

NP:              $\text{NP}_\text{POSS}$    N
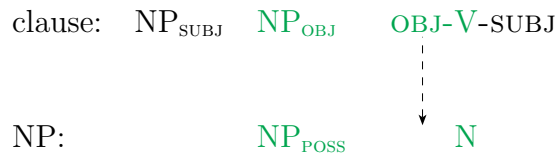
Now the schema in (72) corresponds to a language such as Lower Grand Valley Dani: objects and possessors occur as lexical NPs preceding their governing heads (73-74). As expected, the prefixal object agreement paradigm is recycled as a prefixal possessive paradigm as well. Furthermore, this type of configuration occurs three more times in the survey of accusative languages (Comanche, Hixkaryana, Warao), indicating that it too is fairly common among OV languages.

Lower Grand Valley Dani                (Bromley 1981, p.96, p.93, p.191, p.190)

(73) a. *ap   kut       esi    w-anh-a-tek*
    man white.heron feather 3SG.OBJ-take-3PL.SUBJ-HAB
    'men get white heron feathers'

b. *Ukumhearik-mekke wam*
Ukumhearik-REL    pig

'Ukumhearik's pig'

(74) a. *n-ath-e*
1SG.OBJ-hit-3SG.SUBJ

'he hit me'

b. *n-a-su*
1SG.POSS-PFX-net

'my net'

Language $X_3$ exhibits SVO basic word order in clauses and NP$_{POSS}$-N word order in possessed NPs. This configuration is rarer – the relative order of V-O and NP$_{POSS}$-N violates Greenberg's (1963) Universals 3 and 4 – but now the linear parallel between the two domains makes NOM=POSS extremely straightforward, and thus the nominative agreement paradigm is re-used as a possessor prefix (75).

(75) *Language $X_3$*

clause:    NP$_{SUBJ}$    SUBJ-V-OBJ    NP$_{OBJ}$

NP:        NP$_{POSS}$        N
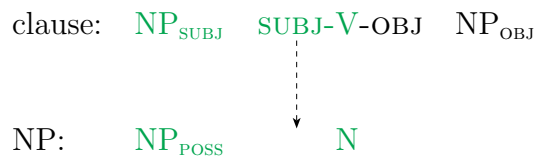
The schema in (75) represents the situation in Cocama. Subjects and possessors occur as lexical NPs preceding their governing heads (76), while pre-clitics may also encode both possessors and subjects (77). Only two languages fits this schema in the sample of accusative languages: Cocama and Maybrat. Again, this relatively lower number is straightforwardly explained by the fact that the configuration in (75) is relatively rare cross-linguistically.

(76) a. *awa=nu    purepe chura-n=kɨ=nu*
     person=PL buy    be.small-NMLZR=DIM=PL

     'the people buy the cute small ones'

    b. *rikua   tapira rimariru iriw=uy*
     reason tapir  grandson return=PST

     'and that's why the tapir's grandson returned'

(77) *ya=mirikua=muki     y=ichari      ya=kaistuma*
    3SG.POSS=wife=COM 3SG.SUBJ=leave 3SG.POSS=yucca.beer

    'with his wife he leaves his yucca beer'

Finally, consider Language $X_4$, where the parallel between domains is inherently disrupted. The language has basic SOV word order (and so technically in terms of linear word order S does occur before V), and possessors occur before the possessed noun, but the problem is that objects occur <u>between</u> subject and verbs (78). Therefore it seems extremely unlikely that the grammar would extend the subject agreement paradigm in this scenario. The linear sequence of NP-agreement-head observed in the examples above (69-77) is simply not available, and this unavailability does seem like a predictor of poor rates of syncretism since only one language in the sample from Chapter 2 actually corresponds to the schema in (78): Maricopa.

(78) *Language $X_4$*

      clause:  NP<sub>SUBJ</sub>   NP<sub>OBJ</sub>    SUBJ-V-OBJ
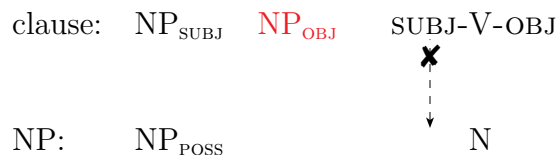
    NP:     NP<sub>POSS</sub>         N

To summarize the discussion in this section, extremely common word order configurations in SVO and SOV languages cross-linguistically seem to naturally lend

themselves to a parallelism between V-O and N-NP$_{POSS}$ (or O-V and NP$_{POSS}$-N). On the other hand, SVO only lends itself to a parallelism that favors NOM=POSS in rare cases, and finally SOV doesn't lend itself to such a parallelism at all. This explanation is consistent with the data from Chapters 2 and 3, in which ACC=POSS patterns occur much more frequently than NOM=POSS.

Thus, along with simple markedness, linear word order may be considered another factor which nudges a language's grammar to re-use an accusative agreement paradigm over a nominative paradigm to encode agreement in the nominal domain. Note however that unlike in Baker (2015), all logically-possible patterns can still occur via this alternate route to syncretism, but due to the pseudo-random nature of this recycling process, some non-ergative syncretisms will nevertheless be more than others cross-linguistically.[19]

## 4.4 No syncretism?

The analyses in Sections 4.2.1 and 4.3 provide an account of the fundamental difference between ergative and non-ergative patterns of syncretism, respectively. The former occur due to a parameterized option whereby the computation of dependent ergative case is extended from the clausal domain to the nominal domain; the latter is characterized as coincidental homophony in the morphological spell-out of case marking or agreement paradigms, which can in turn be influenced by outside factors such as markedness and linear word order.

---

[19] In Section 4.3, some functional arguments were presented to explain the distribution of non-ergative syncretic patterns observed in Chapters 2 and 3. For more general theoretical accounts of morphological syncretism in the generative tradition, see Baerman et al (2005), Bobaljik (2001), Brown & Hippisley (2012), Caha (2013) and Harley (2008).

One question remains, however: how do we model straightforward cases that lack any kind of syncretism between the two domains? Looking again at the distribution in (66), this is by far the most common state of affairs cross-linguistically. In Yakima Sahaptin, for example, the genitive case marker *-mí* does not correspond to the ergative case marker *-nɨm*, or indeed any other bit of morphology in that language's system of case marking (79). In these languages, genitive and/or possessive forms are simply spelled out as their own unique morphological cases or paradigms.

Yakima Sahaptin                                                                (Jansen 2010; p.134, p.297, p.290)

(79) a. *tamánwit-nɨm =nash i-nápayun-ta*
         law-ERG        1SG    3SG.SUBJ-defend-FUT
         'the law will support me'

   b. *i-kwíita-na*              *Spilyáy*
      3SG.SUBJ-go.along-PST  Coyote
      'Coyote was traveling along'

   c. *ɨshchɨ́t kaas-mí*
      path    train-GEN
      'railroad track' (lit. 'path of the train')

Marantz (1991) provides a simple explanation: in his disjunctive hierarchy, default case may be "environment-sensitive", meaning that its realization is conditioned by the specific domain of case computation. For example, default case in the clausal domain is either nominative or absolutive, depending on alignment, as discussed in Section 4.1.1. However, in the nominal domain, default case is genitive, with its own unique morphological spell-out. In (80), the ERG=GEN parameter is set such that dependent ergative case is not assigned in the nominal domain of Yakima, even though the structural requirements would otherwise be met. Instead,

the grammar simply assigns default genitive case in this specific syntactic domain, and its morphological realization is -*mí*. In the clausal domain, default case is absolutive, and its morphological realization is null (79b).

(80) *Default 'genitive' case in the nominal domain*

```
                    DP_i
                   /    \
                  D      nP_i
                        /    \
                       /      \
                      /        DP_k
                     /        /    \
                    n   NP_i     train-GEN
                         |
                         N_i
                         |
                        path
```
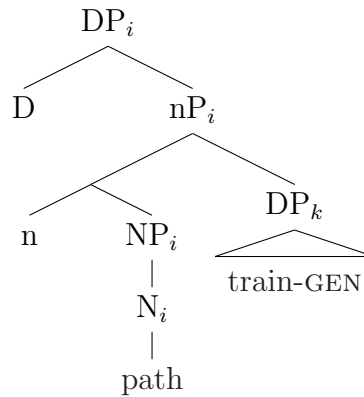
Marantz's analysis implies that the possessed noun head in (80) does not receive case, default or otherwise – otherwise it, too, would receive default genitive case in that domain. A possessed noun head may ultimately receive lexical, dependent or default case according to its role within a larger, clausal construction. Within the nominal domain, however, it is simply not eligible. Note also that default case may not <u>always</u> be "environment-sensitive" in the sense of Marantz. In some languages, the realization of default case is the same in the clausal and nominal domains. For example, in Thai, pronouns are realized in the same form as subjects, objects or possessors (81). I leave it up to future research whether this pattern of essentially unmarked possessors is a cross-linguistically common one.

Thai                                                    (Smyth 2002; p.40, p.40, p.143)

(81) a. *kháw chɔ̂ɔp phǒm*
        3SG  like   1SG
        'he likes me'

b. *phǒm chôɔp kháw*
   1SG  like  3SG

   'I like him'

c. *nîi  mây  chây  bâan  kháw*
   DEM NEG1 NEG2 house 3SG

   'this isn't his house'

The table in (82) summarizes the three different ways that possessive morphology is realized according to the analysis in this Chapter. In a subset of morphologically ergative languages, the mechanisms of ergative case and/or agreement assignment are extended to the nominal domain, according to the parameter in (64). Alternatively, a language's grammar may simply re-use existing case markings or agreement paradigms, but crucially this does not occur as the result of re-applying the relevant algorithms across domains. Finally, as discussed in this last section, when no syncretism at all has occurred, the grammar simply assigns its own dedicated default possessive forms within the nominal domain.

(82) *Summary: three types of possessive morphology in the nominal domain*

| clausal domain | nominal domain |
| --- | --- |
| dependent ERG | dependent ERG |
| dependent ERG | accidental homophony (ERG or ABS) |
| dependent ACC | accidental homophony (NOM or ACC) |
| dependent ERG dependent ACC | no syncretism / default genitive |

## 4.5  Conclusion

This chapter started out by outlining of a syntactic model of case and agreement, based on theories by Marantz (1991), Bobaljik (2008) and Baker (2013). This

two-step, linear model was shown to be effective in modeling clausal data from Chapters 2 and 3, but in addition each component of the model was shown to make good predictions with regards to well-known typological facts. In the next section, that model was straightforwardly applied to the syncretic patterns attested in Chapters 2 and 3. While the general model was shown to be compatible with ergative syncretic patterns (ERG=GEN and ERG=POSS), at the same time it runs into various difficulties with regards to accusative, nominative, and absolutive syncretisms, which are rarer but still attested.

To account for this observed distribution of relatively common ergative syncretic patterns and comparatively rarer non-ergative ones, Baker's (2015) ERG=GEN analysis was formalized as a parameter available only to ergative languages, which predicts the roughly 50% occurrence of ergative syncretisms in the data. Furthermore, it was proposed that the rarer, non-ergative syncretisms are not generated by extending the morphosyntactic computation of dependent case and agreement to the nominal domain, but rather by a somewhat random process of 'recycling', where languages simply re-use existing forms for reasons of economy. In principle, this alternative possibility allows for all logically-possible syncretisms to occur, but in addition the choice of which forms to re-use may be influenced by additional factors such as morphological markedness and linear word order, making certain non-ergative patterns slightly more common than others cross-linguistically.

## Chapter 5

## A NEW PERSPECTIVE ON EXTRACTION ASYMMETRIES

The goal of this chapter is to explore the phenomenon of ergative Ā-extraction asymmetries – sometimes known as syntactic ergativity – and its relationship to morphological ergativity. In Section 5.1, a brief description of the phenomenon is provided, and some basic typological facts are laid out. Specifically, ergative extraction asymmetries only occur in a subset of morphologically ergative languages, and furthermore they may occur only in a subset of Ā-extraction environments. In Section 5.2, an overview of previous syntactic approaches to this phenomenon is provided. These syntactic analyses all share a common thread: the explanation for ergative extraction asymmetries is somehow rooted in the assignment of morphological case. In Section 5.3, a new hypothesis is fleshed out: ergative extraction asymmetries are not motivated by case, but rather they serve a disambiguation function in ambiguous extraction scenarios. This hypothesis makes the strong prediction that languages which would *a priori* produce ambiguous Ā-extraction are exactly those which exhibit these asymmetries, thus rendering them unambiguous. In Section 5.4, the disambiguation hypothesis is shown to produce good results when applied to a large sample of ergative languages. Finally, in Section 5.5, the full ramifications of this hypothesis are explored, and it is argued that similar asymmetries occur in morphologically accusative languages, and furthermore that in some cases morphological alignment and extraction alignment may be mismatched within a single language,

providing further evidence against case as the underlying basis for these extraction asymmetries.

## 5.1 A brief description of ergative extraction asymmetries

According to Chomsky (2000), Ā-extraction occurs when a constituent undergoes phrasal movement to a non-argument position, as in Wh-questions (1a), relative clauses (1b) or focus fronting (1c). The exact structural position of the landing site is often modeled as the specifier of CP, although in some syntactic traditions the left periphery of the clause actually consists of several functional projections, each dedicated to a different kind of Ā-extraction (Rizzi 1997).

(1) a. [$_{CP}$ *what did* [$_{IP}$ *John eat* t ] ]

   b. *the sandwich* [$_{CP}$ OP *that* [$_{IP}$ *John ate* t ] ]

   c. *it was* [$_{CP}$ *a sandwich that* [$_{IP}$ *John ate* t ] ]

Descriptively, ergative extraction asymmetries occur when the transitive object and intransitive subject can be freely extracted via Ā-movement, but some special morphological change marks instances of transitive subject extraction only.[1] As a reminder, the data in (2-3) illustrates morphological ergativity in two languages that we have already discussed: Kaqchikel and Niuean. Kaqchikel exhibits ergative alignment in its system of head-marking because the same morphological paradigm encodes transitive objects and intransitive subjects, while a different paradigm encodes

---

[1] See Deal (2016) and Polinsky (to appear) for overviews of ergative extraction asymmetries, as well as Dixon (1994).

transitive subjects. Niuean exhibits ergative alignment in its system of dependent-marking, because the transitive object and intransitive subject are marked with the ABS marker *e*, while the transitive subject is marked with the ERG marker *he*.

Kaqchikel

(2) a. x-      a-        ru-       tijo  -j
       PERF 2SG.ABS 3SG.ERG teach TR
    'he taught you'

   b. x-      a-        wär
      PERF 2SG.ABS sleep
    'you slept'

   c. x-      Ø-        wär
      PERF 3SG.ABS sleep
    'he slept'

Niuean                                                    (Seiter 1980; p.29, p.28)

(3) a. *ne    kai he   pusi ia    e    moa*
       PAST eat ERG cat    that ABS chicken
    'that cat ate the chicken'

   b. *malona tuai   e    kapiniu ē*
      broken  PERF ABS dish      this
    'this dish is broken'

When it comes to Ā-extraction, Kaqchikel exhibits an ergative extraction asymmetry. Extracting the transitive subject requires the deletion of ergative agreement on the verb (despite still being a semantically transitive clause), and the inclusion of the Agent Focus suffix *-on* (4a, 5a). In contrast, extracting the transitive object (4b, 5b) or the intransitive subject (4c, 5c) entails no special change in morphology. In other words, instances of transitive object and intransitive subject extraction are morphologically consistent with similar clauses lacking extraction, but

the same cannot be said for instances of transitive subject extraction. The Agent Focus suffix *-on* only co-occurs with transitive subject Ā-extraction.

Kaqchikel

(4) a. *achike x-tz'ët-on      ri   achin*
     Wh    PERF-see-**AF**    DET man
     'who saw the man?'

   b. *achike x-u-tz'ët       ri   achin*
     Wh    PERF-3SG.ERG-see DET man
     'who did the man see?'

   c. *achike x-∅-wär*
     Wh    PERF-sleep
     'who slept?'

(5) a. *ri   wakx [  ri x-tz'ët-on    ri   achin  ]*
     DET cow      C PERF-see-**AF** DET man
     'the cow that saw the man'

   b. *ri   wakx [  ri x-u-tz'ët       ri   achin  ]*
     DET cow      C PERF-3SG.ERG-see DET man
     'the cow that the man saw'

   c. *ri   wakx [  ri x-wär      ]*
     DET cow      C PERF-sleep
     'the cow that slept'

Looking at Niuean now, there is a stark contrast. In (6-7), descriptively there is no single piece of morphology that characterizes transitive subject Ā-extraction and transitive subject Ā-extraction only. Morphological case, which is in ergative

alignment, does not differ in any substantial way from what is observed in plain, declarative clauses.

Niuean                                    (Seiter 1980; p.109, p.110; p.110, p.246, p.246; p.246)

(6) a. *ko    hai ne    fifili    a    koe ke  vagahau*
       PRED who N.FUT choose ABS you SUB speak

'who chose you to speak?'

b. *ko    hai  ne    fahi  e    Sione*
   PRED who N.FUT beat ERG Sione

'who did Sione beat?'

c. *ko    hai  ne    nofo he fale   kō*
   PRED who N.FUT live  in  house that

'who lives in that house?'


(7) a. *kehe <u>tama</u>  [  ka   kai e    tau pateta   ]*
       to    child       FUT eat ABS PL  potato

'to the child who is going to eat the potatoes'

b. *mo   e    <u>tagata</u>  [  ne     moto  e    koe  ]*
   with ABS person        N.FUT punch ERG you

'with the person who you punched'

c. *e    <u>tama</u>  [  ne     hau  i    Makefu  ]*
   ABS child       N.FUT come LOC Makefu

'the child who comes from Makefu'

The data in (2-7) thus provides a snapshot of the facts surrounding ergative extraction asymmetries cross-linguistically. Among morphologically ergative languages, which uniquely mark the transitive subject in their morphological alignment, a subset of them will also uniquely mark the transitive subject for Ā-extraction. Other languages, despite being morphologically ergative, will not. Why?

The mystery deepens further when we consider languages such as West Greenlandic. West Greenlandic is a morphologically ergative language because the same dependent-marker encodes transitive objects and intransitive subjects, while a different marker encodes transitive subjects, as in (8).[2]

West Greenlandic                                                    (Bittner & Hale 1996; p.17, p.17)

(8) a. *juuna-p      miiqqa-t      paar(i-v)-a-i*
       Juuna-ERG child-PL.ABS look.after-IND-TR-3SG>3PL

       'Juuna is looking after the children'

    b. *arna-t          mirsur-p-u-t*
       woman-PL.ABS sew-IND-INTR-3PL

       'the women are sewing'

However, when it comes to extraction, the interesting thing is that West Greenlandic only exhibits an asymmetry in a <u>subset</u> of extraction scenarios. For example, in Wh-questions, there is no special morphology that must accompany verbs in transitive subject extraction, or indeed any other kind of extraction (9). Therefore West Greenlandic cannot be said to exhibit an ergative extraction asymmetry in Wh-questions.

West Greenlandic                                                    (Fortescue 1984; p.23, p.16, p.24)

(9) a. *kia        uqaatig-aa*
       who.ERG talk.about-3SG>3SG.INT

       'who talked about it?'

---

[2] In addition, West Greenlandic may exhibit morphological alignment in its system of head-marking, but only in certain configurations; see Chapter 2 for more details.

b. *suna=ana Jaaku-p    siulittaasu-u-vvigi-gaa*
   what=that Jaaku-ERG chairman-be-have.as.place.of-3SG>3SG.PART

   'what is Jaaku chairman of?'


c. *kina tikis-sinnaa-va*
   who come-can-3SG.SUBJ.INT

   'who can come?'


On the other hand, in relative clauses, there is a definite asymmetry. Only when the transitive subject is relativized, the verb must detransitivize to the antipassive with the *-si* morpheme, as in (10a). When the transitive object or the intransitive subject is relativized, the antipassive is not required (10b-10c).

West Greenlandic                                    (Bok-Bennema 1991; p.69, p.65, p.239)

(10) a. *piniartuq*    [  *nannu-mik    tuqut-si-suq*    ]
        hunter.ABS       polar.bear-MOD kill-ANTI-PART

        'the hunter who killed the polar bear'


   b. *nanuq*         [  *Piita-p   tuqu-ta-a*           ]
      polar.bear.ABS    Piita-ERG kill-PART-3SG.SUBJ

      'the polar bear that Piita killed'


   c. [  *ipis-suq*        ]
         be.sharp-PART

      'the thing that is sharp'


This interesting set of facts leads to the driving research questions of this chapter. Why do some morphologically ergative languages pattern like Kaqchikel, exhibiting ergative extraction asymmetries 'across the board', in all extraction scenarios? Why do other morphologically ergative languages pattern like Niuean, not

exhibiting any kind of extraction asymmetry at all? Finally, why do languages of a third type pattern like West Greenlandic, which only exhibits an extraction asymmetry in one domain, but not in another? This intriguing conundrum is summarized in (11) below.

(11) *Morphological ergativity and ergative extraction asymmetries: 3 types*

| | morphological | ergative extraction asymmetry | |
| | ergativity | Wh-question | relative clause |
|---|:---:|:---:|:---:|
| Kaqchikel, Gitksan, Selayarese | ✔ | ✔ | ✔ |
| West Greenlandic, Roviana | ✔ | ✘ | ✔ |
| Niuean, Yukulta, Basque | ✔ | ✘ | ✘ |

## 5.2   Syntactic analyses

In this section, previous attempts to explain the facts in (11) are discussed. While each of these analyses differs in their details, there is a common thread among them: ergative extraction asymmetries are thought to be a by-product of morphological case assignment in the syntax.

### 5.2.1   Aldridge (2004, 2008, 2012)

In a series of papers, Aldridge (2004, 2008, 2012) develops a line of analysis for Tagalog where a combination of several syntactic factors explains ergative extraction asymmetries.[3] Descriptively, Tagalog features an asymmetry because transitive

---

[3]  Aldridge (2004, 2008, 2012) and the other analyses in Section 5.2 refer to extraction asymmetries such at those in (4-5) as "syntactic ergativity". Because the main goal of this chapter is to redefine this phenomenon independently from morphological alignment, the more neutral term "extraction asymmetry" has been adopted. Most of Chapter 5 is concerned with ergative extraction asymmetries, but see also Section 5.5 for accusative extraction asymmetries, which are fully predicted to exist under the disambiguation hypothesis.

objects and intransitive subjects can freely be extracted (12b-12c), but in order for transitive subjects to be extracted, the verb must be realized in its antipassive form (12a).

Tagalog                                                     (Aldridge 2008; p.1446, 1445, p.1446)

(12) a. *sino ang b-um-ili      ng   ida*
        who ABS -INTR.PERF-buy OBL fish
        'who bought the fish?'

     b. *ano   ang b-in-ili      ng   babae*
        what ABS -TR.PERF-buy ERG woman
        'what did the woman buy?'

     c. *sino ang d-um-ating*
        who ABS -INTR.PERF-arrive
        'who arrived?'

   Aldridge proposes to account for these facts in the following manner. In a regular transitive clause, transitive v assigns inherent ergative case to the external argument in its specifier (Legate 2008). Transitive v also assigns structural absolutive case to the internal argument in its base-generated position. The absolutive-marked internal argument must then move to an outer specifier of vP to satisfy an [EPP] requirement, represented in (13) by the solid arrow. The lexical verb then undergoes head-movement all the way to I to derive V-initial surface word order, represented in (13) by the dotted arrow.

Tagalog                                                     (Aldridge 2008; p.1443, p.1445)

(13) *b-in-ili       ng   babae   ang isda*
     -TR.PERF-buy ERG woman ABS fish
     'the woman bought the fish'

256
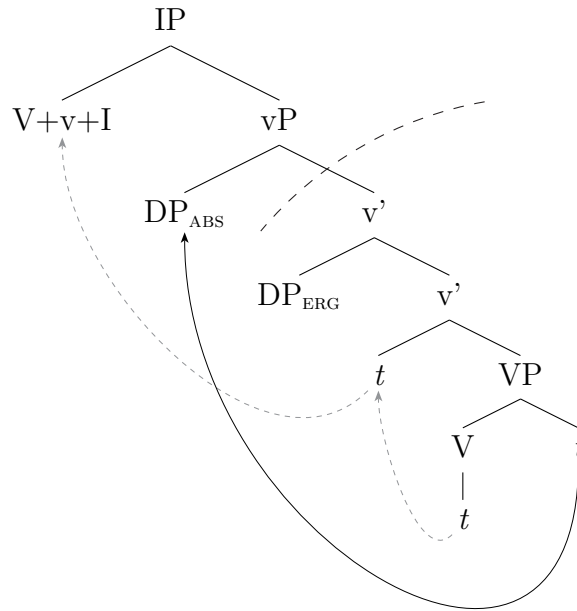
Given that transitive v is a phase head, the asymmetry between transitive subjects and transitive objects thus follows from the Phase Impenetrability Condition (Chomsky 2001). From its position at the phase boundary in the outer specifier of vP, the transitive object can freely extract to a position further up the tree, as in the case of Wh-extraction (12b). However, within the configuration in (13), the transitive subject is not at the edge of the phase boundary, and therefore it is 'trapped', unable to undergo any kind of movement from that position.[4]

In intransitive clauses, the I head assigns absolutive case to the only argument, as in (14). By stipulation, intransitive v does not have an [EPP] requirement, and furthermore it is not phasal. The intransitive subject can thus freely extract from its base-generated position, thus accounting for examples such as (12c).

---

[4] Note, however, that this is not exactly the usual understanding of the PIC; in its original formulation, anything at the phase edge can freely extract, which includes the head in question and any number of specifiers.

Tagalog                                    (Aldridge 2008; p.1443, p.1446)

(14) *d-um-ating*          *ang babae*
     -INTR.PERF-arrive ABS woman

     'the woman arrived'
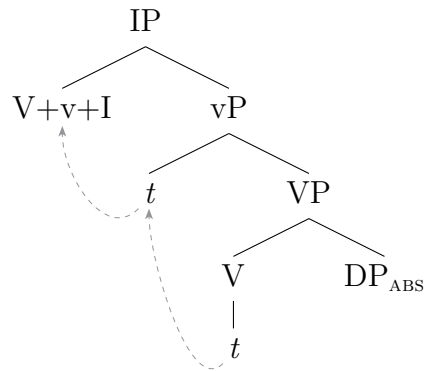


In actual cases of transitive subject extraction, the verb must be realized in its antipassive form, indicated with the infix *-um-*. The verb then assigns lexical oblique case to its internal argument, while the external argument receives absolutive case from I, just like intransitive subjects do in intransitive clauses. Furthermore, because intransitive v is not phasal, the transitive subject can extract from its base-generated position, and thus produce instances of Wh-extraction such as (12a).

Tagalog                                              (Aldridge 2008; p.1446)
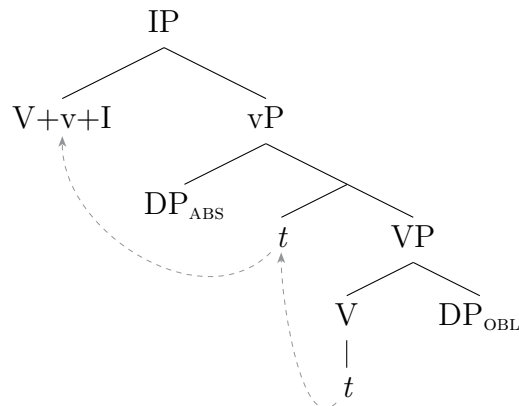
To summarize, ergative extraction asymmetries are explained by the interaction of several syntactic processes: an [EPP] requirement on transitive v heads requires the absolutive-marked transitive object to move to the outer specifier of vP, which blocks movement out of the phase by the structurally-lower transitive subject due to the Phase Impenetrability Condition. If the transitive subject does wish to extract, the clause must be made intransitive first, which in turn obviates the [EPP] requirement responsible for the blocking effect in the first place. Furthermore, if some languages do not exhibit an extraction asymmetry, it is because in those languages the transitive v head lacks an [EPP] feature. Since the absolutive-marked transitive object doesn't have to move to the outer specifier of vP in these languages, there is no blocking effect, and the transitive subject can freely extract from its base-generated position.

### 5.2.2 Coon et al (2014)

Coon et al (2014) seek to extend Aldridge's 'blocking' analysis to account for ergative extraction asymmetries in the Mayan language family. The starting point of the analysis comes from the observation that within that family, some languages linearize the absolutive morpheme immediately after the verb's TAM morpheme ('high ABS'), as in Q'anjobal (15). Other languages linearize the absolutive morpheme after the verb root ('low ABS'), as in Chol (16).

Q'anjob'al                                             (Coon et al 2014; p.190, p.190)

(15) a. *max-ach    y-il-a'*
        ASP-2.ABS 3.ERG-see-TR
        'she saw you'

   b. *max-ach    oq'-i*
        ASP-2.ABS cri-INTR
        'you cried'

Chol                                                        (Coon et al 2014; p.190, p.190)

(16) a. *tyi    y-il-ä-yety*
        ASP 3.ERG-see-TR-2.ABS

    'she saw you'

    b. *tyi    uk'-i-yety*
        ASP cry-INTR-2.ABS

    'you cried'

Furthermore, 'high ABS' languages such as Q'anjob'al tend to exhibit ergative extraction asymmetries, while 'low ABS' languages such as Chol do not. In (17), extraction of the transitive object and the intransitive subject proceeds normally, but only in the case of transitive subject extraction, the verb must lose its ergative agreement, and in addition it must be marked with the Agent Focus morpheme *-on*. However, in Chol, there is no asymmetry. In (18), the same question can correspond to either an instance of transitive subject extraction, or transitive object extraction.

Q'anjob'al                                              (Coon et al 2014; p.180, p.192, p.192)

(17) a. *maktxel max-ach        il-on-i*
        who     ASP-2SG.ABS see-AF-INTR

    'who saw you?'

    b. *maktxel max y-il[-a']        naq winaq*
        who     ASP 3SG.ERG-see-TR CLF man

    'who did the man see?'

    c. *maktxel max way-i*
        who     ASP sleep-INTR

    'who slept?'

(18) *maxki tyi y-il-ä jiñi wiñik*
who    ASP 3SG.ERG-see-TR DET man

'who saw the man?' / 'who did the man see?'

To account for the Q'anjob'al asymmetry in (17), Coon et al propose a similar analysis whereby, in 'high ABS' languages only, the internal argument must move to spec-vP to receive absolutive case from I. As a result, since transitive vP is phasal, the transitive subject is blocked, and it cannot escape from its base-generated position in spec-VoiceP. This 'blocking' effect is schematized in (19) below.

(19) *Modeling the ban on A-extraction in Q'anjob'al* (Coon et al 2014; p.210)
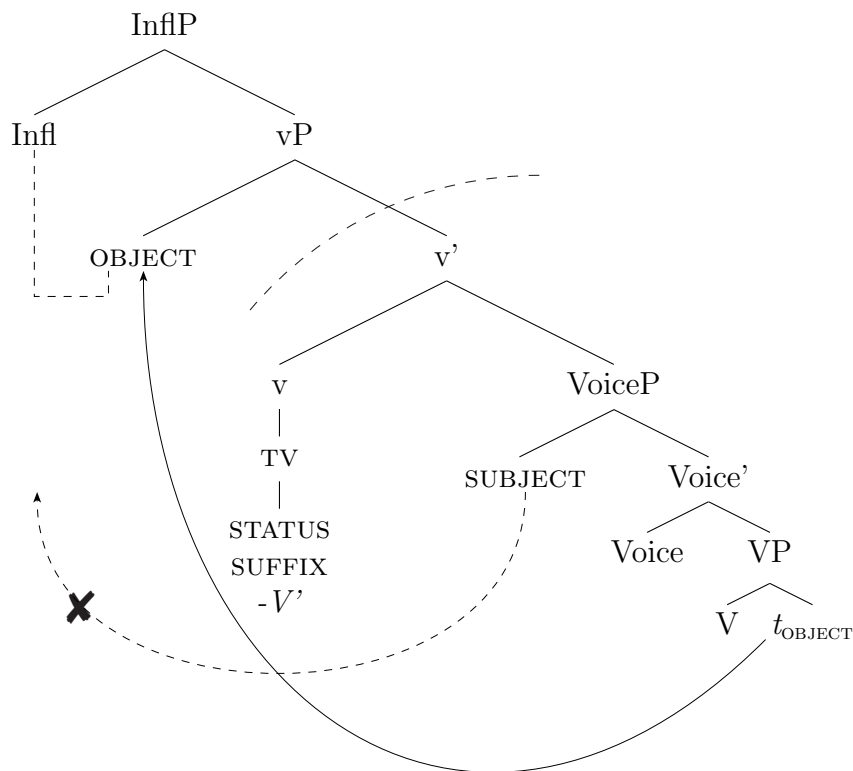
Like for Aldridge, transitive objects and intransitive subjects can freely extract because they either crucially occupy that phase-edge position in spec-vP, or because intransitive v is simply not phasal, respectively. Also like Aldridge, if the transitive subject does want to extract, the verb must be converted to an intransitive construction by way of the Agent Focus morpheme *-on*. This Agent Focus morpheme assigns structural case to the internal argument, leaving the I head free to assign absolutive to the external argument. The internal argument must still move to spec-vP due to an [EPP] requirement, but crucially now the external argument can still extract from its base-generated position in spec-VoiceP, because intransitive vP is not phasal.

(20) *Agent Focus in Q'anjob'al* (Coon et al 2014; p.218)
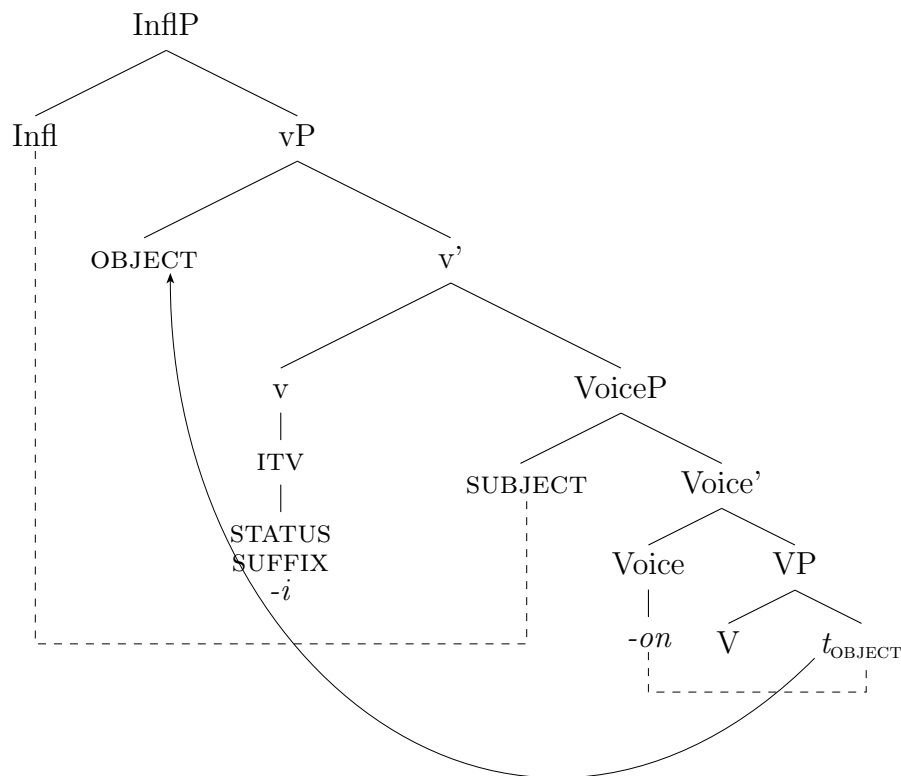
To account for languages with no ergative extraction asymmetry such as Chol (18), Coon et al propose that in those languages, the transitive object doesn't have to move to spec-vP to receive case from I; instead they can be licensed directly by the transitive v head itself. This in turn is consistent with the 'low' location of absolutive morphology within the verbal complex in these languages. In terms of extraction, the absence of a blocking element in spec-vP means that if the transitive subject wants to extract, it can simply first move to the unoccupied phase edge in spec-vP, and then continue on its way to a higher structural position after that.

To summarize, some details differ from Aldridge (2004, 2008, 2012) but the gist of the analysis is the same: ergative extraction asymmetries occur when the transitive subject is blocked from further extraction, due to a combination of [EPP] and phase edge effects at spec-vP. Furthermore, the analysis is at least partially motivated by morphological case. In Coon et al (2014), the fact that the transitive object must move to spec-vP in the first place is designed to capture facts about morpheme linearization, where absolutive morphology occurs immediately after aspectual morphology generated in the I head, but before any other verbal material such as ergative morphology and the lexical verb root.

However, there are reasons to doubt that the mechanics of absolutive case assignment serve as a reliable predictor of ergative extraction asymmetries. The analysis by Coon et al predicts that Mayan 'high ABS' languages should be the ones to exhibit extraction asymmetries, because transitive objects must receive case directly from the I head thus leading to a blocking effect, while 'low ABS' languages should not, since they can receive case *in situ*, thus obviating any need for movement to a blocking position. Unfortunately, some Mayan languages contradict this prediction: in Ixil, transitive subjects cannot freely undergo extraction, despite being a 'low ABS' language morphologically (21).

(21)  a.  *a=in          kat=tzok-on    u     si'-e'*
          F.COP=1SG  CMPL=cut-AF  DEF  firewood-ENC

          'It is I who cut the firewood'

    b.  *a=kuxh-e'          nu-ku-b'an=∅-e'*
          F.COP=only-ENC  INC-1PL.ERG-do-3SG.ABS-ENC

          'it it just this that we did'

    c.  *a=in          kat=vat=in-e'*
          F.COP=1SG  CMPL-sleep-1SG.ABS-ENC

          'it was me who slept'

Similarly, in earlier formulations of her blocking analysis, Aldridge (2004) motivated the movement of the transitive object to the outer specifier of vP by the need to receive absolutive case locally from the I head, instead of an [EPP] requirement as in later versions. The rest of the analysis was essentially the same; from its position in the outer specifier, the transitive object received absolutive case from I but prevented any extraction of the structurally-lower transitive subject in the inner specifier. Aldridge's original analysis therefore made a similar prediction that for all such languages, if the I head assigns absolutive case to the transitive object, a 'blocking' effect should occur, along with an ergative extraction asymmetry.

Again, this is clearly not true for at least one language in the ergative sample of Chapter 2. In Georgian, absolutive case is unavailable for intransitive subjects and transitive objects alike in non-finite environments (22). This is the key diagnostic in Legate (2008) for belonging to the ABS=NOM class of languages, and it is taken as evidence that in Georgian, absolutive case on both transitive objects and intransitive subjects is assigned by the I head.

Georgian (Harris 1981; p.157, p.157)

(22) a. [ *datv-is*     *mok'vla*     *am t'qeši*    ] *ak'rʒalulia*
        bear-GEN killing.NOM this woods.in forbidden.it.is.I.2

   'killing bears in these woods is forbidden'

b. [ *tamad-is*     *damtknareba*  *supraze* ] *uzrdelobaa*
     tamada-GEN yawning.NOM table.on rudeness.it.s.I.2

   'it is rude for the *tamada* to yawn at the table'

And yet, as the data in (23) shows, there is clearly no extraction asymmetry in relative clauses of Georgian, contrary to the original predictions of Aldridge (2004).[5]

Georgian (Foley 2013; p.9, p.9 / Aronson 1991; p.285)

(23) a. *mosc̣avle* [ *romel-ma=c*     *ǵoveltvis icis*       *p̣asuxi* ]
        student    which-ERG=REL always    know.PRES.3SG answer

   'the student which always knows the answer'

b. *kali*    [ *romeli=c*   *gušin*    *Eka-m*    *naxa*       ]
     woman    which=REL yesterday Eka-ERG see.AOR.3SG

   'the woman who Eka saw yesterday'

c. *k'aci* [ *romeli=c*    *lap'arak'-ob-d-a* ]
     man    which=REL he-was-speaking

   'the man who was speaking'

Therefore, analyses such as Aldridge (2004) and Coon et al (2014) predict that morphologically ergative languages in which the transitive object receives its

---

[5] Of course, this Georgian relative clause data is perfectly-well accounted for by the disambiguation hypothesis to be presented in Section 5.3: morphological case makes the nature of the extracted argument extremely clear, and therefore no ergative extraction asymmetry is functionally required. Likewise, the ergative extraction asymmetry in Ixil is predicted by the fact that it *lacks* relevant cues such as morphological case, thus making Ā-extraction potentially ambiguous.

absolutive case from the I head – ABS=NOM languages in Legate's (2008) terms – are the ones that should exhibit extraction asymmetries. However, data from languages such as Ixil and Georgian apparently contradict this prediction.

### 5.2.3 Polinsky to appear (2016, to appear)

Polinsky (2016, to appear) proposes that in languages which exhibit an ergative extraction asymmetry, ergative case is assigned by a null P head, and thus the ban on transitive subject extraction is attributable to an independent ban on both P-stranding and pied-piping of those null heads. In ergative languages which do not exhibit an extraction asymmetry, ergative arguments are realized straightforwardly as DPs, which can therefore extract as normal like transitive objects and intransitive subjects.

(24) *Two types of ergative languages*　　(adapted from Polinsky 2016; p.46, p.48)

a.

```
              vP
            /    \
          PP      v'
         /  \    /  \
    P_ERG  DP  VP    v_[+ACTIVE]
              / \      [ABS]
            DP   V
          [uABS]
```

b.

```
              vP
            /    \
          DP      v'
        [uERG]   /  \
               VP    v_[+ACTIVE]
              / \    [ABS]; [ERG]
            DP   V
          [uABS]
```
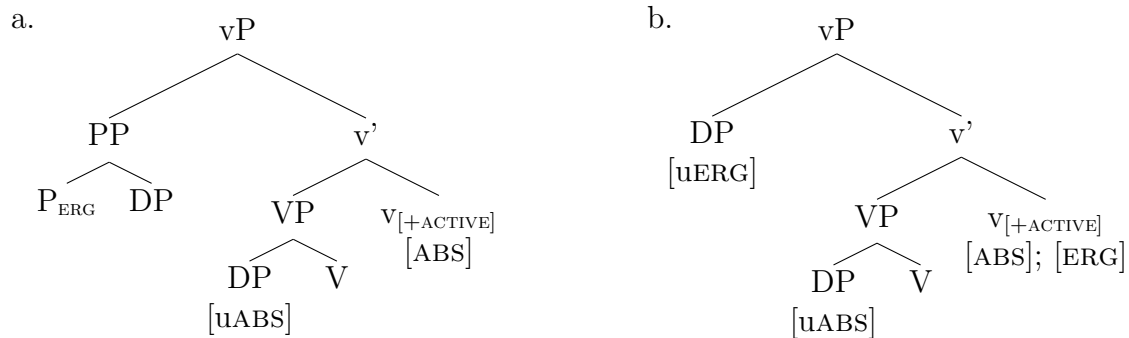
The structures that Polinsky proposes for these two types of ergative languages are in (24). In ERG-as-PP languages, inherent ergative case is assigned by the P head to the DP in its complement, while the PP as a whole still receives it's theta-role from the transitive v head. The transitive v head also assigns a Theta-role and structural absolutive case to the transitive object. In ERG-as-DP languages, the

266

basic configuration is the same, except that the transitive subject receives structural ergative case from the transitive v head.

One of the main motivations for this proposal comes from the observation that ergative alignment is often diachronically descended from passive constructions, where the transitive subject optionally occurs in PP form. In ERG-as-PP languages, the transitive subject PP has been reanalyzed from an optional adjunct to a core argument in the subject position, even if that P is phonologically null. Polinsky points out that there is some precedent for PP subjects cross-linguistically. In non-standard dialects of Russian, transitive objects and intransitive subjects are realized with no overt case marking or prepositions, but transitive subjects feature the preposition *u* 'by', as in (25).

Russian (non-standard dialects)                                        (Polinsky 2016; p.88, p.88)

(25) a. *u   lisicy    unese-n-o        kuročka*
         by fox-GEN carry-PART-N.SG chicken.F.NOM
         'a fox stole a chicken'

   b. *my        ne  spa-t-y        noč'*
      1PL.NOM not sleep-PART-PL night
      'we did not sleep all night'

In this non-standard dialect of Russian, the PP-subject may undergo pied-piping in Ā-extraction scenarios, as in (26). However, according to Polinsky's hypothesis, this is exactly what ergative languages with an ergative extraction asymmetry <u>cannot</u> do: in these languages, all PPs with null P heads cannot extract due to a ban on pied-piping (and P-stranding), which includes ergative PP subjects.

Russian (non-standard dialects)　　　　　　　　　　(Polinsky 2016; p.89, p.89)

(26) a. _muššiny_ [ _u_　_kogo_$_i$　_kanava_　　　t$_i$　_perehopa-n-o_　]
　　　　men　　　　by whom ditch.F.NOM　　　dig-PART-N.SG

　　　'the men who dug the ditch'

　　b. _u_　_kogo_　_ž_　　_èto_　_musor_　　　_vysypa-n-o_?
　　　　by whom EMPH DEIC garbage.M.NOM spill-PART-N.SG

　　　'who scattered all the garbage?'

　　　Furthermore, Polinsky can account for the asymmetry in Ā-extraction across domains from (11) as well. In some languages, restrictions on Ā-extraction are known to vary by domain. For example, in English, relativization of the dative argument in the Double Object construction (27a) is considered less acceptable than Wh-extraction of that same argument (27b). Therefore, by hypothesis, languages which exhibit an asymmetry across extraction domains, such as West Greenlandic in Section 5.1, are ERG-as-PP languages which feature a ban on pied-piping and P-stranding in relative clauses, but not in Wh-question formation.

English　　　　　　　　　　　　　　　　　　　　(Polinsky to appear; p.53, p.53)

(27) a. */?? _the person_$_i$ [ OP$_i$ _they sent_ t$_i$ _a threatening email_ ]

　　b. % _who_$_i$ _did they send_ t$_i$ _a threatening email_?

　　　In principle, Polinsky's theory makes the strong claim that ergative extraction asymmetries are directly explained by a parameterization of how ergative case is assigned. If the ergative argument is assigned inherent ergative case by a null P head as in (24a), it should exhibit a range of PP-like behavior, and by hypothesis an independent ban on P-stranding and pied-piping explains the ergative extraction

asymmetry. If, on the other hand, the ergative argument is assigned structural ergative case as in (24b), no extraction asymmetries are predicted to occur.

In actuality, however, Polinsky's theory is not so easily falsifiable. For example, she cites Tongan as a classic example of an ERG-as-PP language, where ergative arguments exhibit some PP-like properties. By hypothesis, Tongan ergative arguments are headed by a null P, which in turn explains the observed extraction asymmetries because of the independent ban on pied-piping and P-stranding. However, closely-related Niuean's ergative arguments exhibit many of the same PP-like properties as Tongan, and yet no ergative extraction asymmetry is observed. This contradiction is explained away by saying Niuean is a language in diachronic transition – it still retains many of the ergative-as-PP properties, but it is transitioning to ergative-as-DP, which explains the lack of an extraction asymmetry. This diachronic 'escape hatch' undermines the integrity of the theory: Niuean is a language which is predicted to exhibit extraction asymmetries based on the PP-like profile of its ergative arguments, and yet it does not, but this is allowed in the case of languages which are in diachronic transition. The question then becomes: is there any independent evidence for this status of diachronic transition? What predictions does this theory really make?[6]

---

[6] Again, the disambiguation hypothesis in Section 5.3 makes better predictions with regard to these two languages. Niuean is correctly predicted to lack an extraction asymmetry due to its consistent morphological case (see Section 5.4.4). However, Polinsky (2016) writes that in Tongan, ergative and absolutive case morphology are often omitted in rapid, casual speech (p.190). If this is true, then the fact that they exhibit an extraction asymmetry follows naturally: morphological case is not as reliable a cue as in Niuean, thus making Ā-extraction potentially ambiguous!

### 5.2.4 Deal (2016)

Deal (2016) argues for a case-discrimination analysis. In her view, the explanation for ergative extraction asymmetries lies not in the mechanics of case assignment itself, but rather a simple case hierarchy. Borrowing the morphological case hierarchy in (28) from Marantz (1991) and Bobaljik (2008), the analysis states that in languages with no ergative extraction asymmetries, both dependent ergative case and unmarked absolutive case are free to undergo Ā-extraction. However, in languages which exhibit an asymmetry, dependent ergative case is simply pre-specified for the inability to extract, while unmarked absolutive case can still extract freely. This marks a departure from the other analyses in this section; as Deal writes, "the crucial factor is now the morphological case system, rather than the way that subjects are syntactically licensed".

(28) *The case-discrimination hierarchy in Deal (2016)*

<div align="center">

unmarked case ≫ dependent case ≫ lexical/oblique case

</div>

Indeed, Deal's analysis crucially hinges on the operation Agree: "a DP's ability to participate in the operation Agree may be determined in part by its morphological case". In her view, if a head is going to attract an XP to its specifier during the course of Ā-extraction, that head must Agree with the XP first. Furthermore, particular heads can be restricted in only being able to Agree with XPs with certain case values in accordance with the hierarchy in (28). In some languages, unmarked absolutive case and dependent ergative case will both be pre-specified as eligible for that Agree operation, therefore no asymmetry will occur in Ā-extraction. However,

if a language is parameterized such that dependent ergative case is not eligible for Agree prior to Ā-extraction, an ergative extraction asymmetry will occur.

Accounting for the domain asymmetries in (11) then becomes relatively straightforward. The restrictions for which types of arguments can extract are allowed to vary by domain. Therefore, in West Greenlandic Wh-questions, both unmarked absolutive and dependent ergative case can enter into an Agree relationship and therefore undergo Ā-extraction, but in relative clauses, only unmarked absolutive is able to do so.

While simple, this analysis makes very strong predictions. In no language should unmarked case be restricted with regards to extraction, while dependent case is not. In other words, the implicational hierarchy in (28) states that if dependent case can be freely extracted, then unmarked case should always be free to extract as well, regardless of alignment. In Section 5.5, data from both morphologically ergative and morphologically accusative languages will challenge this strong prediction made by Deal's case-discrimination analysis.

Finally, while Deal's analysis makes clear predictions for ergative languages with overt morphological case, it is not so clear what predictions are made when it comes to languages which lack overt case morphology, but still exhibit an extraction asymmetry. Deal considers agreement-based asymmetries such as Kaqchikel (4-5) and Q'anjob'al (17) a separate phenomenon altogether, writing that "various pure head-marking languages use special agreement morphology when transitive subjects are extracted". In her view, the observed change in verbal morphology is simply an overt instance of Wh-agreement, not a restriction on the extraction of specific arguments. In the absence of any explicit theory of how this process takes place,

271

however, we are left without any clear typological predictions with regards to what should and should not be allowed in the purely head-marking class of languages.

### 5.2.5 Interim summary

To varying degrees, each of the theories covered in this section seeks to model ergative extraction asymmetries in some kind of connection with ergative morphology. For Aldridge (2008, 2012), absolutive-marked transitive objects must move to a phase boundary position for [EPP] purposes, which in turn blocks any extraction of the ergative-marked transitive subject. Coon et al (2014) reproduce a similar blocking effect for Mayan 'high ABS' languages, except here the transitive object must move to the edge of the phase boundary in order to receive absolutive case directly from the I head. Polinsky (2016) proposes that the explanation for these extraction asymmetries lies in a parametric difference in the world's morphologically ergative languages: in some languages, ergative arguments are actually PPs with a null head, and the ban on extraction is attributed to a more general ban on pied-piping and P-stranding. Finally, for Deal (2016), the connection between morphological case and ergative extraction asymmetries could not be more direct: simply by virtue of being marked with dependent ergative case, in some languages those arguments cannot undergo Ā-extraction.

In the next section, a very different hypothesis is put forward: the explanation for ergative extraction asymmetries has nothing to do with ergative case, ergative agreement, or the morphosyntactic mechanisms that license these arguments. Instead, ergative extraction asymmetries are predicted by a need for disambiguation: if, *a priori*, Ā-extraction from transitive clauses is ambiguous, then that language

is predicted to exhibit some kind of asymmetry which functionally serves to disambiguate those instances of Ā-extraction.

## 5.3 The disambiguation hypothesis

The typological facts, repeated in (29), are as follows: some languages, such as Kaqchikel, exhibit an ergative extraction asymmetry across all domains; some languages, like Niuean, do not exhibit any such asymmetry; and some languages, like West Greenlandic, only exhibit an asymmetry in a subset of extraction domains (typically in relative clauses only).

(29) *Morphological ergativity and ergative extraction asymmetries: 3 types*

| | **morphological ergativity** | **ergative extraction asymmetry** | |
|---|---|---|---|
| | | Wh-question | relative clause |
| Kaqchikel, Gitksan, Selayarese | ✔ | ✔ | ✔ |
| West Greenlandic, Roviana | ✔ | ✘ | ✔ |
| Niuean, Yukulta, Basque | ✔ | ✘ | ✘ |

The rest of this chapter is devoted to arguing that these asymmetries serve primarily as a **grammaticalized disambiguation strategy** in ambiguous extraction scenarios. The presence of an ergative extraction asymmetry is predicted by a language's typological profile – a combination of basic word order, Wh-strategy, relativization strategy, and morphological case.

In morphologically ergative languages which feature a set of grammatical features such that Ā-extraction will always be clear with regards to which specific argument is being extracted, no extraction asymmetry is predicted to occur. If, however, the typological profile of a language is such that it lacks sufficient cues, and

as a result Ā-extraction of an argument would potentially be ambiguous, then those languages are predicted to exhibit an extraction asymmetry. This new approach can also provide a principled explanation for West Greenlandic-type languages in (29): given that languages may differ in terms of the exact extraction strategy used to form Wh-questions and relative clauses, it follows that if one type of extraction is potentially ambiguous, but the other is not, then only in that first domain will a disambiguating extraction asymmetry be warranted.

Section 5.3.1 provides an introductory discussion of the function of ergative extraction asymmetries in this new framework. Then, the relevant factors that contribute to a language's typological profile are discussed in turn: morphological case (5.3.2), basic word order (5.3.3), Wh-strategy (5.3.4), and relativization strategy (5.3.5). In each case, it is shown that certain configurations produce ambiguous extraction scenarios – except for verb agreement (5.3.6), which is argued to not be a predictor one way or another.

The end result is a theory where a language's typological profile – the sum of relevant grammatical features discussed above – makes a prediction as to whether that language will exhibit an extraction asymmetry. Section 5.3.7 concludes by briefly discussing the sample of ergative languages on which the disambiguation hypothesis will be tested in Section 5.4.

### 5.3.1 What is the function of extraction asymmetries?

While morphologically ergative languages can exhibit all kinds of basic word orders, it seems to be a fact that many of those languages which also exhibit an

ergative extraction asymmetry tend to be V-initial.[7,8] For example, the basic word order in Kaqchikel is VOS, as in (30). Gitksan has VSO basic word order (31) and it also exhibits an ergative extraction asymmetry (32-33). In both Wh-questions and relative clauses of Gitksan, Ā-extraction of the transitive subject requires a 'relativizer' morpheme ʔan immediately following the extracted element.[9]

Kaqchikel (Erlewine 2013; p.2)

(30) *iwïr    x-u-tëj         ri  wäy   ri a Juan*
     yesterday CMPL-3SG.ERG-eat the tortilla Juan
     'yesterday Juan ate the tortilla'

Gitksan (Rigsby 1986; p.260, p.303, p.303, p.303, p.404, p.407, p.405)

(31) *ɬəmoː-yə=s    Tom-t     Mary*
     help-TR=CNN Tom=CNN Mary
     'Tom helped Mary'

(32) a. *naa ʔan=t          qay   ɬəmoo=s  Bruce*
        who A.REL=3SG.ERG.DEP which help=CNN Bruce
        'who helped Bruce?'

---

[7] In Chapter 2, SOV was found to be by far the dominant basic word order (n=24) in that sample of 40 morphologically ergative languages. V-initial basic word order was found in 9 languages, leaving only 6 languages with V-medial basic word order. This distribution is thus consistent with Mahajan's (1997) generalization, which states that morphologically ergative languages tend to be verb peripheral in their basic word order.

[8] See also Polinsky (2016) for an account of the correlation between V-initial word order and syntactic ergativity. In her view, PP ergative arguments are ineligible to move to spec-IP for [EPP] purposes, so in those languages the verb head itself must do so instead. As she points out, however, it is not clear how that would account for V-initial order in an intransitive clause, given that the sole argument is not a PP and thus should be eligible for movement to satisfy the [EPP].

[9] For a discussion of ergative alignment in Gitksan verb agreement, not directly visible in (31), see Section 2.3.2. Also, note that according to Peterson (2015), connectives such as -t (32a) and =ɬ (32c) are not considered true instances of case marking.

b. *naa=ɬ   ɬʷəmoo-yə-n*
who=CNN help-TR-2SG.ERG

'who did you help?'

c. *naa=ɬ   lim-ət*
who=CNN sing-S.REL

'who sang?'

(33) a. *kaʔ-ə-'y=ɬ        <u>kat</u>  [ ʔan-cakʷ-ə=ɬ        naks-t        ]*
see-TR-1SG.ERG=CNN man   A.REL-kill-INCR=CNN spouse-3SG.POSS

'I saw the man who killed his wife'

b. *maɬ-tə-'y       lo:-tit   tim  kʷixʷ-ə=s    John=ɬ   <u>smax</u>  [*
tell-TR-1SG.ERG OBL-3PL FUT shoot-TR=CNN John-CNN bear

*ɬə-kaʔ-n        ]*
DEF-see-2SG.ERG

'I told them that John would shoot the bear that you saw'

c. *tim  t'is-ə=ɬ       xaʔ=ɬ    <u>kat</u>=ɬ   [ paẓ-ət=kə        ]*
FUT hit-TR=CNN slave=CNN man=CNN  run-S.REL=DIST

'the slave will hit the man who ran'

In Ā-extraction scenarios, it is exactly these kinds of V-initial languages that run into problems with potential ambiguity. Imagine a hypothetical language, Language $X_1$, which has the following characteristics: VSO basic word order, no morphological case and Wh-fronting. As the schema in (34) illustrates, the basic typological profile of the language fails to provide enough cues to disambiguate cases of transitive subject and transitive object extraction. In (34a), the transitive subject is extracted to its pre-verbal position, leaving the object in its base-generated, post-verbal position. In (34b), unfortunately, the exact same description applies to transitive object

argument extraction: it moves to its fronted, pre-verbal position, while the subject stays in its base-generated, post-verbal position.

Language X$_1$ (VSO, ✗ m-case, Wh-fronting)

(34) a.  WH$_?$      V    $t$    NP        transitive subject (A) extraction

    b.  WH$_?$      V    NP    $t$        transitive object (O) extraction

    c.  WH$_S$      V    $t$        intransitive subject (S) extraction

Functionally, this is a problem because it renders such simple Wh-questions as 'who saw Tom?' ambiguous with 'who did Tom see?', hardly an ideal scenario. Now imagine a typologically very similar language, Language X$_2$, which marks instances of transitive subject extraction only with an additional morpheme $x$, as in (35). Suddenly, transitive subject and transitive object Wh-questions are no longer ambiguous: the presence of the morpheme $x$ crucially distinguishes (35a) from (35b). As it turns out, the schema for Language X$_2$ corresponds exactly to Kaqchikel, Gitksan, and many other languages which exhibit ergative extraction asymmetries.

Language X$_2$ (VSO, ✗ m-case, Wh-fronting)

(35) a.  WH$_A$    $x$    V    $t$    NP        transitive subject (A) extraction

    b.  WH$_O$      V    NP    $t$        transitive object (O) extraction

    c.  WH$_S$      V    $t$        intransitive subject (S) extraction

In this way, the morphology that characterizes an ergative extraction asymmetry serves to disambiguate in otherwise ambiguous extraction scenarios.[10]  Of course, this is not the first time the claim has been made; the same observation has already been made for languages from well-known ergative families (see Stiebels 2006 for Mayan languages; see Finer 1997 for Selayarese, an Austronesian language). However, the rest of this chapter will be devoted to arguing for this position from a typological perspective.  Looking at a wide range of languages from different language families, I will argue that once we establish where we *would* and *would not* expect to observe an extraction asymmetry on the basis of a language's available cues, the disambiguation hypothesis makes very good predictions cross-linguistically.

### 5.3.2   Morphological case

Functionally, morphological case serves to identify arguments' grammatical roles, and this remains true in Wh-questions. Consider now Language $X_3$, which has the following profile: VSO basic word order, morphological case, and Wh-fronting. Morphological case provides the vital cue which disambiguates transitive subject extraction in (36b) and transitive object extraction in (36c); specifically, ERG case on the Wh-element in the former, and ABS in the latter. Therefore, in any language with overt morphological case, ergative extraction asymmetries are not functionally-motivated and therefore they are predicted not to occur.

---

[10]  Naturally, the functional need for disambiguation would be equally satisfied if transitive object Ā-extraction in (35b) were exceptionally marked instead. See Section 5.5 for instances of exactly that configuration, which is perfectly consistent with the disambiguation hypothesis.  See also discussion in that section of how morphological alignment can influence – but not guarantee – which argument is 'picked out' by an extraction asymmetry.

**Language X₃** (VSO, ✔ m-case, Wh-fronting)

(36)     a.           V   NP$_{\text{ERG}}$   NP$_{\text{ABS}}$           basic clause

         b.   *Wh*$_{\text{ERG}}$   V           NP$_{\text{ABS}}$           transitive subject (A) extraction

         c.   *Wh*$_{\text{ABS}}$   V           NP$_{\text{ERG}}$           transitive object (O) extraction

         d.   *Wh*$_{\text{ABS}}$   V                                       intransitive subject (S) extraction

Note that even though Language X₃ is modeled as VSO (36), that specific basic word order isn't actually contributing to disambiguation. Morphological case is an independent disambiguator: as long as arguments are clearly indicated with morphological case, these languages will always have clear, unambiguous extraction, and therefore they are predicted never to exhibit any ergative extraction asymmetries.

Consider now a slight variant on Language X₃ – one in which the extracted element is not marked with morphological case, but other NPs in the clause still are (this could occur in a variety of Ā-extraction scenarios, such as relative clauses or clefted Wh-constructions). As the schema in (37) shows, this is still a sufficient cue to disambiguate transitive subject Wh-questions from transitive object Wh-questions. In (37b), the Wh-element may be bare, but the immediately following non-questioned element bears absolutive case, identifying the Wh-element as the transitive subject; in the same way, in (37c), the non-questioned element bears ergative case, which in turn identifies the Wh-element as the transitive object. Even though these morphological

cues are slightly less abundant than those in (36), the hypothesis still predicts that they should be sufficient to obviate any need for disambiguating morphology.

Language X$_4$ (VSO, ✔ m-case, Wh-fronting)

(37)  a.      V  NP$_{\text{ERG}}$  NP$_{\text{ABS}}$          basic clause

      b.  *Wh*  V        NP$_{\text{ABS}}$          transitive subject (A) extraction

      c.  *Wh*  V        NP$_{\text{ERG}}$          transitive object (O) extraction

      d.  *Wh*  V                            intransitive subject (S) extraction

The schemas in (36-37) clearly indicate that independently of other grammatical features, morphological case should be a surefire disambiguator, and therefore any language with consistent, overt morphological case is predicted to never exhibit ergative extraction asymmetries.

### 5.3.3  Basic word order

A language's basic word order can also provide a sufficient cue to inherently disambiguate in Ā-extraction scenarios. For example, in (38b), Wh-extraction of the transitive subject results in an Wh V NP sequence, while in (38c), Wh-extraction of the transitive object yields the sequence Wh NP V instead, as the Wh-fronted element jumps ahead of the usually clause-initial transitive subject.

Language $X_5$ (SVO, ✔ Wh-fronting, ✘ m-case)

(38)    a.          $NP_A$   V   $NP_O$          basic clause

        b.          $Wh_A$   V   $NP_O$          transitive subject (A) extraction

        c.    $Wh_O$   $NP_A$   V          transitive object (O) extraction

        d.          $Wh_S$   V          intransitive subject (S) extraction

In this way, basic word order can provide a sufficient cue to disambiguate in extraction scenarios. Crucially however, SVO basic word order in Language $X_5$ is not independently sufficient to disambiguate; in that language, it is the <u>combination</u> of SVO and Wh-fronting which produce two distinct sequences of linear word order in each type of transitive question.

In principle, just like in Wh-questions, basic word order can also disambiguate between relativized transitive subjects and relativized transitive objects. In the schema in (39), Language $X_6$ features post-nominal relative clauses, no morphological case, and an SVO basic word order. As a result, relativization of the transitive subject forms the linear sequence N [ NP V ], while relativizing the transitive object has the distinct sequence N [ V NP ]. In this way, a fixed basic word order can provide disambiguation in relative clauses, as well.

Language $X_6$ (S V O, ✘ m-case, post-N)

(39)    a.       NP$_A$   V   NP$_O$              basic clause

        b.  $\underline{N}_A$  [       V   NP$_O$  ]         transitive subject (A) extraction

        c.  $\underline{N}_O$  [ NP$_A$   V        ]         transitive object (O) extraction

        d.  $\underline{N}_S$  [       V        ]         intransitive subject (S) extraction

### 5.3.4  Wh-strategy

The choice of Wh-extraction strategy is also relevant for the disambiguation hypothesis. For instance, we have already seen that, in the right circumstances, a language which elects to produce questions via Wh-fronting may actually produce ambiguous ones. In Language X$_1$, a combination of Wh-fronting, V-initial word order and a lack of morphological case produces transitive subject extraction which is ambiguous with transitive object extraction (35).

However, if the typological profile is right, even languages with Wh-fronting can produce unambiguous Wh-questions without recourse to ergative extraction asymmetries. For instance, Languages X$_3$ (36) and X$_5$ (38) were shown to benefit from morphological case marking, and V-medial word order, respectively, to produce unambiguous extraction. Therefore Wh-fronting as an extraction strategy may or may not lead to the kind of ambiguity that requires an extra 'something' to disambiguate.

Language X$_7$ (SOV, ✗ m-case, Wh-*in situ*)

(40)    a.  NP$_A$   NP$_O$   V              basic clause

        b.  *Wh*$_A$   NP$_O$   V          transitive subject (A) extraction

    c.   NP$_A$   *Wh*$_O$   V             transitive object (o) extraction

    d.       *Wh*$_S$     V            intransitive subject (s) extraction

In addition to Wh-fronting, Wh-*in situ* is the other extremely common Wh-question formation strategy among the world's languages (Dryer 2013). Clearly, Wh-*in situ* is an inherently unambiguous Wh-question formation strategy. Consider Language $X_7$ in (40), which has the following profile: SOV basic word order, no m-case, and Wh-*in situ*. There is no possible ambiguity in Wh-extraction: if the transitive subject is questioned, as in (40b), it stays in its base-generated, clause-initial position; if the transitive object is questioned, as in (40c), it stays reliably sandwiched between the transitive subject and the verb. Once again, it is a combination of two grammatical features – consistent basic word order, and Wh-*in situ* – which provide sufficient cues in these Wh-extraction scenarios. Therefore, the disambiguation hypothesis predicts that no language with Wh-*in situ* should exhibit an ergative extraction asymmetry in question formation.

### 5.3.5   Relativization strategy

Like for Wh-questions, the choice of extraction strategy when forming relative clauses is relevant for the disambiguation hypothesis. In general, the same principles which determine potential ambiguity in Wh-questions will also apply in the domain of relativization. For example, we have already seen that Language $X_6$ has pre-nominal relative clauses and consistent SVO basic word order, which together produce unambiguous relative clauses (39).

Similarly, consider Language $X_8$, where morphological case is predicted to

always disambiguate the extracted element. The relativization strategy used is post-nominal, externally-headed relative clauses, and in addition there is overt morphological case. In (41b), even though the relativized element doesn't bear any morphological case, the absolutive marker on the embedded, non-extracted element identifies the head of the relative clause as its transitive subject.[11] Likewise, in (41c), the ergative marker on the embedded element identifies the head of the relative clause as its transitive object.

Language $X_8$ (SOV, ✔ m-case, post-nominal)

| (41) | a. | $NP_{ERG}$ | $NP_{ABS}$ | V | | basic clause |
|------|-----|------------|------------|---|---|--------------|
| | b. | <u>N</u> | [ $NP_{ABS}$ | V | ] | transitive subject (A) extraction |
| | c. | <u>N</u> | [ $NP_{ERG}$ | V | ] | transitive object (O) extraction |
| | d. | <u>N</u> | [ | V | ] | intransitive subject (S) extraction |

In addition to externally-headed relativization, some languages opt for a rarer strategy: internally-headed relative clauses. In (42), a relative clause's head does not undergo any overt extraction, and instead it remains *in situ* within the embedded clause. The interpretation of an intransitive internally-headed relative clause is relatively straightforward, as in (42c): given that there is only one argument available, it must be the head of its internally-headed relative clause. In transitive cases, however,

---

[11] Of course, the head of an externally-headed relative clause may bear morphological case, but only as a function of its role within a main clause; it does not constitute a cue towards disambiguation with regards to Ā-extraction. Therefore, under the disambiguation hypothesis, morphological case on the heads of relative clauses is considered irrelevant for the purposes of identifying the relativized element.

an interpretive problem occurs. On the surface, both (42a) and (42b) look identical; even the availability of overt morphological case is no longer helpful.

Language X$_9$ (SOV, ✔ m-case, internally-headed)

(42)    a.   NP$_{ERG}$       NP$_{ABS}$   V                        basic clause

       b.   [   $\underline{NP}_{ERG}$   NP$_{ABS}$   V   ]        transitive subject (A) extraction

       c.   [   NP$_{ERG}$   $\underline{NP}_{ABS}$   V   ]        transitive object (O) extraction

       d.        [   $\underline{NP}_{ABS}$       V   ]        intransitive subject (S) extraction

Therefore, the disambiguation hypothesis predicts that internally-headed relative clauses, by virtue of being inherently ambiguous, will always warrant additional disambiguating material or interpretation strategies.

### 5.3.6   Can verb agreement be a sufficient cue?

Many of the morphologically ergative languages of the world are morphologically ergative by virtue of their system of head-marking. Therefore, we might expect that perhaps verb agreement could serve as a sufficient cue – a counter-predictor of ergative extraction asymmetries – alongside the other factors discussed in Sections 5.3.2-5.3.5.

The schema in (43) illustrates some possible scenarios where verb agreement does indeed help to disambiguate the grammatical role of the extracted element. In principle, extraction of the Wh-element across the verb would lead to ambiguity,

exactly like for Language $X_1$ (34). However, in (43a), the fact that the verb features 1SG absolutive agreement indicates that the post-verbal 1SG pronoun is in fact the transitive object in that clause; it then follows that the Wh-element corresponds to the transitive subject. In exactly the same way, in (43b), the verb's 1SG ergative agreement indicates that the 1SG pronoun is the transitive subject; it then follows that now, the Wh-element corresponds to the transitive object.

Language $X_{10}$ (VSO, ✘ m-case, ✔ verb agreement, Wh-fronting)

(43)  a.  WH$_A$      V-1SGABS      $t$      1SG$_{pro}$

     b.  WH$_O$      V-1SGERG      1SG$_{pro}$   $t$

However, verb agreement may not always be a <u>reliable</u> cue to help disambiguate. Specifically, consider $3 > 3$ transitive scenarios, schematized in (44). The presence of either ergative or absolutive verb morphology – or both! – does not help to identify the grammatical roles of the Wh-element and the 3SG pronoun, because they are both essentially 3SG and thus they can both be encoded by the same verb morphology.

(44)  WH$_?$      V-3SGERG      $t?$      3SG$_{pro}$   $t?$

Since these $3 > 3$ scenarios are bound to be extremely common in language use, verb agreement alone cannot be considered a sufficient cue to obviate the need for an extraction asymmetry under the disambiguation hypothesis. That being said, it does lead to a very important observation: if a language does grammaticalize an

ergative extraction asymmetry to help in scenarios like (44), as the disambiguation hypothesis predicts, it seems very likely that the grammar would extend that extraction asymmetry to <u>all</u> instances of transitive subject extraction, even unambiguous ones!

For example, Language $X_{11}$ features an identical typological profile to Language $X_{10}$ above, except that it does feature an ergative extraction asymmetry as predicted by the disambiguation hypothesis. In (45a), the presence of the morpheme $x$ crucially identifies an extracted transitive subject, while its absence in (45b) indicates an extracted transitive object. Furthermore, the use of this ergative extraction asymmetry has been generalized to all instances of transitive extraction, even those in which verb agreement <u>does</u> help to disambiguate. In (46a), the clause features the morpheme $x$ simply by virtue of being an instance of transitive subject extraction, even though (in this particular instance) verb agreement would have been sufficient to disambiguate.

Language $X_{11}$ (VSO, ✗ m-case, ✔ verb agreement, Wh-fronting

(45)  a.  WH$_A$   $x$   V-3SGABS   $t$   3SG$_{pro}$

    b.  WH$_O$   V-3SGERG   3SG$_{pro}$   $t$

(46)  a.  WH$_A$   $x$   V-1SGABS   $t$   1SG$_{pro}$

    b.  WH$_O$   V-1SGERG   1SG$_{pro}$   $t$

To conclude the discussion in this section, verb agreement alone cannot be considered a sufficient cue to rule out an ergative extraction asymmetry, even though

in a subset of transitive scenarios, verb agreement may indeed disambiguate. As a corollary, it seems highly likely that if an extraction asymmetry is warranted largely due to ambiguous $3 > 3$ transitive scenarios, then that asymmetry will be grammaticalized to occur in all instances of transitive subject extraction, and not just those $3 > 3$ scenarios. In brief, ergative extraction asymmetries might occur in individual sentences which are not themselves ambiguous, but the asymmetry might nevertheless be well-motivated by the language's broader typological profile.[12]

### 5.3.7  An (extended) sample of ergative languages

In Sections 5.3.1-5.3.6, the disambiguation hypothesis was fleshed out. Specific grammatical features, alone or in combination, contribute to a language's typological profile and thus make a prediction as to whether an extraction asymmetry is functionally-motivated in that language. These factors are: morphological case, basic word order, Wh-strategy and relativization strategy. Specifically, the typological profiles that are predicted to exhibit an ergative extraction asymmetry include:

---

[12] The fact that an ergative extraction asymmetry is initially motivated by ambiguous $3 > 3$ scenarios, but eventually extended to all transitive scenarios is not necessarily a guarantee. One might expect that some languages exhibit an extraction asymmetry only in $3 > 3$ scenarios, but not others. This is exactly what happens in Tzotzil, a Mayan language. According to Aissen (1979), transitive subject extraction requires the Agent Focus construction only if the transitive object is third person, as in (i). However, if the object is non-third person, the Agent Focus is no longer required, as in (ii).

Tzotzil                                                                (Aissen 1979; p.459, p.465)

i) *k'usi  i-ti'-on*?                            ii) *k'usi  l-a-s-ti'*?
   what CMPL-eat-AF                              what CMPL-2SG.ABS-3SG.ERG-eat
   'what bit him?'                               'what bit you?'

A configuration such as Tzotzil's is perfectly consistent with the disambiguation hypothesis: in Tzotzil, the ergative extraction asymmetry occurs exactly where it is functionally required. The facts in (i-ii) also indicate that, when checking to see whether a language exhibits an extraction asymmetry, it is best to examine $3 > 3$ sentences at a minimum: the lack of an apparent asymmetry in other transitive scenarios may not necessarily indicate that no asymmetry exists at all.

V-initial languages with Wh-fronting and/or externally-headed relative clauses but no morphological case, V-final languages with Wh-fronting and/or externally-headed relative clauses but no morphological case, and languages which employ the inherently ambiguous internally-headed relative clause strategy. In Section 5.4, these predictions are tested against a sizable sample of morphologically ergative languages.

(47) *Morphologically ergative languages in the extraction asymmetry sample* (n=53)

Chapters 2 & 5:  Abkhaz, Ayutla Mixe, Basque, Burushaski, Canela-Kraho, Cavineña, Epena Pedee, Georgian, Gitksan, Halkomelem, Hindi, Hua, Itzaj, Kaluli, Kapampangan, Kaqchikel, Lezgian, Lhasa Tibetan, Ngiyambaa, Niuean, Paumarí, Seediq, Shipibo, Sinaugoro, Sm'algyax, Sorani Kurdish, Suena, Tiriyó, Trumai, Yakima Sahaptin, Yingkarta, Yukulta, Warlpiri, West Greenlandic (n=34)

Chapter 2 <u>only</u>   Hurrian, Ika, Ku Waru, Tenetehára, Päri, Sumerian (n=6)

Chapter 5 <u>only</u>   Apinajé, Belhare, Cupeño, Diyari, Drehu, Ixil, Katukina, Kuikuro, Karitiâna, Kolyma Yukaghir, Lillooet, Matses, Movima, Q'anjob'al, Roviana, Selayarese, Tseltal, Tukang Besi, Yucatec (n=19)

The ergative sample in Chapter 5 is largely based on the ergative sample in Chapter 2. Whenever possible, languages that were examined for ergative syncretisms in that chapter were also checked for any extraction asymmetries. In total,

34 out of the 40 languages from Chapter 2 also served in the extraction asymmetry sample of this chapter. Those languages are listed in (47).

Some languages from Chapter 2, however, were excluded from this sample in Chapter 5. The reason in all six cases was straightforward: a lack of clear data on Ā-extraction. Those languages' sources were sufficient for determining the existence of syncretism, but not an extraction asymmetry.

In addition, the 34 languages from Chapter 2 were supplemented by 19 new languages. They are also listed in (47). As much as possible, inclusion within this sample was based on the principles discussed in Chapter 2, and the same biases mentioned by Bakker (2011) were avoided (bibliographic bias, genetic bias, areal bias, typological bias and cultural bias).

In total, then, the sample of morphologically ergative languages in this chapter consists of 53 languages. However, when reporting on the results in Section 5.4.7, the success rate of the disambiguation hypothesis will actually be measured in terms of **data points**, and not languages. Here, a "data point" is defined as either a) **Wh-question**, b) **relative clause**, or in some cases c) **focus-fronting data** for one given language. The motivation for this distinction is as follows: in some languages, the disambiguation hypothesis may make a good prediction for its relative clauses, but not its Wh-questions (or vice versa). Therefore it would not be accurate to say that the hypothesis makes either good or bad predictions for that language, but rather the two data points must be teased apart and individually evaluated.

Indeed, in some languages, there may actually be two distinct constructions/ strategies within a single type of Ā-extraction. For example, Hungarian features two very different relativization strategies. In (48a), a post-nominal relative clause is

formed via relative pronoun; in (48b), a pre-nominal relative clause is formed with a verb in its participial form. In this manner, these two relativization strategies would count as separate two data points, despite coming from the same language, and they would each be evaluated individually with regards to the disambiguation hypothesis.

Hungarian (Kenesei et al 1998; p.38, p.46)

(48) a. *a* <u>*könyv*</u> [ *amely-et* *Anna olvas-ott* ] *érdekes* *volt*
DET book which-ACC Anna read-PAST interesting was
'the book that Anna was reading was interesting'

b. *az* [ *Anna által tegnap* *olvas-ott* ] <u>*könyv*</u>
DET Anna by yesterday read-PART book
'the book read by Anna yesterday'

This approach will ultimately allow for a more accurate assessment of the disambiguation hypothesis across this new sample. In total, the dataset comprises 99 total ergative datapoints. On average, clear Wh-question and relativization data were available for each language, however in some cases only one of the two was available, and in other cases, one language has more than one strategy available for a given type of Ā-extraction.

## 5.4 Testing out the disambiguation hypothesis

Now that the predictions of the disambiguation hypothesis have been clearly laid out, and the nature of the sample has been discussed, it is time to evaluate whether the hypothesis does indeed make good predictions. This section will proceed as follows. First, it will be shown that the factors predicting the absence of an ergative extraction asymmetry are quite reliable. In Sections 5.4.1-5.4.3, multiple examples of morphologically ergative languages which lack ergative extraction

asymmetries are presented, as predicted by their typological profile. Then, in Section 5.4.4, examples of actual ergative extraction asymmetries are presented, again as predicted by typological profile of those languages and their for potential ambiguity in Ā-extraction. Section 5.4.5 provides a brief discussion of the Austronesian voice system, and its place within the disambiguation hypothesis. Section 5.4.6 discusses instances where the disambiguation hypothesis makes an incorrect prediction, and finally Section 5.4.7 provides a summary of the quantitative data for the entire dataset.

### 5.4.1 Morphological case as a disambiguator

Shipibo is the first of many languages in the sample for which morphological case successfully predicts the lack of a disambiguation strategy. In (49a), the transitive subject Wh-element is marked with ergative case morphology; in (49b) and (49c), respectively, the transitive object and intransitive subject Wh-elements feature null absolutive morphology. In this way, Wh-questions are easily disambiguated by morphological case in Shipibo. Therefore, because there is no additional morphosyntax that singularly characterizes transitive subject Wh-extraction, we can thus conclude that Shipibo does not exhibit any kind of ergative extraction asymmetry in Wh-questions, exactly as expected under the disambiguation hypothesis.

Shipibo                                                  (Valenzuela 2003; p.378, p.379, p.195)

(49) a. *tso-n-ki       Tsoma      atsa  xeati      xea-ma-ibat-a*
who-ERG-INT Tsoma.ABS yuca drink.ABS drink-CAUS-PAST-PART
'who invited Tsoma to drink yuca beer?'

b. *jawe-ki     Wesna-n    Tsoma     xea-ma-ibat-a*
what.ABS-INT Wesna-ERG Tsoma.ABS drink-CAUS-PAST-PART

'what did Wesna invite Tsoma to drink?'

c. *jawe-rin      nato*
what.ABS-INT2 this

'what is this?'

Yukulta is another language with overt morphological case. In (50a), the *-ya* ergative case marker identifies the Wh-element as being the transitive subject; in (50b), the lack of case morphology on the Wh-element identifies it as corresponding to the absolutive transitive object. Again, there is no ergative extraction asymmetry to report which characterizes transitive subject Wh-extraction in Yukulta.

Yukulta                                           (Keen 1972; p.174, p.175, p.226)

(50) a. *ŋaka-ya   =ṛukanta        kurka ŋitjinta    milyaḻṭa*
who-ERG  =3SG>1SG.PAST take  1SG.POSS spear

'who took my spear?'

b. *ŋaka  =yikanti    yanma    ṛiya*
what  =2SG.FUT today.INT eat

'what will you eat today?'

c. *ŋaka-yana =ṇiŋki    puri*
who-DUB    3SG.FUT come

'I don't know who will come'

Paumarí is yet another language in which case morphology crucially distinguishes transitive subject Wh-questions from transitive object Wh-questions. In (51a), ergative case morphology identifies the Wh-element as the transitive subject

in the clause. In (51b), the sentence is formed based on an alternate OVS clausal structure in which the transitive object is marked with *-ra* accusative case, instead.

Paumarí          (A. Aikhenvald; p.c. / Chapman & Derbyshire 1986; p.242, p.206)

(51) a. *nahina-a bi-na-abina-hi          ida    ojoro*
        who-ERG 3SG.ERG-CAUS-die-THEME.F DEM.F turtle
        'who killed the turtle?'

     b. *nahina-ra vi-fini-vini     adani    isai*
        what-ACC 3PL.SUBJ-fear-TR DEM.PL children
        'what are the children afraid of?'

     c. *nahina ida     soko-a-ja*
        who     DEM.F wash-DETRANS-IMMED
        'who is washing?'

The availability of two alternate clausal types, each exhibiting a different morphological alignment, is what Aikhenvald (2012) refers to as "the Paumarí Puzzle". Nevertheless, this does not qualify as an extraction asymmetry; the Wh-constructions in (51a-51b) correspond in morphosyntax to their basic declarative counterparts in (52a-52b). In other words, the Paumarí data is still consistent with the overall hypothesis because case morphology disambiguates Wh-extraction, and no additional morphosyntax which could be construed as an ergative extraction asymmetry characterizes instances of transitive subject extraction only.

Paumarí                              (Chapman & Derbyshire 1986; p.164, p.196)

(52) a. *Dono-a    bi-ko'diraha-'a-ha          ada    isai hoariha*
        Dono-ERG 3SG.ERG-pinch-ASP-THEME.M DEM.M child other
        'Dono pinched the other boy'

b. *i'oa-ra*        *na-hado-ha*        *ada*     *kodi-abi'i*
<br>tambaqui.fish-ACC CAUS-knife-THEME.M DEM.M 1SG.POSS-father
<br>'my father cut the *tambaqui* fish'

Looking at some relative clause data now, consider the typological profile of Georgian: relative clauses are post-nominal, and both the embedded element and the relative pronoun bear morphological case (53). For example, in (53a), the relative pronoun is inflected with ergative case (*romel-ma* 'which-ERG'), and the embedded argument *ṗasuxi* 'answer' bears null absolutive case, both of which clearly identify transitive subject extraction. As expected, then, no extraction asymmetry is reported in Georgian relative clauses, despite being a morphologically ergative language.

Georgian                     (Foley 2013; p.9, p.9 / Aronson 1991; p.285)

(53) a. <u>*mosçavle*</u> [ *romel-ma=c*     *ġoveltvis icis*        *ṗasuxi* ]
<br>       student     which-ERG=REL always   know.PRES.3SG answer
<br>       'the student which always knows the answer'

     b. <u>*ḳali*</u>     [ *romeli=c*   *gušin*     *Eḳa-m*   *naxa*        ]
<br>       woman    which=REL yesterday Eḳa-ERG see.AOR.3SG
<br>       'the woman who Eḳa saw yesterday'

     c. <u>*k'aci*</u> [ *romeli=c*    *lap'arak'-ob-d-a* ]
<br>       man    which=REL he-was-speaking
<br>       'the man who was speaking'

The four examples in (49-53) all bear the same typological profile as Language $X_3$ from Section 5.3, with morphological case on both the extracted and non-extracted element. In addition, the sample provides examples of languages whose profile matches Language $X_4$, where only the non-extracted element bears case but

where the hypothesis still predicts no asymmetry. Wh-question and relative data from Niuean, repeated here in (54-55), clearly exhibit no extraction asymmetry of any sort.

Niuean                                (Seiter 1980; p.109, p.110; p.110, p.246, p.246; p.246)

(54) a. *ko    hai  ne    fifili   a    koe ke   vagahau*
        PRED who N.FUT choose ABS you SUB speak
        'who chose you to speak?'

   b. *ko    hai  ne    fahi e    Sione*
       PRED who N.FUT beat ERG Sione
       'who did Sione beat?'

   c. *ko    hai  ne    nofo he fale   kō*
       PRED who N.FUT live  in house that
       'who lives in that house?'


(55) a. *kehe <u>tama</u>   [   ka   kai e    tau pateta   ]*
        to   child       FUT eat ABS PL  potato
        'to the child who is going to eat the potatoes'

   b. *mo  e   <u>tagata</u>   [   ne    moto  e    koe   ]*
       with ABS person       N.FUT punch ERG you
       'with the person who you punched'

   c. *e   <u>tama</u>  [   ne    hau  i    Makefu   ]*
       ABS child       N.FUT come LOC Makefu
       'the child who comes from Makefu'

In addition, consider relative clauses in Lhasa Tibetan. In this language, relative clauses are pre-nominal, and yet the presence or absence of morphological case on

the embedded element crucially distinguishes transitive subject relativization (56a) from transitive object relativization (56b). Again, the disambiguation hypothesis makes the correct prediction: because of the availability of case within the relative clause, no additional disambiguating morphology is needed.

Lhasa Tibetan                                                    (Denwood 1999; p.202, p.202, p.200)

(56)  a.  [ *ta:*    *sɛ:bɛ:*        ]  <u>*mɪ*</u>   *tɕi:*
             tiger  kill.NOM.GEN      man   NUM
           'a man who killed a tiger'

      b.  [ *mi:*       *sɛ:bɛ:*        ]  <u>*ta:*</u>   *tɕi:*
             man.ERG   kill.NOM.GEN       tiger  NUM
           'a tiger which was killed by a man'

      c.  [ *kʰasã:*    *le:bə*        ]  <u>*tɛ*</u>
             yesterday  arrive-NOM      the
           'the one who arrived yesterday'

Finally, consider Sorani Kurdish. In Wh-questions, the Wh-element moves to an immediately pre-verbal position; furthermore, the ergative agreement clitic which corresponds to the transitive subject in the past tense must cliticize onto the transitive object according to Thackston (2016). Now, the ergative clitic =y is not an instance of dependent-marking. It changes in form as per the person and number of the transitive subject. However, its morphophonological requirement to cliticize to the transitive object – *aw mirava barzá* 'that tall man' in (57a), *ke* 'who' in (57b) – essentially serves the same disambiguating function as morphological case in the examples above. So while (57) should not be construed as a textbook example of morphological case serving a disambiguation function in Wh-questions, we do make

note that this ergative clitic essentially serves the same purpose, and furthermore that no additional morphology is observed in instances of transitive subject Wh-questions of Sorani Kurdish, exactly as predicted by the disambiguation hypothesis.

Sorani Kurdish                                                                          (W. M. Thackston, p.c.)

(57) a. *dwe      aw   mirov-a   barz-á=y          ke   dît?*
         yesterday DEM man-LINK tall-DEM=3SG.ERG who saw

         'who saw that tall man yesterday?'

    b. *dwe      aw   mirov-a   barz-á   ke=y          dît?*
         yesterday DEM man-LINK tall-DEM who=3SG.ERG saw

         'whom did that tall man see yesterday?'

    c. *dwe      ke   mird?*
         yesterday who died

         'who ran yesterday?'

The examples in (49-57) provide a snapshot of languages that match the typological profile of the hypothetical Languages $X_3$ and $X_4$ in the sample; these are languages in which morphological case alone disambiguates Wh-extraction. The totality of languages in the sample that belong to this set are listed in (58).[13,14]

---

[13] The languages listed in (58) are those in which morphological case <u>alone</u> disambiguates extraction. It does not include languages which are both Wh-*in situ* and feature morphological case, for example.

[14] The subscripts in (58) indicate the existence of more than one data point for that extraction type for that language. For example, Shipibo$_{(1)}$ features externally-headed relative clauses, which are disambiguated by morphological case, but also internally-headed relative clauses, in which case morphological help cannot disambiguate and which are correctly predicted to exhibit an asymmetry (see Section 5.4.4). Yakima Sahaptin$_{(2)}$ features two different externally-headed relative clause constructions, but both of them appear to be disambiguated solely by morphological case.

(58) *Languages of the sample with profiles matching Languages $X_3$ or $X_4$*

Wh-questions: Basque, Cavineña, Diyari, Epena Pedee, Kolyma Yukaghir, Lezgian, Matses, Ngiyambaa, Niuean$_{(1)}$, Paumarí, Roviana, Shipibo, Sorani Kurdish, Tiriyó, Trumai, Tukang Besi$_{(1)}$, Warlpiri, West Greelandic, Yakima Sahaptin, Yingkarta, Yukulta (n=21)

relative clauses: Abkhaz, Apinajé$_{(1)}$, Basque, Belhare$_{(1)}$, Burushaski, Cavineña, Cupeño, Georgian, Hindi, Kaluli, Lezgian, Lhasa Tibetan, Niuean, Paumarí, Shipibo$_{(1)}$, Sinaugoro, Warlpiri, Yakima Sahaptin$_{(2)}$, Yingkarta, Yukulta (n=21)

### 5.4.2 Basic word order as a disambiguator

Recall from Section 5.3.3 that, in the right circumstances, rigid V-medial basic word order <u>alone</u> can potentially serve to disambiguate in Ā-extraction scenarios. This typological profile was schematized as Language $X_5$ for Wh-extraction and $X_6$ for relativization. Only a single language in the entire ergative sample matches this type of profile: Canela-Kraho.[15]

---

[15] Jessica Coon (p.c.) wonders whether claims about the lack of extraction asymmetries in V-medial languages are even testable, given the relative lack of such languages cross-linguistically (see again Mahajan's generalization). In other words, it could be a true statement that V-medial languages never exhibit any asymmetries, but only because there aren't that many in the first place! The present analysis utilizes data from nine morphologically ergative V-medial languages (Drehu, Karitiâna, Katukina, Kuikuro, Paumarí, Tiriyó, Yakima Sahaptin, Yingkarta, Yukulta), out of a total 53 languages. Some do indeed exhibit extraction asymmetries (see Section 5.5) but others do not (see Yukulta and Paumarí in 50-51 above). I leave it up to future work whether the nine languages discussed in this work are indeed representative of all morphologically ergative V-medial languages cross-linguistically; most importantly, the disambiguation hypothesis at least predicts that languages with a profile like Languages $X_5$ and $X_6$ will not exhibit any kind of extraction asymmetry due solely to their basic word order.

The basic word order in Canela-Kraho is actually S Aux O V; nevertheless this fixed basic word order provides a sufficient cue for disambiguation in Wh-questions. For example, in a transitive subject Wh-question, the resulting sequence of elements is Wh Aux NP V, as in (59b). However, generating a transitive subject Wh-question yields the sequence Wh NP Aux V, as in (59c). Thus, just like the verb in Language $X_5$, in Canela-Kraho it is the auxiliary which acts as a fixed 'pivot' in Wh-questions, and by the Wh-element's relative position to it provides the vital cue necessary for disambiguation. As expected, then, there is no special morphology which marks transitive subject extraction only in Canela-Kraho.

Canela-Kraho                                    (Popjes & Popjes 1986; p.130, p.154, p.154)

(59) a. *hũmre te    rop cakwĩn*
         man   PAST dog beat
         'the man beat the dog'

   b. *jũm temã ita   ton*
       who PAST DEM make
       'who made this?'

   c. *ampo ca   ha   krẽ*
       what 2SG FUT eat
       'what will you eat?'

Unfortunately, no language in the ergative sample matches Language $X_6$, in which V-medial word order alone disambiguates in relativization. For illustrative purposes, relative clause data from Maybrat is in (60) below. While Maybrat is actually a morphologically accusative language, it otherwise fits the profile. As

predicted, no additional disambiguating morphology appears in relative clauses of Maybrat.[16]

Maybrat                                                                 (Dol 2007; p.137, p.124, p.258)

(60) a. <u>ait</u>      [ *ro*    *y-eyam*            *tapak*    ] *y-nit*               *po-mna*
         3M.SG    REL 3SG.M.SUBJ-roll tobacco    3SG.M.SUBJ-tell NOM-tell.tale
         'he who rolls a cigarette tells a tale'

     b. <u>aof</u>  [ *ro*    *ana m-fat*              ]
         sago    REL 3PL 3PL.SUBJ-fell
         'the sago tree that they felled'

     c. <u>ana</u> [ *ro*    *m-hu*               *au*   ]
         3PL    REL 3PL.SUBJ-stay DIST
         'those who stayed behind'

### 5.4.3   Wh-*in situ* as a disambiguator

Finally, recall from Section 5.3.4 that Wh-*in situ* was characterized as an inherently unambiguous Wh-strategy. As an independent predictor, all languages with Wh-*in situ* are predicted to be unambiguous, and therefore not exhibit any kind of ergative extraction asymmetry, regardless of any other grammatical features the language may have. In (54), Niuean was shown to lack an ergative extraction asymmetry in its Wh-clefting constructions. Here, in (61), Niuean's alternate Wh-*in situ* strategy is likewise shown to lack any kind of extraction asymmetry.

---

[16] It is not the case that morphologically accusative languages will automatically fail to exhibit an extraction asymmetry. On the contrary, taken to its full extent the disambiguation hypothesis goes beyond alignment, and predicts that similar 'accusative' extraction asymmetries ought to occur in morphologically accusative languages if the circumstances are right. While not the focus of Section 5.4, see Section 5.5 for several examples of extraction asymmetries in morphologically accusative languages, as well as 'mismatches' between morphological alignment and extraction alignment.

(61) a. *fifili    e    hai   a    koe ke    vagahau?*
       choose ERG who ABS 2SG SUB speak

    "who chose you to speak?"

    b. *totō   e    heigoa? ai    fai   mena ke   totō   ai*
       hold ABS what    NEG exist thing SUB hold there

    'hold what? there's nothing there to hold'

    c. *ko    e    heigoa e    mena nā?*
       PRED ABS what   ABS thing that

    'what is that thing?'

Consistent with the disambiguation hypothesis, Drehu (62) and Hindi (63) are two more examples of Wh-*in situ* languages, which do not exhibit any kind of extraction asymmetry despite being morphologically ergative.[17]

---

[17] According to Moyse-Faurie (1983), Drehu actually has two main clause types: VOS, which has ergative alignment in its dependent-marking (i-ii), and SVO, which does not exhibit any kind of morphological alignment at all (iii-iv). The use of these clause types is conditioned by tense/aspect. In any case, Drehu does not exhibit an extraction asymmetry in any of these constructions, as predicted by the disambiguation hypothesis.

Drehu                                    (Moyse-Faurie 1983; p.147, p.147, p.145, p.145)

i. *kola   humuth la    atr   hnen la   joxu*
   PROG kill       DET man ERG   DET chief
   'the chief is killing the man'

ii. *kola  elë  la   tim*
    PROG rise DET water
    'the water is rising'

iii. *angeic a    lep   la   kuli*
     3SG     PRES beat DET dog
     'he beats the dog'

iv. *angeic a    traqa*
    3SG     PRES arrive
    'he arrives'

Drehu                                   (Moyse-Faurie 1983; p.190, p.190, p.83)

(62) a. *hna    lepi   angeic hnei drei*
        PAST beat 3SG    ERG who

     'who beat him?'

     b. *hna    humuthi drei hnen la    joxu*
        PAST kill        who ERG   DET chief

     'who did the chief kill?'

     c. *drei lai*
        who over.there

     'who's over there?'

Hindi                                   (Mohanan 1994; p.208, p.208 / Montaut 2004; p.77)

(63) a. *kis-ne     bahut kaam   kiyaa*
        who-ERG much work.N do.PERF

     'who did a lot of work?'

     b. *raam-ne   kyaa kiyaa*
        Ram-ERG what do.PERF

     'what did Ram do?'

     c. *kaun aayaa*
        who   come.PERF

     'who came?'

    Finally, consider Abkhaz, which exhibits a particular variety of Wh-*in situ*.
It is *in situ* in the sense that Wh-questions are formed by converting an agreement
morpheme to a Wh-morpheme. Thus, in (64a), a transitive subject Wh-question is
formed by generating a Wh-morpheme in the ergative agreement slot which normally

encodes transitive subjects. Similarly, in (64b-64c), transitive object and intransitive subject Wh-questions respectively are formed by generating the Wh-morpheme in the absolutive slot. As it turns out, there is no asymmetry in (64) which could be construed as an asymmetry. The questions in (64) are simply inherently unambiguous.

Abkhaz                                                    (Hewitt 1989; p.11, p.15, p.10)

(64)  a.  *yə-z-fa-x'à-da*
          3SG.ABS-who-eat-PERF-INT
          'who has already eaten it?'

      b.  *yə̀-b-dər-wa-da*
          whom-2SG.ERG-know-DYN-INT
          'whom do you know?'

      c.  *y-àa-da*
          who-came-INT
          'who came?'

To be fair, it is not clear whether any of the examples in (61-64) actually constitute true Ā-extraction; it is, after all, Wh-*in situ*. That being said, the facts are still consistent with the overall hypothesis: here are some morphologically ergative languages which do not exhibit any extraction asymmetries, and this fact can be directly attributed to the unambiguous nature of the Wh-*in situ* strategy. In fact, it should come as no surprise that among the Wh-*in situ* languages in the sample, none of them exhibit any kind of extraction asymmetry, exactly as predicted by the disambiguation hypothesis. These languages are listed in (65) below.

(65) *Languages of the sample with profiles similar to Language $X_7$*

Wh-questions:   Abkhaz, Burushaski, Cupeño, Drehu, Georgian, Hindi, Hua,
                Niuean$_{(1)}$, Sinaugoro, Suena (n=10)

### 5.4.4   Ergative extraction asymmetries in the data

At this point, we have covered a wide range of languages in the sample that are predicted not to exhibit any kind of ergative extraction asymmetry. According to the disambiguation hypothesis, they are predicted not to because their typological profile is such that $\bar{\text{A}}$-extraction is not ambiguous. In this section, we will cover a wide range of languages which <u>are</u> predicted to exhibit ergative extraction asymmetries, specifically <u>because</u> $\bar{\text{A}}$-extraction would otherwise be ambiguous.

In addition to Kaqchikel (4-5), many Mayan languages exhibit ergative extraction asymmetries. Like Kaqchikel, their typological profile is such that they do not benefit from unambiguous extraction. They are V-initial, they do not feature morphological case, and content questions are formed by Wh-fronting. As predicted by the disambiguation hypothesis, transitive subject extraction is marked in some salient way in Q'anjob'al (66), Yucatec (67) and Ixil (68). In Q'anjob'al and Ixil, detransitivization of the verb and the addition of the Agent Focus morpheme are required, while in Yucatec, only detransitivization and omission of TAM morphology takes place. Regardless of their particular instantiation, these are all classic examples of ergative extraction asymmetries, often characterized as "syntactic ergativity" in the literature.

Q'anjob'al                                        (Coon et al 2014; p.225, p.192, p.192)

(66) a. *maktxel max il-on*[-i]        *naq winaq*
        who      ASP  see-**AF**-INTR  CL   man

   'who saw the man?'

   b. *maktxel max y-il-*[*-a'*]        *naq winaq*
        who      ASP  3SG.ERG-see-TR  CL   man

   'who did the man see?'

   c. *makxel max way-i*
        who      ASP  sleep-INTR

   'who slept?'

Yucatec                                  (Gutiérrez-Bravo & Monforte 2009; p.84, p.86

                                                        / Bricker 1979; p.133)

(67) a. *máax* ___ *il-ik-∅*        *María*
        who          see-IND-3SG.ABS Maria

   'who sees Maria?'

   b. *ba'ax t-u*          *maan-aj-∅*        *le*    *ko'olelo'*
        what  CMPL-3SG.ERG buy-MOD-3SG.ABS DEM woman

   'what did the woman buy?'

   c. *máax b'in-∅*        *k'íiwik*
        who    go-3SG.ABS square

   'who went to the square?'

Ixil                                              (Blunk 2008; p.30, p.26, p.26)

(68) a. *a=in*          *kat=tzok-on*    *u*    *si'-e'*
        F.COP=1SG   CMPL=cut-**AF**  DEF firewood-ENC

   'It is I who cut the firewood'

b. *a=kuxh-e'*        *nu-ku-b'an=∅-e'*
    F.COP=only-ENC INC-1PL.ERG-do-3SG.ABS-ENC
    'it it just this that we did'

c. *a=in*        *kat=vat=in-e'*
    F.COP=1SG CMPL-sleep-1SG.ABS-ENC
    'it was me who slept'

Note that the Ixil data in (68) are actually examples of focus fronting, but still the point remains the same: only in the event of transitive subject focus (68a) must the verb lose its ergative agreement morphology and be marked with Agent Focus. Note also that from a functional point of view, the actual form of the extraction asymmetry is arbitrary. The asymmetry takes the form of an Agent Focus morpheme in Q'anjob'al and Ixil, but in Yucatec it merely takes the shape of an omission.[18] From a functional perspective, they are serve the same purpose: transitive subject Ā-extraction is signaled by an unexpected departure from the normal morphosyntax of plain, declarative transitive sentences.

Of course, the predictions made by the disambiguation hypothesis are not specific to the Mayan language family. It predicts that <u>any</u> language should exhibit an extraction asymmetry, regardless of language family – just as long as the right typological profile is met. Selayarese (Austronesian), Gitksan (Tshimshianic) and Lillooet (Salishan) fit the bill: they are V-initial, they feature Wh-fronting, but they lack morphological case.

In Selayarese, verbs in declarative clauses typically feature an absolutive agreement suffix, and an ergative agreement prefix if necessary (Finer 1997). However,

---

[18] Halkomelem also marks its extraction asymmetry by omission of agreement morphology; see Appendix C (36).

only in transitive subject Wh-questions can the absolutive suffix remain on the verb (69a). In other words, in both transitive object (69b) and intransitive subject Wh-questions (69c), the verb <u>must</u> drop its absolutive agreement marker despite obviously having absolutive arguments in the clause. This is a textbook example of an ergative extraction asymmetry because with regards to absolutive morphology, S and O arguments pattern together to the exclusion of A arguments.

Selayarese                                             (Finer 1997; p.689, p.689 / D. Finer, p.c.)

(69) a. *inai la-erang-i     loka-ñjo*
        who 3.ERG-bring-3.ABS banana-DEF
        'who brought the bananas?'

   b. *apa  la-taro   i   Baso ri lamari*
      what 3.ERG-put DET Baso in cupboard
      'what did Baso put in a cupboard?'

   c. *inai ak-kelong*
      who INTR-sing
      'who sang?'

From the perspective of our current hypothesis, the sole purpose of the constraint on absolutive morphology in (69b) is to disambiguate from the corresponding subject Wh-question 'what put Baso in a cupboard?', implausible as it may be. In fact, if absolutive morphology were realized in (69b), that's exactly what the question would mean.

Like in Yucatec, the Selayarese data is also quite revealing in that it again suggests that the exact disambiguation mechanism is rather arbitrary. The asymmetry doesn't necessarily have to take the form of a new, unexpected morpheme in

transitive subject Ā-extraction; it can alternatively take the form of an unexpected omission or addition in transitive object and intransitive subject Ā-extraction instead. Crucially, it is the morphological <u>difference</u> between transitive subject and transitive object Wh-questions that provides the vital cue, even though semantically they are both transitive in exactly the same way. This level of flexibility – that the theory is not bound by a specific morphological realization of the asymmetry – is a strength of the disambiguation hypothesis, and yet at the same time it is something that the syntactic analyses from Section 5.2 fail to capture altogether. Those analyses would have a hard time explaining why, for instance, it is Selayarese transitive objects and intransitive subjects that appear to be 'constrained' in Ā-extraction, and not the transitive subject.

Gitskan and Lillooet are two other morphologically ergative languages which have the same typological profile as Selayarese and the Mayan languages in (66-68). As predicted by the disambiguation hypothesis, only in the case of transitive subject Wh-extraction must some special identifying morphology occur. In Gitksan, it takes the form of a 'relativizer' morpheme *ʔan* immediately following the Wh-word; in Lillooet, a special *-talí* verbal affix occurs in lieu of normal transitive subject agreement morphology.

Gitksan                                                          (Rigsby 1986; p.303, p.303, p.303)

(70) a. *naa ʔan=t            qay   łəmoo=s  Bruce*
         who **A.REL**=3SG.ERG.DEP which help=CNN Bruce
         'who helped Bruce?'

    b. *naa=ł       łʷəmoo-yə-n*
         who=CNN help-TR-2SG.ERG
         'who did you help?'

309

c. *naa=ł    lim-ət*
   who=CNN sing-S.REL

   'who sang?'

(71) a. *šwat kʷu áċx-ən-táli   ta    nkyáp-a*
   who IRR see-TR-**TOP** DET coyote-DET

   'who saw the coyote?'

b. *štam kʷu áċx-ən-aš*
   what IRR see-TR-3SG.ERG

   'what did she see?'

c. *šwat skíczaʔ qwatsáts*
   who  mother leave

   'whose mother left?'

The examples in (69-71) illustrate extraction asymmetries in Wh-questions of Selayarese, Gitksan and Lillooet. However, it should not be surprising that these languages also display the same asymmetries in their relative clauses too. Since the same basic principles still apply, relative clauses in these languages also present a fertile ground for ambiguity in the absence of any additional disambiguating morphology: as post-nominal relative clauses, in some sense they involve movement 'across' the clause-initial verb, and case morphology is still unavailable. As expected, then, the exact same disambiguating strategies observed in Wh-questions are also enlisted to disambiguate in relativization, as in (72-74).

(72) a. <u>*pakoko*</u> [ *to-la-pallu-ñjo-i*        *jukuʔ-na*     ]
      farmer     REL-3.ERG-cook-DEF-<span style="color:red">3.ABS</span> fish-3.POSS

      'the farmer who cooked his fish'


    b. <u>*jukuʔ*</u> [ *nu-la-pallu-ñjo*       i     *Ali* ]
      fish     REL-3.ERG-cook-DEF DET Ali

      'the fish Ali cooked'


    c. <u>*tedong*</u> [ *nu-ak-kelong-ñjo*    ]
      buffalo    REL-INTR-sing-DEF

      'the buffalo who sang'


Gitksan                                    (Rigsby 1986; p.404, p.407, p.405)


(73) a. *kaʔ-ə-'y=ł*           <u>*kat*</u> [ <span style="color:red">*ʔan-cak$^w$-ə=ł*</span>      *naks-t*        ]
      see-TR-1SG.ERG=CNN man   <span style="color:red">A.REL</span>-kill-INCR=CNN spouse-3SG.POSS

      'I saw the man who killed his wife'


    b. *mał-tə-'y*        *lo:-tit*    *tim* *k$^w$ix$^w$-ə=s*    *John=ł*    <u>*smax*</u> [
      tell-TR-1SG.ERG OBL-3PL FUT shoot-TR=CNN John-CNN bear
      *łə-kaʔ-n*         ]
      DEF-see-2SG.ERG

      'I told them that John would shoot the bear that you saw'


    c. *tim* *t'is-ə=ł*      *xaʔ=ł*    <u>*kat=ł*</u>    [ *pax-ət=kə*        ]
      FUT hit-TR=CNN slave=CNN man=CNN    run-S.REL=DIST

      'the slave will hit the man who ran'


Lillooet                                   (Roberts 1999; p.285, p.289, p.286)


(74) a. *ti*    *sqáycw-a* [ *ti*    *tup-un'-<span style="color:red">táli</span>-a*    *s-Bill*    ]
      DET man-DET    DET hit-TR-<span style="color:red">TOP</span>-DET NOM-Bill

      'the man that hit Bill'

b. *nilh wi* <u>*s-John*</u> <u>*múta?*</u> <u>*s-Bill*</u> [ *i* *ats'x-en-án-a* *lhkúnsa* ]
FOC PL.DET NOM-John CONJ NOM-Bill PL.DET see-TR-1SG.NOM-DET now

'it was John and Bill that I saw'[19]

c. *tsukw t'u?* <u>*s?ents*</u> [ *ti* *tsícw-a* ]
only PART 1SG.EMP DET go-DET

'only I went; the one who went is me'

All of the examples thus far in Section 5.4.4 come from V-initial languages. However, Ayutla Mixe illustrates the point that V-initial basic word order is not a requirement, and the right conditions for an ergative extraction asymmetry can indeed be met in a V-final language. Ayutla Mixe features SOV basic word order, Wh-fronting, but no morphological case. In principle, fronting either the transitive subject or the transitive object would produce an ambiguous Wh NP V linear word order. As predicted, in instances of transitive subject Wh-extraction only, some unexpected changes in verb morphology take place. The prefix that normally encodes 3rd person transitive subjects goes from *y-* to *t-*, as in (75a), and in addition the verbal mood marker detransitivizes from *-yp* to *-p*. Compare with the extracted transitive object example in (75b), which features normal transitive clause verb morphology.

Ayutla Mixe                        (Romero-Méndez 2009; p.459, p.460, p.458)

(75) a. *pëën yë'ë* *pu'uy jam* *t-jëët-p*
who DEM.M board DEIC.D 3.ERG-saw-INDEP

'who is sawing the board?'

---

[19] Example (74b) is unusual in that Roberts's translation seems to indicate a cleft construction more than a relative clause. Nevertheless, (74b) is clearly listed as an example of a relative clause in his discussion of extraction asymmetries; perhaps, on analogy with the translation for (74c), an equivalent translation for (74b) might be 'John and Bill are the ones that I saw'.

b. *tii=ja'a*     *m-itsy*            *jam*    *tsyäm y-tun-yp*
what=DEM.D 2.POSS-younger.sibling DEIC.D now 3.ERG-do-INDEP.TR

'what is your brother doing right now?'

c. *pë̈ën men-p*
who   come-INDEP

'who is coming?'

Again, the unexpected morphological changes in (75a) ostensibly serve to distinguish that sentence's transitive subject Wh-meaning from its closely-related transitive object counterpart, 'who is the board sawing?'. And while these data points provide for potential confounds with silly or completely implausible alternative meanings, it is not hard to imagine how the same point still holds in different circumstances.[20] In all these languages which exhibit an ergative extraction asymmetry in Wh-questions, it is the unique morphology itself which crucially disambiguates between minimal pairs such as 'who saw the man?' and 'who did the man see?'.

The morphological nature of the asymmetry in Ayutla Mixe – an unexpected change in both agreement and mood morphology in (75a) – further illustrates the heterogeneity of ergative extraction asymmetries cross-linguistically. These facts are perfectly compatible with the disambiguation hypothesis, which doesn't care how, exactly, the asymmetry will be realized. The disambiguation hypothesis simply dictates that some obvious morphological difference must differentiate between instances

---

[20] Indeed, the specific Ayutla Mixe examples in (75) reinforce the point first made in Section 5.3.6: sometimes, the highly grammaticalized morphology that characterizes an asymmetry will occur in scenarios where it may not have been required. In (75a), it is fairly clear that a transitive event of sawing involving a board and an unnamed person will most likely entail the person sawing the board in some capacity. In other words, (75a) is a situation where real-world knowledge *probably* obviated the need for overt disambiguation. Nevertheless, grammaticalization of the ergative extraction asymmetry in Ayutla Mixe means that it will sometimes be deployed in a redundant fashion in scenarios such as these.

of transitive subject and transitive object Ā-extraction. However, it bears repeating that no single syntactic analysis from Section 5.2 can single-handedly capture this broad range of morphological asymmetries; in each case, the theory predicts that extraction asymmetries should be characterized by a single, consistent morphological exponent, and typically in instances of transitive subject extraction only.

Also in the Austronesian family, Roviana is a very interesting case. In main clauses, ergative-absolutive alignment is observed on the case marking of proper NPs, pronouns and NPs which contain a numeral, as in (76). Corston-Oliver (2003) thus characterizes the alignment in main clauses as "split-ergativity according to NP type" because other types of NPs lack alignment altogether. In any case, in Wh-questions the disambiguation hypothesis predicts that no extraction asymmetry should occur, given the fairly widespread availability of this morphological case. Indeed, this is exactly what occurs: no special morphology or omission can be said to characterize transitive subject Wh-extraction in (77a) compared to other types of Wh-extraction in (77b-77c).

Roviana                                  (Corston-Oliver 2003; p.281, p.281, 279

/ Waterhouse 1928; p.39, p.233, p.146)

(76) a. *seke-i-a*       *e*    *zima se*   *maepeza*
     hit-TR-3SG.OBJ ERG Zima ABS Maepeza
     'Zima hit Maepeza'

    b. *doɣor-i-a*     *rau*     *si*    *asa*
     see-TR-3SG.OBJ 1SG.ERG ABS 3SG.ABS
     'I saw him'

c. *meke doɣor-i-a      ri       si   keke iɣana*
   and    see-TR-3SG.OBJ 3PL.ERG ABS one  fish

'and they saw a fish'

(77) a. *esei hoper-i-a        sa    igana hie*
     who spear-TR-3SG.OBJ DEF fish    DEM

'who speared this fish?'

b. *na   sa    si   kaqu tepa-i-a          rau*
   DEF what FOC FUT  ask.for-TR-3SG.OBJ 1SG.ERG

'what shall I ask it?'

c. *esei si    zukezuke   hoi*
   who FOC torch.RED there

'who is that with a torch?'

Interestingly, by some quirk of grammar, in subordinate clauses only arguments are not marked for case in Roviana, and any traces of alignment are lost. Exactly in this environment, and exactly as predicted, a special agreement morpheme is required in instances of transitive subject relativization. The exact shape of this special agreement morpheme is conditioned by the embedded, non-extracted transitive object in the relative clause. Therefore, this special agreement suffix is realized as *-na* because the non-extracted transitive object in (78a) is third person singular, however if that transitive object were first person singular, the suffix would be *-gu* instead, for example (Corston-Oliver 2003).

Roviana                                              (Corston-Oliver 2003; p.275, p.275, p.275)

(78) a. *hierana sa   koreo* [ *sapu tupa-na         e      zone* ]
     this     DEF boy    REL  punch-3SG.NSUF PERS John

'this is the boy that punched John'

b. *hierana sa* <u>*koreo*</u> [ *sapu tupa-i-a*      *e*    *zone* ]
    this     DEF boy     REL punch-TR-3SG.OBJ PERS John
    'this is the boy that John punched'

c. *hierana sa*    <u>*tie*</u>    [ *sapu kote taloa* ]
    this     DEF man    REL FUT leave
    'this is the man who is going away'

To be clear: it appears highly likely that in Roviana, because morphological case is suddenly no longer available in relative clauses, the language has innovated a special requirement in transitive subject relative clauses to help distinguish sentences like (78a) from (78b). In this respect, Roviana relative clauses are perfect minimal pairs with Niuean ones; in the latter, morphological case is still available as a cue, and thus no extraction asymmetry results!

Furthermore, it seems difficult to accurately characterize the exact morphological nature of the *-na* suffix in (78a). Corston-Oliver labels it a "nominal suffix" which indexes the O argument. This is not an issue in the disambiguation hypothesis: it only predicts that a morphological asymmetry <u>should</u> occur in this specific environment, but not <u>what</u> that morphological difference should specifically be. Again, the syntactic analyses from Section 5.2 cannot capture the morphological diversity of these asymmetries; by their very nature, they are each committed to a single morphosyntactic process, in instances of transitive subject Ā-extraction only, which corresponds to a single, consistent morphological exponent in turn.

To summarize the results thus far: the disambiguation hypothesis can successfully account for a wide range of ergative extraction asymmetries, with heterogeneous morphological instantiations, in languages such as Kaqchikel, Selayarese, Ayutla Mixe and Roviana. In contrast, syntactic analyses cannot provide a similar

breadth of coverage. This growing evidence suggests that the underlying cause of these extraction asymmetries is not some single, underlying morphosyntactic mechanism related to case assignment; rather, it is functional in origin, and therefore its morphological realization is consequently rather open-ended.

The table in (79) lists all the languages in the sample which pattern like Language $X_2$: they do not benefit from a felicitous combination of morphological case, word order or extraction strategies, and as such they require an ergative extraction asymmetry, in the form of unexpected morphology, to unambiguously indicate instances of transitive subject Ā-extraction.[21]

(79) *Languages of the sample with profiles matching Language $X_2$*

Wh-questions: Ayutla Mixe, Gitksan, Ixil, Kapampangan, Kaqchikel, Lillooet, Q'anjob'al, Seediq, Selayarese, Sm'algyax, Yucatec$_{(1)}$ (n=11)

relative clauses: Gitksan, Halkomelem, Ixil, Kapampangan, Kaqchikel, Lillooet, Matses, Movima, Q'anjob'al, Roviana, Seediq, Selayarese, Sm'algyax, Tiriyó, Trumai, Yucatec$_{(1)}$ (n=16)

At this point we have covered some basic examples of ergative extraction asymmetries in the sample; the next few examples are a little bit more involved. Ultimately, however, the disambiguation hypothesis still makes correct predictions.

---

[21] Note, however, that two languages from the sample (Ixil and Tseltal) also match the profile of Language $X_2$, but still do not exhibit any kind of extraction asymmetry. These cases of incorrect predictions by the disambiguation hypothesis are discussed in Section 5.4.6. Note also that some languages match the profile of Language $X_2$, but resolve the ambiguity as an accusative extraction asymmetry: see Yucatec durative aspect (139) and Kolyma Yukaghir attributive relative clauses (64) in Appendix C.

Trumai is a language which *a priori* seems to have sufficient cues that would obviate the need for an ergative extraction asymmetry. Indeed, Trumai marks its arguments with morphological case, as in (80). The only problem is that transitive verbs can come in two distinct frames, which are lexically-conditioned: some transitive verbs occur in a A-ERG O V frame (80a), and other transitive verbs occur in an A V O-DAT frame (80b).[22]

Trumai                                                    (Guirardello 1999; p.91, p.91)

(80) a. *hai-ts      kodechich  disi ka in*
        1SG-ERG snake       kill  FOC/TENSE
        'I killed the snake'

     b. *ha    fa   kodechïch-atl ka in*
        1SG kill snake-DAT    FOC/TENSE
        'I killed the snake'

The end result is that while Trumai may indeed have morphogical case in its typological profile, this morphological case is not a particularly <u>reliable</u> cue, especially in relative clauses. Compare the 1SG argument in (80a) and (80b). They are both semantically the agent of the clause, but due to those verbs' different, lexically-conditioned syntactic frames, the agent is not reliably marked with one consistent morphological case.

---

[22] According to Guirardello (2010), verbs that use the A V O-DAT frame tend to be verbs of perception ('see', 'listen', 'smell'), mental activity ('think', 'believe', 'like') and verbs that express habitual events with conventionalized patients ('eat', 'drink', 'cook', 'bake'). And yet the generalization is not absolute, two verbs that mean 'to kill' in (80) occur in different syntactic frames. See Guirardello (1999, 2010) for more details.

To complicate matters further, relative clauses in Trumai show a great degree of flexibility: as reported by Guirardello (1999), they can occur post-nominally (81a), pre-nominally (81b) or even appositively (81c) – in the latter case, head and relative clause do not necessarily have to be adjacent in linear order.

Trumai                                                      (Guirardello 1999; p.417, p.416, p.417)

(81) a. <u>axos</u> [ *huma-t'*       *ke*   ] *yi wapta*
       child   take.bath-NMLZR RLZR   YI fall
       'the child who took a bath fell'

    b. [ *ma-t'*     *ke*  ] <u>*kiki*</u> *yi otl*   *kawa chïin*
       eat-NMLZR REL   man YI sleep go   FOC/TENSE
       'the man who ate went to sleep'

    c. <u>*kiki*</u> *huma-ktsu* [ *ma-t'*     *ke*  ]
       man take.bath    eat-NMLZR REL
       'the man who ate went to take a bath'

Given the facts in (80-81), relative clauses in Trumai arguably do not benefit from many cues at all – morphological case is unreliable, and so too is relative clause word order. Therefore, perhaps it should come as no surprise that Trumai requires a special particle *chïk*, used in transitive subject relative clauses, to act as an additional helpful cue in relativization (82a). When transitive objects or intransitive subjects are relativized, the particle *ke* is used instead (82b-82c). This is another example of an ergative extraction asymmetry in relative clauses, and when the sum of relevant

factors is considered, it seems fairly well-motivated in terms of the disambiguation hypothesis.[23]

Trumai                                                    (Guirardello 2010; p.219, p.219, p.218)

(82) a. *ha   hu'tsa chï in      <u>axos</u>-a-tl      [ ha    aton mud husa-t'    chï-k ]*
        1SG  see     FOC/TENSE child-EV-DAT  1SG  pet   neck tie-NMLZR REL
        'I saw the boy who tied my pet'

   b. *ha   hu'tsa chï in      <u>axos</u>-a-tl      [ hai-ts    midoxos-t'a ke  ]*
        1SG  see     FOC/TENSE child-EV-DAT  1SG-ERG call-NMLZR  REL
        'I saw the boy whom I called'

   c. *ha   hu'tsa chï in      <u>axos</u>-a-tl      [ esa-t'        ke  ]*
        1SG  see     FOC/TENSE child-EV-DAT  dance-NMLZR REL
        'I saw the boy who danced'

When it comes to Wh-questions, descriptively Wh-extraction of the transitive subject does not entail the use of the *chïk* relativizer (83a). Why should this conform to our disambiguation hypothesis? Perhaps it has to do with the inherent structural difference between Wh-questions and relative clauses in this language. In

---

[23] The one possible wrinkle regarding extraction asymmetries in Trumai relative clauses is that those formed from dative-marked arguments, as in i) below, are also marked with the *chïk* relativizer. Therefore, the distribution of these relativizers is S, O → *ke* and A, DAT → *chïk*, which is hardly a typical ergative alignment. Nevertheless, the overall point still stands: it seems plausible that the lack of <u>reliable</u> cues in Trumai relative clauses motivates dedicated morphemes to help disambiguate the different extraction possibilities.

Trumai                                                    (Guirardello 2010; p.219)

i) *ha   hu'tsa chï in      <u>axos</u>-a-tl      [ ha   fa    fa-t'        chï-k ]*
   1SG  see     FOC/TENSE child-EV-DAT  1SG  beat beat-NMLZR REL

   'I saw the boy whom I beat'

(83), transitive Wh-questions in Trumai feature both arguments coded with respect to their role in the clause. For example, in (83a), the Wh-word features ergative case, while the non-questioned element bears null absolutive case, making identification of grammatical roles within the ERG-ABS alignment straightforward. Furthermore, there is only one Wh-strategy – Wh-fronting – and it is used consistently.[24]

Trumai (R. Guirardello; p.c.)

(83) a. *te    yi-k    in    ha    midoxos*
       who YI-ERG FOC 1SG call
       'who is calling me?'

   b. *te    yi in    Kumaru-k    padi ke*
      who YI FOC Kumaru-ERG wait KE
      'whom is Kumaru waiting for?'

   c. *te    yi in    sa    ke*
      who YI FOC dance KE
      'who is dancing?'

Going back to relative clauses, the fact that they do exhibit an extraction asymmetry indicates that a sort of compounding effect may be taking place. In Section 5.3.2, the disambiguation hypothesis stated that even in cases of Ā-extraction where only the non-extracted element bears morphological case, this should be enough of a cue for unambiguous extraction (Language $X_4$), and this was certainly

---

[24] Note that transitive object and intransitive subject questions in (83) still feature a *ke* morpheme, despite the lack of *chĭk* in transitive subject Wh-questions. According to Guirardello (p.c.), this *ke* morpheme is distinct from the *ke* relativizer in relative clauses, despite likely having the same synchronic origin. This is because the *ke* in (83) occurs in a wider range of syntactic environments – see Guirardello (1999) Section 5.1.2 for more details.

true for languages such as Niuean and Lhasa Tibetan. But perhaps X$_4$-type morphological case, in conjunction with unreliable morphological alignment (ERG-ABS vs ABS-DAT) and multiple relativization strategies produce a threshold effect where embedded clause morphological case is no longer enough. Although convoluted, this explanation is consistent with the disambiguation hypothesis, and furthermore it provides an explanation for another West Greenlandic-type language from Section 5.1: one in which relative clauses exhibit an ergative extraction asymmetry, but Wh-questions do not.

Tiriyó is another language in which a combination of factors – very free word order and appositive relative clauses – arguably explain the presence of an extraction asymmetry in relative clauses. According to Carlin (2004), the most unmarked word order is OVS, as in (84a), although a closely-related SVO variant with an overt 3>3 prefix *n-* is also common (84b). Also, a different OVS construction occurs with non-finite morphology and a 'dummy' agreement prefix *t-*, and it is frequently associated with remote past/narrative past semantics (84c). Crucially, in this last construction, transitive subjects are marked with the suffix *-ja*.

Tiriyó                                                    (Carlin 2004; p.481, p.481, p.220)

(84)  a.  *wïtoto*         *∅-i-pijo-ja-n*              *pahko*
          human.being 3<3-TR-whip-PRES-NCERT 1POSS.father
          'my father is beating the Amerindian'

      b.  *wïtoto*         *n-i-pijo-ja-n*              *pahko*
          human.being 3>3-TR-whip-PRES-NCERT 1POSS.father
          'the Amerindian is beating my father'

c. *kana  t-ënee-se        i-ja*
   fish   COREF-bring-NF 3-GOAL
   'he brought fish'

In Wh-questions, Tiriyó makes use of its several available clausal constructions judiciously. In transitive subject (85a) and intransitive subject (85c) Wh-questions, the grammar tends to use one of the fully finite constructions in (84a-84b). However, in transitive object Wh-questions, the grammar tends to opt for the non-finite construction with the clearly identified, post-verbal transitive subject marked by the *-ja* suffix.[25]

Tiriyó                                              (Carlin 2004; p.253, p.116, p.235)

(85) a. *a-kï        n-i-pono-ja-n*
        WH-ANIM 3>3-TR-tell-PRES-NCERT
        'who is telling it?'

   b. *a-kï-jamo-hpe       tï-wë-e        i-ja,    mëhparë-ton*
      WH-ANIM-PL-INDEF COREF-shoot-NF 3-GOAL game.animals-PL
      'what on earth kinds of game animals did he shoot?'

   c. *a-kï        n-e:-ja-n*
      WH-ANIM 3>3-come-PRES-NCERT
      'who is coming?'

Like Paumarí, Tiriyó has several clausal frames available, and it associates them with a specific kind of Wh-extraction to head off any potential ambiguity.

---

[25] The exact nature of the suffix *-ja* is not entirely clear. Both Carlin (2004) and Meira (1999) seem to agree it may have been ergative in origin diachronically, but now its use has been extended to recipients and causees in other contexts. Meira adds that the suffix only ever occurs on "sentient" participants, and notes a similar syncretism between sentient ergative and sentient dative forms in Circassian languages.

These do not constitute a completely unexpected change in morphology – rather the language is simply making use of existing morphosyntax from plain, declarative sentences and thus this is not considered an ergative extraction asymmetry.

In relative clauses, however, a potential compounding effect similar to Trumai is observed. It appears as if the combination of free word order, multiple clausal frames and appositive relative clauses warrants extra disambiguating morphology. As a result, a dedicated relativizer *-ne* is used in for transitive subject relativization (86a); transitive object and intransitive subject relative clauses features their own dedicated relativizers as well (86b-86c). Indeed, the examples in (86a-86b) are particularly revealing. If they didn't benefit from those dedicated relativizers, the relative clause in (86a) could just as easily mean 'the person who the dog hit'; likewise, the one in (86b) could just as easily mean 'the dog who hit my father'.

Tiriyó                                                    (Meira 1999; p.573, p.573, p.574)

(86) a. *an-po=n-ai*        *wïtoto*, [ *kaikui i-tuuka-ne-npë*        ]
    WH-LOC=3S$_A$-COP person   dog   3-hit-**A.NMLZR**-PAST
    'where is the person who hit the dog?'

   b. *kaikui ë-waarë*, [ *pahko i-n-tuuka-hpë*        ]
    dog   2-COGN [ father 3-O.NMLZR-hit-PAST
    'do you know the dog that my father hit?'

   c. *mëërë* [ *ëturu-ketï-npë*        ] *ji-tuuka*
    3AN   talk-S.NMLZR-PAST  1O-hit
    'that one, who had been talking, hit me'

Finally, recall from the discussion in Section 5.3.5 that internally-headed relative clauses were considered an inherently ambiguous relativization strategy. In the

sample, a trio of languages does indeed use this rare strategy: in Shipibo, Belhare and Apinajé, relative clauses may be formed in which the head is not overtly extracted from the relative clause itself (87-89). The interpretation of these internally-headed relative clauses is fairly straightforward in the case of intransitives: in the three (b) examples below, only one core argument is even available, thus it is interpreted as the relative clause's head. However, in transitive clauses, a real problem presents itself: which argument counts as the head? It turns out that in all three languages, a constraint operates which prevents transitive subjects from being interpreted as the head of an internally-headed relative clause, thus resolving the interpretation problem.

Shipibo                                                      (Valenzuela 2003; p.482, p.483)

(87) a. [ *pitso-n*      *bake*      *natex-a*    ] *-tonin-ra joshin        pike*
         parakeet-ERG child.ABS bite-PART    ERG-EV  banana.ABS eat-CMPL

   'the child the parakeet bit ate he banana'

   * 'the parakeet that bit the child ate the banana'

   b. [ *mi-bé*     *ainbo*      *jo-a*            ] *-ra no-n      onan-yama-ke*
         2SG-COM woman.ABS come-PART-ABS    EV 1PL-ERG know-NEG-CMPL

   'we don't know the woman who came with you'

Belhare                                                          (Bickel 2003; p.564, p.564)

(88) a. [ *tombhira-ŋa* *wa*      *seiʔ-s-u=na*                    ] *chitt-he-m*
         lynx-ERG      chicken kill-TR.PERF-3SG.ABS=DET   find-PAST-1PL.ERG

   'we found the chicken that the lynx had killed'

   * 'we found the lynx that had killed the chicken'

b. [ _maʔi_    khiu-ʔ=na          ] misen-niu-t-u-ga                    i
   person quarrel-NON.PAST=DET   know-NON.PAST-3SG.ABS-2SG.ERG INT

   'do you know the person who is quarrelling?'

Apinajé                                                (Oliveira 2005; p.282, p.283)

(89) a. [ _ic-tɛ_   a-mə̃ _bɨ_  j-arẽɲ    ja        ] na   tɛ   krĩ   õ   kamə̃ pa
        1-ERG 2-DAT man RP-tell.NF DEF.ART   REAL HAB village one INSV  live

        'this man I'm telling you about lives in the other village'

   b. [ _di_    čəm         ja        ] na    prɛ   ra   iɲ-mə̃ a-j-arẽ
        woman RP-stand.NF DEF.ART   REAL PAST ASP 1-DAT 2-RP-tell

        'this woman standing there had already told me about you'

This, too, counts as an example of an ergative extraction asymmetry, albeit an unusual one: the generalization is that in these three languages, only absolutive arguments may be interpreted as arguments outside of the relative clause. Furthermore, we now have a potential explanation for more West Greenlandic-type languages, which only exhibit an asymmetry in a subset of domains. Recall that Shipibo Wh-questions do not exhibit any kind of asymmetry at all (see 49 in Section 5.4.1). This is because, in that domain, abundant cues are available to identify the correct extracted argument. Similarly, externally-headed transitive subject relative clauses are not constrained in any way, or marked with any special morphology. In (90), the transitive subject may easily extract to form a pre-nominal, externally-headed relative clause, in stark contrast with its internally-headed counterpart in (87a). Of course, this is fully predicted by the disambiguation hypothesis: in (90), morphological case provides an excellent cue for extraction. However, because Shipibo sometimes makes use of an alternative, potentially-confusing relativization strategy – internally-headed

relative clauses – Shipibo only requires an asymmetry <u>in that particular domain</u>, as needed.

Shipibo                                                          (Valenzuela 2002; p.67)

(90) [ *bake*       *natex-a*   ] <u>*pitso-n-ra*</u>        *joshin*      *pi-ke*
       child.ABS bite-PART      parakeet-ERG-EV banana.ABS eat-CMPL
     'the parakeet that bit the child ate the banana'

Again, it is not clear how any of the syntactic analyses from Section 5.2 could explain the difference between an internally-headed transitive subject relative clause of Shipibo (87a) and an externally-headed one (90). In principle, the assignment of morphological case should be isomorphic in both clause types, and if ergative extraction asymmetries are predicated on the assignment of morphological case, then we wouldn't expect any difference between them at all.

Furthermore, many of those syntactic analyses formulate the asymmetry as a constraint on the extraction of ergative-marked transitive subjects only, but note that oblique arguments are also ineligible to serve as heads of internally headed-relative clauses. Neither Shipibo comitative arguments (87b) nor Apinajé dative arguments (89a) can be interpreted as heads in these constructions. This is unexpected under most syntactic accounts in Section 5.2, but perfectly consistent with the disambiguation hypothesis: if oblique arguments <u>were</u> eligible, it would confuse the interpretation of those internally-headed relative clauses in the same way that ergative-marked arguments would, and thus they are rendered ineligible for straightforward, functional reasons.

To summarize this section, each of the languages in (66-90) exhibit an ergative extraction asymmetry. According to the disambiguation hypothesis, the asymmetries occur because those languages don't benefit from a typological profile which produces unambiguous instances of Ā-extraction. The exact morphological nature of the asymmetry varies from language to language, which is not a problem for the disambiguation hypothesis; however, the syntactic analyses from Section 5.2 will struggle to account for this wide range of phenomena with a just single morphosyntactic explanation.[26]

### 5.4.5 The Austronesian voice system and the disambiguation hypothesis

Two Austronesian languages in the sample merit special discussion due to their voice system. In Seediq and Kapampangan, transitive declarative clauses can occur in two distinct frames: a VOS 'Agent Focus' construction (91a, 92a), and a VSO 'Patient Focus' construction (91b, 92b). In addition to differences in basic word order, case, and agreement, these constructions require special morphological changes in verbs themselves, which identify the verb as being in Agent Focus or Patient Focus (compare *mekan-puqun*, *sumulat-isulat*).[27]

Seediq                                                                (Holmer 1996, p.66, p.66)

(91) a. *mekan wawa ka huling*
       eat.AF meat KA dog

       'the dog is eating meat'

---

[26] For additional data, see Appendix C; namely, Matses relative clauses which feature a similar profile to Tiriyó and Trumai (77), Tukang Besi internally-headed relative clauses (124), and several other languages with the profile of Language $X_2$.

[27] Of course, not every Austronesian language still exhibits voice like its Proto-Austronesian ancestor. Verbs in Niuean (3-6), Drehu (62), Selayarese (69) and Roviana (76-78) do not bear any overt morphology which could be identified as belonging to the Austronesian voice system.

b. *puqun huling ka wawa*
   eat.PF dog    KA meat

'a dog will eat the meat'

Kapampangan <span style="float:right">(Mirikitani 1972, p.116, p.116)</span>

(92) a. *sumulat ya ng poesia ing lalaki para king babai*
   write.AF 3SG ACC? poem ABS boy    BEN     girl

'the boy will write a poem for the girl'

b. *isulat ne ning lalaki ing poesia king mestra*
   write.PF 3SG>3SG ERG boy    ABS poem DAT teacher

'the boy will write the poem to the teacher'

In both these languages, the voice system is already manifested independent of Ā-extraction. However, the voice system arguably serves a disambiguating function in Ā-extraction across the clause-initial verb. In (93-96), verbs must feature Agent Focus morphology when the transitive subject is extracted, and likewise they must feature Patient Focus morphology when the transitive object is extracted. Does this count as an ergative extraction asymmetry?

Seediq <span style="float:right">(Holmer 1996, p.86, p.86, p.74, p.74)</span>

(93) a. *ima wada qmita Pawan?*
   who PRET see.AF Pawan

'who saw Pawan?'

b. *ima wada qtaun na ka Pawan?*
   who PRET see.PF NA KA Pawan

'who did Pawan see?'

(94) a. *seediq* [ *mnapa*      *laqi* ]
person    carry.AF.PRET child

      'the person who carried a child'

    b. *laqi* [ *napa*      *na ka seediq* ]
child    carry.PF.PRET NA KA person

      'the child carried by the person'[28]

Kapampangan                         (Mirikitani 1972, p.181, p.181, p.190, p.193)

(95) a. *ninu ing sinulat*
who DEF wrote.AF

      'who is the (one who) wrote?'

    b. *nanu ing isulat mu*
what DEF write.PF 2SG

      'what is the (thing) you will write?'

(96) a. *ing babaing* [ *tuturung*    *Inglis*    *king anak ku*      ]
DEF woman    teaching.AF English to    child 1SG.POSS

      'the woman who is teaching English to my child'

---

[28] In (93-96), *ka* is a definiteness particle, according to Holmer (1996). It typically marks a definite clause-final object in the Patient Focus construction (93b), or a definite clause-final subject in the Agent Focus construction (93a). Occasionally, however, it can also mark a clause-final object in the Agent Focus construction, as long as the subject takes the form of a clitic, as in (i).

Seediq                                                  (Holmer 1996; p.67, p.66)

i. *qmnita*     *=ku ka boyak*             ii. *puqun na huling ka wawa*
see.AF.PRET 1SG KA boar                 eat.PF NA dog    KA meat

  'I saw the boar (but I didn't see the hunters)'        'the dog will eat the meat'

In addition, the particle *na* marks a definite subject in the Patient Focus construction, as in (ii). It cannot be considered an actual ergative marker, though. Compare with (91b): an indefinite transitive subject in the Patient Focus is bare.

b. *panyulat me    ing <u>lapis</u>* [ *a   ibye   ku   keka* ]
<br>write    2SG>3SG DEF pencil   REL give.PF 1SG 2SG

'you write with the pencil I gave you'

The voice system in these languages is pervasive. It occurs in some shape or form in every clause, whether or not Ā-extraction has taken place. However, the fact remains that voice morphology does seem to serve a disambiguating function, providing a vital cue to the listener as to which argument has been extracted. Does this count as an extraction asymmetry? Recall that in Paumarí, it was observed that the grammar makes use of two clausal frames, and using the morphological case associated with them was not considered to qualify as an extraction asymmetry.

Here, however, there are reasons to consider voice systems of Seediq and Kapampangan as legitimate instances of ergative extraction asymmetries. As the diagram in (97) shows, voice morphology is consistently associated with <u>all</u> types of Ā-movement, and with significant changes to declarative word order. In Seediq and Kapampangan, sentences without Ā-extraction are V-initial. However, if Ā-extraction does take place, the resulting clause is no longer V-initial, and the accompanying voice morphology serves a crucial disambiguation function. It just so happens that, in addition to being strongly associated with Ā-movement, the voice system has a default status such that it also occurs in plain, declarative V-initial clauses with no Ā-movement as well.

(97) *Clausal structure in Paumarí and Seediq: a comparison*

|  |  |  |  |  | |  |  |  |
|---|---|---|---|---|---|---|---|---|
| Paumarí | declarative | S-ERG | V | O | *or* | O-ACC | V | S |
|  | Wh-question | Wh-ERG | V | O |  | Wh-ACC | V | S |

|  |  |  |  |  |  |  |  |  |
|---|---|---|---|---|---|---|---|---|
| Seediq | declarative |  | $V_{AF}$ | O | S | *or* | $V_{PF}$ | S O |
|  | <u>all</u> Ā-extraction | S | $V_{AF}$ | O |  | O | $V_{PF}$ | S |

Therefore, for the purposes of this survey, voice morphology in Seediq and Kapampangan are indeed considered instances of ergative extraction asymmetry, with the caveat that it exceptionally occurs in plain, declarative clauses as well. Furthermore, the disambiguation hypothesis makes a good prediction for these two languages. If the crucial voice morphology didn't occur in (93-96), the identity of the extracted argument would be unclear. In other words, it is <u>not</u> the case that <u>extra</u> morphology occurs in instances of transitive subject extraction only, in addition to the existing voice morphology.

### 5.4.6 Limitations of the disambiguation hypothesis

So far, we have presented ample evidence that the main predictor for ergative extraction asymmetries is ambiguous Ā-extraction. The disambiguation hypothesis thus makes good predictions two ways: if a language's typological profile provides enough cues such that extraction will never be ambiguous, no extraction asymmetry is predicted to occur. As described in Sections 5.4.1-5.4.3, the disambiguation hypothesis correctly predicts the lack of an extraction asymmetry for 53 data points in the total dataset of 99.

Furthermore, if a language's profile lacks sufficient cues, such that Ā-extraction will potentially be ambiguous, the hypothesis predicts that an extraction asymmetry should indeed occur. As described in Section 5.4.4, the correct prediction of an ergative extraction asymmetry is made for 31 data points in the dataset. This leaves fifteen remaining datapoints to be accounted for the dataset; in this next section, the discussion will cover instances where the disambiguation unfortunately makes an incorrect prediction. These incorrect predictions come in two varieties: instances where an extraction asymmetry was predicted to occur, but it did not (5.4.6.1), and instances where extraction asymmetries still occur, but they do not seem to be predicted by the need for disambiguation (5.4.6.2-5.4.6.3).

### 5.4.6.1 Some languages simply tolerate ambiguity!

Ambiguity is a well-known property of language. In fact, it even comes in different varieties, such as structural ambiguity (98a) and lexical ambiguity (98b). The fact that ambiguity exists in likely every known natural language indicates that, to a certain degree, it is not an unmanageable problem.

(98) a. *I saw the man with the telescope.*

   b. *bank:* (n.) an institution where one can place and borrow money and take care of financial affairs <u>or</u> (n.) an edge of river, lake, or other watercourse

It should come as no surprise, then, that in some languages, extraction ambiguity is tolerated as well. For example, Itzaj is a Mayan language with the same typological profile as Kaqchikel and the other Mayan languages from Section 5.4.4. As such, the disambiguation hypothesis makes the prediction that it should feature

some kind of unexpected morphology to distinguish instances of transitive subject and transitive object extraction. And yet, as the data in (99) shows, the ambiguity is simply tolerated, and speakers must instead rely on context to obtain the correct interpretation.

Itzaj                                                                    (Hofling 2000; p.422/ A. Hofling, p.c.)

(99) a. *maax t-uy-il-aj*
      who   CMPL-3SG.ABS-see-CMPL.TR
      'who saw him?' or 'who did he see?'

    b. <u>*b'alum*</u> [ *a'*   *t-u-kin-s-aj*             *a'*   *winik-e*   ]
      jaguar   DET CMPL-3SG.ABS-die-CAUS-DEP.TR DET man-TOP
      'a jaguar that killed the man' or 'a jaguar that the man killed'

In the ergative Ā-extraction dataset, only two languages exhibit this kind of tolerance for ambiguity: Itzaj and Tseltal. In both cases, the ambiguity tolerance extends to both Wh-questions and relative clauses, for a total of four data points where the disambiguation hypothesis makes an incorrect prediction.

### 5.4.6.2 The motivation for an extraction asymmetry may be morpho-phonological

In at least one language, the motivation for an ergative extraction asymmetry doesn't seem to be disambiguation, but rather morphophonological in nature. Katukina is an SVO language in which the ergative marker *na=* modifies the transitive subject immediately preceding it, but it must cliticize to the verb immediately following it, as in (100a). In this rigid SUBJECT-*na*=VERB configuration all elements must remain "strictly adjacent", according to Queixalós (2010) – as (101a-101d)

demonstrate, adjuncts may occur quite freely in the clause, but they may not occur between the transitive subject and the transitive verb.

Katukina                    (Queixalós 2010; p.217, p.217, p.240, p.240, p.240, p.240)

(100) a. *pi:da   na=ti     paiko*
          jaguar ERG=kill grandfather
      'the jaguar killed grandfather'

      b. *tyuku  wa:pa*
          die    dog
      'the dog died'

(101) a. *niama pi:da   na=ti    paiko*
          then   jaguar ERG=kill grandfather
      'then the jaguar killed grandfather'

      b. *pi:da   na=ti    niama paiko*
          jaguar ERG=kill then   grandfather
      'then the jaguar killed grandfather'

      c. *pi:da   na=ti    paiko       niama*
          jaguar ERG=kill grandfather then
      'then the jaguar killed grandfather'

      d. * *pi:da   niama na=ti    paiko*
            jaguar then   ERG=kill grandfather
      'then the jaguar killed grandfather'

Given this apparently inviolable constraint, it is not surprising that the following asymmetry occurs in the domain of clefting. For O and S arguments, the clefted argument occurs clause-initially, immediately followed by the focus particle

*kana* (102b-102c). However, for A arguments, the occurrence of *kana* between the transitive subject and the transitive verb would violate the morphophonological constraint in question; as a result, when the transitive subject is clefted, the verb must be detransitivized with antipassive morphology, which in turn no longer licenses ergative case marking, and thus the constraint is by-passed, as in (102a).

Katukina                                     (Queixalós 2010; p.258, p.245, p.245)

(102)  a. *itiyan kawahiri kana wa-duni     tyon*
          this     cat       FOC  ANTI-catch rat
       'it's this cat that caught the rat'

    b. *a-obatyawa      kana Aro na=nuhuk kariwa*
       3SG.POSS-wife FOC  Aro ERG=give white.man
       'it's his own wife that Aro gave to the white man'

    c. *Maranmaran na=tyo       kana tona    tyo*
       Maranmaran GEN=daughter FOC  go.away EXCL
       'it's Maranmaran's daughter that went away!'

The asymmetry in (102) is clearly an example of an ergative extraction asymmetry; however, because it is not motivated by disambiguation, this counts as instance where the current hypothesis does not make the correct prediction. Indeed, this data indicates that while disambiguation does serve as the <u>main</u> predictor of an ergative extraction asymmetry, there may exist independent constraints which motivate them in a minority of languages.

### 5.4.6.3   Unknown independent constraints

Recall from Section 5.1 that West Greenlandic exhibits an ergative extraction asymmetry in the domain of relative clauses (103) but not Wh-questions (104). West

Greenlandic's typological profile, specifically the availability of morphological case, predicts that no such asymmetry should occur, in either relative clauses or Wh-questions. In principle, an absolutive-marked object within a transitive relative clause should provide a sufficient cue that extraction of the transitive subject has taken place. If that is indeed the case, then how do we account for the data in (103), a transitive subject relative clause where the verb must occur with antipassive morphology, and the transitive object with oblique case?

West Greenlandic                                    (Bok-Bennema 1991; p.69, p.65, p.239)

(103) a. *piniartuq*       [   *nannu-mik*       *tuqut-si-suq*       ]
         hunter.ABS        polar.bear-MOD kill-**ANTI**-PART
      'the hunter who killed the polar bear'

     b. *nanuq*           [   *Piita-p*     *tuqu-ta-a*               ]
         polar.bear.ABS       Piita-ERG kill-PART-3SG.SUBJ
      'the polar bear that Piita killed'

     c. [   *ipis-suq*            ]
            be.sharp-PART
      'the thing that is sharp'

West Greenlandic                                    (Fortescue 1984; p.23, p.16, p.24)

(104) a. *kia*        *uqaatig-aa*
         who.ERG talk.about-3SG>3SG.INT
      'who talked about it?'

     b. *suna=ana  Jaaku-p     siulittaasu-u-vvigi-gaa*
         what=that Jaaku-ERG chairman-be-have.as.place.of-3SG>3SG.PART
      'what is Jaaku chairman of?'

c. *kina tikis-sinnaa-va*
   who come-can-3SG.SUBJ.INT

   'who can come?'

I can only provide two highly speculative explanations: first, it may well be that a case-marking language such as West Greenlandic has a constraint on two consecutive morphologically-unmarked arguments. This scenario would occur, for instance, if a relativized transitive subject acts as intransitive subject of the main clause, as in (105). Another possibility, schematized in (106), is that allowing absolutive-marked transitive objects in relative clauses may lead to Garden Path effects if the head of the relative clause is subject of a main clause transitive verb itself. In this scenario, the first two arguments and the transitive verb would be parsed as belonging to the transitive main clause, leading to a Garden Path effect when the actual transitive object and transitive verb unexpectedly follow after that.[29]

(105) [ $N_{ABS}$ [ $N_{ABS}$ V ] ] V

(106) [ $N_{ERG}$ [ $N_{ABS}$ V ] ] $N_{ABS}$ V

In any case, these suggestions remain highly speculative, and they are not supported by any kind of hard data. At this point, it simply suffices to say that the

---

[29] Note however that an analysis of Garden Path avoidance won't work for a language such as Chukchi. In that language, relative clauses are pre-nominal and yet the rest of the facts remain the same (i-ii). Thanks to Maria Polinsky for bringing this to my attention.

Chukchi                                        (Polinsky to appear; p.22, p.13)

i. [ *məlgr-epə  ine-kune-lʔ-ən*        ] *ənpənačg-ən*         ii. [ *ənpənačg-e    kənnə-lʔ-ən*       ] *milger*
   gun-ABL   ANTI-buy-PART-ABS   old.man.ABS                   old.man-ERG buy-PART-ABS   gun.ABS

   'the old man that bought a gun'                               'the gun that the old man bought'

338

disambiguation hypothesis does not correctly predict the occurrence of an extraction asymmetry in West Greenlandic relative clauses. With further research, an independent constraint (such as the morphophonological one proposed for Katukina) may be uncovered to explain these West Greenlandic facts, and other recalcitrant areas where the disambiguation hypothesis incorrectly predicts no asymmetry. In sum, the disambiguation hypothesis makes this type of incorrect prediction for five total data points in the dataset; the summary of all types of incorrect predictions is provided in (107).

(107) *Data points (n=9) where the disambiguation hypothesis makes an incorrect prediction*

| extraction asymmetry is predicted, but does not occur | |
|---|---|
| Itzaj | Wh-questions, relative clauses |
| Tseltal | Wh-questions, relative clauses |

| extraction asymmetry is <u>not</u> predicted, but occurs anyways | |
|---|---|
| Katukina | focus fronting, relative clauses |
| Tukang Besi | Wh-questions$_{(1)}$, relative clauses$_{(1)}$ |
| West Greenlandic | relative clauses |

### 5.4.7 The ergative sample by the numbers

The table in (108) provides some quantitative analysis for the ergative sample in Chapter 5. Data is broken down by type of Ā-extraction and basic word order type. The second column lists the number of data points per category. The third

column indicates whether the disambiguation hypothesis makes a good prediction with regards to Ā-extraction, and finally the fourth column indicates what percentage of data points actually exhibit an extraction asymmetry in each category.

(108) *The disambiguation hypothesis & the ergative sample: results*

| | # data points | % correct | | % asymmetry | |
|---|---|---|---|---|---|
| <u>all</u> ergative data | 99 | 90/99 | = 90.9% | 42/99 | = 42.4% |
| Wh-questions | 50 | 46/50 | = 92.0% | 16/50 | = 32.0% |
|   V-initial | 18 | 15/18 | = 83.3% | 12/18 | = 66.7% |
|   V-medial | 10 | 9/10 | = 90.0% | 3/10 | = 30.0% |
|   V-final | 22 | 22/22 | = 100.0% | 1/22 | = 4.5% |
| relative clauses | 49 | 44/49 | = 89.8% | 26/49 | = 53.1% |
|   V-initial | 19 | 16/19 | = 84.2% | 16/19 | = 84.2% |
|   V-medial | 7 | 6/7 | = 85.7% | 2/7 | = 28.6% |
|   V-final | 23 | 22/23 | = 95.7% | 8/23 | = 34.8% |

The main observation is that the disambiguation hypothesis does seem to make very good predictions overall, being correct 91.8% of the time in this sample. Furthermore, that high rate of correct predictions stays consistent when broken down by Ā-extraction type: the disambiguation hypothesis makes a correct prediction 92% of the time for Wh-questions, and also 89.8% of the time for relative clauses. Therefore, globally-speaking, it does seem like the disambiguation hypothesis can reliably predict which languages will an extraction asymmetry and which won't.

Some additional observations are in order. When looking at the data from the perspective of basic word order, the disambiguation hypothesis does seem to fare better for V-final languages, and worse for V-initial languages. Consistently, in each extraction type, the success rate for V-initial languages is below the 90% mark, but for V-final languages it is above the 90% mark. This is due in part to the four Mayan data points, discussed in Section 5.4.6.1, where extraction ambiguity was simply tolerated.

Globally, extraction asymmetries occur at a rate of 42.4% across the sample. When separated by extraction type, they occur more often in relativization (53.1%) than in Wh-extraction (32%). Why should this be the case? An explanation may be found in the types of strategies available. For example, there are only two Wh-extraction strategies (Wh-fronting and Wh-*in situ*), and one of them is inherently unambiguous (Wh-*in situ*). Therefore it follows that if a high proportion of Wh-questions will always be unambiguous, they will not require any disambiguating morphology.

Relative clauses, on the other hand, feature a wider variety of extraction strategies cross-linguistically. Some of them are more likely to lead to interpretation difficulties (appositive relative clauses), and others are by definition ambiguous (internally-headed relative clauses). Furthermore, looking at the most common type of relativization strategy (externally-headed relative clauses), their bi-clausal nature may present extra difficulties for disambiguation. It is typical in an externally-headed relative clause for the unextracted element to bear morphological case, but on the other hand, the relativized head must bear case which reflects its role in the main clause. In effect, this leads to a loss of half the morphological cues available in

341

this type of Ā-extraction. In isolation, the disambiguation hypothesis predicts that morphological case on only the non-extracted element is sufficient for unambiguous extraction, but in specific cases (ex.: Trumai) it was argued that the bi-clausal nature of relative constructions leads to a compounding effect, which does require additional disambiguating morphology after all.

If this approach is on the right track, bi-clausal relative constructions stand in contrast with Wh-extraction, which is typically monoclausal and in which both arguments – extracted and non-extracted – can bear morphological case for maximum potential disambiguation. This is speculative, but it suggests that additional confounding factors (multiple morphological alignments, extremely unreliable word order, etc) cannot push Wh-questions past the 'threshold' where an extraction asymmetry is warranted, simply because both arguments can always clearly be case-marked.

The data in (108) also reveals that actual rates of extraction asymmetries may be lower than previously reported. The figure of 42.4% reported for this sample does not seem compatible with the claim in Polinsky (to appear) that "in sum, most morphologically ergative languages represented in the WALS sample behave like Tongan [i.e. exhibit an extraction asymmetry], not like Basque [i.e. do not exhibit an extraction asymmetry]". This claim is repeated in Polinsky (2016), where, based on an ergative and accusative sample of 32 languages each, Polinsky concludes that syntactic ergativity is extremely common among ergative languages, while syntactic 'accusativity' is extremely rare, to such an extent that morphologically languages are posited to be fundamentally different in that they have an "independently available element of language design" (p.25). In the next section, it will be argued that syntactic accusativity is actually quite common, and furthermore that it occurs at

comparable rates too, as predicted by a disambiguation hypothesis which doesn't make alignment-specific predictions.

The chart in (109) is a visual representation of all 99 data points in the ergative sample, broken down by subtype according to the discussion in Sections 5.4.1-5.4.6. Green slices represent data points where the disambiguation hypothesis correctly predicted no extraction asymmetry. Orange slices represent data points where the disambiguation hypothesis correctly predicted an extraction asymmetry. The blue slice represents data points where the disambiguation hypothesis correctly predicted an extraction asymmetry in a morphologically ergative language, but what occurred was actually an accusative extraction asymmetry; these special cases are discussed in Section 5.5.2 below. Finally, the red slice represents data points where the disambiguation hypothesis made an incorrect prediction.

(109) *The 98 data points in the ergative sample by subtype*
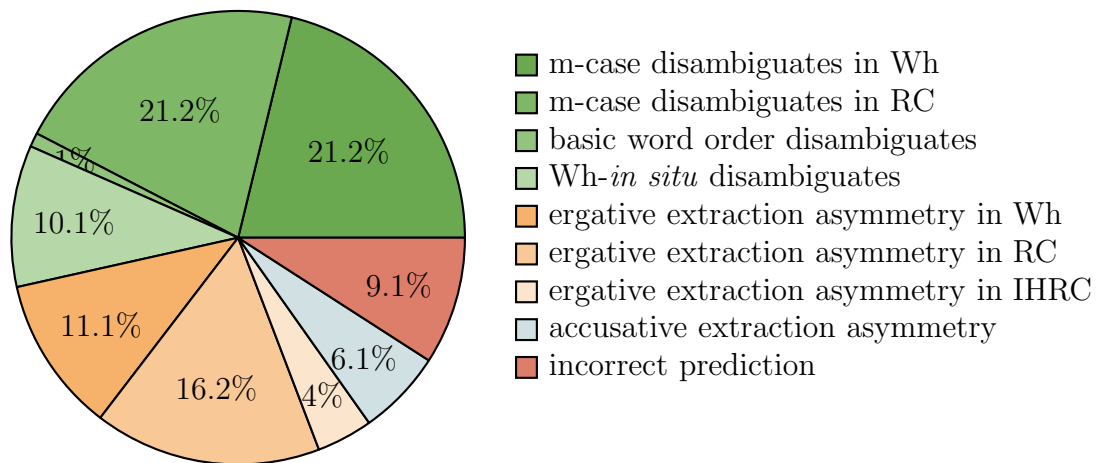
## 5.5 The disambiguation hypothesis goes beyond alignment

The main claim made in this chapter is that the best predictor of an extraction asymmetry is the lack of reliable cues which will help identify the extracted element in Ā-extraction scenarios. As a result, and contrary to the syntactic analyses in Section 5.2, the link between ergative morphological alignment and extraction is not necessarily a given. Therefore, if this approach is on the right track, we might expect to find instances of ergative extraction asymmetries in morphologically accusative languages, as long as the typological profile is right. Furthermore, we might expect to find 'accusative' extraction asymmetries, but in morphologically ergative languages too. In the rest of this section, we will present evidence of just that.

### 5.5.1 Ergative extraction, accusative morphology

There is some evidence that morphologically accusative languages may exhibit ergative extraction asymmetries, if the circumstances are right. For example, Nehan is an OVS Austronesian language with strictly accusative alignment: as the data in (110a-110b) shows, a verbal prefix agrees with both the transitive and intransitive subject. In addition, a verbal suffix may agree with transitive objects, as in (110a). Arguments themselves are not marked by case.

Nehan (Glennon 2014; p.19, p.44)

(110) a. *o      keketik  k-e-eit=isi=r                    kilon*
       ART.PL child   PERF-3SG.SUBJ-bite=PL.P=LIG coconut.crab
      'the coconut crab bit the children'

   b. *k-e-la              manasa=mpe inggon*
       PERF-3SG.SUBJ-go now=EMP    3SG
      'he went now indeed'

In Wh-questions, however, Nehan exhibits an extraction asymmetry that can only be characterized as ergative. In (111b-111c), the extraction of either the transitive object or the intransitive subject proceeds straightforwardly, with no change in verbal morphology. However, only in the case of transitive subject Wh-questions, verbs must be intransitivized by converting to the antipassive: in (111a), the transitive object is demoted to oblique status as marked by the preposition *tang*, and the suffix which corresponds to the direct object is no longer realized.

Nehan                                                        (Glennon 2014; p.107, p.107, p.107)

(111) a. *mai=s*     *k-a-kuse*                *tang* *Pita*
          who=3PL PERF-3PL.SUBJ-hold OBL Peter
          'who arrested Peter?'

      b. *mai k-a-kuse=in=r*                     *polis*
         who PERF-3PL.SUBJ-hold=3SG=LIG police
         'who did the police arrest?'

      c. *mai k-e-la*
         who PERF-3SG.SUBJ-go
         'who went?'

Of course, this makes perfect sense given Nehan's typological profile. Wh-words are fronted, producing Wh V NP word order in both transitive subject and transitive object Wh-questions, and arguments are not marked with morphological case.[30] In this regard, Nehan is simply the accusative doppelganger to a language

---

[30] Since default word order in Nehan is O V S, *a priori* one would expect Wh O V linear word order in transitive subject Wh-extraction. The observed Wh V O word order in (111a) implies verb raising in Wh-questions in the sense of Pollock (1989). In any case, the fact remains that Wh V NP word order in both types of transitive Wh-questions, in conjunction with the lack of morphological case, present an ambiguous extraction scenario consistent with the principles in Section 5.3.

like Selayarese. And in order to distinguish between potentially ambiguous cases of Wh-extraction in Wh-questions, it marks the transitive object extraction with a conspicuous change in verb morphology.

Nehan relative clauses, too, show the same ergative extraction asymmetry morphology. In (112b), the verb morphology for a relativized transitive object is similar to that in plain, declarative sentences. However, for relativized transitive subjects, the verb must be detransitivized and the transitive object must be demoted to oblique status (112a).

Nehan                                                              (Glennon 2014; p.111, p.111)

(112) a. <u>*kilon*</u>=*ene* [ *ge-k-e-eit*   *baka uoum totoguo* ]
   coconut.crab REL-PERF-3SG-bite first ahead OBL.1SG
   'the coconut crab here which already bit me'

  b. <u>*kilon*</u>=*ene* [ *ge-k-u-haluh*=*i*=*o*      ]
   coconut.crab REL-PERF-1SG-hit=SG.OBJ=1SG.SUBJ
   'the coconut crab here which I hit'

Bella Coola, a Salishan language, provides further evidence for ergative extraction phenomena in a morphologically accusative language. Bella Coola is a VSO language, whose system of head-marking is in nominative-accusative alignment, as in (113a-113b).[31]

---

[31] According to Forrest (1994), Bella Coola has reanalyzed the ergative alignment of 3rd person agreement inherited from Proto-Salish. Like in other Salish languages, however, the agreement pattern for non-3rd persons remains nominative-accusative, but with hierarchical constraints. For example, the passive construction must be used when 2nd person participant is being acted upon by a 3rd person participant. In any case, the end result is that head-agreement in Bella Coola is fully nominative-accusative; see the appendix in Forrest (1994) for more details.

Bella Coola (Forrest 1994; p.151, p.151)

(113) a. *k'x-i-s*        *ti-ʔimlk-tx*     *ci-xnas-cx*
see-3SG.OBJ-3SG.SUBJ ART-man-ART ART-woman-ART

'the man sees the woman'

b. *ƛ'ap-s*     *ti-ʔimlk-tx*
go-3SG.SUBJ ART-man-ART

'the man goes'

Again, Bella Coola provides the ideal typological profile for ambiguity: V-initial word order, no morphological case, and Wh-fronting/post-nominal relative clauses. Again, the disambiguation hypothesis makes the correct prediction that this is exactly the kind of language which would require extra morphology to distinguish transitive subject extraction from transitive object extraction. In (114a) and (115a), respectively, extraction of the transitive subject entails an unexpected change of 3>3 transitive verb morphology to *-t*, instead of the normal *-i-s* sequence of (113a). In other cases of extraction, no unexpected changes in agreement morphology occur.[32]

Bella Coola (Davis & Saunders 1997; p.78, p.78, p.78, p.99, p.98, p.100)

(114) a. *stamks ti-χm-t*       *ti-q́umsxiwaɬ-tx*
what    ART-bite-3SG>3SG ART-cat-ART

'what bit the cat?'

---

[32] Note also that according to Forrest (1994), intransitive 3SG *-s* is not "fully obligatory", and may be dropped for stylistic reasons, as in (114c) and (115c). Davis & Saunders (1997) suggest that the use of 3SG intransitive *-s* is conditioned by what they call 'peripherality'; if the 3SG intransitive argument has been previously mentioned in discourse, it will be marked with *-s*; if it has not, the marking will be null.

b. *stamks ti-χm-i-s*                *ti-waćˊ-tx*
   what    ART-bite-3SG.OBJ-3SG.SUBJ ART-dog-ART

   'what did the dog bite?'

c. *stamks a-ćkta-∅*
   what    PRE-happen-3SG.SUBJ

   'what happened?'

(115) a. *ti-<u>nusʔūlχ</u>*   [   *ti-k'x-t*              *ti-ʔaq̓ʷlikʷ-tx*        ]
        ART-thief          ART-see-3SG>3SG ART-policeman-ART

   'the thief who saw the policeman'

b. *ti-<u>nusʔūlχ</u>*   [   *ti-ʔaq̓ʷ-i-s*                 *ti-ʔaq̓ʷlikʷ-tx*        ]
   ART-thief          ART-lock.up-3SG.OBJ-3SG.SUBJ ART-policeman-ART

   'the thief whom the policeman locked up'

c. *ti-<u>nusʔūlχ</u>*   [   *ti-λ'ikm-∅-tx*            ]
   ART-thief          ART-run-3SG.SUBJ-ART

   'the thief who is running'

It is easy to see from the data in (114-115) that the unique *-t* morphology in cases of transitive subject extraction is the only thing preventing them from being interpreted as transitive object extraction. In other words, if (114a) featured regular transitive agreement morphology, it would mean 'what did the cat bite?; likewise, if (115a) featured regular transitive agreement morphology, it would mean 'the thief whom the policeman saw'.

Cocama is yet another language with nominative-accusative morphology, but which exhibits an ergative extraction asymmetry. The extent of accusative alignment in the language is limited to third person pronouns (other pronouns and NPs show neutral alignment). In (116-117), the same nominative pronoun encodes transitive

subjects and intransitive subjects; on the other hand, a separate enclitic pronoun encodes transitive objects.[33]

Cocama (Vallejos Yopán 2010; p.513, p.513, p.513, p.513)

(116) a. *uri   chikari=ura*
      3SG look.for=3SG.OBJ

    'he/she looks for he/she/it'

    b. *uri   tsatsatsɨma*
      3SG scream

    'he/she screams'

(117) a. *ya   chikari=ay*
      3SG look.for=3SG.OBJ

    'he/she looks for he/she/it'

    b. *ya   tsatsatsɨma*
      3SG scream

    'he/she screams'

In Wh-questions, Cocama consistently makes use of its V-medial basic word order to distinguish between transitive subject Wh-questions and transitive object Wh-questions. In (118a-118b), the former are identified by Wh V O linear sequencing, while in (118c-118d) the sequence Wh A V identifies the latter.

Cocama (Vallejos Yopán 2010; p.545, p.259, p.545, p.259, p.723)

---

[33] The data in (116-117) illustrates Cocama's speaker-oriented pronominal system. The pronouns in (116) correspond to a male speaker, and the pronouns in (117) correspond to a female speaker.

(118) a. *awa=tipa yumi uni=uy      ami-tsu*
        who=Q    give   water=PST1 grandfather-DAT

    'who gave water to grandfather?'

   b. *awa ayuka=n=uy*
      who hit=2SG=PST1

    'who hit you?'

   c. *awa=ka=tipa mijiri  yumi uni=uy*
      who=LOC=Q Miguel give   water=PST1

    'to whom did Miguel give the water?'

   d. *awa n=umi=uy*
      who 2SG=see=PST1

    'whom did you see?'

   e. *awa=tipa ene*
      who=Q    2SG

    'who are you?'


However, when it comes to Cocama's relative clauses, things get a little more complicated. Relative clauses are verb-final, and in addition they can either be pre-nominal or post-nominal. Given that Cocama lacks morphological case, *a priori* the sequence N [ N V ] can correspond to either a relativized transitive subject or transitive object. So too can the sequence [ N V ] N correspond to either type of relative clause. Therefore it makes perfect sense, given the paucity of inherent cues in the domain of relative clauses, that Cocama makes use of a dedicated transitive subject relativizer *-tara*, as in (119a) and (120a). A separate relativizer *-n* serves for both transitive objects and intransitive subjects (119b-119c, 120b-120c).

Cocama    (Vallejos Yopán 2010; p.593, p.590, p.585, p.593, p.590, p.589)

(119) a. [ *nana      erura-tara*    ] *napitsara muna tsa=tukini=uy*
           pineapple bring-REL.A     man       steal  1SG=hammock=PAST

      'the man who brought the pineapple stole my hammock'

    b. *tana erutsu-ka* [ *rana yumi-n*      ] *karamina*
        1PL  bring-REI   3PL  give-NMLZR   corrugated.iron

      'we carry out the corrugated iron that they donate'

    c. *Victor* [ *ikara-n*      ] *awa*
        Victor   sing-NMZLR   person

      'Victor is a person that sings'

(120) a. *yawara* [ *tsa=mɨmɨra karuta-tara* ] *yapana=uy*
          dog       1SG=SON    bite-REL.A      run=PAST

      'the dog that bit my son escaped'

    b. *tsa*  *mɨmɨra* [ *yawara karuta-n*    ] *yapana=uy*
        1SG  son        dog    bite-NMLZR   run=PAST

      'my son that the dog bit escaped'

    c. *ya*  *uchima-ta*   *tsuwɨ* [ *pua-n*      ]
        3SG  go.out-CAUS  blood    rot-NMLZR

      'he takes out the blood which is decomposed'

In sum, the Cocama data in (116-120) illustrates a morphologically accusative language which is the counterpart to Trumai on the ergative side: Wh-questions lack any kind of extraction asymmetry, but due to their inherent complexity, relative clauses *do* exhibit an extraction asymmetry in that domain only.

Finally, morphologically accusative languages can make use of the internally-headed relativization strategy, as well. Southern Pomo is one such language: in

(121a-121b), the head of the relative clause occurs in its normal clause-internal position, and yet its head is interpreted as an argument of the main clause.

Southern Pomo                                                    (Walker 2013; p.448, p.449)

(121) a. [ *Ɂaṭ:i-kʰe*        *ċihṭa*     *Ɂač͓ʰ:a-w=wan*        ] *dohlo-ḱ*
        3SG.COREF-POSS  bird.game  catch-PERF=DET.OBJ   take.off-DIR

    '(he) took off his own game that (he) had caught'

   b. [ *ma-k:a-ċ*                    *š:ba:ṭʰaw*  *maṭʰ:i*  *miṭ:i=:čon* ] *Ɂuhṭehṭe-w*
        3.COREF-grandmother-GS poor         blind    lie=PAT        tell-PERF

    '(they) told their poor blind grandmother who was lying (there)'

However, transitive internally-headed relative clauses still face the same problem: of the two core arguments, which one should effectively be 'promoted' to main clause argument status? Recall from (87-89) that the solution in Shipibo, Belhare and Apinajé was to simply arbitarily ban transitive subjects from serving as heads of internally-headed relative clauses. It turns out that the exact same constraint applies in Southern Pomo. Despite being morphologically accusative in alignment, Southern Pomo internally-headed relative clauses exhibit the same kind of interpretation restriction because transitive subjects alone cannot serve as their heads (N. Walker, p.c.).

    To summarize this section: it appears that ergative extraction asymmetries can indeed occur in morphologically accusative languages. In fact, we have found typological counterparts in accusative languages to many of the ergative languages discussed in Section 5.4.4. In Nehan and Bella Coola, syntactic ergativity is 'across the board' like in Kaqchikel, Gitksan, Lillooet and Selayarese. In Cocama, the asymmetry is confined to relative clauses, likely due to the unreliability of morphological

case and word order in that domain – exactly like in Trumai. Finally, in Southern Pomo, an ergative interpretation constraint is observed in internally-headed relative clauses, just like in Shipibo, Belhare and Apinajé.

Furthermore, the data in (110-121) presents a serious challenge to the syntactic analyses of Section 5.2. Insofar as Deal (2016) predicts that the morphologically unmarked argument should never be restricted, these examples are clear counterexamples, because it is the nominative-marked transitive subject alone which is barred from simple extraction, while the accusative-marked transitive object is not. It is also unclear how 'blocking' accounts such as Aldridge (2004, 2008, 2012) and Coon et al (2014) can account for these ergative extraction asymmetries, given that they are predicated on the assignment of absolutive morphological case, but the languages in (110-121) are morphologically accusative. Finally, Polinsky's (2016) null P analysis cannot account for this data either; by hypothesis, the null P configuration in her account is confined to morphologically ergative languages.

### 5.5.2 Accusative extraction, ergative morphology

The findings in the previous section are exactly what one would expect if extraction asymmetries as a phenomenon are actually independent from morphological alignment. Going even further – is it possible to find extraction asymmetries that we might characterize as 'syntactic accusativity', but in ergative languages?

Two unrelated Amazonian languages may indeed provide evidence for this configuration. Kuikuro is a Carib language with OVS basic word order and which exhibits ergative morphological alignment in both its head-marking and dependent-marking. In (122a-122b), pronominal transitive objects and intransitive subjects are

marked as a prefix on the verb; in addition, overt transitive subjects are marked with the *-heke* ergative suffix.

Kuikuro                                                    (Franchetto 2010; p.124, p.124)

(122)  a.  *u-ahetinhomba-tagü i-heke*
        1.ABS-help-CONT    3-ERG
        'he is helping me'

    b.  *u-ünkgü-tagü*
        1.ABS-sleep-CONT
        'I am sleeping'

In cleft constructions, the focused argument moves to the front of the sentence and is marked with an affirmative suffix *-ha*. For transitive and intransitive subjects, this proceeds straightforwardly, as in (66a) and (66c), respectively. However, for transitive subjects, something interesting happens: the object is clefted, but due to some independent requirement, the normally post-verbal transitive subject argument must be promoted to a pre-verbal position, as in (66b). In these instances of transitive object extraction, and only in these instances, the verb must be marked with what Franchetto (2010) calls an 'object marker' *ng-*.

Kuikuro                                                    (Franchetto 2010; p.145, p.145, p.144)

(123)  a.  *u-ingãtzu-ha*     *ekise-i hikutaha enge-ni-mbüngü*
        1.POSS-sister-AFF 3-COP turtle    eat-AGNR-SUBS
        'it was my sister who ate the turtle'

    b.  *hikutaha-ha ege-i*    *u-ingãtzu*   *ng-enge-tagü*
        turtle-AFF   DIST-COP 1.POSS-sister OM-eat-CONT
        'it was a turtle that my sister was eating'

354

c. *u-ingãtzu-ha    ekise-i t-iniluN-ta-tinhü-i*
   1.POSS-sister-AFF 3-COP 3.ANA-cry-CONT-PART-PNR-COP
   'it was my sister who was crying'

This object marker seems like a great candidate for an accusative extraction asymmetry as per the disambiguation hypothesis. Notice that due to the subject's movement to a pre-verbal position in (66b), that object marker is the only thing indicating that clefting of the object has taken place! In other words, if the verb was not marked with the object marker in (66b), that sentence would mean 'it was a turtle that was eating my sister' instead!

A very similar process happens in Karitiâna, a Tupian language. This language shows ergative alignment in its system of head-marking. In (124a-124d), the verbal prefix agrees with either the transitive object or the intransitive subject.

Karitiâna                              (Storto 1999; p.157, p.157, p.157, p.157)

(124) a. *yn   a-ta-oky-j         an*
        1SG 2SG.ABS-DECL-hurt-IRR 2SG
        'I will hurt you'

   b. *a-ta-opiso-t          an*
      2SG.ABS-DECLR-listen-N.FUT 2SG
      'you listened'

   c. *yjxa    ∅-na-ahee-t          iso*
      1P.INCL 3SG.ABS-DECL-blow-N.FUT fire
      'we blew the fire'

   d. *∅-naka-hỹrỹja-t        taso*
      3SG.ABS-DECL-sing-N.FUT man
      'the man sang'

In Wh-questions, the Wh-element is moved to the front of the clause, as in (125). For transitive and intransitive subjects, the derivation proceeds straightforwardly; in the former case, the absolutive prefix agrees with the non-Wh transitive object, and in the latter case, the intransitive particle *i-* takes up the absolutive 'slot'. However, in instances of transitive object Wh-extraction, after the Wh-word has fronted, the non-questioned transitive subject is 'promoted' to the absolutive slot, as in (125b). This quirk of grammar creates the exact same interpretive problem as in Kuikuro! To resolve the problem, the grammar of Karitiâna requires instances of transitive subject Wh-extraction to be accompanied with what Storto (1999) calls the "object focus" morpheme *ti-*.

Karitiâna                                  (Storto 1999; p.200, p.198 / Everett 2007; p.325)

(125) a. *morã a-sokõ'i-j      ano*
        Wh   2SG.ABS-tie.up-IRR 2SG

      'who is going to tie you up?'

   b. *mõrãmõn a-ti-hĩrã̃*
      what     2SG.ABS-O.FOC-smell

      'what did you smell?'

   c. *mora-mon i-oty-t*
      WH-COP   PART-bathe-N.FUT

      'who is it that bathed?'

In much the same way as in Kuikuro, this object focus morpheme is the only thing that prevents a question like (125b) from being interpreted as 'what smelled you?'. And because this morpheme targets transitive object extraction specifically,

it too should be considered an instance of an accusative extraction asymmetry (or 'syntactic accusativity') in a morphologically ergative language.

In total, there are four morphologically ergative languages in the dataset which exhibit an accusative extraction asymmetry – Karitiâna (Wh-questions and relative clauses), Kolyma Yukaghir (attributive relative clauses only), Kuikuro (focus fronting) and Yucatec (durative aspect only) – making up the last six data points in the ergative sample of Chapter 5.[34] The data from Karitiâna and Kuikuro are presented in this section because of the unusual asymmetry between morphological alignment and extraction alignment, but note however that all six of these data-points still count as successfully-predicted extraction asymmetries for the purposes of the disambiguation hypothesis: the lack of available cues makes for potentially ambiguous Ā-extraction in these languages, just like for the other morphologically ergative languages discussed in Section 5.4.4. The fact that is it the accusative-marked transitive object that is marked for extraction does not matter; the purposes of disambiguation are still served by this 'asymmetrical' strategy.

Finally, one last example of a potential accusative extraction asymmetry within a morphologically ergative language must be discussed. Heaton et al (2016) report that in Kaqchikel, an ergative extraction asymmetry in the form of the Agent Focus morpheme *-on* is indeed commonly attested in the domain of Wh-questions, consistent with the data from Section 5.1. However, experimental results show that in an elicited production task, a large proportion of speakers of Kaqchikel do not mark relativized transitive subject constructions with Agent Focus morphology, but rather with fully normal transitive verb morphology, as in (126a). Instead, those

---

[34] See Appendix C for accusative extraction asymmetries in Kolyma Yukaghir and Yucatec.

speakers opt to mark verbs in relativized transitive object constructions with passive morphology, as in (126b), especially if the relativized object is animate.

Kaqchikel (Heaton et al 2016; p.5, p.39, p.40)

(126) a. *ri* <u>*achin*</u> [ *ri* *n-∅-u-tij* *ri* *wotz'otz'* ]
      DET <u>man</u> REL INC-3SG.ABS-3SG.ERG-eat DET pork.rind
    'the man who is eating pork rinds'

   b. *ri* <u>*xtän*</u> [ *ri* *n-∅-tz'ub'-äx* *r-oma* *ri* *ixöq* ]
      DET <u>girl</u> REL INC-3SG.ABS-kiss-PASS 3SG.POSS-OBL DET woman
    'the girl who is being kissed by the woman'

This new relative clause data of Kaqchikel departs from what has been commonly attested in the past, where Agent Focus morphology is described as occurring in both Wh-questions and relative clauses uniformly (see, for example, Brown et al 2006). Indeed, this difference is described in Heaton et al (2016) as a relatively recent innovation among younger speakers – the relativization patterns described in (126) are commonly used by speakers in the 20-30 age group, but almost not at all by those in the 41-50 age group.

In any case, this recent phenomenon is interesting for at least two reasons. First, it provides direct evidence for the disambiguation hypothesis: in speakers which have lost disambiguating Agent Focus morphology in transitive subject relative clauses, it appears that relativizing transitive objects now requires passive morphology to compensate, fully shifting from an ergative extraction asymmetry to an accusative one in the domain of relative clauses, as need dictates. In other words,

it is not the case that Agent Focus morphology was lost in transitive subject relative clauses, but no corresponding change has occurred in transitive object relative clauses to compensate.[35]

Second, it provides more evidence of an accusative extraction asymmetry in a morphologically accusative language. Intriguingly, it actually provides evidence for a split in syntactic ergativity – among younger speakers of Kaqchikel, Wh-questions continue to exhibit an ergative asymmetry, but relative clauses appear to be shifting towards an accusative one.

This new data from Kuikuro, Karitiâna and Kaqchikel in (122-126) contributes to the growing evidence against the syntactic analyses in Section 5.2. Again, the morphological basis for each of those syntactic analyses leads to clear, incorrect predictions. In those analyses, the constraints on Ā-extraction occur as a direct result of either absolutive or ergative morphological case assignment, which then prevents the transitive subject from extracting in turn. However, for the new data presented in this section, morphological case does not seem to be the correct predictor, since it is the absolutive-marked transitive object which cannot extract instead, while the ergative-marked transitive subject and the absolutive-marked intransitive subject can extract freely.

---

[35] It is particularly telling that, as reported by Heaton et al (2016), passive morphology is much more likely to occur if that relativized O argument is animate. This makes perfect sense: pragmatics can often disambiguate with an inanimate argument, as in _house_ [ _saw man_ ], which is extremely unlikely to refer to 'the house that saw the man', regardless of morphology, but obviously 'the house that the man saw'. However, pragmatic factors cannot help to disambiguate a scenario such as _woman_ [ _saw man_ ], which can equally well correspond to 'the woman that the man saw' and 'the woman that saw the man'. And it is exactly with these relativized, animate O situations that passive morphology is occurring among younger speakers of Kaqchikel.

### 5.5.3 Completing the paradigm - accusative extraction, accusative morphology

So far, we have presented data of ergative asymmetries in languages of both ergative and accusative morphological alignment, and also of accusative asymmetries in languages of ergative alignment. If indeed this hypothesis is on the right track, we should also expect to find accusative extraction phenomena in morphologically accusative languages, completing the paradigm. Again, there exist at least two languages which are good candidates for this cell of the paradigm.

In Nuuchahnulth, pronominal clitics are in nominative-accusative alignment because they can only correspond to transitive or intransitive subjects, as in (127).

Nuuchahnulth                                                (Davidson 2002; p.311, p.339)

(127) a. *ḥa:ḥu:p-a:=ʔa:qλ='aλ=a:=aḥ*              *suẃa*
       instruct-CONT=INTENT=TEMP=IND=1SG 2SG
       'I shall instruct you'

   b. *tu:ḥ-uk=a:=aḥ*          *siẏa*
      afraid-DUR=IND=1SG 1SG
      'I for my part am afraid'

In terms of relativization strategies, Nuuchahnulth makes use of headless relative clauses – these are relative clauses which lack an overt head, as in (128). In (128b), however, a special morpheme *-čił* occurs in headless transitive object relative clauses only. The minimal pair in (128a-128b) illustrates how that one morpheme serves to distinguish between a headless transitive subject construction and a headless transitive object construction. This *-čił* morpheme seems like a straightforward candidate for an accusative extraction asymmetry.

(128) a. [ *yaqʷ=mit=ʔi:tq*        *qaḥ-sa:p*            *t'ux-šiƛ*                    ]
         one.who=PAST=DEF dead-CAUS.PERF kill.from.hiding-PERF

   'the one who had killed him from ambush'

   b. [ *yaqʷ-<span style="color:red">čił</span>=mit=ʔi:tq*        *qaḥ-sa:p*            *t'ux-šiƛ*                ]
         one.who-<span style="color:red">do.to</span>=PAST=DEF dead-CAUS.PERF kill.from.hiding-PERF

   'the one he had killed from ambush'

   c. [ *yaqʷ-ćuq-(y)a:=ʔi:tq*                  ]
         that.which-in.mouth=CONT=DEF

   'that which is in one's mouth'

Sandawe is a Khoe language which exhibits accusative morphology in two ways: subjects may be marked with the nominative suffix *-á:*, and verbs may feature accusative agreement with the transitive object, as in (129). However, neither of these are a requirement, as both forms of accusativity are optional. For example, in (130a) an intransitive sentence doesn't feature the nominative suffix, in (130b) a transitive sentence doesn't feature the nominative suffix either, and in (130c) a transitive verb lacks the transitive object marker.

(129) a. *búrì̥-á:*        *gìtł'é-ː̀*      *|ʰòrŏ̃:-sé-é*
         mouse-NOM clothing-SP have.hole-CAUS-3M.SG.OBJ

   'a mouse has made hole in a piece of clothing'

   b. *ⁿ|ĩ̂*      *tʃʰíà-á:*  *ʔíé-ː́*    *kópòkópò*
      body all-NOM stay-& shake

   'the whole body was shaking'

361

(130) a. *hèwéʔgâ-à*      *kéùtò-꞊ˊ |í-à*      *tʰâ*
       and.so-3M.SG.CL pig-SP    come-CONN run

       'and so the pig came running'

    b. *bô-xé:-꞊ˊ*      *hèwéxé: kʷá:*      *nìnéwì*    *m̀fâlmè-à*
       word-PL-SP DEM     NARR.3M.SG Nineveh king.GEN-3M.SG.CL

       *!'ŏ:-é-wà*
       get-3M.SG.OBJ-MULT

       'these words reached the king of Nineveh'

    c. *sá:*        $^n$*|úm̥sú-꞊ˊ-sṳ̀-á:*    *tɬ'àbísó:-sà*      *!'ŏwé*
       NARR.3F.SG wife-SP-3F.SG-NOM stomach-3F.SG.CL get

       'and then the wife became pregnant'

It may indeed be because of the unreliability of these cues, both in morphological case and in head-marking, that relative clauses in Sandawe exhibit an accusative extraction asymmetry. In (131a) and (131b), respectively, transitive and intransitive subjects are relativized via a possessive construction, according to Eaton (2010). However, in order to produce a relativized transitive object, a wholly different strategy is used. For example, in (132a), because the relativized transitive object is 3rd person masculine singular, the verb must feature the relativizer *-ì*; in (132b), a relativized inanimate plural transitive object requires the relativizer *-xé:*. The exact relativizer is conditioned by animacy and number of the object. Again, this is an instance of an accusative extraction asymmetry, because relativization of the transitive object is singled out by unique morphosyntax.

Sandawe                                (Eaton 2010; p.179, p.179, p.180, p.181)

(131) a. *hòsó* [ *mêlìtä*   *jàʔbé-sí-sò-꞊ˊ-sò*       ]
       3PL    boat-in work-POSS-3PL-SP-3PL

       'they who worked in the boat'

b. [ $g^w á$:-bò-sí-è-$\overset{\backslash}{\overset{\cdot}{\cdot}}$ ]
   thirst-POSS-3M.SG-SP

'he who has thirst'

(132) a. <u>hábúsà-$\overset{\backslash}{\overset{\cdot}{\cdot}}$</u> [ tʃí-á: pó-é-**ì**-$\overset{\backslash}{\overset{\cdot}{\cdot}}$
   condition  1SG-NOM 2SG.OBJ-3SG.M.OBJ-**PRO**-SP ]

'the condition which I gave you'

b. <u>bô-xé:-$\overset{\backslash}{\overset{\cdot}{\cdot}}$</u> [ tʃí-á: pó-Ɂ-wá:-**xé**:-$\overset{\backslash}{\overset{\cdot}{\cdot}}$ ]
   word-PL-SP  1SG-NOM 2SG.OBJ-3PL.INAN.OBJ-**PL**-SP

'the words I gave you'

To summarize the discussion on extraction asymmetries and alignment: in Section 5.4.4, multiple examples were shown of ergative extraction asymmetries in morphologically ergative languages. In this section, examples of ergative extraction in morphologically accusative languages, accusative extraction in ergative languages, and accusative extraction in morphologically accusative languages were shown too, which altogether provide a compelling argument that extraction asymmetries may in fact be independent from morphological alignment. Each of the four attested logical possibilities in this new paradigm is illustrated in (133) below.

(133) *Extraction asymmetries and morphological alignment: all four logical possibilities are attested*

|  | morphologically ergative | morphologically accusative |
|---|---|---|
| **ergative asymmetry** | many examples | Nehan, Bella Coola, Cocama, Southern Pomo |
| **accusative asymmetry** | Kuikuro, Karitiâna | Nuuchahnulth, Sandawe |

An analysis of extraction asymmetries in an expanded dataset of morphologically accusative languages is beyond the scope of this chapter; however, just looking at the data from the languages in Chapter 3 is quite revealing. Of those 32 languages for which clear extraction data was available, yielding 67 total data points, some interesting patterns emerge.

(134) *The disambiguation hypothesis & the accusative sample*

| | # data points | % correct | | % asymmetry | |
|---|---|---|---|---|---|
| <u>all</u> accusative data | 67 | 61/67 | = 91.0% | 27/67 | = 40.3% |
| | | | | | |
| Wh-questions | 30 | 29/30 | = 96.7% | 4/30 | = 13.3% |
| V-initial | 6 | 5/6 | = 83.3% | 4/6 | = 66.6% |
| V-medial | 10 | 10/10 | = 100.0% | 0/10 | = 0.0% |
| V-final | 14 | 14/14 | = 100.0% | 0/22 | = 0.0% |
| | | | | | |
| relative clauses | 37 | 32/37 | = 86.5% | 23/37 | = 62.2% |
| V-initial | 8 | 7/8 | = 87.5% | 6/8 | = 75.0% |
| V-medial | 14 | 11/14 | = 78.6% | 7/14 | = 50.0% |
| V-final | 15 | 14/15 | = 93.3% | 10/15 | = 66.6% |

In (134), the disambiguation hypothesis was calculated to be correct for those morphologically accusative data points 91% of the time, a figure remarkably similar to the 90.9% rate for the expanded ergative sample. Furthermore, some of the same trends from the ergative sample reappear: for instance, the disambiguation hypothesis appears to be more reliable for V-final languages for for V-initial ones, in both Wh-extraction and relativization. Furthermore, the overall rate of extraction

asymmetries was calculated at 40.3%, also a remarkably similar figure to the 42.4% rate reported for the expanded ergative sample. Extraction asymmetries are likewise more common in relative clauses than in Wh-extraction, although in this case the difference is even larger: 13.3% for Wh-extraction datapoints, 62.2% for relativization data points.

(135) *Extraction asymmetries in morphologically accusative languages*

| | |
|---|---|
| accusative extraction asymmetry: | Aguaruna, Awa Pit, Fongbe, Hixkaryana, Hungarian, Indonesian, Jola Bandial, Lele, Manipuri, Maori, Maricopa, Sandawe, Nuu-chahnulth, Sandawe, Somali, Turkish |
| ergative extraction asymmetry: | Bella Coola, Cocama, Nehan, Southern Pomo |

A proper analysis of extraction asymmetries in morphologically accusative languages should warrant a larger sample size. I leave it up to future work whether these generalizations would hold up to an expanded dataset. However, if the numbers reported for this smaller accusative sample are on the right track, they are extremely encouraging for the disambiguation hypothesis. Exactly as expected, extraction asymmetries occur at roughly the same rate regardless of alignment, and in addition the hypothesis can correctly predict these asymmetries at roughly the same rate regardless of alignment. In the next section, we conclude this chapter with a discussion of the impact of the findings presented in this chapter, and how they challenge existing views on the relationship between morphological alignment and extraction asymmetries.

### 5.5.4  In perspective - final comparison with competing syntactic approaches

This chapter has presented arguments that the best cross-linguistic predictor of extraction asymmetries, both ergative and accusative, is actually whether transitive Ā-extraction is potentially ambiguous in a given language. Competing syntactic accounts, outlined in Section 5.2, contend that the explanation for these extraction asymmetries somehow involves the assignment of morphological case. However, these competing analyses suffer from a range of problems, pointed out throughout the chapter, and thus they are summarized below.

- minor theoretical issues
    - not all ABS=NOM languages have an ergative extraction asymmetry (5.2.2)
    - some 'low ABS' languages have an ergative extraction asymmetry (5.2.2)
    - ERG-as-PP analyses lack synchronic falsifiability (5.2.3)
- cannot account for morphological diversity of ergative extraction asymmetries
    - ex.: Yucatec, Selayarese, Ayutla Mixe (5.4.4)
- cannot account for domain asymmetries between Wh-questions and RCs
    - ex.: Roviana, Trumai, Tiriyó (5.4.4)
- cannot account for internally-headed RC restrictions
    - ex.: Shipibo, Belhare, Apinajé (5.4.4)
- cannot account for ergative extraction asymmetries in accusative languages
    - ex.: Nehan, Bella Coola, Cocama, Southern Pomo (5.5.1)
- cannot account for accusative extraction asymmetries in ergative languages
    - ex.: Kuikuro, Karitiâna (5.5.2)

For these reasons, the disambiguation hypothesis is considered a superior account of extraction asymmetries cross-linguistically. That being said, the following counter-argument may be made against the disambiguation hypothesis: if indeed ergative extraction asymmetries are independent of morphological alignment, why is it that in the vast majority of morphologically ergative languages which do exhibit an asymmetry, it is in fact the *ergative* argument which is restricted in some way? Wouldn't one expect a 50/50 split, all things being equal, with a much higher proportion of Kuikuro/Karitiâna-type languages which are morphologically ergative but in which exhibit an accusative asymmetry? Doesn't the actual distribution in Sections 5.4.4 and 5.5 indicate some kind of close relationship between morphological case/alignment and extraction asymmetries?[36]

This is a fair point, and I can only speculate, but I would suggest that if the disambiguation hypothesis is on the right track, and languages really do innovate special morphology purely for disambiguating purposes, maybe it makes sense to pick out the argument which is already (independently) marked by some other part of the grammar (morphological alignment) as special in extraction, too. In other words, if disambiguation needs require either the A or O argument to be marked in some special way, but the morphology of the language *already singles out A* in some way, perhaps the grammar is more likely to 'piggyback' on this marked status of A in extraction as well, resulting in an ergative extraction asymmetry over an accusative one. Marking the O argument, although it achieves the same goal of disambiguation, may be viewed as more 'costly' since it may also entail marking the S argument superfluously, for which ambiguity is never an issue in the first place.

[36] Thanks to Jessica Coon for raising this point.

### 5.5.5 Is there any room for formal accounts of extraction asymmetries at all?

To be clear, the argument made in the preceding section is that the disambiguation hypothesis is superior to the syntactic accounts of Section 5.2 from a *typological* perspective. That is to say, when looking at the set of morphologically ergative languages globally, the functional need for disambiguation serves as a reliable predictor for extraction asymmetries (both ergative and accusative). In this regard, the disambiguation hypothesis is only in direct competition with a *subset* of those syntactic analyses; for example, only Deal (2016) and Polinsky (2016) make universal, cross-linguistic claims as to the underlying causes of syntactic ergativity across the board, claims which appear to be incompatible with many of the findings in this chapter for the reasons outlined above.

On the other hand – and in all fairness – language-specific analyses such as Aldridge (2004, 2008, 2012) and Coon et al (2014) do not explicitly make such claims. The scope of those analyses is restricted to a subset of Austronesian and Mayan languages, respectively. And in that regard they still remain quite useful, providing a model of grammar that generates grammatical and only grammatical sentences in those specific languages, regardless of what the cross-linguistic facts may be.[37]

To illustrate with just one example: as discussed in Section 5.2.2, Coon et al (2014) provide a syntactic analysis which accounts for ergative extraction asymmetries in languages such as Q'anjob'al (136a). But that is not all that their analysis accounts for. In addition, their analysis accounts for bi-clausal "crazy antipassive"

---

[37] Thanks to Bob Frank for raising this point at the LSA 2017 Annual Meeting.

constructions in which a modal verb selects for a non-finite embedded clause, as in (136b).

Q'anjob'al                                                                (Coon et al 2014; p.180, p.221)

(136) a. *maktxel max-ach     il-on-i*
          who     ASP-2SG.ABS see-AF-INTR
          'who saw you?'

      b. *lanan* [ *hach      hin-laq'-on-i*              ]
          PROG    2SG.ABS 1SG.ERG-hug-AF-INTR
          'I am hugging you'

Recall that under their 'blocking' analysis, the Agent Focus *-on* must be generated to assign case to the internal argument, thus allowing the external argument to escape from vP and undergo Ā-extraction (20). Via the exact same mechanism, Coon et al's analysis captures the fact that the *-on* morpheme also licenses the transitive object in these non-finite embedded clauses. Because no finite I head is available to provide absolutive case, the Agent Focus morpheme must be enlisted for the transitive object to be licensed in this environment as well (137).

In this way, a formal analysis such as Coon et al (2014) can capture several facts at once in a language like Q'anjobal, including its ergative extraction asymmetry, the "crazy antipassive" construction, as well as simple morpheme linearization facts in what they term 'high ABS' Mayan languages. Furthermore, such analyses are deliberately constructed to be compatible with existing, independently-valid theories in the syntactic literature. So, to answer the question posed in the header of this subsection: of course there is still room for formal accounts of extraction

asymmetries. It depends what the scope of the exercise is. For the big picture, the disambiguation hypothesis seems useful, but when zooming in to specific languages and language families, formal syntactic accounts undoubtedly contribute to our understanding of the grammatical nuts and bolts that generate acceptable sentences in those languages.

(137) *"Crazy Antipassive" in Q'anjob'al* (Coon et al 2014; p.222)

```
                          vP
                      ___/  \___
                    v          VoiceP
                    |        __/    \__
                   -i      PRO        Voice'
                  INTR              __/    \__
                               Voice          VP
                                 |          _/  \_
                                -on        V      hach
                                           |    2SG.ABS
                                          laq'
                                          hug
```
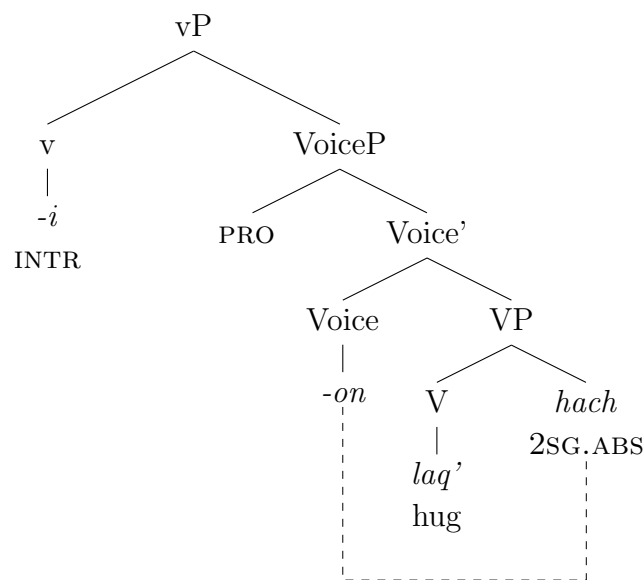
## 5.6   Conclusion

This chapter aimed to explore a certain mystery regarding ergative extraction asymmetries: some morphologically ergative languages have them, some don't, and some have them only in a subset of extraction environments. A disambiguation-based theory was developed which makes clear predictions: if the extraction of a transitive argument (subject or object) would *a priori* result in ambiguity, it is highly likely that that language will also exhibit an ergative extraction asymmetry

to disambiguate; however if that language already has abundant cues at its disposal such that transitive extraction is not ambiguous, it will not.

By looking at data from Wh-questions and relative clauses across a range of languages and language families, support for this hypothesis was provided. Furthermore, a tentative explanation was provided for languages such as Shipibo, which exhibit an asymmetry only in the domain of relative clauses: it is because relativization is more likely to be an inherently ambiguous strategy that some languages only require an asymmetry in that context.

In the latter half of the paper, the hypothesis was extended to morphologically accusative languages, showing that similar disambiguation phenomena are attested in those languages, too. In fact, all four permutations of extraction asymmetries (transitive subject *vs* transitive object) and morphological alignment (morphologically ergative *vs* morphologically accusative) were shown to be attested, strongly suggesting that, from a global perspective, the link between these two parts of the grammar is not as direct as previously thought. Some counter-examples to the disambiguation hypothesis were provided – but overall, across a large sample of ergative languages and a smaller sample of accusative ones, it was shown that the hypothesis makes very good predictions. The chapter then ended with a summary and discussion of the arguments against existing syntactic analyses of extraction asymmetries, which are judged inadequate to cover their full range cross-linguistically.

## Chapter 6

## GENERAL CONCLUSIONS, FUTURE RESEARCH

At the outset, the stated goal of this dissertation was to explore the phenomenon known as ergativity via three driving questions: what is the nature of the ergative patterns syncretism known as ERG=GEN and ERG=POSS? What is the best way to model these patterns of syncretism in the grammar? What is the relationship between morphological ergativity and syntactic ergativity? In this short section, we will review the main findings as they relate to each of those research questions, and highlight any directions for future research that seem relevant.

### What is the nature of the ergative patterns syncretism known as ERG=GEN and ERG=POSS?

As defined in Chapter 2, ERG=GEN and ERG=POSS are morphological patterns that shows a high degree of overlap in the expression of agreement and case in verbal and nominal forms; specifically, in the morphology that encodes transitive subjects and possessors of nouns. The main finding of the typological survey in that chapter is that these patterns of syncretism appear to be quite common among the world's morphologically ergative languages, occurring in over half the languages in the sample. Furthermore, they appear to transcend major typological parameters, such as marking type, basic word order and headedness, suggesting that they are a

property of 'ergativity' broadly-speaking and not simply the by-product of a more specific grammatical configuration.

Obviously, the most important task at this point is to check, by expanding the number of languages studied, whether the generalizations from Chapter 2 still hold up in a larger sample. Chapter 2's sample of 40 languages, encompassing 28 distinct language families, is certainly a good start, but of course more is always better. Many languages encountered later in the writing of the dissertation – namely, in Chapter 5 – seem to indicate a pattern of ergative syncretism (Katukina, Movima), while others do not (Drehu, Tukang Besi).

Another promising area of research barely touched upon in Chapter 2 involves ERG=GEN/ ERG=POSS and their relationship to alienable *vs* inalienable possession. Recall that for five of the languages exhibiting ERG=GEN or ERG=POSS, it was observed that they actually contained a more specific pattern where it was only *alienable* possession morphology in the nominal domain that overlapped with transitive subject morphology in the verbal domain. These preliminary findings are certainly intriguing, and again, more data is needed to see whether this generalization is a actually robust one, or just a blip on the radar.

In the same way, the sample of accusative languages in Chapter 3 ought to be supplemented to either confirm or disprove the claim that similar syncretic patterns involving accusative or nominative morphology are simply less common cross-linguistically. Some of the specific claims made in that chapter should be examined more closely in a broader sample, as well. Unlike the ergative patterns from Chapter 2, which manifest in all types of morphologically ergative languages (i.e. 'across the board'), the survey in Chapter 3 found that patterns such as ACC=POSS and

NOM=POSS are much more common than their dependent-marking counterparts ACC=GEN and NOM=GEN.

**What is the best way to model these patterns of syncretism in the grammar?**

To begin answering this question, the analysis in Chapter 4 used as its foundation a combined model of case and agreement based on previous work by Marantz (1991), Bobaljik (2008) and Baker (2013). This combined model, schematized in (1), is predicated on the assumption that case feeds agreement, and it was endorsed not only because it can account for a wide range of case and agreement data in the clausal domain, but because it captures some important typological generalizations as well.

(1) *Bobaljik-Baker model of case and agreement*

syntactic structure $\rightarrow$ case algorithm $\rightarrow$ agreement algorithm

However, not all the components of the combined model are drawn directly from the sources cited above. Indeed, some minor innovations were added, and it remains to be seen whether those contributions are theoretically well-motivated. For instance, the combined model's treatment of 'double agreement' is theorized to take place in the reverse order from the computation of case. In Marantz (1991), marked dependent case is generated first if the requirements are met, and then unmarked default case takes place last in the derivation. In this model, double agreement is generated by first computing unmarked agreement for that alignment (absolutive or nominative), and then computing marked agreement last (ergative or accusative).

This assumption accounts for the data in Chapters 2 and 3, but for now it lacks any independent support.

Another innovation implemented within the model is the assumption that null morphological case is computed first to feed agreement in languages suck as Kaqchikel. Again, this assumption is a logical necessity if agreement truly is always derived from case, however a simpler alternative is to straightforwardly compute agreement directly in these languages, skipping the previous step of generating null morphological case (see Coon 2017). If this alternative route is taken, and agreement is thus not always derived from case, is it still possible to capture the typological generalizations from Bobaljik (2008) in some other way?

Another question of theoretical importance concerns the actual locus of agreement, both in the clausal and nominal domains. Baker (2013) asserted that the locus of agreement in the clausal domain is parameterized such that nominative agreement always occurs in the I head, but in some languages, accusative agreement also occurs in I, and in others it is realized in the head of transitive vP. In Chapter 4 it was simply assumed that ergative agreement, like its marked counterpart in accusative languages, always occurs in v based on non-finite evidence from Kaqchikel. This assumption predicts that if only one type of agreement is available in languages with ergative alignment, it will always be ergative. It remains to be seen whether this typological generalization holds up over time. Furthermore, the analysis in Chapter 4 assumed that the locus of ergative agreement in nominals was the n head, the presumed counterpart to v in that domain, and that absolutive agreement may be modeled in D. Again, there is no theoretical precedent for a D head assigning agreement within the nominal domain, but the question remains: if languages such as

Tiriyó and Tenetehára exhibit ABS=POSS, where is ABS generated in nominals?

Finally, the main new claim made in Chapter 4 was that while Baker's (2016) analysis can account for ERG=GEN and ERG=POSS patterns, all the other attested syncretic patterns in the data (ACC=GEN, ACC=POSS, NOM=POSS, ABS=POSS) are generated in a fundamentally different way by the grammar. It was suggested that languages may opt to 'recycle' case or agreement paradigms from elsewhere in the grammar, in a pseudo-random way that may be influenced by non-syntactic factors. This is a major theoretical assumption, opening up a new range of possibilities for the architecture of grammar, and it remains to be seen whether independent evidence for such processes provide support for such an idea.

## What is the relationship between morphological ergativity and syntactic ergativity?

The main claims from Chapter 5 are as follows: first, that ergative extraction asymmetries serve primarily a disambiguating function in potentially ambiguous Ā-extraction scenarios. Quantitative data from that chapter, drawn from a sample of 53 ergative languages, shows that while these asymmetries occur in 42.4% of morphologically ergative languages, the disambiguation hypothesis correctly predicts whether or not a language will exhibit an asymmetry over 90% of the time.

The disambiguation hypothesis is also meant to account for the fact that extraction asymmetries may only occur in a subset of extraction scenarios, and furthermore than their distribution is not random: some languages will exhibit an asymmetry in both Wh-questions and relative clauses, some only in relative clauses, but no language appears to exhibit an extraction asymmetry only in Wh-questions

(2).

(2) *Morphological ergativity and ergative extraction asymmetries: 3 types*

|  | **morphological** | **ergative extraction asymmetry** | |
|---|---|---|---|
|  | **ergativity** | Wh-question | relative clause |
| Kaqchikel, Gitksan, Selayarese | ✔ | ✔ | ✔ |
| West Greenlandic, Roviana | ✔ | ✘ | ✔ |
| Niuean, Yukulta, Basque | ✔ | ✘ | ✘ |

Obviously, a natural direction for future research is to expand the sample of ergative languages to see whether the generalizations hold: is syntactic ergativity best thought of as a disambiguation strategy? Does syntactic ergativity manifest itself in only a minority of morphologically ergative language, contra the claims in Polinsky (2016)? And is the reported typological asymmetry in (2) valid?

Pushing the disambiguation hypothesis even further, in Chapter 5 it is argued that the relationship between morphological ergativity and ergative extraction asymmetries is not as close as previously thought, and perhaps it ought to be severed altogether. If this is indeed the case, we should expect to observe equivalent accusative asymmetries in morphologically accusative languages, and even in morphologically ergative languages too. Some evidence of exactly these predicted configurations is presented in Chapter 5, but again it behooves future research efforts to 'keep digging' to see whether additional examples of these previously-unattested configurations can be unearthed.

# REFERENCES

Aikhenvald, A. 2012. *The Languages of the Amazon.* Oxford: Oxford University Press.

Aissen, J. 1996. "Pied-Piping, Abstract Agreement, and Functional Projections in Tzotzil". *Natural Language and Linguistic Theory.* 14: 447-491.

Aissen, J. 1979. "Agent Focus and Inverse in Tzotzil". *Language.* 75:3. 451-484.

Aldridge, E. 2012. "Antipassive and ergativity in Tagalog". *Lingua.* 122:3. 192-203.

Aldridge, E. 2008. "Phase-based account of extraction in Indonesian". *Lingua.* 118:10, 1440-1469.

Aldridge, E. 2007. "Minimalist Analysis of Ergativity". *Sophia Linguistica.* 55: 123-142.

Aldridge, E. 2004. *Ergativity and word order in Austronesian languages.* PhD dissertation. Ithaca, NY: Cornell University Press.

Alexiadou, A. 2001. *Functional structure in nominals: nominalization, and ergativity.* Amsterdam: John Benjamins.

Andersen, T. 1988. "Ergativity in Päri, a Nilotic OVS language". *Lingua* 75. 289-324.

Anderson, S. 1985. "Inflectional morphology." In T. Shopen, ed., *Language typology and syntactic description: grammatical categories and the lexicon.* 3: 150-201. Cambridge: University of Cambridge Press.

Anderson, S. 1976. "On the notion of subject in ergative languages". In C. Li, ed., *Subject and Topic.* New York: Academic Press. 1-23.

Andvik, E. 1999. *Tshangla Grammar*. PhD dissertation. University of Oregon: University of Oregon Press.

Aronson, H. 1991. "Modern Georgian". In A. C. Harris, ed., *The indigenous languages of the Caucasus*, Vol. 1. Delmar, N.Y.: Caravan. 219-312.

Ayres, G. 1991. *La gramatica Ixil.* La Antigua, Guatemala: Centro de Investigaciones Regionales de Mesoamérica.

Baerman, M., D. Brown & G. Corbett. 2005. *The syntax-morphology interface: A study of syncretism.* Cambridge: Cambridge University Press.

Baker, M. 2015. *Case: its principles and its parameters.* Cambridge: Cambridge University Press.

Baker, M. 2013. "On Agreement and its Relationship to Case: Some Generative Ideas and Results". *Lingua.* 130. 14-32.

Baker, M. 1988. *Noun incorporation.* University of Chicago Press: University of Chicago.

Baker, M. & J. Bobaljik. To appear. "On inherent and dependent theories of ergative case." In J. Coon, D. Massam & L. Travis, eds. *The Oxford Handbook of Ergativity.* Oxford: Oxford University Press.

Baker, M. & N. Vinokurova. 2010. "Two modalities of case assignment in Sakha". *Natural Language and Linguistic Theory* 28:593-642.

Bakker, D. 2011. "Language Sampling". In J. J. Song, ed., *The Oxford Handbook of Linguistic Typology.* Oxford: Oxford University Press. 100-127.

Bassène, A.-C. 2007. *Morphosyntaxe du jóola banjal.* Cologne: Rüdiger Köppe Verlag.

Bauer, W. 1993. *Māori.* London: Routledge.

Beddor, P. & R. A. Krakow. 1999. "Perception of coarticulatory nasalization by speakers of English and Thai: Evidence for partial compensation". *Journal of the Acoustical Society of America* 106. 2868-2887.

Bendor-Samuel, D. 1972. *Hierarchical structures in Guajajara.* Norman: Summer Institute of Linguistics of the University of Oklahoma.

Bhat, D. & M. S. Ningomba. 1997. *Manipuri grammar.* In *LINCOM Studies in Asian Linguistics.* Vol. 4. Munich: LINCOM EUROPA.

Bickel, B. 2003. "Belhare". In G. Thurgood & R. J. LaPolla, eds., *The Sino-Tibetan Languages.* London: Curzon Press. 546-570.

Bittner, M. & K. Hale. 1996. "Ergativity: toward a theory of a heterogeneous class". *Linguistic Inquiry.* 27:531-604.

Blake, B. 1994. *Case.* Cambridge: Cambridge University Press.

Blau, J. 1980. *Manuel de kurde.* Paris: Librarie Charles C. Klincksieck.

Blunk, W. B. 2008. *The Agent-Focus construction in Ixil Maya: a descriptive/formal analysis.* M.A. thesis. Austin: University of Texas Press.

Bobaljik, J. 2008. "Where's phi? Agreement as a post-syntactic operation." In D. Harbour, D. Adger & S. Béjar, eds., *Phi-Theory: Phi features across interfaces and modules.* Oxford: Oxford University Press, 295-328.

Bobaljik, J. 2001. "Syncretism without paradigms: Remarks on Williams 1981, 1994". In G. Booij & J. van Marle, eds., *Yearbook of morphology.* Dordrecht: Kluwer.

Bobaljik, J. 1993. "On ergativity and ergative unergatives". *MIT Working Papers in Linguistics.* 19. Cambridge, MA: MIT Working Papers in Linguistics. 45-88.

Bobaljik, J. & P. Branigan. 2006. "Eccentric agreement and multiple case checking". In A. Johns, D. Massam & J. Ndayiragije, eds., *Ergativity: emerging issues.* Dordrecht: Springer.

Bobaljik, J. & S. Wurmbrand. 2008. "Case in GB/Minimalism". In A. Malchukov & A. Spencer, eds., *The Oxford Handbook of Case.* Oxford: Oxford University Press.

Bok-Bennema, R. 1991. *Case and Agreement in Inuit*. Berlin: Foris.

Bošković, Ž. 2007. "On the Locality and Motivation of Move and Agree: An Even More Minimal Theory". *Linguistic Inquiry*. 38: 589-644.

Bricker, V. 1979. "Wh-questions, relativization and clefting in Yucatec Maya". In L. Martin, ed., *Papers in Mayan Linguistics*. Columbia, MI: Lucas Brothers. 107-136.

Broadbent, S. 1964. *The Southern Sierra Miwok language*. Berkeley: University of California Press.

Bromley, H. M. 1981. *A grammar of Lower Grand Valley Dani*. Canberra: Research School of Pacific Studies, Australian National University.

Brown, D. & A. Hippisley. 2012. *Network morphology: a defaults-based theory of word structure*. Cambridge: Cambridge Studies in Linguistics.

Brown, R. M., J. M. Maxwell, W. E. Little, & A. Bauer. 2006. *La ütz awäch? Introduction to Kaqchikel Maya language*. Austin: University of Texas Press.

Butt, M. 2006. *Theories of case*. Cambridge: Cambridge University Press.

Caha, P. 2013. "Explaining the structure of case paradigms by the mechanisms of nanosyntax". *Natural Language & Linguistic Theory*. 31: 1015-1066.

Camp, E. 1985. "Split ergativity in Cavineña". *International Journal of American Linguistics* 51. 38-58.

Campbell, L. 1985. *The Pipil language of El Salvador*. Berlin: Mouton de Gruter.

Carlin, E. B. 2004. *A grammar of Trio*. New York: P. Lang.

Chapman, S. & D. C. Derbyshire. 1986. "Paumarí". In D. C. Derbyshire & G. K. Pullum, eds., *Handbook of Amazonian Languages*. Vol. 3. Berlin: Mouton de Gruyter. 161-354.

Charney, J. 1993. "A grammar of Comanche". In R. J. DeMallie & D. R. Parks, eds., *Studies in the anthropology of North American Indians*. Lincoln, NE: University of Nebraska Press.

Chomsky, N. 2001. "Derivation by phase". In M. Kenstowicz, ed., *Ken Hale: A life in language*. Cambridge, MA: MIT Press. 1-52.

Chomsky, N. 2000. "Minimalist inquiries: the framework". In R. Martin, D. Michaels & J. Uriagereka, eds., *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, Mass: MIT Press. 89-155.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris Publications.

Cinque, G. 1995. *Italian syntax and Universal Grammar*. Cambridge: Cambridge University Press.

Clemens, L. & J. Coon. 2016. "Deriving verb-initial order in Mayan". Manuscript. http://ling. auf.net/lingbuzz/003067. Accessed August 31, 2016.

Coler, M. 2014. *A grammar of Muylaq' Aymara*. Leiden: Brill.

Comrie, B. 2013. "Alignment of Case Marking of Full Noun Phrases". In M. Dryer & M. Haspelmath, eds., *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/ chapter/98, accessed August 16, 2016.)

Comrie, B. 1978. "Ergativity". In W. Lehmann, ed., *Syntactic typology: Studies in the Phenomenology of Language*. Austrin: University of Texas Press. 329-394.

Comunidad Lingüística Q'anjob'al. 2005. *Gramática descriptiva Q'anjob'al*. Guatemala: ALMG.

Coon, J. To appear. "Mayan morphosyntax". *Language and Linguistics Compass*. Special volume on Mayan linguistics.

Coon, J. To appear. "Little-$v°$ agreement and templatic morphology in Ch'ol". To appear in *Syntax*.

Coon, J. 2017. "Two types of ergative agreement: implications for case". To appear in *MITWPL*.

Coon, J. 2013. *Aspects of Split Ergativity*. Oxford: Oxford University Press.

Coon, J. 2010. "VOS as predicate fronting in Chol Mayan". *Lingua* 120: 345-378.

Coon, J. & M. Adar. 2013. "Ergativity". In M. Aronoff, ed., *Oxford Bibliographies in Linguistics*. New York: Oxford University Press.

Coon, J., D. Massam & L. Travis. To appear. *The Oxford Handbook of Ergativity*. Oxford: Oxford University Press.

Coon, J., P. M. Pedro & O. Preminger. 2014. "The role of case in A-bar extraction asymmetries". *Linguistic Variation* 14:2. 179-242.

Corston-Oliver, S. 2003. "Core arguments and the inversion of the nominal hierarchy in Roviana". In J. Dubois, L. Kumpf & W. J. Ashby, eds., *Preferred argument structure: grammar as architecture*. Amsterdam: John Benjamins. 273-300.

Creider, C. & J. Creider. 1989. "A grammar of Nandi". In F. Rottland & M. L. Bender, eds., *Nilo-Saharan Linguistic Analyses and Documentation*. Vol. 4. Hamburg: Helmut Buske Verlag.

Curnow, T. 1997. *A grammar of Awa Pit (Cuaiquer): an indigenous language of south-western Colombia*. Phd dissertation. Canberra: Australian National University.

Davidson, M. 2002. *Studies in Southern Wakashan (Nootkan) Grammar*. PhD dissertation. Buffalo: SUNY Buffalo.

Davis, H., D. Gardiner & L. Matthewson. 1993. "A comparative look at Wh-questions in Northern Interior Salish". In *Papers for the 28th International Conference on Salish and Neighbouring Languages*. Seattle: University of Washington. 79-95.

Davis, P. & R. Saunders. 1997. *A grammar of Bella Coola*. Missoula: University of Montana Occasional Papers in Linguistics.

Deal, A. R. 2016. "Syntactic Ergativity: Analysis and Identification". *Annual Review of Linguistics.* 2:165- 185.

Deal, A. R. 2013. "Possessor raising". *Linguistic Inquiry.* 44: 391-432.

Dench, A. 1998. *Yingkarta.* Munich: LINCOM EUROPA.

Denwood, P. 1999. "Tibetan". In T. Bynon and D. Bennett, eds., *London Oriental and African Language Library.* Vol. 3. Amsterdam: John Benjamins Publishing Company.

Derbyshire, D. 1985. *Hixkaryana and linguistic typology.* Dallas: SIL Publications in Linguistics.

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

Dixon, R.M.W. 1972. *The Dyirbal language of North Queensland.* Cambridge: Cambridge University Press.

Dol, P. 2007. *A grammar of Maybrat : a language of the Bird's Head Peninsula, Papua Province, Indonesia.* Canberra: Pacific Linguistics.

Donaldson, T. 1980. *Ngiyambaa, the language of the Wangaaybuwan.* Cambridge, Cambridge University Press.

Dryer, M. 2013. "Order of Subject, Object and Verb". In M. Dryer & M. Haspelmath, eds., *The World Atlas of Language Structures Online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/ chapter/81, accessed September 9, 2016.)

Dryer, M. 2013. "Position of Interrogative Phrases in Content Questions". In M. Dryer & M. Haspelmath, eds., *The World Atlas of Language Structures Online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/ chapter/93, accessed August 16, 2016.)

Eaton, H. 2010. *A Sandawe grammar.* Dallas: SIL International.

Erlewine, M. Y. 2013. "Dissociating the syntax and morphological realization of Kaqchikel Agent Focus". In M. Kenstowicz, ed., *Studies in Kaqchikel Grammar*. 25-50.

Evans, N. 1995. *A grammar of Kayardild: with historical-comparative notes on Tangkic*. Berlin: Mouton de Gruyter.

Everett, C. 2007. *Patterns in Karitiana: articulation, perception and grammar*. PhD dissertation. Houston: Rice University Press.

Finer, D. 1998. "Sulawesi Relatives, V-Raising, and the CP-Complement Hypothesis". *Canadian Journal of Linguistics*. 43:283-306.

Finer, D. 1997. "Contrasting A'-Dependencies in Selayarese". *Natural Language and Linguistic Theory*. 15: 677-728.

Finer, D. 1994. "On the nature of two A'-positions in Selayarese". In N. Corver & H. van Riemsdijk, eds., *Studies on Scrambling: movement and non-movement approaches to free word-order phenomena*. Berlin: Mouton de Gruyter.

Foley, S. 2013. *The syntax of Georgian relative clauses*. Senior honors thesis. New York: NYU.

Foley, W. 2000. "The languages of New Guinea". *Annual Review of Anthropology*. 29: 357-404.

Forman, M. L. 1971. *Kapampangan grammar notes*. Honolulu: University of Hawaii Press.

Forrest, L. 1994. "The de-transitive clauses in Bella Coola". In T. Givón, ed., *Voice and Inversion*. Amsterdam: John Benjamins.

Fortescue, M. 1984. *West Greenlandic*. London: Croom Helm.

Franchetto, B. 2010. "The ergativity effect in Kuikuro". In S. Gildea & F. Queixalós, eds., *Ergativity in Amazonia*. 121-158.

Frank, P. S. 1985. *A grammar of Ika: a dissertation in linguistics*. PhD dissertation. University of Pennsylvania: University of Pennsylvania Press.

Frajzyngier, Z. 2001. *A grammar of Lele*. Stanford: CSLI Publications.

Friederici, A., A. Hahne, & A. Mecklinger. 1996. " The temporal structure of syntactic parsing: early and late effects elicited by syntactic anomalies". *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22:5, 1219-1248.

Galloway, B. 1993. "A grammar of Upriver Halkomelem". In *University of California publications in linguistics*. Vol. 96. Berkeley: University of California Press.

Givón, T. 1976. "Topic, Pronoun, and Grammatical Agreement". In C. Li, ed., *Subject and Topic*. New York: Academic Press. 149-188.

Glennon, J. 2014. *Syntactic ergativity in Nehan*. MA Thesis. Graduate Institute of Applied Linguistics.

Glinert, L. 1989. *The grammar of Modern Hebrew*. Cambridge: Cambridge University Press.

Göksel, A. & C. Kerslake. 2005. *Turkish: a comprehensive grammar*. London: Routledge.

Gordon, L. 1986. "Maricopa morphology and syntax". In *University of California publications in linguistics*. Vol. 108. Berkeley: University of California Press.

Guillaume, A. 2008. *A grammar of Cavineña*. New York: Mouton de Gruyter.

Greenberg, J. H. 1963. "Some universals of grammar with particular reference to the order of meaningful elements". In J. H. Greenberg, ed., *Universals of language*. London: MIT Press.

Grosh, A. & S. Grosh. 2004. *Essentials for the Kaluli language*. Ukarumpa, Papua New Guinea: Summer Institute of Linguistics.

Guirardello, R. 2010. "Ergativity in Trumai". In S. Gildea & F. Queixalós, eds., *Ergativity in Amazonia*. 121-158.

Guirardello, R. 1999. *A reference grammar of Trumai*. PhD dissertation. Rice University: Rice University Press.

Gutiérrez-Bravo, R. & J. Monforte. "Focus, agent focus and relative clauses in Yucatec Maya". In H. Avelino, J. Coon & E. Norcliffe, eds., *New Perspectives in Mayan Linguistics*. Cambridge: MIT Working Papers in Linguistics. 83-96.

Haegeman, L. 1991. *Introduction to Government and Binding Theory*. Oxford: Blackwell.

Haiman, J. 1980. *Hua, a Papuan language of the eastern highlands of New Guinea*. Amsterdam: J. Benjamins.

Hale, K. 1983. "Warlpiri and the grammar of non-configurational languages". *Natural Language and Linguistic Theory*. 1, 5-47.

Harley, H. 2008. "When is a syncretism more than a syncretism? Impoverishment, metasyncretism, and underspecification". In D. Adger, S. Bejar & D. Harbour, eds., *Phi Theory: Phi-features across modules and interfaces*. Oxford: Oxford University Press.

Harms, P. L. 1994. *Epena Pedee syntax*. Arlington & Dallas: Summer Institute of Linguistics and the University of Texas at Arlington Publications in Linguistics.

Harris, A. 1981. *Georgian syntax: a study in relational grammar*. Cambridge: Cambridge University Press.

Harrison, C. H. 1986. "Verb prominence, verb initialness, ergativity and typological disharmony in Guajajara". In D. C. Derbyshire & G. K. Pullum, eds., *Handbook of Amazonian Languages*. Vol. 1. Berlin: Mouton de Gruyter. 407-439.

Haspelmath, M. 1993. *A grammar of Lezgian*. New York: Mouton de Gruyter.

Heaton, R., K. Deen & W. O'Grady. 2016. "The status of syntactic ergativity in Kaqchikel". *Lingua*. 170: 35-46.

Hendrick, R. 2007. "Syntactic labels and their derivations". In S. Karimi, V. Samiian & W. K. Wilkins, eds., *Phrasal and Clausal Architecture: Syntactic derivation and interpretation*. Amsterdam: John Benjamins.

Hewitt, B. 1989. *Abkhaz*. London: Routledge.

Hiraiwa, K. 2000. "On Nominative-Genitive Conversion". In O. Matushansky & E. Guerzoni, eds., *A Few from Building E-39: MIT Working Papers in Linguistics* 39. 66-124.

Hofling, C. 2000. *Itzaj Maya Grammar*. Salt Lake City: University of Utah Press.

Holisky, D. A. 1987. "The case of the intransitive subject in Tsova-Tush. *Lingua*. 71, 103-132.

Holmberg, A. & T. Hróarsdóttir. 2003. "Agreement and movement in Icelandic raising constructions". *Lingua*. 113: 997-1019.

Holmer, A. 1996. *A parametric grammar of Seediq*. Lund: Lund University Press.

Holst, J. H. 2014. *Advances in Burushaski linguistics*. Tübingen : Narr Verlag.

Imanishi, Y. 2014. *Default ergative*. PhD dissertation. MIT: MIT Press.

Jansen, J. 2010. *A grammar of Yakima Ichishkíin/Sahaptin*. PhD dissertation. University of Oregon: University of Oregon Press.

Johns, A. 2000. "Ergativity: a perspective on recent work". In L. Cheng & R. Sybesma, eds., *The First GLOT International State-of-the-Article Book: The Latest in Linguistics*. The Hague: Mouton de Gruyter.

Johns, A. 1992. "Deriving ergativity". *Linguistic Inquiry*. 23:1, 57-87.

Jónsson, J. G. 1996. *Clausal architecture and case in Icelandic*. PhD dissertation. Amherst: University of Massachussetts, Amherst.

Keen, S. 1972. *A description of the Yukulta language - an Australian Aboriginal language of north-west Queensland*. MA Thesis. Monash University.

Kenesei, I., R. M. Vago & A. Fenyvesi. 1998. *Hungarian.* London: Routledge Descriptive Grammars.

Koopman, H. & D. Sportiche. 1991. "The position of subjects". *Lingua* 85: 211-258.

Kramer, R. 2014. "Clitic doubling or object agreement: the view from Amharic". *Natural Language and Linguistic Theory.* 32. 593-634.

Kuno, S. 1973. *The structure of the Japanese language.* Cambridge: MIT Press.

Laka, I. 1996. *A brief grammar of Euskara, the Basque language.* Ms., Euskal Herriko Unibertsitatea, Vitoria-Gasteiz. http://www.ehu.eus/documents/2430735/0/A-brief-grammar-of-euskara.pdf, accessed February 2, 2017.

Lander, Y. 2008. "Varieties of genitive". In A. Malchukov & A. Spencer, eds., *The Oxford Handbook of Case.* Oxford: Oxford University Press.

Lefebvre, C. & A.-M. Brousseau. 2002. *A grammar of Fongbe.* Berlin: Mouton de Gruyter.

Legate, J. A. To appear. "The locus of ergative case." In J. Coon, D. Massam & L. Travis, eds. *The Oxford Handbook of Ergativity.* Oxford: Oxford University Press.

Legate, J. A. 2008. "Morphological and abstract case". *Linguistic Inquiry.* 39:1, 55-101.

Legate, J. A. 2002. *Warlpiri: Theoretical Implications.* PhD dissertation. Cambridge, MA: MIT Press.

Levin, B. 1983. *On the nature of ergativity.* PhD dissertation. Cambridge, MA: MIT Press.

Li, C. & R. Lang. 1979. "The syntactic irrelevance of an ergative case in Enga and other

Loogman, A. 1965. "Swahili grammar and syntax". In G. Grosschmid, ed., *Duquesne Studies African Series.* Vol. 1. Pittsburgh: Duquesne University Press.

MacKay, C. 1999. *A grammar of Misantla Totonac.* Salt Lake City: University of Utah Press.

Mahajan, A. 1997. "Universal Grammar and the Typology of Ergative Languages". In eds. A. Alexiadou & T. A. Hall, *Studies on Universal Grammar and Typological Variation.* Amsterdam: John Benjamins.

Mahajan, A. 1990. *The A/A-bar distinction and movement theory.* PhD dissertation. Cambridge: MIT Press.

Marantz, A. 1991. "Case and Licensing". In G. Westphal, B. Ao & H.-R. Chae, eds., *Proceedings of ESCOL '91.* Ithaca, NY: Cornell Linguistics Club. 234-253.

Marantz, A. 1984. *On the nature of grammatical relations.* Cambridge, MA: MIT Press.

Marantz, A. 1981. *On the nature of grammatical relations.* PhD dissertation. Cambridge, MA: MIT Press.

Markman, V. 2010. "Case theory: a historical overview". *Language and Linguistics Compass.* 4: 846-862.

Maslova, E. 2003. *A grammar of Kolyma Yukaghir.* Berlin: Mouton de Gruyter.

Massam, D. 2000. "VSO and VOS: Aspects of Niuean word order". In A. Carnie & E. Guilfoyle, eds., *The syntax of verb initial languages.* Oxford: Oxford University Press.

Massam, D., C. Gorrie & A. Kellner. 2006. "Niuean determiners: everywhere and nowhere". In C. Gurskie & M. Radisic, eds., *The proceedings of the 2006 Canadian Linguistics Association Annual Conference.*

McCarus, E. N. 2009. "Kurdish". In G. Windfuhr, ed., *The Iranian languages.* Routledge: London.

McGinnis, M. 1998. *Locality in A-movement.* Cambridge: MIT Press.

McGregor, W. 2010. "Optional ergative case marking systems in a typological-semiotic perspective". *Lingua*. 120: 1610-1636.

McGregor, W. 2006. "Focal and optional ergative marking in Warrwa". *Lingua*. 116: 393-423.

Meira, S. 1999. *A grammar of Tiriyó*. PhD dissertation. Rice University: Rice University Press.

Merlan, F. & A. Rumsey. 1991. *Ku Waru : language and segmentary politics in the western Nebilyer Valley, Papua New Guinea*. Cambridge: Cambridge University Press.

Mirikitani, L. 1972. "Kapampangan syntax". *Oceanic Linguistics Special Publications* 10. Hawai'i: University of Hawai'i Press.

Mohanan, T. 1994. *Argument structure in Hindi*. Stanford: CSLI Publications.

Montaut, A. 2004. *A grammar of Hindi*. Munich: LINCOM EUROPA.

Moravcsik, E. 1978. "Agreement". In J. Greenberg, ed., *Universals of Human Language IV: Syntax*. Stanford University Press: Stanford.

Moravcsik, E. 1974. "Object-verb agreement". *Working papers on language universals* 15. Stanford: Committee on Linguistics.

Moyse-Faurie, C. 1983. *Le drehu: langue de Lifou*. Paris: Société d'Études Linguistiques et Anthropologiques de France.

Mulder, J. 1994. "Ergativity in Coast Tsimshian (Sm'algya̲x)". In *University of California publications in linguistics*. Vol. 124. Berkeley: University of California Press.

Munshi, S. 2015. *A Grammatical Sketch of Hunza Burushaski*. http://ltc.unt.edu/∼sadafmunshi/Burushaski/language/grammatical_sketch.pdf, accessed October 1, 2016.

Nichols, J. 1986. "Head-marking and dependent-marking grammar". *Language*. 62: 56-119.

Noonan, M. 1992. *A grammar of Lango.* Berlin: Mouton de Gruyter.

Oliveira, C. 2005. *The language of the Apinajé people of Central Brazil.* PhD dissertation. University of Oregon: University of Oregon Press.

Oniga, R. & N. Schifano. 2014. *Latin: a linguistic introduction.* Oxford: Oxford University Press.

Otsuka, Y. 2000. *Ergativity in Tongan.* PhD dissertation. Oxford: Oxford University Press.

Overall, S. 2007. *A grammar of Aguaruna.* PhD dissertation. Melbourne: Research Centre for Linguistic Typology, La Trobe University.

Palancar, E. 2008. "Varieties of ergative". In A. Malchukov & A. Spencer, eds., *The Oxford Handbook of Case.* Oxford: Oxford University Press.

Pesetsky, D. & E. Torrego. 2011. "Case". In C. Boeckx, ed., *The Oxford Handbook of Linguistic Minimalism.* Oxford: Oxford University Press.

Peterson, T. To appear. "Ergativity across Tsimshianic". In J. Coon, D. Massam, & L. Travis, eds., *The Oxford handbook of ergativity.* New York: Oxford University Press.

Polinsky M. To appear. "Syntactic ergativity". In M. Everaert & H. van Riemsdijk, eds., *The Blackwell Companion to Syntax.* New York: Wiley-Blackwell.

Polinsky M. 2016. *Deconstructing ergativity: two types of ergative languages and their features.* Oxford: Oxford University Press.

Popjes, J. & J. Popjes. 1986. "Canela-Kraho". In D. C. Derbyshire & G. K. Pullum, eds., *Handbook of Amazonian Languages.* Vol. 1. Berlin: Mouton de Gruyter. 128-199.

Pollock, J.-Y. 1989. "Verb Movement, Universal Grammar, and the Structure of IP". *Linguistic Inquiry.* 20: 365-424.

Preminger, O. 2014. *Agreement and its failures.* Cambridge: MIT Press.

Queixalós, F. 2010. "Grammatical relations in Katukina-Kanamari. In S. Gildea & F. Queixalós, eds., *Ergativity in Amazonia.* 235-283.

Rigsby, B. 1986. *Gitksan grammar.* St. Lucia, Queensland: University of Queensland Press.

Rizzi, L. 1997. "The fine structure of the left periphery". In L. Haegeman, ed., *Elements of Grammar.* Dordrecht: Kluwer.

Roberts, T. 1999. "Grammatical Relations and Ergativity in St'át'imcets (Lillooet Salish)". *International Journal of American Linguistics.* 65: 275-302.

Roberts, T. 1994. *Subject and Topic in St'át'imcets (Lillooet Salish).* MA thesis. Vancouver: University of British Columbia.

Rodrigues, A. 2000. "*Ge-Pano-Carib* X *Jê-Tupí-Karib*: sobre relaciones lingüísticas prehistóricas en Sudamérica". In L. Miranda, ed., *Actas del I Congreso de Lenguas Indígenas de Sudamérica.* Tome 1. Lima: Universidad Ricardo Palma. 95-104.

Romero-Figeroa, A. 1997. *A reference Grammar of Warao.* Munich: LINCOM Europa.

Romero-Méndez, R. 2009. *A reference grammar of Ayutla Mixe (Tukyo'm Ayuujk).* PhD dissertation. SUNY Buffalo: SUNY Press.

Saeed, J. 1999. "Somali". In T. Bynon and D. Bennett, eds., *London Oriental and African Language Library.* Vol. 10. Amsterdam: John Benjamins Publishing Company.

Sapir, J. D. 1965. *A grammar of Diola-Fogny: a language spoken in the Basse-Casamance region of Senegal.* Cambridge: Cambridge University Press.

Seiter, W. 1980. *Studies in Niuean Syntax.* New York: Garland Publishing, Inc.

Siewerska, A. 2004. *Person.* Cambridge: Cambridge University Press.

Silverstein, M. 1976. "Hierarchies of features and ergativity". In R.M.W. Dixon, ed., *Grammatical categories in Australian languages.* New Jersey, NJ: Humanities Press. 121-171.

Simeoni, A. 1978. *Päri: a Luo language of southern Sudan.* Bologna: Editrice Missionaria Italiana.

Smyth, D. 2002. *Thai: an essential grammar.* London: Routledge.

Sneddon, J. 1996. *Indonesian: a comprehensive grammar.* London: Routledge.

Stiebels, B. 2006. "Agent Focus in Mayan languages". *Natural Language and Linguistic Theory.* 24: 501-570.

Storm, H. 2003. *A handbook of Japanese grammar.* Munich: LINCOM EUROPA.

Storto, L. R. 1999. *Aspects of a Karitiana Grammar.* PhD dissertation. MIT: MIT Press.

Tauberschmidt, G. 1999. *A grammar of Sinaugoro.* Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, the Australian National University.

Tauberschmidt, G. 1992. "Transitivity and ergativity in Sinauḡoro". In *Language and Linguistics in Melanesia*, 23:179-191.

Tomlin, R. 1986. *Basic word order: functional principles.* London: Croom Helm.

Valenzuela, P. 2003. *Transitivity in Shipibo-Konibo grammar.* PhD dissertation. University of Oregon: University of Oregon Press.

Valenzuela, P. 2002. *Relativization in Shipibo-Konibo: a typologically-oriented study.* Munich: LINCOM EUROPA.

Vallejos Yopán, R. 2010. *A grammar of Kokama-Kokamilla*. PhD dissertation. University of Oregon: University of Oregon Press.

Walker, N. A. 2013. *A grammar of Southern Pomo, an indigenous language of California*. PhD dissertation. Santa Barbara: UC Santa Barbara Press.

Waterhouse, J. 1928. *A Roviana and English dictionary*. Revised and enlarged by L. M. Jones, 1949. Sydney: Epworth.

Wilbur, J. 2014. "A grammar of Pite Saami". In M. Haspelmath, ed., *Studies in diversity linguistics*. Vol. 5. Berlin: Language Science Press.

Wilson, D. 1974. "Suena Grammar". In *Workpapers in Papua New Guinea languages*. Vol. 8. Ukarumpa, Papua New Guinea: Summer Institute of Linguistics.

Wiltschko, M. 2006. "On 'ergativity' in Halkomelem Salish". In A. Johns, ed., *Ergativity*. 197-227.

Wood, J. 2012. *Icelandic Morphosyntax and Argument Structure*. PhD dissertation. New York: NYU Press.

Woolford, E. 2006. "Lexical Case, inherent Case, and argument structure". *Linguistic Inquiry*. 37:1, 111-130.

Yuan, M. 2013. *A phasal account of ergativity in Inuktitut*. MA thesis. Toronto: University of Toronto Press.

**THE SAMPLE OF ERGATIVE LANGUAGES**
(raw data from the ergative sample)

### Abkhaz

*family*: Northwest Caucasian

*source*: Hewitt (1989)

*pattern*: ERG=POSS (full overlap)

(1) a. *a-gəgšʷə̀g sə-r-bè-yt'*        (Hewitt 1989; p.150, p.140, p.129)
      DET-beast 1SG.ABS-3PL.ERG-see-FIN

      'the beasts saw me'

    b. *a-jə́yas nərcʷ yə-n-xò-yt'*
      DET-river on.that.side.of 3PL.ABS-PREV-live-FIN

      'they live on that side of the river'

    c. *r-an a-yʷnə̀ də-yʷnà-n*
      3PL.POSS-mother DET-house 3SG.ABS-house-STAT.PAST

      'their mother was in the house'

(2)

|  | head-marking (Hewitt 1989) |
|---|---|
| ergative absolutive | pp.101-103 |
| possessive | p.116 |

As discussed in Chapter 2, the Abkhaz data is compatible with both ERG=POSS and DAT=POSS syncretic patterns.

**Ayutla Mixe**

*family*: Mixe-Zoque

*source*: [Romero-Méndez (2009)](#)

*pattern*: ERG=POSS (full overlap)

(3) a. *Carlos tu'uk uk  y-ex-yp*                                 (Romero-Méndez 2009;
       Carlos one   dog 3.ERG-see-INDEP.TR                          p.300, p.210, p.94)

   'Carlos saw a dog'

   b. *tsu'uts-p    yë'ë    uk*
      bite-INDEP DEM.M dog

   'this dog bites'

   c. *y-wet*
      3.POSS-cloth

   'his cloths'

(4)

|            | head-marking [(Romero-Méndez 2009)](#) |
|------------|-----------------------------------------|
| ergative absolutive | p.295 |
| possessive | p.258 |

Note that head-agreement in Ayutla Mixe is complicated by several factors. Alignment is only ergative in 3 > 3 scenarios, and in addition a person hierarchy is

397

in effect since only one argument may be marked on the verb at any time. Finally, Ayutla Mixe distinguishes between independent and dependent clause types, which also conditions different agreement paradigms. See Romero-Méndez (2009) Section 8.2 for details.

**Basque**

*family*: language isolate

*sources*: Hualde & Ortiz de Urbina (2003), Laka (1996)

*pattern*: none

(5)  a.  *zazpi gizone-k ekarri dute piano-a*          (Laka 1996; p.62 p.23, p.34)
         seven man-ERG brought AUX piano-DET
         'seven men have brought the piano'

    b.  *ume-a      etorri  da*
        child-DET arrived AUX
        'the child has come'

    c.  *margolari ezagun hor-ren   erretrato-a*
        painter    known that-GEN portrait-DET
        'that well-known painter's portrait'

(6)

|            | dependent-marking | head-marking (Laka 1996) |
|------------|-------------------|--------------------------|
| ergative   | p.62              | pp.92-100                |
| absolutive | p.61              |                          |
| possessive | p.34              | -                        |

398

**Burushaski**

*family*: language isolate

*sources*: Munshi (2015), Willson (1996), Holst (2014), Anderson (2007)

*pattern*: ERG=GEN (partial overlap)

(7)  a. *hamal-e*　　*ḍarbeṣ*　*i-cʰarkan-uman*
　　　neighbour-**ERG** Darbes 3M.OBJ-beat-3PL.SUBJ

　　　'the neighbours beat Darbes'

　　b. *hiles gucʰar-imi*　　　　　　　　　(Munshi 2015; p.29, p.12, p.17)
　　　boy　walk-3M.SUBJ

　　　'the boy walked'

　　c. *in-e*　　　*laqpis*
　　　3SG-**GEN** handkerchief

　　　'his handkerchief'

(8)

|  | dependent-marking (Munshi 2015) | pronouns |
|---|---|---|
| ergative |  | p.59 |
| absolutive | p.46 | p.50 |
| possessive |  |  |

Note that head-agreement in Burushaski is in accusative alignment, see Holst (2014) pp.40-48. See also Willson (1996) for some idiosyncratic exceptions to this accusative alignment in verb agreement.

**Canela-Kraho**

*family*: Gê

*source*: [Popjes & Popjes (1986)](#)

*pattern*: ERG=POSS (full overlap)

(9) a. *i-te*      *a-pupun*          (Popjes & Popjes 1986; p.147, p.132, p.169)
1SG-PAST 2SG-see

'I saw you'

   b. *a-jōt*
2SG-sleep

'you slept'

   c. *i-jõ*      *wapo*
1SG-POSS knife

'my knife'

(10)

|  | head-marking<br>(Popjes & Popjes 1986) |
|---|---|
| ergative<br>absolutive | pp.128-133 |
| possessive | pp.168-169 |

Unlike alienable possession (9c), which is an example of ERG=POSS, in-alienable possession exhibits an ABS=POSS pattern. See Popjes & Popjes (1986) pp.168-169.

**Cavineña**

*family*: Tacanan

*sources*: Guillaume (2008), Camp (1985)

*pattern*: ERG=GEN (minor overlap)

(11) a. *tueke-dʸa  tu-ha-tu  medu-kʷare      e-kʷe  babi-či*
    next-EMP 3-GEN-3 claw-REM.PAST 1-GEN grandfather.DIM
    'next it clawed my grandfather'

b. *dutʸa apuna-tu kʷa-haka-ma hu-kʷare     meta babi-ra*     (Camp 1985;
    all    night-3   go-stop-NEG  do-REM.PAST night hunt-to     p.52, p.41, p.39)
    'every night he always went to hunt'

c. *tume tawi-ya-ke-dʸa-tu       Munišu-ha    sapatu tiru-kʷare*
    then  sleep-PRES-when-EMP-3 Minishu-GEN shoe    burn-REM.PAST
    'then while he was sleeping, Munishu's shoe burned'

(12)

| | dependent-marking (Guillaume 2008) |
|---|---|
| ergative absolutive possessive | p.41, p.52 |

Transitive clauses are not typically in an ERG=GEN configuration in Cavineña; only in specific 3 > 3 scenarios may the construction in (11a) be used. In addition, the use of ergative dependent-marking is conditioned by several factors. For instance, if a pronoun is acting as transitive subject upon a lexical NP, no

401

ergative morphology occurs. In addition, a person hierarchy is in effect, such that only lower-ranked transitive subject acting upon a higher-ranked transitive object triggers ergative morphology. See Camp (1985) for fuller discussion on all these points.

## Epena Pedee

*family*: Embera

*source*: Harms (1994)

*pattern*: none

(13) a. *ɨmɨkʰĩraa-rã́-pa hápa      hidiu-hi-dá*      (Harms 1994; p.10, p.10, p.34)
       man-PL-ERG     canoe.ABS drag-PAST-PL

       'men dragged the canoe'

    b. *wẽrá-rã    če-hi-dá      barrée*
       woman-PL come-PAST-PL downstream

       'the women came downstream'

    c. *juan  carlos  akʰṍre pa-hí*
       Juan Carlos father  AUX-PAST

       'Juan was Carlos's father'

(14)

|  | dependent-marking (Harms 1994) |
|---|---|
| ergative absolutive | pp.9-10 |
| possessive | p.14, p.58 |

Epena Pedee verbs agree with the transitive or intransitive subject in number, thus making head-marking alignment accusative. See Harms (1994) pp.55-57.

**Georgian**

*family*: Kartvelian

*sources*: Harris (1981), Aronson (1991)

*pattern*: none

(15) a. *glex-ma      datesa          simind-i*          (Harris 1981; p.1, p.43, p.101)
      peasant-ERG he-sowed-it-II-I corn-ABS

      'the peasant sowed corn'

   b. *namcxva-i  gamocxva*
      pastry-ABS it-baked-II-2

      'the pastry baked'

   c. *kv-is      saxl-i     avašene*
      stone-GEN house-ABS I-built-it-II-I

      'I built a stone house'

(16)

| | dependent-marking (Aronson 1991) | pronouns |
|---|---|---|
| ergative absolutive possessive | pp.228-229 | pp.239-240 |

Georgian verbs inflect for subject, object and indirect object, and the alignment is nominative-accusative. See Harris (1981) pp.29-30 for details.

**Gitksan**

*family*: Tsimshianic (Interior Tsimshian subgroup)

*sources*: Rigsby (1986), Peterson (2015)

*pattern*: ERG=POSS (full overlap)

(17) a. *ɬəmo:-yə-t*      *'nu:'m*      (Rigsby 1986; p.261, p.261, p.257, p.323)
       help-TR-**3SG.ERG** 1PL.ABS
       'he helped us'

    b. *ɬəmo:-yə-'m*      *'nit*
       help-TR-1PL.ERG 3SG.ABS
       'we helped him'

    c. *sa:*    *paẋ 'nit*
       away ran 3SG.ABS
       'he ran away'

    d. *kimxti-t*
       sister-**3SG.POSS**
       'his sister'

(18)

|  | head-marking (Rigsby 1986) |
|---|---|
| ergative absolutive possessive | pp.412-414 |

Note also that in dependent clauses, alignment is still ergative, however here the 'Series II' pronoun encodes transitive objects and intransitive subjects, and a new 'Series I' pronominal paradigm encodes transitive subjects. See Peterson (2015) for more discussion of alignment in Tsimshianic.

**Halkomelem**

*family*: Salishan

*sources*: Galloway (1993), Wiltschko (2006)

*pattern*: ERG=POSS (minor overlap)

(19) a. *q'ó:y-t-es*      *te*    *Strang* *te*    *sqelá:w*               (Wiltschko 2006; p.197, p.197;
        kill-TR-3ERG DET Strang DET beaver                        Galloway 1993; p.183)

        'Strang killed the beaver'

    b. *í:mex*    *te*    *Strang*
        walking DET Strang

        'Strang is walking'

    c. *s-tálə-s*

        NOM-wife-3.POSS

        'his wife'

(20)

|  | head-marking (Galloway 1993) |
| --- | --- |
| ergative | p.176 |
| absolutive | p.177 |
| possessive | p.179 |

      Note that Halkomelem verb agreement is only in ergative alignment for third person arguments; see discussion in Wiltschko (2006).

**Hindi**

*family*: Indo-European (Indo-Aryan subgroup)

*sources*: Montaut (2004), Mohanan (1994)

*pattern*: none

(21) a. *ravii-ne    roṭii    $k^h$aayii*          (Mohanan 1994; p.103, p.71, p.60)
Ram$_M$-ERG bread$_F$ eat.PERF.F.SG

'Ram had eaten bread'

b. *Raam        giraa*
Ram$_M$.ABS fall.PERF.M.SG

'Ram fell hard'

c. *raam-ne   bacce-kaa  naam  pukaaraa*
Ram-ERG child-GEN name  call-PERF

'Ram called the child's name'

(22)

|             | dependent-marking (Montaut 2004) | head-marking |
|-------------|:------------------:|:------------:|
| ergative absolutive | p.179 | p.140 |
| possessive | p.64 | - |

**Hua**

*family*: Trans-New Guinea (Kainantu-Goroka subgroup)

*source*: Haiman (1980)

*pattern*: none

406

(23) a. *Buro' fu-mu    ke*          (Haiman 1980; p.361, p.119, p.238)
     Buro   pig-ERG see.3SG.SUBJ
     'the pig saw Buro'

   b. *ma    de    frie*
     DEM man die.3SG.SUBJ
     'this man died'

   c. *de-ma'    fu*
     man-GEN pig
     'the man's pig'

(24)

|  | dependent-marking (Haiman 1980) |
| --- | --- |
| ergative | pp.228-232 |
| absolutive | p.228 |
| possessive | pp.238-240 |

Note that use of the ergative suffix is optional; see discussion in Haiman (1980) p.361. Also, note that verbs agree with both subjects and objects in nominative-accusative fashion. Transitive and intransitive subjects are marked by changes in the verb stem's vowel as well as verb suffixes; see discussion in Haiman pp.47-48 and summary table p.79. Object agreement takes the form of a prefix, see pp.371-373.

**Hurrian**

*family*: Hurro-Urartian

*sources*: Bush (1964), Wegner (1999)

*pattern*: none

407

(25) a. *ta̱he-š    erbi id-i-a*
man-ERG dog hit-TR-3SG.ERG                                / Bush 1964; p.129)

'the man strikes the dog'

b. *tahe un-a*
man come-INTR

'the man comes'

c. *maθriyan-ne-fe    omin-ne-fe    ewren-ne*
Egyptian-SG-GEN land-SG-GEN lord-SG

'king of the Egyptian land'

(26)

|  | dependent-marking (Wegner 1999) | head-marking (Bush 1964) |
|---|---|---|
| ergative | | p.206 |
| absolutive | p.43 | pp.253-255 |
| possessive | | p.116 |

**Ika**

*family*: Chibchan

*source*: Frank (1985)

*pattern*: none

(27) a. *in'gui tšeirua-se'-ri   wakuma-ri guako-u-na*              (Frank 1985;
one    man-ERG-TOP skunk-TOP kill-AUX-DIST               p.49, p.65, p.56)

'a man killed a skunk'

408

b. *mouga-ri awa'rei zoža-na*
two-TOP  below   go-DIST

'two men went below'


c. *nivi zei   tutusoma*
1PL  GEN  hat

'our hats'


(28)

| | dependent-marking (Frank 1985) |
|---|---|
| ergative absolutive | pp.49-50 |
| possessive | pp.55-57 |


Ika features an allomorph of the genitive marker *zei*, the prefix *zä-*, which attaches to possessed nouns; see discussion in Frank (1985) p.57. In addition, verbs agree with both subjects and objects, producing a head-marking system in accusative alignment. See pp.67-70 for discussion of verbal head-marking in Ika.


**Itzaj**

*family*: Mayan (family subgroup)

*source*: Hofling (2000)

*pattern*: ERG=POSS (full overlap)


(29) a. *k-    uy-     il  -aj -o'on*          (Hofling 2000; p.37, p.37, p.37, p.25)
COM 3SG.ERG see TR 1PL.ABS

'he saw us'


409

b. *tal   -o'on*
   come 1PL.ABS

   'we came'

c. *tal   -ij*
   come 3SG.ABS

   'he came'

d. *uy-       äj- kax*
   3SG.POSS- M  chicken

   'his chicken'

(30)

|  | head-marking (Hofling 2000) |
|---|---|
| ergative | p.36 |
| absolutive | p.37 |
| possessive | p.36 |

Note that Itzaj verb agreement is only in ergative alignment in completive aspect. In incompletive aspect, verb agreement is in accusative alignment. See Hofling (2000) pp.36-38.

**Kaluli**

*family*: Trans-New Guinea (Bosavi subgroup)

*source*: Grosh & Grosh (2004)

*pattern*: ERG=GEN (minor overlap)

(31) a. *Iba-ya: siabulu-wo: Hana o:lia: Sala o:lia: ma:no: sofa:*
    Iba-**ERG** potato-TOP Hanah ACCOMP Sarah ACCOMP food   cook.PAST

    'Eva cooked sweet potatoes for Hannah and Sarah'

(Grosh & Grosh 2004; p.63, p.60, p.13)

410

b. *kabi fa:   gulufo:    ko:lo: a:-difa:*
   axe  handle break.PAST CONJ REP-put.PAST

   'the axe handle broke so I put it on again'

c. *tili-dabu     kalu-keisale gio:, Gode-ya: nanog di-a:bi*
   pull-hear.PAST man-woman 2PL God-GEN work  do-IMP.FUT

   'you believers must do God's work!'

(32)

| | dependent-marking | pronouns |
| | (Grosh & Grosh 2004) | |
|---|---|---|
| ergative | p.65 | p.37 |
| absolutive | | |
| possessive | p.13 | |

## Kapampangan

*family*: Austronesian (Philippines subgroup)

*sources*: Mirikitani (1972), Forman (1971)

*pattern*: ERG=POSS (full overlap)

(33) a. *isulat    (ne) ning lalaki ing  istorya*     (Mirikitani 1972; p.77, p.77, 45)
        write.OV he-it ERG boy   ABS story

        'the boy will write the story'

b. *sumulat (ya) ing lalaki*

   write.AV he ABS lalaki

   'the boy will write'

411

c. *lapis  ning anak ini*
   pencil GEN child this

   'this is the child's pencil'

(34)

|  | dependent-marking | pronouns (Forman 1971) |
|---|---|---|
| ergative | pp.53-54 | p.62 |
| absolutive | pp.57-59 | pp.60-61 |
| possessive | pp.53-54 | p.62 |

In addition to individual pronouns, there are also portmanteau subject + object pronouns; see Forman (1971) pp.63-65. Naturally, those forms lack alignment. Also, note that ergative alignment (in both pronouns and dependent-marking) is contingent on the VSO Object Focus construction. In the alternative VOS Agent Focus construction, both the transitive subject and the intransitive subject are marked with the same forms, thus making the alignment accusative.

**Kaqchikel**

*family*: Mayan

*source*: author's fieldwork (but see also Brown et al 2006)

*pattern*: ERG=POSS (partial overlap)

(35) a. *x-    i-    ru-    tijo  -j*
       PERF 1SG.ABS 3SG.ERG teach TR
       'she taught me'

b. *x- i- wär*         c. *x- Ø- wär*
   PERF 1SG.ABS sleep          PERF 3SG.ABS sleep
   'I slept'                'she slept'

d. *ri ru- wakx*
   DET 3SG.POSS cow
   'her cow'

(36)

|  | head-marking<br>(Brown et al 2006) |
| --- | --- |
| ergative | p.49 |
| absolutive | p.29 |
| possessive | p.31 |

**Ku Waru**

*family*: Trans-New Guinea (Chimbu-Wahgi subgroup)

*source*: Merlan & Rumsey (1991)

*pattern*: none

(37) a. *na-ni kera laim-yl tud*
      1SG-ERG bird cassowary-DEF hit.PERF.1SG.SUBJ

      'I killed the cassowary'

    b. *na pukur*                   (Merlan & Rumsey 1991;
      1SG go.PROG.1SG.SUBJ            p.340, p.327, p.338)

      'I'm going'

    c. *Don-nga lku si nyim*
      Don-GEN house crowdedness be.PERF.3SG.SUBJ

      'Don's house was crowded'

(38)

| | head-marking<br>(Merlan & Rumsey 1991) |
|---|---|
| ergative<br>absolutive | pp.339-340 |
| possessive | p.335 |

Ku Waru verbs agree with transitive and intransitive subjects, producing nominative-accusative alignment. See Merlan & Rumsey (1991) p.326. Note also that while the transitive clause in (37a) exhibits ergative dependent-marking on the transitive subject, there are some exceptions. See pp.337-339 for a discussion of different clause types.

**Lezgian**

*family*: Northeast Caucasian

*source*: Haspelmath (1993)

*pattern*: ERG=GEN (minor overlap)

(39) a. *gada-di   utanmišwil-äj wiči-n    wil-er  čünüx-iz alaqʰ-na*
     boy-ERG shame-INEL   self-GEN eye-PL hide-INF strive-AOR
     'the boy tried to hide his eyes out of shame'

   b. *kac      stol-di-n   k'anikaf    xkec'-na*            (Haspelmath 1993;
      cat.ABS table-GEN from.under go.out-AOR            p.104, p.170, p.84)
      'the cat came out from under the table'

   c. *koridor-di-n cl-a*
      hall-GEN      wall-INESS
      'on the wall of the hall'

414

(40)

| | dependent-marking (Haspelmath 1993) |
|---|---|
| ergative | pp.74-75 |
| absolutive | p.83 |
| possessive | p.84 |

As discussed in Chapter 2, ERG=GEN in Lezgian constitutes a borderline case, given that in principle the two forms are distinguished by the extra morpheme *-n*, but according to Haspelmath (1993), this segment is often deleted, thus resulting in homophonous ergative and genitive dependent-marking morphology.

**Lhasa Tibetan**

*family*: Sino-Tibetan (Tibetic subgroup)

*sources*: Denwood (1999), Delancey (1984)

*pattern*: ERG=GEN (full overlap)

(41) a. ⌟ŋɛ:     ⌟jɪgɪ   ⌝taŋdʑɔ      (Denwood 1999; p.197, p.196, p.123, p.101)
   1SG.ERG letter send-VOL
   'I can easily send a letter'

    b. ⌝kʰɔ̃:     ⌟ŋa ⌝tɑ̃: ⌝naŋbəre:
   3SG.HON 1SG send do-LINK-AUX
   'he sent me'

    c. ⌟ŋa ⌟naŋlə     ⌟jø:
   1SG home-LOC exist
   'I was at home'

d. ⌐ŋɛ:      ⌐lagbə
   1SG.GEN hand
   'my hand'

(42)

|            | dependent-marking (Denwood 1999) |
|------------|----------------------------------|
| ergative   | pp.193-197                       |
| absolutive | p.192                            |
| possessive | p.101                            |

As noted in Chapter 2, the ERG=GEN pattern in Lhasa Tibetan is not always consistent, and as such it was characterized as a 'borderline' case.

**Ngiyambaa**

*family*: Pama-Nyungan (Wiradhuric subgroup)

*source*: Donaldson (1980)

*pattern*: ERG=GEN (minor overlap)

(43) a. *ga:-nhi      =naŋ-gal mayiŋ-gu    gana:-ga*        (Donaldson 1980; p.90,
       carry-PAST =ABS-PL person-ERG shoulder-LOC          p.219, p.117, p.230)
       'the men carried them on (their) shoulders'

   b. *mu:n-miyi =lu        dhi:rba-nhi mayi*
      all-TR        3SG.ERG know-PAST person
      'he knew everybody'

   c. *dhiŋga:-ŋa:n            mayi   ga-ɽa*
      meat-skilled.at.catching person be-PRES
      '(she) is a person who is always catching meat'

416

d. *mayiŋ-gu  ŋiya*
   person-DAT law
   'blacks' law'

(44)

|  | dependent-marking [(Donaldson 1980)](#) | pronouns |
|---|---|---|
| ergative absolutive possessive | p.82 | pp.122-129 |

Like many other Pama-Nyungan languages, the ERG=GEN pattern in Ngiyam-baa can also be described as DAT=GEN.

**Niuean**

*family*: Austronesian (Polynesian subgroup)

*source*: [Seiter (1980)](#)

*pattern*: ERG=GEN (minor overlap)

(45) a. *ne  kai  he  pusi ia  e  moa*          (Seiter 1980; p.29, p.28, p.34)
      PST eat ERG cat  that ABS chicken
      'that cat ate the chicken'

   b. *malona tuai  e  kapiniu ē*
      broken  PERF ABS dish     this
      'this dish is broken'

   c. *e  kapiniu he  kulī*
      ABS dish    GEN dog
      'the dog's dish'

417

(46)

| | head-marking (Seiter 1980) |
|---|---|
| ergative absolutive | pp.28-29 |
| possessive | pp.34-35 |

Note that two different sets of case markers encode common nouns and proper nouns. Furthermore, the ergative case marker for proper nouns *e* is homophonous with the absolutive case marker for common nouns.

**Päri**

*family*: Nilotic

*sources*: Andersen (1988), Simeoni (1978)

*pattern*: ERG=POSS (full overlap)

(47) a. *ùbúr dháagɔ̀ á-yáaɲ'-è̀*           (Andersen 1980; p.293, p.292
    Ubur woman CMPL-insult-3SG.ERG           / Simeoni 1978; p.76)

    'Ubur insulted the woman'

  b. *dháágɔ̀ á-ŋɛ̀ɛ-ɔ̀*
    woman CMPL-laugh-SUF

    'the woman laughed'

  c. *tyend-e*
    foot-3SG.POSS

    'his feet'

(48)

|  | head-marking | |
|---|---|---|
|  | (Andersen 1988) | (Simeoni 1978) |
| ergative absolutive | p.297 | - |
| possessive | - | p.35 |

Note that suffixed vowels in Päri are subject to Advanced Tongue Root (ATR) harmony: they must always agree with the [± ATR] value of the stem vowel. Note also that while, in OVS constructions, Päri features an ergative dependent marker *-ì*, it is not clear from either Andersen (1988) or Simeoni (1978) whether Päri features an equivalent genitive dependent marker.

**Paumarí**

*family*: Arauan

*source*: Chapman & Derbyshire (1986)

*pattern*: none

(49) a. *Dono-a    bi-ko'diraha-'a-ha           ada    isai  hoariha*
Dono-ERG 3SG.ERG-pinch-ASP-THEME.M DEM.M child other

'Dono pinched the other boy'

(Chapman & Derbyshire 1986;
p.164, p.195, p.254)

b. *ihamahi-'a-ha            ada    Dono*
be.angry-ASP-THEME.M DEM.M Dono

'Dono was angry'

c. *voroni-'i-hi       ida    kidi-hado*
fall-ASP-THEME.F DEM.F 3SG.POSS-knife

'his knife fell'

(50)

|  | dependent-marking | head-marking |
|---|---|---|
|  | (Chapman & Derbyshire 1986) | |
| ergative absolutive | p.164 | p.287 |
| possessive | - | pp.254-256 |

Paumarí exhibits a consistent ergative split based on person. Both dependent-marking and head-marking are only ergative in alignment for third persons: non-third person transitive subjects are not marked with the suffix *-a*, and the pre-verbal agreement 'slot' is nominative for non-third persons as well. Note also that in addition to SVO (49a), Paumarí features an alternative OVS construction in which the pre-verbal transitive object is marked with accusative case *-ra*. See Chapman & Derbyshire (1986) for more details on all these points.

**Seediq**

*family*: Austronesian (Formosan subgroup)

*source*: Holmer (1996)

*pattern*: ERG=GEN (full overlap)

(51) a. *wada =ku =na qtaun*        (Holmer 1996; p.69, p.57, p.51)
        PRET =1SG.NOM =3SG.ERG see.PF
        'he saw me'

    b. *wada dehuk heya*
        PRET arrive 3SG
        'he arrived'

c. *qalux ka    uban =na*
   black DEF hair  =3SG.GEN

   'its hair is black'

(52)

|  | head-marking (Holmer 1996) |
|---|---|
| ergative absolutive possessive | pp.31-32 |

In Seediq, sentences can either occur in the VSO Patient Focus construction or the VOS Agent Focus construction. See discussion in Holmer (1996) pp.57-58.

**Shipibo**

*family*: Panoan

*source*: Valenzuela (2003)

*pattern*: ERG=GEN (partial overlap)

(53) a. *jiwi-n-ra    Sani rishki-ke*        (Valenzuela 2003; p.325, p.203, p.132)
        tree-ERG-EV Sani hit-CMPL

        'the tree hit Sani'

   b. *isá-ra    noya-i*
        bird-EV fly-INC

        'the/a bird is flying'

   c. *nokon    tita-n        chomo*
        1SG.GEN mother-GEN jar

        'my mother's jar'

421

(54)

| | dependent-marking (Valenzuela 2003) | pronouns |
|---|---|---|
| ergative absolutive | pp.323-326 | p.185 |
| possessive | p.225 | |

**Sinaugoro**

*family*: Austronesian (Papuan subgroup)

*sources*: Tauberschmidt (1999), Tauberschmidt (1992)

*pattern*: ERG=GEN (full overlap)

(55) a. *tau-na*    *bua*    *e*    *vini-gu-to*         (Tauberschmidt 1992;
man-ERG betelnut 3SG.SUBJ give-1SG.OBJ-PERF     p.181, p.181, p.189)

    'the man gave me a betelnut'

   b. *mero e*      *ḡani-ḡani-ni*
boy   3SG.SUBJ eat-RED-IMPF

    'the boy is eating'

   c. *au-na*    *motuka e*    *rakava-to*
1SG-GEN car    3SG.SUBJ bad-PERF

    'my car got damaged'

(56)

|  | dependent-marking (Tauberschmidt 1999) |
|---|---|
| ergative absolutive | p.71 |
| possessive | pp.59-61 |

Verbs in Sinaugoro agree with subjects and objects in nominative-accusative alignment, see Tauberschmidt (1999) pp.23-33 for more details.

## Sm'algyax

*family*: Tsimshianic (Coast Tsimshian subgroup)

*sources*: Mulder (1994), Peterson (2015)

*pattern*: ERG=POSS (partial overlap)

(57) a. *ap'ax*     *-d*     *-u*        (Mulder 1994; p.58, p.101,
     remember 3.ABS 1SG.ERG        p.62, p.36, p.34 p.76)

     'I remembered him'

   b. *ada k'a*      *-t'aa-t*
     and for.a.while sit-3.ABS

     'and he sat for a while'

   c. *hak'o* *-(y)u*
     back    1SG.POSS

     'my back'

(58) a. *ha'ligoot -sga awta*     *-ga* [ *dza ła*    *al*    *dzak -sga*    *sts'ool -ga* ]
     thought ERG porcupine ABS    C    PAST EMPH dead ABS.SUB beaver DEM

     'porcupine thought that beaver was dead'

   b. *t'aa*    *-ga sm'ooygit-ga*
     be.SG ABS chief-DEM

     'there was a chief'

   c. *naks*    *-ga na*    *-wayk*    *-t*      *-ga*
     spouse GEN POSS brother 3SG.POSS DEM

     'their brother's wife' (lit. 'spouse of their brother')

(59)

|  | dependent-marking | head-marking |
|---|:---:|:---:|
|  | (Mulder 1994) | |
| ergative absolutive | p.33 | p.50 |
| possessive | p.44 | p.63 |

Note that, at first glance, the head-marking syncretism in Sm'algyax looks like overlap between 'objective' (i.e. absolutive) and possessive agreement. However, when the transitive object is in the third person, as an exception the transitive subject is <u>also</u> encoded with 'objective' morphology. The end result is that the same morphology that encodes transitive subjects when the object is in the third person also encodes possessors. See discussion in Mulder (1994) pp.57-59.

Note also that in terms of dependent-marking, Sm'algyax displays an ABS= GEN pattern for common nouns (both 'present' and 'absent'), while for proper nouns, as well as all forms in the subjunctive mood, the pattern is actually ERG=GEN. See full discussion of predicative connectives in Mulder (1994) Section 2.2.1.

**Sorani Kurdish**

*family*: Indo-European (Iranian subgroup)

*sources*: McCarus (2009), Blau (1980), Thackston (2015)

*pattern*: ERG=POSS (full overlap)

(60) a. *bīnī-yān-in*　　　　　　　(McCarus 2009; p.617, p.609 / Blau 1980; p.63)
see-3PL.ERG-2PL.ABS

'they saw you (PL)'

b. *hāt-in*

come-2PL.ABS

'you (PL) came'

c. *qamik-yān*

thumb-3PL.POSS

'their thumb'

(61)

| | head-marking (Blau 1980) |
|---|---|
| ergative absolutive | p.55 |
| possessive | p.63 |

**Suena**

*family*: Trans-New Guinea (Binandere subgroup)

*source*: Wilson (1974)

*pattern*: none

(62) a. *mama na-so    ugama    zo    meni gamu-nu-a*
      father 1SG-GEN crocodile DET ERG   eat-3SG.SUBJ.REM.PAST-IND
      'a crocodile ate my father'

   b. *ema zo    tup-i-a*                                    (Wilson 1974;
      man DET come-3SG.SUB.PRES-IND                 p.78, p.40, p.80)
      'a man is coming'

   c. *bama    zai-ra    suna*
      woman old-GEN dog
      'the old woman's dog'

(63)

| | dependent-marking (Wilson 1974) |
|---|---|
| ergative absolutive | p.78 |
| possessive | p.16, p.80 |

426

Suena verbs agree with transitive and intransitive subjects, producing accusative alignment. See Wilson (1974) pp.59-60 for details.

**Sumerian**

*family*: language isolate

*sources*: Edzard (2003), Hayes (1997) , Michalowski (1980)

*pattern*: ERG=POSS (minor overlap)

(64) a. *lugal-e     e           mu-n-du-∅*
　　　 king-ERG temple.ABS PERF-3SG.ERG-build-3SG.ABS
　　　 'the king built the temple'

　　 b. *lugal      i-tuš-∅*　　　　　　　　(Michalowski 1980; p.91,
　　　 king.ABS PERF-sit-3SG.ABS　　　　　 p.92; Edzard 2003; p.29)
　　　 'the king sat down'

　　 c. *lugal-a-ni*
　　　 owner-V-3SG.POSS
　　　 'his owner'

(65)

|  | dependent-marking | head-marking |
|---|---|---|
| ergative absolutive | p.88 (Michalowski 1980) | pp.89-94 (Michalowski 1980) |
| possessive | p.30 (Edzard 2003) | p.30 (Edzard 2003) |

The ERG=POSS pattern in Sumerian only occurs in the third person, noting however that Sumerian distinguishes between personal and non-personal forms.

In (64a-64c), the form *-n(i)* occurs with a human transitive subject or possessor; alternatively, the form *-b(i)* corresponds to non-human transitive subjects or possessors. Note also that while Sumerian does exhibit dependent-marking, no ERG=GEN pattern occurs.

**Tenetehára**

*family*: Tupian

*sources*: Bendor-Samuel (1972), Harrison (1986)

*pattern*: ABS=POSS (full overlap)

(66) a. *he-rurywete*
     1SG-be.happy
     'I am happy'

   b. *he-petek*
     1SG-beat
     'he beats me'

   c. *he-mukaw*
     1SG-gun
     'my gun'

(Harrison 1986; p.421/
Bendor-Samuel 1972; p.91, p.106)

(67)

|  | head-marking (Harrison 1986) |
|---|---|
| ergative absolutive possessive | pp.419-424 |

428

Note that the verbal head-marking system in Tenetehára is constrained by a person hierarchy such that only one argument is encoded via agreement on the verb. In addition, this system is perhaps best characterized as having active-stative alignment. Unaccusative transitive subjects and transitive objects are encoded by the same morphological paradigm (66a-66b), but unergative intransitive subjects and transitive subjects are encoded by another paradigm. See Harrison (1986) Section 8 for discussion.

**Tiriyó**

*family*: Carib

*source*: Meira (1999)

*pattern*: ABS=POSS (full overlap)

(68) a. *ë-emamina*                                           (Meira 1999; p.283, p.290, p.201)
      2SG.ABS-play

      'you have played'

  b. *ë-eta*
      2SG.ABS-hear

      'it has heard you'

  c. *ë-pata*
      2SG.POSS-village

      'your village'

(69)

| | dependent-marking | head-marking |
|---|---|---|
| | (Meira 1999) | |
| ergative absolutive | pp.334-335, p.512 | pp.283-284 |
| possessive | - | pp.515-519 |

Like in Tenetehára, Tiriyó's system of head-marking is constrained by a person hierarchy such that, whenever possible, Speech Act Participants (SAP) will always be encoded in the verb's sole prefix 'slot'. Also, at first glance the system appears to be active-inactive, but Meira (1999) ultimately rejects such a split-S analysis because the two intransitive verbal classes do not consistently pattern as unergative and unaccusative. In the end, Tiriyó verb agreement is best analyzed as ergative in alignment, but with a substantial lexical sub-class which encodes the intransitive subject with an idiosyncratic paradigm bearing some resemblance to clearly ergative morphology used for transitive verb agreement.

**Trumai**

*family*: language isolate

*source*: Guirardello (1999)

*pattern*: none

(70) a. *hai-ts    kodechich disi ka in*
     1SG-ERG snake    kill FOC/TENSE
    'I killed the snake'

  b. *axos* [ *huma-t'*    *ke* ] *yi wapta*
     child  take.bath-NMLZR REL  YI fall

(Guirardello 1999; p.91, p.416, p.77)

'the child who took a bath fell'

c. *hai-kte*     *ole*     *wa*
   1SG-GEN manioc plantation

   'my manioc plantation'

(71)

|  | head-marking (Guirardello 1999) |
|---|---|
| ergative absolutive | pp.257-258 |
| possessive | pp.76-78 |

## Warlpiri

*family*: Pama-Nyungan (Ngarrkic subgroup)

*sources*: Legate (2002), Nash (1985), Simpson (1991)

*pattern*: none

(72) a. *ngarrka-ngku ka      wawirri     panti-rni*          (Legate 2002; p.17,
         man-ERG      PRES.IMPF kangaroo.ABS spear-N.PAST            p.31, p.47)

   'the man is spearing the kangaroo'

   b. *ngaju-rna        parnka-ja*
      1SG.ABS-1SG.NOM run-PAST

   'I ran'

   c. *maliki jakamarra-kurlangu paka-rnu*
      dog    Jakamarra-GEN      hit-PAST

   'he hit Jakamarra's dog'

(73)

432

|  | dependent-marking (Simpson 1991) |
|---|---|
| ergative | p.11 |
| absolutive | p.10 |
| possessive | p.58 |

Like in many Pama-Nyungan languages, bound pronouns are in nominative-accusative alignment and must cliticize to the first word in the clause; see Legate (2002) pp.126-127.

**West Greenlandic**

*family*: Eskimo-Aleut

*sources*: Bok-Bennema (1991), Sadock (2003), Fortescue (1984)

*patterns*: ERG=GEN (full overlap), ERG=POSS (partial overlap)

(74) a. *piniartu-p   nanuq      tukut-taa*
       hunter-ERG polar.bear kill-IND.TR.3SG>3SG
       'the hunter killed the polar bear'

(Bok-Bennema 1991;
p.72, p.72, p.72)

  b. *Piita tikip-puq*
     Piita arrive-IND.3SG.SUBJ
     'Piita has arrived'

  c. *piniartu-p   irnir-a*
     hunter-GEN son-3SG.POSS
     'the hunter's son'

(75)

|  | dependent-marking | head-marking |
|---|---|---|
|  | (Bok-Bennema 1991) |  |
| ergative | pp.72-73 | |
| absolutive | pp.71-72 | pp.192-201 |
| possessive | pp.72-73 | |

In addition to an ERG=GEN pattern, West Greenlandic also exhibits a partial ERG=POSS pattern, in those constructions where head-marking is indeed in ergative alignment. See Bok-Bennema (1991) Chapter 5 for discussion of head-marking in West Greenlandic.

**Yakima Sahaptin**

*family*: Penutian

*source*: Jansen (2010)

*pattern*: none

(76) a. *tamánwit-nɨm =nash i-nápayun-ta*                    (Jansen 2010;
    law-ERG        1SG    3SG.SUBJ-defend-FUT       p.134, p.297, p.290)
    'the law will support me'

    b. *i-kwíita-na*              *Spilyáy*
       3SG.SUBJ-go.along-PAST  Coyote
       'Coyote was traveling along'

    c. *ɨshchɨ́t kaas-mí*
       path    train-GEN
       'railroad track' (lit. 'path of the train')

(77)

| | dependent-marking (Jansen 2010) |
|---|---|
| ergative absolutive | pp.128-129 |
| possessive | p.162 |

Yakima Sahaptin features two different ergative dependent markers: the suffix -*nɨm*, used in transitive 3 > 1/2 scenarios (76a), and the suffix -*yin*, used n transitive 3 > 3 scenarios. This implies that Yakima Sahaptin dependent marking is characterized by split ergativity, given that other transitive scenarios don't trigger ergative morphology. See Jansen (2010) pp.129-129 for more details.

Also, Yakima Sahaptin head-marking is in accusative alignment, but in transitive clauses, verb agreement morphology may be fusional. See Jansen (2010) p.81 and pp.127-128 for further discussion.

**Yingkarta**

*family*: Pama-Nyungan (Kartu subgroup)

*source*: Dench (1998)

*pattern*: none

(78) a. *thuthu-ngku jarti-lanyi mantu*  (Dench 1998; p.19, p.29, p.51)
      dog-ERG    eat-PRES  meat
      'the dog is eating the meat'

  b. *thuthu ngayu    murla-lpiya*
     dog     1SG.DAT die-INCH
     'my dog died'

c. *wanthala kuka wura-wu*
   where    meat dog-DAT

   'where is the dog's meat?'

(79)

| | dependent-marking (Dench 1998) |
|---|---|
| ergative | p.16, p.54 |
| absolutive | p.54 |
| possessive | p.16 |

Like in many Pama-Nyungan languages, genitive morphology is syncretic with dative forms, not ergative ones. Also, first and second person arguments may be encoded by bound pronouns which are in nominative-accusative alignment and which typically cliticize to the first word in a clause; see Dench (1998) pp.33-37.

**Yukulta**

*family*: Pama-Nyungan (Tangkic subgroup)

*source*: Keen (1972)

*pattern*: none

(80) a. *tir-iya    kanta    pa:tja mantuwara*    (Keen 1972; p.121, p.244, p.133)
    snake-ERG TR.PAST bite    boy

    'the snake bit the boy'

b. *wartja  yinka tula:tula:tja  mantuwara*
   quickly PAST descend     boy

   'the boy got down quickly'

c. *tjawitja kantu ṇi-wan-tji    nal-i*
   run      blood 3SG-GEN-LOC head-LOC

   'blood runs over his head'

**THE SAMPLE OF ACCUSATIVE LANGUAGES**
(raw data from the accusative sample)

**Aguaruna**

*family*: Jivaroan

*source*: Overall (2007)

*pattern*: ACC=GEN (partial overlap)

(1) a. *amɨ mi-na   duʃi-mitika-ha-mɨ-i*                    (Overall 2007; p.303,
       2SG 1SG-ACC laugh-CAUS-1SG.OBJ.IMPF-2SG.SUBJ-DECL       p.506, p.260, p.130)
       'you are making me laugh'

   b. *amɨ wi-tʃau-aita-ku-mɨ-ĩ-ka*                          *ʃiiha*
      2SG go.PERF-NEG.REL-COP-SIM-2SG.SUBJ-DS-COND well
      *anɨa-sa-nu*              *puhu-mai-inu-aita-ha-i*
      be.happy-SBD-1.SUBJ.SS live-POT-NR-COP-1SG.SUBJ-DECL
      'if you had not gone, I would be happy'

   c. *ami-na apahui tuki   puhu-wa=nu        yaĩ-pa-ka-ti*
      2SG-ACC god    always live-3SG.SUBJ=ANA.REL help-2SG.OBJ-INTS-JUSS
      'may God, who lives forever, help you'

   d. *ami-na   apa*
      2SG-GEN father
      'your father'

438

(2)

| | dependent-marking | head-marking |
| | (Overall 2007) | |
|---|---|---|
| nominative | pp.214-215 | pp.361-365 |
| accusative | | p.315 |
| possessive | pp.217-218 | p.199 |

Aguaruna's system of dependent-marking is nominative-accusative; however if the transitive subject is first person plural, second person singular, or second person plural, the transitive object is not marked with the *-na* accusative suffix. There is one exception to this generalization: if the subject is second person singular, and the object is first person singular, than that transitive object <u>is</u> marked with *-na* (1a). See Overall (2007) p.216 for details.

In addition, Aguaruna's system of head-marking is also nominative-accusative, and subjects are consistently encoded on the verb; however, only first person singular, first person plural and second person singular are encoded via accusative suffix on the verb; see Overall (2007) p.315.

## Awa Pit

*family*: Barbacoan

*source*: Curnow (1997)

*pattern*: none

(3) a. *uspa-na   Santos-ta   tɨt-shi-a-mtu-y*
　　　3PL-TOP Santos-ACC cut-DESID-PL.SUBJ-IMPF-NONLOCUT
　　　'they want to stab Santos'

　　b. *uspa-na   ɨ-mɨtaz-tu-y*　　　　　　　　(Curnow 1997; p.66,
　　　3PL-TOP go-INCEP.PL.SUBJ-IMPF-NONLOCUT　　　　p.171, p.142)
　　　'they are about to go'

c. *na-na*    *Santos-kasa Demetrio-wa*   *yal*    *i-ta-w*
1SG-TOP Santos-COM Demetrio-GEN house go-PAST-LOCUT.SUBJ

'I went with Santos to Demetrio's house'

(4)

| | dependent-marking | pronouns | head-marking |
|---|---|---|---|
| | (Curnow 1997) | | |
| nominative accusative | p.65 | p.86 | pp.192-199 |
| possessive | pp.141-142 | | - |

In Awa Pit, finite verbs are marked for person. They mark a binary Locutor/ Non-Locutor distinction; in declarative clauses, the Locutor corresponds to first person singular, and Non-Locutor corresponds to all other persons. Thus, if the 1SG Locutor is involved in any way in the sentence, it must be marked over the Non-Locutor. This is a hierarchical system, and in the past tense the alignment of this agreement morphology is nominative-accusative. For example, in a past tense declarative clause, if the 1SG argument is either the transitive or intransitive subject, it must be marked by *-ta-w* on the verb (3c); if the 1SG argument is the transitive object, it must be marked by *-ti-s*. Non-past tense head-marking is in neutral alignment: it merely tracks the presence or absence of the Locutor regardless of grammatical role. See Curnow (1997) Section 8.2 for discussion of this unique head-marking system.

In addition, note that in Awa Pit, only human and referential lexical NP objects are marked with the accusative marker *-ta*. See Curnow (1997) p.65.

**Aymara**

*family*: Aymaran

*sources*: Coler (2014), Klose (2015)

*pattern*: none

(5) a. *rumulu-x(a)  Akhawan(a)-∅  jiwa-ya-w(a)-ch(i)-(i)*
    Rumulu-ᴛᴏᴘ Akhawana-ᴀᴄᴄ die-ᴄᴀᴜs-ᴘᴀʀᴛ-ᴄɴᴊ-3sɢ.sᴜʙᴊ

    'Romulu must've killed Akhawana'

(Coler 2014; p.89, p.147, p.610)

  b. *uka  wallpa-x(a) kawki-r(u) k'awna-sk(a)-i-sti(i)*
    that hen-ᴛᴏᴘ   where-ᴀʟʟ  lay.egg-ᴘʀᴏɢ-3sɢ.sᴜʙᴊ-ɪɴᴛ

    'where is that hen laying eggs?'

  c. *khiti-n(a) uta-pa-r(u)    sara-tan-st(i)*
    who-ɢᴇɴ   house-3.ᴘᴏss-ᴀʟʟ go-1ɪɴᴄʟ>3sɢ.ꜰᴜᴛ-ɪɴᴛ

    'to whose house will we go?'

(6)

|  | dependent-marking | head-marking |
| --- | --- | --- |
|  | (Coler 2014) | |
| nominative accusative | p.168 | p.261 |
| possessive | p.610 | p.609 |

In Aymara, accusative dependent-marking is manifested as "suppression of the preceding vowel" according to Coler (2014) p.62 and p.204. In addition, note that Aymara's head-marking system is only partially in nominative-accusative alignment. Intransitive events, or transitive events featuring a third person object feature a single

441

nominative suffix on the verb; however, any transitive event involving either first or second persons is encoded via fusional morphology. See discussion in Coler (2014) Chapter 8.

**Barasano**

*family*: Tucanoan

*source*: [Jones & Jones (1991)]

*pattern*: none

(7) a. *bũ-re    sĩa-gɨ-bĩ        yai*            (Jones & Jones 1991;
       2SG-ACC kill-M.SG-3SG.M.SUBJ wildcat         p.65, p.16, p.4)
       'the wildcat will probably kill you'

  b. *bõa-beti-rũgũ-ã-bĩ*
     work-NEG-HAB-PRES-3SG.M.SUBJ
     'he never works'

  c. *Sabidõ ya-wi*
     Sabino POSS-house
     'Sabino's house'

(8)

|  | dependent-marking | head-marking |
|---|---|---|
|  | (Jones & Jones 1991) | |
| nominative | p.65 | p.73 |
| accusative | | - |
| possessive | - | p.4 |

In Barasano, definite transitive objects are typically marked by the accusative suffix *-re*, but indefinite ones are not. See Jones & Jones (1991) pp.65-66.

**Cocama**

*family*: Tupian

*source*: <u>Vallejos Yopán (2010)</u>

*pattern*: NOM=POSS (full overlap)

(9) a. *ya   chikari=ay*        (Vallejos Yopán 2010; p.513, p.513, 718)
       3SG look.for=3SG.OBJ

   'he/she looks for he/she/it'

   b. *ya    tsatsatsɨma*
      3SG scream

   'he/she screams'

   c. *ya=mirikua=muki     y=ichari        ya=kaistuma*
      3SG.POSS=wife=COM 3SG.SUBJ=leave 3SG.POSS=yucca.beer

   'with his wife he leaves his yucca beer'

(10)

|  | pronominal clitics (VallejosYopán 2010) |
| --- | --- |
| nominative accusative | p.201 |
| possessive | p.716 |

In Cocama, only 3rd person clitic pronouns display a nominative-accusative pattern: transitive subject and intransitive subject arguments are realized as proclitics to the verb, while transitive object arguments are realized as enclitics. See Vallejos Yopán (2010) p.201. Also, the pronominal system is gendered according to the speaker; the forms in (9) are uttered by a female speaker, but both female and male paradigms constitute a NOM=POSS pattern.

443

**Comanche**

*family*: Uto-Aztecan (Numic subgroup)

*sources*: Charney (1993), Robinson & Armagost (1990)

*patterns*: ACC=GEN (partial overlap), ACC=POSS (partial overlap)

(11) a. *tena-pɨ-tsa waʔi-pɨ-ʔa puni*       (Charney 1993; p.202, p.94,
      man-ABS-TOP woman-ABS-ACC see      p.56, p.144, p.81, p.127, p.87)

      'the man sees the woman'

   b. *ɨ-tsaa-tɨi-tsa miʔa-tɨ*
      your-good-friend-TOP go-GEN.ASP

      'your good friend is leaving'

   c. *Mia-ʔa-tsa satiʔi*
      Mia-GEN-TOP dog

      'Mia's dog'

(12) a. *hakaniti ɨnnɨ nápukuwáá-ʔa tsahani-ka*
      how.ACC 2SG car-ACC drive-STAT

      'what kind of car do you drive?'

   b. *hakah-nai ɨnnɨ kɨma-yu*
      where-DIR 2SG come-PROG

      'where are you coming from?'

   c. *ɨ-pɨa-tai-tuʔi nɨɨ*
      2SG.OBJ-leave-go.out-IRR

      'I'm going to divorce you'

   d. *ɨ-ámawóó-mati nɨɨ kɨh-kaʔa-tuʔi*
      2SG.POSS-apple-PART 1SG teeth-break-IRR

      'I'm going to bite off a piece of your apple'

444

(13)

| | dependent-marking | pronouns (Charney 1993) | head-marking |
|---|---|---|---|
| nominative | pp.54-55 | p.98 | - |
| accusative | pp.53-54 | - | p.99 |
| possessive | pp.56-57 | | |

The exact form of Comanche accusative dependent-marking is lexically-conditioned, but despite this variation ACC=GEN syncretism is highly consistent; see Charney (1993) Section 3.1.4. Note also that Comanche has what Charney calls an 'absolutive' suffix, which occurs on nouns in isolation (i.e. that are not in a compound, or incorporated) or that are topicalized.

In addition, transitive and intransitive subjects are encoded by free pronouns, but direct objects and possessors are encoded by clitic prefixes. See discussion in Charney (1993) Chapter 4.

**Daga**

*family*: Trans-New Guinea (Dagan subgroup)

*source*: Murane (1974)

*pattern*: none

(14) a. *apan yampo orup den  mokare yaw-an-e*      (Murane 1974; p.136,
man  three  girl  with snake  see-3PL.SUBJ-MED      p.136, pp.44-45, p.32)
'the three men and the girl saw a snake'

b. *oaenapan Otare kaisepen menan oam gugunawa namu onam-on*
people  Otare cut.INF for  Friday  first  come-3PL.SUBJ
'in order to cut Otare the people came first on Friday'

445

c. *yawa-mo-n*
   see-3PL.OBJ-3SG.SUBJ

   'he saw them'

d. *nani-mu*
   hand-3PL.POSS

   'their hands'

(15)

|  | head-marking (Murane 1974) |
|---|---|
| nominative | pp.47-54, pp.63-64 |
| accusative | p.44, pp.63-64 |
| possessive | p.32 |

**Fongbe**

*family*: Niger-Congo (Volta-Niger subgroup)

*source*: Lefebvre & Brousseau (2002)

*pattern*: none

(16) a. *ùn     mɔ̀ Kɔ̀kú*                    (Lefebvre & Brousseau 2002;
      1SG.NOM see Koku                          p.63, p.139, p.63 p.68)

      'I see Koku'

   b. *ùn     wá, é     kó     tɔ́*
      1SG.NOM come 3SG.NOM already go.out

      'I came but he had already left'

c. *Kɔ̀kú mɔ̀ mì*
   Koku see 1SG.ACC

   'Koku saw me'

d. *xwé   cὲ*
   house 1SG.GEN

   'my house'

(17)

|  | dependent-marking | head-marking (Lefebvre & Brousseau 2002) |
|---|---|---|
| nominative accusative | - | pp.62-63 |
| possessive | pp.44-46 | pp.67-68 |

Fongbe clitic pronouns are analyzed as head-marking in this survey due to their restricted distribution: they must always occur either before or after the verb, unlike their free pronoun counterparts. See discussion in Lefebvre & Brousseau (2002) pp.62-63.

**Hebrew**

*family*: Afro-Asiatic (Semitic subgroup)

*sources*: Glinert (1989), Cole (1976)

*pattern*: ACC=POSS (minor overlap)

(18) a. *sha'-alt      éfo    hayi-ti*                    (Glinert 1989; p.275,
        ask-2SG.M.SUBJ where be-1SG.SUBJ                    p.125, p.52 , p.52)

        'you asked where I was'

447

b. *kshe-tsilts-alt,*      *ani*   *bidiyuk*   *difdáf-ti*         *bo*
when-ring-2SG.M.SUBJ 1SG just     leaf.through-1SG.SUBJ P+it

'when you rang, I was just leafing through it'

c. *lehazir-Ha*
warn.INF-2SG.M.OBJ

'to warn you'

d. *iyum-Ha*
threat-2SG.M.POSS

'your threat'

(19)

| | dependent-marking | head-marking (Glinert 1989) |
|---|---|---|
| nominative | - | p.470 |
| accusative | p.157 | p.52, p.499 |
| possessive | p.24 | pp.30-31 |

According to Glinert (1989), there are several strategies to encode possession in Hebrew, and only a subset of them make use of the possessive suffix paradigm. See p.24 for an overview. Note also that only definite transitive objects are introduced by the case marker *et*, p.157. Finally, note that while accusative morphology suffixing directly to the verb (18c) is grammatical, it is considered very formal and even archaic (Sandy Abu Adas, p.c.). More commonly, the accusative suffix paradigm concatenates with the case-marker *et*, p.499.

448

**Hixkaryana**

*family*: Carib

*source*: [Derbyshire (1985)](#)

*pattern*: ACC=POSS (full overlap)

(20) a. *n-ahosɨ-ye*                           *kamara*            (Derbyshire 1985; p.32,

3.SUBJ-grab-DIST.PAST.CMPL jaguar                  p.31, p.32, p.110)

'the jaguar grabbed him'

    b. *n-eweh-yatxhe*                  *worɨskomo komo*

3.SUBJ-take.a.bath-COLL.N.PAST women     COLL

'the women are taking a bath'

    c. *toto y-ahosɨ-ye*                  *kamara*

man 3.OBJ-grab-DIST.PAST.CMPL jaguar

'the jaguar grabbed the man'

    d. *toto y-owa-nɨ*

man 3.POSS-chest-POSS

'the man's chest'

(21)

| | head-marking<br>(Derbyshire 1985) |
|---|---|
| nominative<br>accusative | p.188 |
| possessive | p.199 |

Hixkaryana's system of head-marking is in nominative-accusative alignment, but a hierarchical set of constraints operates such that typically only one argument can be encoded on the verb. Thus in (20a) the verb features nominative agreement which corresponds to the transitive subject because the transitive object is elided; however in (20c) the verb must feature accusative head-marking which corresponds to the overt, pre-verbal transitive object. See discussion in Chapter 2, as well as in Derbyshire (1985) pp.187-194.

**Hungarian**

*family*: Uralic (Finno-Ugric subgroup)

*sources*: Kenesei et al (1998), Kiss (2002)

*pattern*: NOM=POSS (partial overlap)

(22) a. *a    lány  ír-ja                    a    level-et*                    (Kenesei et al 1998; p.195,
     the girl   write-3SG.SUBJ.DEF the letter-ACC                    p.57, p.231, p.209)

    'the girl is writing the letter'

  b. *Péter  el-olvas-∅                    egy   könyvet*
    Peter  PFX-read-3SG.SUBJ.INDEF DET book

    'Peter will read a book'

  c. *a    lány  áll-∅*
    the girl   stand-3SG.SUBJ.INDEF

    'the girl is standing'

  d. *Pál   hibá-ja                tudatlanság-ból fakad-t*
    Paul mistake-3SG.POSS ignorance-ELA   stem-PAST.3SG.SUBJ.INDEF

    'Paul's mistake stemmed from his ignorance'

(23)

|            | dependent-marking | head-marking |
|            | (Kenesei et al 1998) | |
|------------|-------------------|--------------|
| nominative | p.195             | pp.289-290   |
| accusative | p.197             |              |
| possessive | -                 | pp.216-218   |

451

As discussed in Chapters 3 and 4, Hungarian features two nominative agreement paradigms. One paradigm encodes transitive subjects when the object is indefinite (22b), as well as intransitive subjects (22c). Another encodes transitive subjects when the object is definite (22a). The possessive head-marking paradigm overlaps with parts of either nominative paradigm, resulting in a partial NOM=POSS pattern. See Kenesei et al (1998) pp.289-290 for details.

**Indonesian**

*family*: Austronesian (Malay subgroup)

*source*: Sneddon (1996)

*pattern*: ACC=POSS (full overlap)

(24) a. *saya harus mem-beli    sepatu baru*
       1SG  must  *meN*-buy.AV shoe   new
       'I must buy new shoes'

(Sneddon 1996; p.20, p.134, p.170, p.171)

   b. *saya tinggal di rumah ini*
       1SG  live.OV P  house  DET
       'I live in this house'

   c. *Narti me-nunggu-ku*
       Narti *meN*-wait.for.AV-1SG.OBJ
       'Narti is waiting for me'

   d. *rumah-ku*
       house-1SG.POSS
       'my house'

(25)

| | head-marking (Sneddon 1996) |
|---|---|
| nominative accusative possessive | p.170 |

If an active voice transitive verb is prefixed with *meN-*, it can also take bound object suffix (24a). These same suffixes can also encode possession, thus resulting in an ACC=POSS pattern. A similar set of bound prefixes may encode agents of passive verbs in passives of type two. See Sneddon (1996) Section 2.71 for details.

**Japanese**

*family*: Japonic

*sources*: Kuno (1973), Storm (2003), Hiraiwa (2000)

*pattern*: none

(26) a. *Yamada-san ga    pan   o    tabema-shita*                    (Storm 2003;
       Mr. Yamada NOM bread ACC eat-PERF                          p.62, p.79, p.121)

       'Mr. Yamada ate the bread'

   b. *Ima Mita-san ga    ofisu ni   imasu*
       Ms. Mita      NOM office LOC now

       'Ms. Mita is in the office now'

   c. *watashi no    heya*
       1SG      GEN room

       'my room'

(27)

453

|  | dependent-marking |
|  | (Storm 2003) |
| --- | --- |
| nominative | p.62 |
| accusative | p.75 |
| possessive | p.74 |

**Jola Bandial**

*family*: Niger-Congo (Atlantic-Congo subgroup)

*source*: Bassène (2007)

*pattern*: ACC=POSS (full overlap)

(28) a. *ni-jug-ol*        *figen*    *Sállagi*        (Bassène 2007; p.81,
      1SG.SUBJ-see-3SG.OBJ yesterday Sallagi        p.172, p.92, p.54)

      'I saw him yesterday in Sallagi'

    b. *ínje u-m-e*       *tiyaŋ*    *no-robo-e*
      1SG COP-CL1-DEM outside 1SG.SUBJ-sit.down-TAM

      'I sat down outside'

    c. *Atejo na-sen-om*        *si-rálam*
      Atejo 3SG.SUBJ-give-1SG.OBJ 4-money

      'Atejo gave me the money'

    d. *ga-aɲen-om*
      9-POST-hand-1SG.POSS

      'my hand'

(29)

|  | head-marking (Bassène 2007) |
|---|---|
| nominative | pp.74-75 |
| accusative | p.92 |
| possessive | p.54 |

There are two nominative subject paradigms, Series I and Series II. See Bassène (2007) pp.74-75. Neither of them overlaps with possessive head-marking morphology.

**Kannada**

*family*: Dravidian

*source*: Sridhar (1990)

*pattern*: none

(30) a. *na:ge:ndra   vishala:kṣi-yannu   maduveya:-danu*
     Nagendra   Vishalakshi-ACC   marry-3SG.M.SUBJ.PAST
     'Nagendra married Vishalakshi'

    (Sridhar 1990; p.160, p.132)

  b. *ravi   malagi-da*
     Ravi sleep-3SG.M.SUBJ.PAST
     'Ravi slept'

  c. *huḍuga-na   haṇa:-vannu   kadiya-be:ḍa*
     boy-GEN   money-ACC   steal-PROH
     'don't steal the boy's money'

(31)

|  | dependent-marking | head-marking |
|  | (Sridhar 1990) |  |
|---|---|---|
| nominative |  | p.221 |
| accusative | p.156 |  |
| possessive |  | - |

According to Sridhar (1990), accusative dependent-marking *-annu* is only obligatory with human referents, otherwise it is optional. See pp.86-87 for discussion.

## Kayardild

*family*: Pama-Nyungan

*sources*: Evans (1995), Round (2009)

*pattern*: none

(32) a. *ngada     kurri-ju midijin-kuru-wuru     dangka-wuru*
     1SG.NOM see-POT medicine-PROP-MPROP person-MPROP

     'I will see the doctor' (lit 'medicine-having person')

  b. *ri-in-ki          warrku-ya ngada     wirdi-ja      ngakan-ki*
     east-from-LOC sun-LOC   1SG.NOM remain-ACT sandbank-LOC

     'all morning I remained on the sandbank'

  c. *wuu-ja    ngijin-ji    wadu*               (Evans 1995; p.146,
     give-IMP 1sg-MLOC smoke            p.140, p.169, p.151)

     'give me a smoke!'

  d. *dathin-karra maku-karra kularrin-da   kurrka-th*
     that-GEN      woman-GEN brother-NOM take-ACT

     'that woman's brother took it'

(33)

| | dependent-marking (Evans 1995) | pronouns |
|---|---|---|
| nominative | pp.136-138 | |
| accusative | pp.145-148, pp.328-330 | p.202 |
| possessive | pp.143-144, pp.150-151 | |

In addition to overt nominative case (32d), Kayardild also features 'proprietive' (32a) and 'locative' (32c) cases which seem to encode transitive objects fairly consistently. See discussion in Evans (1995) pp.145-148 and pp.328-330. Pronouns also occur with proprietive and locative case, but these case markers concatenate to the oblique pronominal stem instead of the nominative stem.

**Koasati**

*family*: Muskogean

*source*: Kimball (1991)

*pattern*: none

(34) a. *okitalkkí-k ittinsá:wa-n kawwí-:ci-hawa-:s*
ice-NOM    branch-ACC snap.PL.TR-MULT-AUDIT-PAST.
'one can hear the ice snapping the branches all around'

b. *ittó híssi-k    lá:na-t    aɬí:y*        (Kimball 1991; p.76,
tree hair-NOM be.yellow-CNN go          pp.90-91, p.433)
'the leaves are turning yellow'

c. *hociɬí im-laykí*
star    3.POSS-dung
'meteor' (lit. 'star's dung')

457

(35)

| | dependent-marking | head-marking |
|---|---|---|
| | (Kimball 1991) | |
| nominative | pp.390-391 | |
| accusative | pp.392-395 | - |
| possessive | - | pp.432-435 |

Koasati has an active-stative system of verbal head-marking, thus making it ineligible for any kind of nominative or accusative syncretism. See Kimball (1991) Chapter 3 for discussion of verb agreement.

**Kolyma Yukaghir**

*family*: Yukaghir

*source*: Maslova (2003)

*pattern*: none

(36) a. *met tet-ul    juø*                                    (Maslova 2003; p.10,
      1SG 2SG-ACC see.TR.1SG                                    p.175, p.481, p.169)

      'I saw you'

   b. *met qollume tiŋ lebie-get kewe-j-možī̆-je*

      1SG soon DEM earth-ABL go-PERF-PRSP-INTR.1SG

      'I have to leave this earth soon'

   c. *kin   met-kele edies'-u-m*
      who 1SG-ACC call-∅-TR.3SG

      'who is calling me?'

d. *met      numø-ge   el-jaqa-teje*
   1SG.GEN house-LOC NEG-arrive-FUT-INTR.1SG

   'I will not reach my house'

(37)

|  | dependent-marking | pronouns | head-marking |
|---|:---:|:---:|:---:|
|  | (Maslova 2003) | | |
| nominative | pp.89-90 | | p.140 |
| accusative | pp. 93-95 | p.234 | |
| possessive | p.77 | | - |

Note that Kolyma Yukaghir is ineligible for a NOM=GEN pattern by definition because both nominative and genitive case are null. Note also that Kolyma Yukaghir features special focus morphology (both in dependent-marking and head-marking) which is ergative in alignment; see description of 'predicative case' in Maslova (2003) pp.91-93, p.140.

**Lango**

*family*: Nilotic (Eastern subgroup)

*source*: Noonan (1992)

*pattern*: ACC=POSS (partial overlap)

(38) a. *án  à-cámò     dèk*                    (Noonan 1992; p.119,
      1SG 1SG.SUBJ-eat stew                    p.119, p.35, p.78)

      'I ate stew'

   b. *án  à-dòk*
      1SG 1SG.SUBJ-go.back

      'I went back'

c. *ò-nὲn-á*
3SG.SUBJ-see-1SG.OBJ

'he saw me'

d. *tyĕn-á*
leg1SG.POSS

'my leg'

(39)

| | head-marking (Noonan 1992) |
|---|---|
| nominative | p.91 |
| accusative | p.96 |
| possessive | p.78 |

**Latin**

*family*: Indo-European

*source*: Oniga & Schifano (2014)

*pattern*: none

(40) a. *caesar      exercitu-m reduxi-t*                    (Oniga & Schifano 2014;
Caesar.NOM army-ACC lead.back.PERF-3SG.SUBJ          p.182, p.175, p.246)

'Caesar led back the army'

b. *Terentiu-s    ambula-t*
Terence-NOM walk.PRES-3SG.SUBJ

'Terence is walking'

c. *Vergil-i    descriptio*
Virgil-GEN description.NOM

'Virgil's description'

(41)

| | dependent-marking | pronouns | head-marking (Oniga & Schifano 2014) |
|---|---|---|---|
| nominative accusative | pp.59-84 | p.93 | p.112 |
| possessive | | | - |

Latin features five noun declensions, corresponding to many different dependent-marking paradigms; see Oniga & Schifano (2014) pp.59-84. With very few exceptions, there is no systematic overlap between possessive and either nominative or accusative forms.

**Lele**

*family*: Afro-Asiatic (Chadic subgroup)

*source*: Frajzyngier (2001)

*pattern*: ACC=POSS (partial overlap)

(42) a. *gol-gi gé*                                          (Frajzyngier 2001;
      see-2SG.M.OBJ 3PL.SUBJ          p.112, p.37, p.187, p.16)

      'they saw you'

  b. *bè gé*

      give 3PL.SUBJ

      'they gave'

  c. *tamá-ŋ    jè   tè-gè            hire*
      woman-DEF IMPF chase.away-3PL.OBJ often

      'the woman often chased them away'

d. *dìngàw  dí-gè*
ferocity GEN.PL-**3PL.POSS**
'their ferocity'

(43)

| | head-marking (Frajzyngier 2001) |
|---|---|
| nominative | p.100 |
| accusative | p.109 |
| possessive | pp.62-70 |

Nominative pronominal clitics follow the verb for third person subjects, but precede the verb for all other persons. See Frayzyngier (2001) p.100. Accusative pronominal clitics usually immediately follow the verb, producing V O S word order, but in some configurations the word order may change to V S O. See pp.111-112.

## Lower Grand Valley Dani

*family*: Trans-New Guinea (West Trans-New Guinea subgroup)

*source*: Bromley (1981)

*pattern*: ACC=POSS (full overlap)

(44) a. *w-ath-in*　　　　　　　　　　　　　　　　(Bromley 1981; p.192,
1SG.OBJ-hit-2SG.SUBJ　　　　　　　　　　　　p.25, p.191, p.190)
'you hit him'

b. *his-in*
swell.up-2SG.SUBJ
'you swell up'

462

c. *h-ath-e*
   2SG.OBJ-hit-3SG.SUBJ

   'he hit you'

d. *h-oppase*
   2SG.POSS-father

   'your father'

(45)

| | head-marking<br>(Bromley 1981) |
|---|---|
| nominative | p.192 |
| accusative | p.191 |
| possessive | p.190 |

Note that Lower Grand Valley Dani also features a dependent-marker *-en/-nen* which Bromley (1981) labels ergative. It does optionally mark the transitive subject with ablative semantics (i.e. the 'source' of the action), but in addition that suffix serves many other non-ergative functions. It can encode i) instrumental semantics, ii) it can refer to the source/cause in an <u>intransitive</u> construction, iii) it is used in a 'facilitative' construction, iv) or as a time adjunct (i.e. 'the time when...') and finally v) in a narrative construction when a given event X is viewed as the 'source' of a subsequent event Y. See discussion pp.266-268. For these reasons, the *-en/-nen* suffix is not considered a true ergative marker for the purposes of this survey.

463

**Manchu**

*family*: Tungusic

*source*: Gorelova (2002)

*pattern*: none

(46) a. *i    boo    be    weile-mbi*          (Gorelova 2002; p.166, p.164, p.176)
    3SG house ACC build-IMPF

    'he builds a house'

  b. *muke  inengdari  sain  i    eye-mbi*
    water every.day good GEN flow-IMPF

    'water every day flows well'

  c. *boo    i    ejen*
    house GEN master

    'master of the house'

464

(47)

|  | dependent-marking (Gorelova 2002) | pronouns |
|---|---|---|
| nominative | p.163 | |
| accusative | p.166 | p.216 |
| possessive | pp.175-176 | |

**Manipuri**

*family*: Sino-Tibetan

*source*: Bhat & Ningomba (1997)

*pattern*: none

(48) a. *pulis-nə huranbə-bu pha-re*
  police-NOM thief-ACC catch-PERF

  'the police have caught the thief'

 b. *əŋaŋ-nə kəppi*
  child-NOM cried

  'the child cried'

 c. *oja-gi məca-du siŋŋi*
  teacher-GEN son-that clever

  'the teacher's son is clever'

(Bhat & Ningomba
1997; p.63, p.103, p.78)

(49)

465

| | dependent-marking [Bhat & Ningomba 1997] |
|---|---|
| nominative accusative | p.52 |
| possessive | p.125 |

Manipuri's system of dependent-marking is in accusative alignment; however, in some circumstances the nominative suffix *-nə* is omitted if the subject is less agentive or less animate. Additionally, the accusative suffix *-bu* may be omitted if the transitive object is inanimate. See discussion in Bhat & Ningomba (1997) pp.193-195.

**Maricopa**

*family*: Yuman

*source*: Gordon (1986)

*pattern*: NOM=POSS (full overlap)

(50) a. *m-wik-k*                            (Gordon 1986; p.18, p.19, p.19, p.31)
       2.SUBJ-help-REAL
       'you helped him'

    b. *m-ashvar-k*
       2.SUBJ-sing-REAL
       'you sang'

    c. *ny-wik-k*
       1>2-help-REAL
       'I helped you'

d. *m-mpur*
   2.POSS-hat

   'your hat'

(51)

| | dependent-marking | head-marking |
|---|---|---|
| | (Gordon 1986) | |
| nominative | p.37 | p.21 |
| accusative | p.41 | |
| possessive | - | p.30 |

As discussed in Chapter 3, verbal head-marking in Maricopa is not always unambiguously in accusative alignment; some transitive scenarios are encoded via fusional morphology (50c). See Gordon (1986) Section 1.11 for details.

**Maori**

*family*: Austronesian (Polynesian subgroup)

*source*: Bauer (1993)

*pattern*: none

(52) a. *kua  tae    mai   ngaa   manuhiri*  (Bauer 1993; p.266, p.268, p.108)
     TAM arrive hither the.PL visitor

    'the visitors have arrived'

  b. *ka    kapo   au   i    te   puu*
     TAM snatch 1SG ACC the gun

    'I snatched the gun'

  c. *te   kurii  a    te   tamaiti*
     the dog   GEN the child

    'the child's dog'

(53)

| | dependent-marking (Bauer 1993) |
|---|---|
| nominative | pp.266-267 |
| accusative | pp.267-268 |
| possessive | pp.107-108, pp.209-212 |

Maori marks a distinction between 'A-possession', which encodes semantics of alienable possession, dominance and control, and 'O-possession', which encodes semantics of inalienable possession and subordination. See discussion in Bauer (1993) pp.203-212. Neither of these configurations results in syncretism with the clausal domain, however.

**Maybrat**

*family*: Papuan

*source*: Dol (2007)

*pattern*: NOM=POSS (minor overlap)

(54) a. *y-po*       *ku*    *kiniah*            (Dol 2007; p.77, p.51, p.75, p.63)
3SG.M.SUBJ-hold child small
'he holds the small child'

b. *y-ros*
3SG.M.SUBJ-stand
'he stands'

c. *m-ape*           *ait*
3SG.U.SUBJ-carry.on.back 3SG.M
'she carries him on her back'

469

d. *y-ana*
   3SG.M-POSS-head
   'his head'

(55)

|  | head-marking (Dol 2007) |
|---|---|
| nominative | p.62 |
| accusative | - |
| possessive | p.62 |

According to Dol (2007), only a subset of Maybrat verbs encode the subject via prefix, see Section 3.1.1. In addition, NOM=POSS overlap only extends to inalienable possession, because alienable possession is derived via a different construction with no syncretic pattern; see Dol (2007) Section 5.2.

## Mian

*family*: Trans-New Guinea (Ok subgroup)

*source*: Fedden (2007)

*pattern*: none

(56) a. *né   kóbó   ka-teme-b-i=be*
       1SG 2SG.M 2SG.OBJ-see-IMPF-1SG.SUBJ=DECL
       'I am looking at you'

(Fedden 2007; p.248, p.115, p.281, p.222)

   b. *né   bomanomo Boutlantema taman*
      1SG tomorrow   Boutlantema valley
      *unaa-amab-i=be*
      go-FUT.NANPL.SUBJ-1SG.SUBJ=DECL
      'tomorrow I will go to the Boutlantema valley'

c. *é   skilón=laak  nan-tama-n-e-bio=be*
   3SG foot=down  1SG.OBJ-bite-PST-3SG.M.SUBJ-PAST=DECL

   'it bit me down in my leg'

d. *né   imak=e,           Nialiaaleb=e*
   1SG husband=SG.M Nioliaaleb=SG.M

   'my husband, Nialiaaleb'

(57)

|  | head-marking (Fedden 2007) |
|---|---|
| nominative | p.250 |
| accusative | pp.254-255 |
| possessive | - |

In Mian, the direct object is encoded as a prefix only in a small subset of verbs; see Fedden (2007) pp.254-255 for details.

**Misantla Totonac**

*family*: Totonacan

*source*: MacKay (1999)

*pattern*: none

(58) a. *utun  ta-staa*                  (MacKay 1999; p.127, p.128, p.95, p.109)
      3PL   3PL.SUBJ-sell.TR

      'they sell X'

   b. *utun  ta-paš*
      3PL   3PL.SUBJ-bathe

      'they bathe'

c. *kit   ik-laa-ti̦hwan*
   1SG 1SG.SUBJ-3PL.OBJ-look.for.TR

   'I look for them'

d. *iš-qa̦qa̦-lukut-na̦-ka̦n*
   3.POSS-ear-bone-PL-POSS.PL

   'their horns'

(59)

| | head-marking (MacKay 1999) |
|---|---|
| nominative accusative | p.122 |
| possessive | pp.349-351 |

**Nandi**

*family*: Nilotic (Southern subgroup)

*source*: Creider & Creider (1989)

*pattern*: none

(60) a. *kí:-a-nge:t          kipet*          (Creider & Creider 1989;
        PAST-1SG.SUBJ-wake Kibet              p.93, p.93, p.117, p.40)

        'I woke Kibet up'

     b. *ki:-á-ngê:t            áne:*
        PAST-1SG.SUBJ-wake 1SG

        'I woke up'

c. *ka-kás-â*
   PAST-hear-3SG.SUBJ.1SG.OBJ

   'he heard me'

d. *te:-nyu:*
   cow-1SG.POSS

   'my cow'

(61)

|  | head-marking<br>(Creider & Creider 1989) |
|---|---|
| nominative<br>accusative | pp.97-98 |
| possessive | p.40 |

Nandi transitive and intransitive subjects may be marked with special nominative tonal marking, see discussion in Creider & Creider (1989) pp.41-43 and pp.168-169. Note also that subject agreement takes the form of a prefix (60a-60b), but object agreement is always fusional with subject agreement, and instead it takes the form of a post-verbal suffix (60c).

**Northern Sami**

*family*: Uralic (Sami subgroup)

*source*: Wilbur (2014)

*pattern*: ACC=GEN (minor overlap)

(62) a. *ma-v          dǻn        sida*                          (Wilbur 2014; p.120,
        what-ACC.SG 2SG.NOM want.2SG.PRES                       p.244; p.214; p.89)

     'what do you want?'

473

b. *suovade          dån?*
   smoke.2SG.PRES 2SG.NOM

   'do you smoke?'

c. *båtsoj    máhtta        duv      nala  båhte-t*
   reindeer can.3SG.PRES 2SG.GEN upon come-INF

   'the reinder can attack you' (lit. 'come upon you')

d. *men  ådtjo           sáme          giela-v           ságasti-t  duv*
   but  may.2SG.PAST Saami.GEN.SG language-ACC.SG speak-INF 2SG.GEN
   *åbeni-j*
   sibling-COM.PL

   'but were you allowed to speak the Saami language with your siblings?'

(63)

|            | dependent-marking | head-marking (Wilbur 2014) |
|------------|:-----------------:|:--------------------------:|
| nominative |                   |                            |
| accusative |       p.93        |           p.160            |
| possessive |                   |             -              |

**Nuuchahnulth**

*family*: Wakashan

*source*: Davidson (2002), Nakayama (2001)

*pattern*: none

(64) a. *ńaacsaa='aλ=maa=aḥ      suẃa*          (Davidson 2002; p.110,
        see=TEMP=IND=1SG.SUBJ 2SG                p.108, p.22, p.307)

        'I see you'

474

b. *ʔi:ḥʷ=ʾaλ=maa=aḥ*      *haanaʔa=ʾaλ*

    very=TEMP=IND=1SG.SUBJ play.lehal=TEMP

    'I was playing *lehal* (a gambling game) in a big way'

c. *naʔ:taḥ=ʾλ=ʾi:s*

    listen=TEMP=IMP.2SG>1SG

    'listen to me!'

d. *ʔaλ-qimɬ=uk=qaa=s*          *kuunaa*

    two-many.round-objects=POSS=DEF=1SG.POSS schooner

    'my two schooners"

(65)

| | clitic pronouns (Davidson 2002) |
|---|---|
| nominative accusative possessive | p.264 |

Typically, transitive and intransitive subjects are encoded by a 'Set 1' suffix cliticizing to the first word in the clause, while transitive objects are realized as free pronouns (64a-64b). In some configurations, however, both the transitive subject and the transitive object can be simultaneously encoded via fusional clitic (64c). See Davidson (2002) Section 4.3.1 for more details. Note also that 1st and 2nd persons must be encoded via clitic if possible, resulting in the use of inverse morphology if the transitive subject is third person (see Section 4.2.2). Finally, note that clitics from 'Set 2b' encode possession.

**Pipil**

*family*: Uto-Aztecan (Aztecan subgroup)

*sources*: [Campbell (1985)](), [King (2012)]()

*pattern*: none

(66) a. *ti-nech-ita-k*                                    (Campbell 1985; p.56,
        2SG.SUBJ-1SG.OBJ-see-PAST                      p.56, p.54, p.43)

        'you saw me'

     b. *ti-kuch-ki*
        2SG.SUBJ-sleep-PAST

        'you slept'

     c. *ni-mits-ita-k*
        1SG.SUBJ-2SG.OBJ-see-PAST

        'I saw you'

     d. *mu-chi:l*
        2SG.GEN-chili

        'your chili pepper'

(67)

|            | head-marking (Campbell 1985) |
|------------|------------------------------|
| nominative | p.54, p.56                   |
| accusative | p.55                         |
| possessive | pp.42-43                     |

Note that Pipil features what Campbell (1985) calls an 'absolutive' suffix, but it is not case marking. It simply concatenates to an otherwise bare noun root (like in Comanche). If the noun root is not bare (i.e. part of a compound, with possessive prefix), it is not affixed with the 'absolutive' suffix.

**Sandawe**

*family*: Khoisan

*source*: Steeman (2011), Eaton (2010)

*pattern*: none

(68) a. búrǐ-á:       gìtɬ'é-ː̀     |ʰòrǒ̌:-sé-é                    (Eaton 2010;
       mouse-NOM clothing-SP have.hole-CAUS-3SG.M.OBJ      p.63, p.136, p.17)

   'a mouse has made a hole in a piece of clothing'

   b. ⁿ|ǐ̂    tʃʰíà-á:  ʔíé-ː́  kópòkópò
      body all-NOM stay-& shake

   'the whole body was shaking'

   c. bô-xé:-ː̀     hèwéxé: kʷá:          nìnéwì   m̀fâlmè-à
      word-PL-SP DEM       NARR.3SG.M Nineveh king.GEN-3SG.M.CL
      !'ŏ:-é-wà
      get-3SG.M.OBJ-MULT

   'these words reached the king of Nineveh'

(69)

|            | dependent-marking (Eaton 2010) | pronouns (Steeman 2011) | head-marking |
|------------|------------|------------|------------|
| nominative | pp.135-136 | p.96 | - |
| accusative | pp.50-52 | - | p.172 |
| possessive | - | pp.80-81 | p.80, p.87 |

Sandawe also features subject + mood portmanteau clitics, see Steeman (2011)
p.96. Also, according to Eaton (2010), the accusative dependent marker *-ts'ì* is op-
tional, and in complementary distribution with accusative head-marking on the verb.
See Eaton (2010) p.50 for details.

**Sierra Miwok**

*family*: Utian

*sources*: Broadbent (1964), Freeland (1951)

*pattern*: none

(70) a. *naŋŋaʔ halki-i                hikaahy-j*
       man    hunt.PRES.IMPF-3SG.SUBJ deer-ACC

       'the man is hunting the deer'

    b. *naŋŋaʔ ʔynny-y*                    (Broadbent 1964; p.130, p.132;
       man    come.PRES.IMPF-3SG.SUBJ              Freeland 1951, p.19)

       'the man is coming'

    c. *čukùu-ŋ šúukii-š*
       dog-GEN tail-3SG.POSS

       'the dog's tail'

(71)

|  | dependent-marking | head-marking |
|---|---|---|
|  | (Broadbent 1964) | |
| nominative | p.49 | |
| accusative | pp.49-50 | p.43 |
| possessive | pp.50-51 | |

Nominative verbal head-marking is encoded by Series 1 suffixes, and possessive head-marking is encoded by Series 3 suffixes; see Broadbent (1964) Table 6 p.43. Note also that verbal head-marking fusionally encodes both subject and object if the object is either first person or second person.

**Somali**

*family*: Afro-Asiatic (Cushitic subgroup)

*source*: Saeed (1999)

*pattern*: none

(72) a. *wàa-ay          idin        ark-een*                          (Saeed 1999; p.74,
        DECL-3PL.SUBJ 2PL.OBJ see.PAST-3PL.SUBJ              p.56, p.73, p.217)

      'they saw you (PL)'

   b. *naagó-kíi     wàa-ay           yimad-een*
      woman-DET DECL-3PL.SUBJ come.PAST-3PL.SUBJ

      'the women came'

   c. *wàa-uu           keen-ay*
      DECL-3SG.M.SUBJ bring.PAST-3SG.M.SUBJ

      'he brought them'

   d. *hargaha   adhi-gooda*
      skin.DET sheep.and.goats-3PL.M.GEN

      'the skins of their sheep and goats'

(73)

|              | pronouns | head-marking |
|              | (Saeed 1999) ||
|--------------|----------|--------------|
| nominative   | p.72     | pp.85-93     |
| accusative   | -        | p.73         |
| possessive   | -        | p.115        |

According to Saeed (1999), nominative clitics are optional, and they can cliticize to non-verbal elements. Object clitics are not optional, and always occur immediately preceding the verb (except for third person transitive objects, which are null). In addition, there is a second set of object clitics which is reserved for ditransitive clauses. See Section 4.2.2. for details.

**Southern Pomo**

*family*: Pomoan

*source*: Walker (2013)

*pattern*: none

(74) a. *ču:maṭ'=wam:u     hoʔ:o=wi     biʔk'i-k:'i-w     šiʔmi=ʔwan*
　　　gray.squirrel=NOM teeth=INSTR gnaw-ITER-PERF bow=ACC
　　　'the squirrel gnawed it with his teeth, the bow'

　　b. *kʰaʔbe=ʔwam:u ʔiy:o=ṭow   či:yo-w*  (Walker 2013; p.366
　　　rock=NOM        under=ABL stay-PFV      p.366, p.362)
　　　'Rock (Man) sat below'

　　c. *ču:maṭ'=čo:kʰe     šiʔmi=ʔwan*
　　　gray.squirrel=GEN bow=ACC
　　　'Squirrel's bow'

(75)

|            | dependent-marking | pronouns |
|------------|------------------|----------|
|            | (Walker 2013) |          |
| nominative | pp.366-367 |          |
| accusative | pp.367-368 | p.226 |
| possessive | pp.361-362 |          |

In addition to unambiguous nominative *=wam:u* and accusative *=ʔwan* dependent-marking, Southern Pomo features additional suffixes such as *-yey* and *-čon*, which tend to correspond to arguments with high and low agentivity, respectively. See Walker (2013) p.381, and also Table 47 p.394 for a summary of alignment and grammatical marking strategies in Southern Pomo.

In addition, note that there exist head-marking suffixes for first person singular and second person singular only. Walker considers them 'enigmatic' since the rest of the paradigm is unavailable. See discussion pp.328-329 for details.

**Swahili**

*family*: Bantu

*sources*: Polomé (1967), Vitale (1981), Deen (2002)

*pattern*: none

(76) a. *ni-na-ki-som-a*               *ki-tabu*        (Deen 2002; p.43, p.19, p.27
        1SG.SUBJ-PRES-OBJ$_7$-read-IND 7-book        / Polomé 1967; p.107)
        'I am reading the book'

    b. *ni-mèsha-fík-a*
        1SG.SUBJ-PERF.COMP-arrive-IND
        'I have already arrived'

    c. *a-li-ni-pig-a*
        3SG.SUBJ-PAST-1SG.OBJ-hit-IND
        'he hit me'

    d. *m-walimu w-a-ngu*
        1-teacher   1-PART-1SG.POSS
        'my teacher'

(77)

|  | head-marking | |
| --- | --- | --- |
|  | (Deen 2002) | (Polomé 1967) |
| nominative | pp.22-23 | - |
| accusative | pp.26-27 | |
| possessive | - | p.107 |

**Turkish**

*family*: Turkic

*source*: Göksel & Kerslake (2005)

*pattern*: NOM=POSS (minor overlap)

(78) a. *her   gün çikolata   yi-yebil-ir-im*
    every day chocolate eat-PSB-AOR-1SG.SUBJ
    'I could eat chocolate every day'

(Göksel and Kerslake 2005;
 p.141, p.48, p.143, p.24)

   b. *gid-ebil-ir-im        de*
    go-PSB-AOR-1SG.SUBJ also
    'and I can indeed go'

   c. *beni anlama-mak-ta    inat     ed-iyor*
    1SG understand-SUB-LOC obstinate AUX-IMPF
    'she obstinately refuses to understand me'

   d. *hal-im*
    condition-1SG.POSS
    'my condition'

483

(79)

| | dependent-marking | head-marking |
|---|---|---|
| | (Göksel & Kerslake 2005) | |
| nominative | p.154 | pp.82-84 |
| accusative | p.128, p.156 | |
| possessive | pp.161-162 | p.66 |

**Warao**

*family*: language isolate

*source*: Romero-Figeroa (1997)

*pattern*: ACC=POSS (full overlap)

(80) a. *raukaba      sinar-ine*                                    (Romero-Figeroa 1997;
      slashed.field measure-1SG.SUBJ                              p.78, p.65, p.23, p.5)

      'I measured my slashed field'

  b. *Wauta omi   naru-ki-tia-ine*
      Wauta PRIV go-INTENT-HAB-1SG.SUBJ

      'I am going to see Wauta as usual'

  c. *sina ma-yehe-bu-te-ra*
      who 1SG.OBJ-call-ITER-N.PAST-INT

      'who calls me repeatedly?'

  d. *ma-rahe haya-te*

      1SG.POSS-brother run-N.PAST

      'my brother runs'

(81)

| | dependent-marking | pronouns | head-marking (Romero-Figeroa 1997) |
|---|---|---|---|
| nominative | | | pp.64-65 |
| accusative | - | | pp.66-67 |
| possessive | p.91 | - | pp.43-44 |

Warao head-marking and free pronouns are in nominative-accusative alignment, but only for a subset of persons (1SG, 2SG and 1PL). For other persons in the paradigm, no dedicated nominative or accusative forms exist, and instead OSV basic word order must be relied upon to identify grammatical roles. See Romero-Figeroa (1997) pp.64-67 for more details.

## Appendix C

## EXTRACTION ASYMMETRIES IN
## MORPHOLOGICALLY ERGATIVE LANGUAGES

(raw Wh-question data) / (raw relative clause data)

### Abkhaz

*family*: Northwest Caucasian

*source*: Hewitt (1989)

*extraction asymmetry*: none

(1) a. *yə-z-fa-x'à-da*
    3SG.ABS-who-eat-PERF-INT

    'who has already eaten it?'

    (Hewitt 1989; p.11, p.15, p.10, p.36, p.36, p.36)

b. *yə̀-b-dər-wa-da*
    who-2SG.ERG-know-DYN-INT

    'whom do you know?'

c. *y-àa-da*
    who-came-INT

    'who came?'

(2) a. [ *a-ph°ə̀s*    *də-z-šə̀-z*      ] *a-xàc°a*
    DET-woman 3SG.ABS-who-kill-NON.FIN    DET-man.PL
    *∅-aa-wè-yt'*
    3PL.ABS-come-DYN-FIN

    'here come the men who killed the woman'

486

b. [ *a-xàc'a*    *yə̀-y-ba-z*          ] *a-ph˚ə̀s*
     DET-man who-3SG.ERG-see-NON.FIN    DET-woman

     'the woman whom the man saw'

c. [ *rc'ay˚ə̀-s*    *yə̀-q'o-w*          ] *a-ph˚ə̀s*
     teacher-PRED who-be-STAT.NON-FIN    DET-woman

     'the woman who is a teacher'

(3)

| Wh-questions | relative clauses |
|:---:|:---:|
| Hewitt (1989) ||
| pp.10-15 | pp.35-38 |

Abkhaz features the following Wh-extraction strategy: Wh-questions are formed by converting a verbal agreement morpheme to a Wh-morpheme. Thus, in (1a), a transitive subject Wh-question is formed by generating a Wh-morpheme in the ergative agreement slot which normally encodes transitive subjects. Similarly, in (1b-1c), transitive object and intransitive subject Wh-questions respectively are formed by generating the Wh-morpheme in the absolutive slot. The same strategy is also used in relativization (2). In all cases, Ā-extraction is unambiguous, and as a result there is no extraction asymmetry, as predicted by the disambiguation hypothesis. See Hewitt (1989) pp.101-103 for details on Abkhaz's system of head-marking.

**Apinajé**

*family*: Jê (Northwest subgroup)

*source*: Oliveira (2005)

*extraction asymmetry*: ergative (internally-headed relative clauses only)

487

(4) a. [ *ic-tɛ* *a-mə̃* <u>*bɨ*</u> *j-arẽɲ* *ja* ] *na* *tɛ* *krĩ* *õ* *kamə̃* *pa*
    1-ERG 2-DAT man RP-tell.NF DEF.ART REAL HAB village one INSV live

    'this man I'm telling you about lives in the other village'

   b. [ <u>*di*</u> *čəm* *ja* ] *na* *prɛ* *ra* *iɲ-mə̃* *a-j-arẽ*
    woman RP-stand.NF DEF.ART REAL PAST ASP 1-DAT 2-RP-tell

    'this woman standing there had already told me about you'

(5)

| relative clauses |
| :---: |
| [Oliveira (2005)](Oliveira%20(2005)) |
| |
| pp.281-286 |

(Oliveira 2005;
p.282, p.283)

Internally-headed relative clauses are restricted in the sense that only transitive objects or intransitive subjects can serve as heads. In the alternate externally-headed strategy, there is no such restriction.

**Ayutla Mixe**

*family*: Mixe-Zoque

*source*: [Romero-Méndez (2009)](Romero-Méndez%20(2009))

*extraction asymmetry*: ergative

(6) a. *pëën* *yë'ë* *pu'uy jam* *t-jëët-p*
    who DEM.M board DEIC.D 3.ERG-saw-INDEP

    'who is sawing the board?'

   (Romero-Méndez 2009;
   p.459, p.460, p.458)

   b. *tii=ja'a* *m-itsy* *jam* *tsyäm* *y-tun-yp*
    what=DEM.D 2.POSS-younger.sibling DEIC.D now 3.ERG-do-INDEP.TR

    'what is your brother doing right now?'

c. *pëën men-p*
   who  come-INDEP
   'who is coming?'

(7)

| Wh-questions |
| :---: |
| Romero-Méndez (2009) |
| pp.457-461 |

      Due to its lack of morphological case, V-final basic word order and Wh-fronting, in principle transitive Wh-questions in Ayutla Mixe will feature ambiguous Wh NP V linear word order. As predicted by the disambiguation hypothesis, in transitive subject Wh-extraction only, the verb must undergo an unexpected change in both agreement and mood morphology in order to provide a clear cue that the transitive subject has indeed been extracted.

**Basque**

*family*: language isolate

*sources*: Hualde & Ortiz de Urbina (2003), Laka (1996), de Rijk (1972)

*extraction asymmetry*: none

(8) a. *nor-k    ikusi du  gizon-a*      (Laka 1996; p.22, p.22 / D. Rubio,
    who-ERG seen  has man-DET     p.c. / de Rijk 1972; p.65, p.67, p.88)
    'who has seen the man?'

    b. *zer    eman dio gizon-a-k    ume-a-ri*
     what given has man-DET-ERG child-DET-DAT
    'what has the man given to the child?'

489

c. *nor erori da kale-a-n*
who fallen is street-DET-LOC

'who fell in the street?'

(9) a. [ *guztia daki-en* ] *irakasle-a ez da gaur etorri-ko*
everything know-REL teacher-DET NEG is today come-FUT

'the teacher who knows everything will not come today'

b. [ *gizon-a-k jo du-en* ] *txakur-ra gaiztoa da*
man-DET-ERG hit has-REL dog-DET bad is

'the dog that the man hit is bad'

c. [ *etorri da-n* ] *osaba aberatsa da*
come is-REL uncle.DET rich is

'the uncle who has come is rich'

(10)

| Wh-questions | relative clauses |
|---|---|
| Laka (1996) | de Rijk (1972) |
| pp.22-23 | pp.65-73 |

**Belhare**

*family*: Sino-Tibetan (Kiranti subgroup)

*source*: Bickel (2003)

*extraction asymmetry*: ergative (internally-headed relative clauses only)

(11) a. [ *tombhira-ŋa* <u>*wa*</u>    *seiʔ-s-u=na* ] *chitt-he-m*
        lynx-ERG    chicken kill-TR.PERF-3SG.ABS=DET  find-PAST-1PL.ERG

      'we found the chicken that the lynx had killed'

    \* 'we found the lynx that had killed the chicken'

  b. [ <u>*maʔi*</u>  *khiu-ʔ=na* ] *misen-niu-t-u-ga*      *i*
      person quarrel-NON.PAST=DET  know-NON.PAST-3SG.ABS-2SG.ERG INT

      'do you know the person who is quarrelling?'

(12)                                                             (Bickel 2003;
                                                         p.564, p.564)

| relative clauses |
| :---: |
| Bickel (2003) |
| p.558, p.564, p.566 |

Internally-headed relative clauses are restricted in the sense that only transitive objects or in- transitive subjects can serve as heads. In the alternate externally-headed strategy, there is no such restriction.

**Burushaski**

*family*: language isolate

*sources*: Munshi (2006), Munshi (2015), Willson (1996), Anderson (2007)

*extraction asymmetry*: none

(13) a. *in*     *men-e*   *e-sqan-imi*                     (Munshi 2006; p.148,
        3SG.ABS who-ERG 3SG.M.OBJ-kill-3SG.M.SUBJ.PAST    p.148, p.148 / Munshi
                                          2015; p.40, p.40)

      'which man killed this man?'

b. *ine     men-an     e-sqan-imi*
   3SG.ERG who-INDEF 3SG.M.OBJ-kill-3SG.M.SUBJ.PAST

   'which man did this man kill?'


c. *alma:ri-lu      besen bi*
   cupboard-LOC what  be.PRES.3SG

   'what is in the cupboard?'


(14) a. [ *chil    min-um    ine* ] <u>*hir*</u>
          water drink-ADJ REL   man

   'the man who drank water'


b. [ *hir-e       min-um    ite̱* ] <u>*chil*</u>
     man-ERG drink-ADJ that   water

   'the water that the man drank'


c. (intransitive relative clause data not available)


(15)

| Wh-questions | relative clauses |
|---|---|
| (Munshi 2006) | (Munshi 2015) |
| pp.147-150 | pp.38-39 |


**Canela-Kraho**

*family*: Jê (Northwest subgroup)

*source*: Popjes & Popjes (1986)

*extraction asymmetry*: none


(Popjes & Popjes 1986; p.154, p.154, p.154)

492

(16) a. *jũm temã ita  ton*
     who PAST DEM make

   'who made this?'

   b. *ampo ca  ha  krẽ*
      what 2SG FUT eat

   'what will you eat?'

   c. *jũm mã ata*
      who    DEM

   'who is that?'

(17)

| Wh-questions |
| :---: |
| Popjes & Popjes (1986) |
| pp.153-158 |

In Canela-Kraho, basic word order helps to disambiguate between transitive Wh-questions without recourse to additional morphology: transitive subject Wh-questions are Wh Aux NP V (16a), and transitive object Wh-questions are Wh NP Aux V (16b).

**Cavineña**

*family*: Tacanan

*sources*: Guillaume (2008), Camp (1985)

*extraction asymmetry*: none

(18) a. *ai=ra   =tu   tiru-sha-wa=ama   Lizardu=ja   arusu tee*
WH=ERG =3SG burn-CAUS-PERF=NEG Lizardu=GEN rice garden

'who prevented Lizardu's rice garden from burning?'

(Guillaume 2008; p.70, p.102, p.491, p.757, p.764, p.748)

b. *ai   =tu-ke   =mi mare-wa*
WH 3SG-ABS 2SG shoot.at-PERF

'what did you shoot at?'

c. *ai   sere   tume*
WH intestine there

'what sort of intestines are those?'

(19) a. *tume =tukwe ani-kware   <u>bina</u> [ i-ke   susu-ti-ya=ke   ]*
there EV sit-REM.PAST bat 1SG-ABS suck-go-IMPF=LIG

'there was a vampire bat that was going to suck me (during my sleep)'

b. *e-wane=ke=ra =pa   udu-kware   <u>jae</u> [ tu-ra*
3-wife=3=ERG =REP cook.on.rack-REM.PAST fish 3SG-ERG
*a-aje-ya=ke   ]*
affect-DISTR-IMPF=LIG

'his wife was cooking the fish that he was catching'

c. *beta wekaka iwa-kware   <u>camion=kwana</u> [ e-kueti-u=ke   ]*
two day wait.for-REM.PAST truck=PL POT-pass-POT=LIG

'we waited two days for trucks that could possibly pass'

(20)

| Wh-questions | relative clauses |
| --- | --- |
| Guillaume (2008) | |
| pp.100-103 | pp.747-770 |

Similar to Matses and Trumai, overt case marking makes the interpretation of Wh-questions straightforward (18), but relative clauses are trickier. According to Guillaume (2008), relative clauses can be pre-nominal, post-nominal, headless or internally-headed. As a result, the grammar of Cavineña features a soft constraint such that transitive subject relative clauses tend to be avoided: "there are tendencies – but no absolute requirement – for the [common argument] to be in non-A function within the relative clause" (p.747). Being just a 'soft' constraint – and given that transitive subject relative clauses <u>do</u> exist (19a) – this cannot count as an actual asymmetry; however it is interesting to note that the complexity of Cavineña relative clauses appear to put them just on the <u>cusp</u> of warranting a grammaticalized constraint that would count as an actual extraction asymmetry, like in Trumai or Matses.

**Cupeño**

*family*: Uto-Aztecan (Northern subgroup)

*source*: Hill (2005)

*extraction asymmetry*: none

'who could it be spoiling my medicine?'

b. *e'=et*      *haxi̱-y*    *pe̱-yka*        *neq-qa*
   2SG=2SG.ABS who-ACC 3SG.POSS-behind come-PRES.SG
   'who are you walking behind?'

c. *hax-im=el*       *nalmeyu-qat-im*         (Hill 2005; p.382, p.380,
   who-PL=3PL.ABS sing-IMM.FUT-PL               p.379, p.416, p.419)
   'who are gonna sing?'


(22) a. *mu=ku'ut ivi̱-ngax*        *pe'* *pem-hi̱wchu*    <u>*sewe-t*</u>          [
   and=REP PROX.DEM-ABL DET 3PL.SUBJ-know rattlesnake-ABS
   *mi=pe-qe̱'en-ve*              ]
   3PL.OBJ=3SG.SUBJ-bite-SUB
   'and it is said from this they knew that Rattlesnake was the one who had

   bitten them'

b. *axwe̱-ch-i*      *pe'*   <u>*isi-ly*</u>      [ *ne-tewi̱-ve*        ]
   DEM-ABS-ACC DET coyote-ABS    1SG.SUBJ-see-SUB
   'that's the same coyote I saw before'

c. (intransitive relative clause data not available)

(23)

| | Wh-questions | relative clauses |
|---|---|---|
| Cupeño (2005) | | |
| | pp.377-387 | pp.364-365, pp.415-418 |

     Like other Uto-Aztecan languages, Cupeño features an 'absolutive' nominal suffix which doesn't actually encode alignment, but rather it suffixes to nouns which are not possessed. See Hill (2005) Section 5.1. Note also that Cupeño is only ergative-absolutive by virtue of its system of clitic pronouns. Free pronouns, lexical NPs and verbal head-marking are all in nominative-accusative alignment. See Hill (2005) pp.77-81 for details.

## Diyari

*family*: Pama-Nyungan (Karnic subgroup)

*source*: Austin (1981)

*extraction asymmetry*: none

(24) a. *miṉa-li    ŋaṉa      ṇanda-ṇa wara-yi*    (Austin 1981; p.54, p.192, p.234)
       what-ERG 1SG.ACC hit-PART  AUX-PRES

   'what hit me?'

   b. *miṉa yundu     wayi-ṇa     wara-yi    ṯana kupa wakaṛa-ṇaṉṯu*
      what 2SG.NOM cook-PART AUX-PRES 3PL  child come-IMPL.DS

   'what were you cooking before the children came along?'

   c. *waṛaṉa ṯika     ŋana-yi*
      who      mistake be-PRES

   'who is mistaken?'

(25)

| Wh-questions Austin (1981) |
|---|
| p.151 |

## Drehu

*family*: Austronesian (Southern Oceanic subgroup)

*source*: Moyse-Faurie (1983)

*extraction asymmetry*: none

(Moyse-Faurie 1983; p.190, p.190, p.83)

497

(26) a. *hna   lepi   angeic hnei drei*
PAST beat 3SG    ERG who

 'who beat him?'

b. *hna    humuthi drei hnen la    joxu*
PAST kill         who ERG DET chief

 'who did the chief kill?'

c. *drei lai*
who over.there

 'who's over there?'

(27)

| Wh-questions |
| :---: |
| Moyse-Faurie (1983) |
| pp.189-194 |

   In addition to a VOS clausal frame with ergative-absolutive morphological case, there is also an SVO clausal frame with no alignment whatsoever. In either case, Drehu does not exhibit any kind of extraction asymmetry, as predicted by the disambiguation hypothesis, because Wh-questions are *in situ*.

### Epena Pedee

*family*: Embera

*source*: Harms (1994)

*extraction asymmetry*: none

(28) a. *kʰái-pa pʰáta kʰo-hí-ma*     (Harms 1994; p.122, p.122, p.125)
   who-ERG plantain eat-PAST-INT
   'who ate the plantains?'

  b. *kʰã́a-ta kʰo-hí-ma josé-pa*
   what-FOC eat-PAST-INT José-ERG
   'what did José eat?'

  c. *kʰã́a-tu hã-gá wãyáa bɨ-páa a-hí*
   what-CONTR that-like pass.by be-UNCERT say-PAST
   'he asked, "what might it be that is going through there like that?"'

(29)

| Wh-questions |
| --- |
| Harms (1994) |
| pp.122-128 |

**Georgian**

*family*: Kartvelian

*sources*: Harris (1981), Aronson (1991), Foley (2013)

*extraction asymmetry*: none

(30)  a. *Nato-s    vin     aǩoca*                    (Foley 2013; p.27, p.27 / Aronson
      Nato-DAT who.ERG kiss.3SG.SUBJ.AOR              1991; p.285, p.285, p.285)

      'who kissed Nato?'

   b. *nato-m    vi-s     aǩoca*
      Nato-ERG who-DAT kiss.3SG.SUBJ.AOR

      'who did Nato kiss?'

   c. (intransitive Wh-data not available)

(31)  a. <u>*k'ac-i*</u>    [ *romel-ma-c*     *c'ign-i*    *da-c'er-a*  ]
      man-ABS    which-ERG-REL book-ABS he-wrote-it

      'the man who wrote the book'

   b. <u>*k'ac-i*</u>    [ *romel-i-c*     *v-nax-e*  ]
      man-ABS    which-ABS-REL I-saw-him

      'the man whom I saw'

c. <u>k'ac-i</u>   [ *romel-i-c*      *lap'arak'-ob-d-a* ]
   man-ABS   which-ABS-REL he-was-speaking

   'the man who was speaking'

(32)

| Wh-questions | relative clauses |
|:---:|:---:|
| Aronson (1991) ||
| p.304 | pp.284-288 |

## Gitksan

*family*: Tsimshianic (Interior Tsimshian subgroup)

*sources*: Rigsby (1986), Peterson (2015)

*extraction asymmetry*: ergative

(33) a. *naa ʔan=t*                *qay*   *ɬəmoo=s*  *Bruce*
     who A.REL=3SG.ERG.DEP which help=CNN Bruce

     'who helped Bruce?'

   b. *naa=ɬ*      *ɬʷəmoo-yə-n*                (Rigsby 1986; p.303, p.303,
      who=CNN help-TR-2SG.ERG                p.303, p.404, p.407, p.405)

      'who did you help?'

   c. *naa=ɬ*      *lim-ət*
      who=CNN sing-S.REL

      'who sang?'

(34) a. *kaʔ-ə-'y=ɬ*                <u>kat</u>  [ *ʔan-cakʷ-ə=ɬ*            *naks-t*              ]
     see-TR-1SG.ERG=CNN man   A.REL-kill-INCR=CNN spouse-3SG.POSS

     'I saw the man who killed his wife'

501

b. *maɫ-tə-'y*      *lo:-tit*    *tim* *kʷixʷ-ə=s*      *John=ɫ*    <u>*smax*</u>   [
tell-TR-1SG.ERG OBL-3PL FUT shoot-TR=CNN John-CNN bear

   *ɫə-kaʔ-n*           ]
DEF-see-2SG.ERG

'I told them that John would shoot the bear that you saw'

c. *tim*   *t'is-ə=ɫ*        *x̣aʔ=ɫ*      <u>*kat=ɫ*</u>      [ *pax̣-ət=kə*       ]
FUT hit-TR=CNN slave=CNN man=CNN    run-S.REL=DIST

'the slave will hit the man who ran'

(35)

| Wh-questions | relative clauses |
|:---:|:---:|
| Rigsby (1986) ||
| pp.301-306 | pp.403-408 |

In Gitksan, whenever a transitive subject undergoes Ā-extraction, it must be followed by the relativizer *ʔan*. In addition, agreement within the clause reverts to the dependent paradigm. See Rigsby (1986) pp.287-288 for details on transitive subject extraction, and pp.411-414 for agreement in independent and dependent constructions.

## Halkomelem

*family*: Salishan

*sources*: Galloway (1993), Wiltschko (2006), Suttles (2004)

*extraction asymmetry*: ergative

(36) a. *niʔ*   *cən* *k̓ʷéc-nəxʷ* <u>*k̓ʷθə*</u> [ *niʔ*   *ċéw-ət*   \_\_\_   *təwx̣̓a* ]     (Suttles 2004;
AUX 1SG look-TR    DET   AUX help-TR     3SG        p.80, p.80, p.75)

'I saw the one who helped him'

b. *niʔ   cən   k̓ʷéc-nəxʷ   k̓ʷθə* [ *nem̓ ceʔ   c̓éw-ət-əs        k̓ʷənáɬtən* ]
   AUX 1SG look-TR   DET   AUX FUT help-TR-3.ERG 3PL

   'I saw the one they will help'

c. *k̓ʷθə* [ *ni   q̓áy* ]
   DET   AUX die

   'the one who died'

(37)

| relative clauses |
| --- |
| Galloway (1993) |
| pp.451-453 |

The nature of the ergative extraction asymmetry in Halkomelem relative clauses is as follows: in normal circumstances, third person transitive subjects are encoded via the verbal suffix *-əs*, as in (36b). However, in transitive subject relative clauses only, this person marker is omitted (36a), despite ostensibly featuring a third person transitive subject. In the absence of any case morphology or other available cues, this omission is what prevents a sentence like (36a) from being interpreted as 'I saw the one that he helped', a transitive object relative clause.

**Hindi**

*family*: Indo-European (Indo-Aryan subgroup)

*sources*: Montaut (2004), Mohanan (1994)

*extraction asymmetry*: none

(Mohanan 1994; p.208, p.208, p.209
/ Montaut 2004; p.77, p.247, p.247)

(38) a. *kis-ne   bahut kaam   kiyaa*
who-ERG much work.N do.PERF

'who did a lot of work?'

b. *raam-ne   kyaa kiyaa*
Ram-ERG what do.PERF

'what did Ram do?'

c. *kaun aayaa*
who   come.PERF

'who came?'

(39) a. *maĩ <u>Sītā</u> se bāt karū̃ga* [ *jo   is   mãmle mẽ sab kuch   jāntī hai   * ]
1SG Sita P   will talk      REL DEM matter in   everything know be.PRES

'I will go and talk to Sita, who knows everything in this affair'

b. *vah  <u>eksarsais</u>* [ *jo   raam-ne   kiyaa   * ] *bahut muškil   tʰaa*
DEM exercise    REL Ram-ERG do.PERF   very   difficult be.PAST

'the exercise that Ram did was very difficult'

c. *kuch  moṭĩ-moṭĩ <u>aurtẽ</u>   bāt  kar rahī   thĩ̃* [ *jo   khūb sajĩ-djajĩ thĩ̃  * ]
some fat-RED    woman talk do  PROG IMPF   REL fully arranged IMPF

'some big fat ladies, who were all dressed-up, were speaking'

(40)

| Wh-questions | relative clauses |
|---|---|
| Montaut (2004) ||
| pp.266-267 | pp.235-242, pp.247-250 |

**Hua**

*family*: Trans-New Guinea (Kainantu-Goroka subgroup)

*source*: Haiman (1980)

*extraction asymmetry*: none

(41) a. *kzo'-mamu' na   dekeza-mo ri-ro-na*
who-ERG    DEM arrow-TOP take-PERF-3SG.SUBJ.ANT
*h'ahu-bro-yo-na               hi-ga-na            ma  kva'*
stick.in-TR-PERF-3SG.SUBJ.ANT do-MED-3SG.SUBJ.ANT DEM fashion
*d-gasi-e*
1SG.OBJ-do-3SG.SUBJ

'who was it who stuck the poison arrow and caused this thing to happen to

me?'

   b. *d-vo-o             hu-na           da'auvana*
1SG.POSS-cousin-VOC say-3SG.SUBJ.ANT what
*d-gasi-ne*
1SG.OBJ-do-3SG.SUBJ

'he said "cousin, what have you done to me?"'

   c. *kzo-ga-e*                              (Haiman 1980; p.481, p.477, p.276)
who-2SG-PRED

'who are you?'

(42)

| Wh-questions |
|---|
| Haiman (1980) |
| pp.276-278, pp.401-402 |

505

**Itzaj**

*family*: Mayan (Yucatecan subgroup)

*source*: [Hofling (2000)](#)

*extraction asymmetry*: none

(43) a. *maax t-uy-il-aj*                                     (Hofling 2000; p.422, p.422,

who     CMPL-3SG.ABS-see-CMPL.TR                   p.471 / A. Hofling, p.c.)

'who saw him?' or 'who did he see?'

b. *maax tan-u-k'ay*       *b'aay-loj*

who     DUR-3SG.ABS-sing like-DIST

'who is singing like that?'

(44) a. <u>*b'alum*</u> [ *a'*    *t-u-kin-s-aj*                     *a'*    *winik-e*    ]

jaguar    DET CMPL-3SG.ABS-die-CAUS-DEP.TR DET man-TOP

'a jaguar that killed the man' or 'a jaguar that the man killed'

b. *t-a'*     <u>*b'ej*</u> [ *k-u-b'el*      *ich kaj-ej*     ]

on-DET road    INC-3SG.ABS-go into town-TOP

'the road that goes into town'

(45)

| Wh-questions | relative clauses |
|:---:|:---:|
| Hofling (2000) | |
| pp.421-428 | pp.468-485 |

      Transitive Ā-extraction is ambiguous in Itzaj, resulting in an incorrect prediction for the disambiguation hypothesis.

**Ixil**

*family*: Mayan (Mamean subgroup)

*source*: Ayres (1991), Blunk (2008)

*extraction asymmetry*: ergative

(46) a. *a=in        kat=tzok-on   u    si'-e'*
       F.COP=1SG CMPL=cut-AF DEF firewood-ENC

       'It is I who cut the firewood'

   b. *a=kuxh-e'        nu-ku-b'an=∅-e'*
       F.COP=only-ENC INC-1PL.ERG-do-3SG.ABS-ENC

       'it it just this that we did'

   c. *a=in        kat=vat=in-e'*        (Blunk 2008; p.30, p.26, p.26 /
       F.COP=1SG CMPL-sleep-1SG.ABS-ENC     Ayres 1991; p.91, p.92, p.93)

       'it was me who slept'

(47) a. *u   <u>kaay</u>  [ va' tzeq'-on u    tx'i' ]*
       DEF horse   C    kick-AF DEF dog

       'the horse who kicked the dog'

   b. *u   <u>tze'</u>  [ va' n-un-sa'          ]*
       DEF stick   C    INC-1SG.ERG-want

       'the stick that I wanted'

   c. *kab'al-e'* [ *va' nim chitu'* ]
       house-PL   C    big   CHITU

       'the houses that are very big'

(48)

| Wh-questions | relative clauses |
| --- | --- |
| (Blunk 2008) | |
| pp.25-32 | |

508

**Kaluli**

*family*: Trans-New Guinea (Bosavi subgroup)

*source*: Grosh & Grosh (2004)

*extraction asymmetry*: none

(49) a. <u>*kalu*</u> [ *ni*       *ga-lo:*      *afa*   *di*       *a:ma:*     ] *mogago: dimido:*
        man    1SG.GEN woman-REL steal take.PAST that-ERG    bad       do.PAST
        'the man who stole my wife did very badly'

    b. *Mendi fa:la:dow-a:sa:ga:* [ *sowa:-ya:-lo:*    *olo:*   *a:no:* ] <u>*ena*</u>    *ba:d-ale*
        Mendi arrive-SEQ         snake-ERG-REL shoot that    there see-REP
        'having arrived at Mendi, (they) looked at the place the snake had bitten'

    c. <u>*ko:su*</u> [ *tamin-ami-lo: mio:*      *a:no:* ] *o:m-ba*     (Grosh & Grosh 2004;
        plane    before-LOC-REL come.PAST that      same-NEG       p.67, p.28, p.31)
        'that is not the plane that came before'

(50)

| relative clauses |
| :---: |
| Grosh & Grosh (2004) |
| p.41, p.86 |

**Kapampangan**

*family*: Austronesian (Philippines subgroup)

*sources*: Mirikitani (1972), Forman (1971)

*extraction asymmetry*: accusative

(51) a. *ninu ing sinulat*  (Mirikitani 1972, p.181,
    who  DEF  wrote.AF  p.181, p.190, p.193, p.192)

    'who is the (one who) wrote?'

  b. *nanu ing isulat mu*
    what  DEF  write.PF  2SG

    'what is the (thing) you will write?'

  c. (intransitive Wh-data not available)


(52) a. *ing babaing* [ *tuturung Inglis king anak ku* ]
    DEF woman  teaching.AF English P  child 1SG.POSS

    'the woman who is teaching English to my child'

  b. *panyulat me ing lapis* [ *a ibye ku keka* ]
    write  2SG>3SG DEF pencil  REL give.PF 1SG 2SG

    'you write with the pencil I gave you'

  c. *ding tau* [ *ng mumukyat king buntuk ning Arayat* ]
    DET people  REL climb.AF  P  mountain P  Arayat

    'the people who climb to the mountain of Arayat'

(53)

| Wh-questions | relative clauses |
|---|---|
| (Mirikitani 1972) | |
| pp.177-185 | pp.189-196 |

A classic Austronesian extraction asymmetry, which is technically accusative
in nature, since in this respect transitive and intransitive subject Ā-extraction pattern
together in opposition to transitive object extraction.

510

**Kaqchikel**

*family*: Mayan (K'ichee' subgroup)

*source*: author's fieldwork (but see also Brown et al 2006)

*extraction asymmetry*: ergative

(54) a. *achike x-tz'ët-on        ri    achin*
WH    PERF-see-**AF**    DET man

'who saw the man?'

b. *achike x-u-tz'ët            ri    achin*
WH    PERF-3SG.ERG-see DET man

'who did the man see?'

c. *achike x-∅-wär*
WH    PERF-sleep

'who slept?'

(55) a. *ri    wakx    [    ri x-tz'ët-on    ri    achin    ]*
DET cow        C PERF-see-**af** DET man

'the cow that saw the man'

b. *ri    wakx    [    ri x-u-tz'ët            ri    achin    ]*
DET cow        C PERF-3SG.ERG-see DET man

'the cow that the man saw'

c. *ri    wakx    [    ri x-wär            ]*
DET cow        C PERF-sleep

'the cow that slept'

(56)

|  Wh-questions | relative clauses |
| :---: | :---: |
| (Brown et al 2006) | |
| p.181 | p.182 |

**Karitiâna**

*family*: Tupian

*sources*: Everett (2007), Storto (1999)

*extraction asymmetry*: accusative

(57) a. *morã a-sokõ'i-j          ano*          (Storto 1999; p.200, p.198, p.192 /
      Wh   2SG.ABS-tie.up-IRR 2SG      Everett 2007; p.325, p.339, p.385)

      'who is going to tie you up?'

  b. *mõrãmõn a-ti-hĩr̃ã*
     what     2SG.ABS-O.FOC-smell

     'what did you smell?'

  c. *mora-mon i-oty-t*
     WH-COP   PART-bathe-N.FUT

     'who is it that bathed?'

(58) a. <u>*taso*</u> [ *ɲõnso  mĩː* ] *naka-mĩː-t*    *owã*
      man   woman hit   N.SAP-hit-N.FUT child

     'the man who hit the woman hit the child'

  b. *y-py-so'oot-on*          *yn  <u>eremby</u>*   [ *João ti-sokõ'ĩ*    ] *-t*
     1SG.ABS-ASSERT-see-N.FUT 1SG hammock  João O.FOC-tie.up  OBL

     'I saw the hammock that João tied up'

c. *kĩmbɨednã na-okot* <u>*ɲõnso*</u> [ *terektereŋ-tɨka* ]
   dog           N.SAP-bite.N.FUT  woman    dance-PROG

   'the dog bit the woman who is dancing'

(59)

| Wh-questions | relative clauses |
|:---:|:---:|
| Everett (2007) | Storto (1999) |
| pp.323-326 | pp.130-132 |

The ambiguity in Karitiâna Wh-questions crucially hinges on the fact that the verbal prefix slot ostensibly encodes absolutive arguments, and yet, in the event of transitive object Wh-extraction, it exceptionally agrees with the transitive subject argument (57b). In principle, this exceptional agreement pattern would produce the same surface word and morpheme order as transitive subject Wh-extraction (57a). In actuality, transitive object Wh-extraction is disambiguated via a dedicated 'O.FOC' morpheme *ti-*.

Transitive relative clauses are also potentially ambiguous, but for a different reason. According to Storto (1999), embedded clause word order is verb-final (p.119); in conjunction with the loss of agreement in embedded clauses, this produces transitive relative clauses which would feature the same linear sequence NP [ NP V ] (58a-58b). Again, the dedicated 'O.FOC' morpheme *ti-* clearly identifies instances of transitive object relativization.

**Katukina**

*family*: Katukinan

*source*: Queixalós (2010)

*extraction asymmetry*: ergative

(60) a. *itiyan kawahiri kana wa-duni    tyon*  (Queixalós 2010; p.258, p.245,
      this    cat        FOC  ANTI-catch rat       p.245, p.258, p.246, p.246)

    'it's this cat that caught the rat'

  b. *a-obatyawa    kana Aro na=nuhuk kariwa*
     3SG.POSS-wife FOC  Aro ERG=give white.man

    'it's his own wife that Aro gave to the white man'

  c. *Maranmaran na=tyo        kana tona    tyo*
     Maranmaran GEN=daughter FOC  go.away EXCL

    'it's Maranmaran's daughter that went away!'


(61) a. *i-jik          nyan piya [ wa-dahudyi-nin  Hanani ]*
     1SG.ERG-know DEIC man    ANTI-bring-DEP Hanani

    'I know the man who brought Hanani'

  b. *yo-hik          nyan [ Nodia na=dahudyi-nin ] tukuna*
     1SG.ERG-know DEIC   Nodia ERG=bring-DEP   Indian

    'I know the Indian that Nodia brought'

  c. *yo-hik          nyan [ waokdyi-nin ] anyan piya*
     1SG.ERG-know DEIC   arrive-DEP    this    man

    'I know the man who arrived'


(62)

|  focus-fronting | relative clauses |
| :---: | :---: |
| Queixalós (2010) | |
| pp.244-245, | p.246, |
| pp.257-259 | pp.257-259 |

As discussed in Chapter 5, the motivation for the ergative extraction asymmetry in Katukina does not appear to be disambiguation. Rather, it seems to be based on the morphophonological requirement for the ergative case marker *na=* to always cliticize directly onto the transitive verb. As a result, Katukina's ergative extraction asymmetry is not predicted by the disambiguation hypothesis.

**Kolyma Yukaghir**

*family*: Yukaghir

*source*: Maslova (2003)

*extraction asymmetry*: accusative (attributive relative clauses only)

(63) a. *kin   mit-kele   qamie*
     who 1PL-ACC help
     'who has helped us?'

    (Maslova 2003; p.481, p.143, p.143, p.418, p.421, p.418)

  b. *lem-dik   kes'ī-me*
     what-ABS bring-2SG.ERG
     'what have you brought?'

  c. *kin-tek   jaqte-l*
     who-ABS sing-INTR
     'who is singing?'

(64) a. [ *purk-in*     *šoromo lē-je*     ] <u>*šoromo*</u> *ō-d'ek*    *ell'e*
       seven-ATTR person eat-ATTR ] person COP-2SG CONTR

     'you are a person who has eaten seven people, aren't you?'

   b. *kes'ï-k*       [ *tit*    *īde-met*               ] <u>*anil*</u>
     bring-2SG.IMP    2SG catch-CAUS-PASS.ATTR.2PL    fish

     'bring the fish you have caught'

   c. *tāt*    *āj*    *qon-de-ge*      *čuge-ge* *irk-in*    [ *qodō-d'e* ] <u>*šaqale-k*</u>
     CONN CONN go-3SG.SUBJ-DS way-LOC one-ATTR    lie-ATTR    fox-ABS
     *juø-l'el-mele*
     see-INF-3SG.ERG

     'when he was still going, he saw a fox lying on the way'

(65)

| Wh-questions | relative clauses |
| --- | --- |
| Maslova (2003) | |
| pp.480-486 | pp.416-426 |

     Kolyma Yukaghir has two alignments in its system of morphological case: it may mark transitive objects with overt accusative case (63a), or it may mark transitive objects and intransitive subjects with overt absolutive case (63b-63c). The latter is required in certain scenarios, including Wh-questions, which ensures that the latter will never be ambiguous. See Maslova (2003) Section 12.2 for more on absolutive alignment in Kolyma Yukaghir.

     In relative clauses, Kolyma Yukaghir distinguishes between attributive relative clauses (64a, 64c) and passive attributive relative clauses (64b). This is an

accusative extraction asymmetry. In terms of the disambiguation hypothesis, the asymmetry seems well-motivated because in this environment morphological case is not a reliable cue; unlike in Wh-questions, embedded elements do not require overt case morphology (see bare, embedded transitive object in 64a). Therefore, the grammar of Kolyma Yukaghir makes use of dedicated relativizers to disambiguate transitive attributive relative clauses.

Finally, see also Maslova's (2003) discussion of nominal relative clauses, pp.424-425. Although data is less clear, this alternate strategy appears to have an accusative extraction asymmetry as well.

## Kuikuro

*family*: Cariban (Southern Carib subgroup)

*source*: Franchetto (2010)

*extraction asymmetry*: accusative

(66) a. *u-ingãtzu-ha     ekise-i hikutaha enge-ni-mbüngü*     (Franchetto 2010;
      1.POSS-sister-AFF 3-COP turtle   eat-AGNR-SUBS      p.145, p.145, p.144)

      'it was my sister who ate the turtle'

   b. *hikutaha-ha ege-i     u-ingãtzu     ng-enge-tagü*
      turtle-AFF    DIST-COP 1.POSS-sister OM-eat-CONT

      'it was a turtle that my sister was eating'

   c. *u-ingãtzu-ha     ekise-i t-iniluN-ta-tinhü-i*
      1.POSS-sister-AFF 3-COP 3.ANA-cry-CONT-PART-PNR-COP

      'it was my sister who was crying'

(67)

In Kuikuro, Ā-extraction of the transitive object requires the normally post-verbal transitive subject to move to the immediately pre-verbal position; Franchetto (2010) refers to it as being "promoted" in this scenario (p.145). As a result of this independent requirement, the grammar must make use of the prefix *om-* to specifically indicate these instances of transitive object Ā-extraction to help disambiguate from instances of transitive subject=A-extraction.

## Lezgian

*family*: Northeast Caucasian

*source*: [Haspelmath (1993)](#)

*extraction asymmetry*: none

(68) a. *am    ni       kuk'war-na*
it.ABS who.ERG broke-AOR
'who broke it?'

     (Haspelmath 1993; p.421, p.421, p.193, p.340, p.340, p.340)

    b. *wuna    ana  wuč     iji-zwa*
2SG.ERG there what.ABS do-IMPF
'what are you doing there?'

    c. *im      wuč     insan ja*
this.ABS what.ABS person COP
'what kind of person is this?'

(69) a. [ *itim-ri*      *čül-ler-a*      *iji-zwa-j*      ] *wiri* <u>*k'walax-ar*</u>
        man-ERG.PL field-PL-INESS do-IMPF-PTP    all    work-PL

        *dišehli-jr-in*      *xiw-e*      *hat-na*
        woman-PL-GEN neck-INESS fall-AOR

     'all the work that the men used to do in the fields fell on the women'

b. [ *pačahdi-n xazina*    *čünüx-aj* ] <u>*uğri-jar*</u> *čun*     *ja*
   king-GEN treasury steal-AOP    thief-PL we.ABS COP

     'we are the thieves who stole the king's treasury'

c. [ *q̓ʰfe-j*         ] <u>*jac*</u> *žanawur-ri*   *req'-e*      *kuk'war-na*
   go.away-AOP     bull wolf-ERG.PL way-INESS tear-AOR

     'the bull which had gone away was killed by wolves on the way'

(70)

| Wh-questions | relative clauses |
|---|---|
| Haspelmath (1993) ||
| pp.421-423 | pp.340-342 |

## Lhasa Tibetan

*family*: Sino-Tibetan (Tibetic subgroup)

*sources*: Denwood (1999), Delancey (1984)

*extraction asymmetry*: none

(71) a. [ *ta:*   *sɛ:bɛ:*      ] <u>*mɪ*</u>   *tɕi:*      (Denwood 1999; p.202, p.202, p.200)
      tiger kill.NOM.GEN    man NUM

     'a man who killed a tiger'

b. [ *mi:*      *sɛ:bɛ:*      ] <u>*ta:*</u>   *tɕi:*
<span style="color:gray">man.ERG kill.NOM.GEN    tiger NUM</span>

'a tiger which was killed by a man'

c. [ *kʰasã:*      *le:bə*      ] <u>*tɛ*</u>
<span style="color:gray">yesterday arrive.NOM    the</span>

'the one who arrived yesterday'

(72)

|          relative clauses          |
| :--------------------------------: |
|          Denwood (1999)            |
|            pp.200-202              |

**Lillooet**

*family*: Salishan (Interior Salish subgroup)

*sources*: van Eijk (1985), Roberts (1999), Roberts (1994), Davis et al (1993)

*extraction asymmetry*: ergative

(73) a. *šwat kʷu áċx-ən-táli   ta    nkyáp-a*
        who  IRR see-TR-**TOP** DET coyote-DET

   'who saw the coyote?'

   b. *štam kʷu áċx-ən-aš*                    (Davis et al 1993; p.86, p.83 / Roberts 1994;
      what IRR see-TR-3SG.ERG                  / p.30 /Roberts 1999; p.285, p.289, p.286)

   'what did she see?'

   c. *šwat skíczaʔ qwatsáts*
      who  mother leave

   'whose mother left?'

(74) a. *ti    sqáycw-a [ ti    tup-un'-táli-a     s-Bill      ]*
        DET man-DET   DET hit-TR-**TOP**-DET NOM-Bill

   'the man that hit Bill'

521

b. *nilh wi* <u>*s-John*</u> <u>*múta?*</u> <u>*s-Bill*</u> [ *i* *ats'x-en-án-a* *lhkúnsa* ]
   FOC PL.DET NOM-John CONJ NOM-Bill PL.DET see-TR-1SG.NOM-DET now

   'it was John and Bill that I saw'

c. *tsukw t'u?* <u>*s?ents*</u> [ *ti* *tsícw-a* ]
   only DISC 1SG.EMP DET go-DET

   'only I went; the one who went is me'

(75)

| Wh-questions | relative clauses |
|---|---|
| Davis et al (1993) | Roberts (1999) |
| pp.79-95 | pp.283-291 |

**Matses**

*family*: Panoan (Mayoruna subgroup)

*source*: Fleck (2003)

*extraction asymmetry*: ergative (relative clauses only)

(76) a. *tsundan cun cano buan-o-sh*
    who.ERG 1.GEN canoe take-PAST-3.SUBJ

    'who took my canoe?'

(Fleck 2003; p.834, p.981, p.981, p.1024, p.1022, p.299)

b. *atoda debi-n bë-ac-o-sh*
   what Davy-ERG bring-INFER-PAST-3.SUBJ

   'what did Davy bring?'

c. tsuda *mibi ne-e-∅*
   who.ABS 2.ABS be-N.PAST-1>2.INT

   'who are you?'

(77) a. _cun_ _papa_ _pado-n_ [ _cun_ _tita_ _bed-quid_ ] _chedo-n_
1.GEN father deceased-ERG 1.GEN mother grab-A.NMLZR group-ERG

_uidënua-onda-sh_
hold-DIST.PAST-3.SUBJ

'my late father who captured my mother, and his group, grabbed him'

b. _batachued-n_ _mani_ _che-quid_ [ _matses-n_ _chococa-aid_ ]
tayra-ERG plantain eat-HAB Matses-ERG bury-P.NMLZR

'tayras eat bananas that were planted by Matses'

c. _mani_ [ _sin-aid_ ] _is-ash_ _dadpen_ _cuesban_ _cho-quid_ _nuntan_
plantain ripe-P.NMLZR see-after many bat come-HAB inside

'after seeing ripened plantains, many bats come inside the house'

(78)

| Wh-questions | relative clauses |
|---|---|
| Fleck (2003) ||
| pp.980-983 | pp.1019-1026 |

Relative clauses in Matses do not benefit from a stable set of identifying cues. Like in Trumai, they may be pre-nominal, post-nominal, headless or appositive. Furthermore, some transitive verbs occur in an 'unmarked' ABS-ABS frame, and in addition the third person free pronoun is null for all grammatical roles. Therefore it seems likely that all these facts together create a 'compounding effect', which makes relative clauses inherently more difficult to interpret, and which motivates the two dedicated nominalizers _-quid_ and _-aid_ in accordance with the disambiguation hypothesis.

See Fleck (2003) pp.1019-1026 for more discussion of relative clauses. See also discussion pp.295-300: in non-past tenses, *-quid* is used for intransitive relativization, which in turn corresponds to nominative-accusative alignment for those dedicated nominalizers.

In Wh-questions, however, identification of the extracted argument is more straightforward, and thus no additional morphology is warranted, as predicted by the disambiguation hypothesis. In sum, the extraction asymmetry in relative clauses of Matses is very similar that in Trumai discussed in Chapter 5.

**Movima**

*family*: language isolate

*source*: Haude (2010), Haude (2006)

*extraction asymmetry*: ergative

(79) a. *i'nes* <u>*senyo:ra*</u> [ *di'* *kwey* *vel-na* *no-kis* *asna=∅* ]
 DET.F lady      REL ANTI look.after-DIR OBL-DET.N home=1SG.POSS
 'the lady that looks after my house'

   b. *isos*   <u>*wa:ka*</u> [ *di'* *chikaye=is* *neyru=s* *bet'i* ]
 DET.PL COW      REL find.DIR=3PL.ABS here=DET grassland
 'the cows which they had found in this grassland'

   c. *koro'*   *kos*   <u>*si:doj*</u>   [ *di'* *a:mon* *no-koł* *bałim=∅* ]
 DEM.N DET.N monkey    REL enter    OBL-DET.N field=1SG.POSS
 'there is a monkey that has entered my field'

                                                                (Haude 2010; p.298,
(80)                                                             p.296, p.296)

| relative clauses |
|:---:|
| [Haude (2010)](#) |
| pp.296-299 |

There are two major grammatical constraints operating in Movima. First, Movima features a hierarchical system such that the more 'salient' argument must always occur in the immediately post-verbal position. By default, if the transitive subject is more salient, VSO word order occurs and the verb is marked with 'direct' morphology. However, if the more salient argument is the transitive object, word order is VOS and the verb is marked with 'inverse' morphology.

In addition, in principle only the argument that is not immediately post-verbal (the "obviate" argument in Haude 2010) can be relativized; this is not an issue with a less salient transitive object or an intransitive subject. They are straightforwardly relativized with either direct, or null intransitive morphology on the verb, respectively (79b-79c). Nor is it a problem for a less salient transitive subject; relativization will occur with inverse morphology on the verb.

However, when a conflict occurs between the hierarchical salience requirements and the obviate relativization constraint, Movima needs a way to relativize more salient transitive subjects and less salient transitive objects. To do so, the grammar resolves the conflict by converting the verb to antipassive form via the morpheme *kwey*, allowing for the Ā-extraction in these particular scenarios. In (79a), a more salient transitive subject has been relativized. This is a good prediction for the disambiguation hypothesis: in the absence of any other cues, and due to the

grammatical constraints in the language, if that relative clause's verb didn't feature antipassive morphology, it would mean 'the lady that my house looks after' instead.

Note that morphological alignment is ergative in Movima in the following way: in the pragmatically-unmarked, VSO direct construction, transitive objects and intransitive subjects are encoded via "external cliticization". Transitive subjects only are encoded via "internal cliticization". See discussion in Haude (2010) pp.289-292 and Table 3 p.292 for summary of external *vs* internal clitic properties.

## Ngiyambaa

*family*: Pama-Nyungan (Wiradhuric subgroup)

*source*: Donaldson (1980)

*extraction asymmetry*: none

(81) a. *ŋa:ndu-wa=nu:*       *bumiyi*       (Donaldson 1980; p.264, p.261, p.264)
       who.ERG-EXCL=2.OBL hit.PAST
       'who hit you?'

   b. *minja-wa:=ndu*       *dha-yi*
       what.ABS-EXCL=2.NOM eat-PAST
       'what did you eat?'

   c. *ŋa:ndi-wa:*     *waŋa:y manabi-nji*
       who.ABS-EXCL NEG     hunt-PAST
       'who didn't go hunting?'

(82)

| Wh-questions Donaldson (1980) |
| --- |
| pp.259-270 |

526

**Niuean**

*family*: Austronesian (Polynesian subgroup)

*source*: Seiter (1980)

*extraction asymmetry*: none

(83) a. *fifili   e   hai  a   koe ke  vagahau?*
     choose ERG who ABS 2SG SUB speak

(Seiter 1980; p.106, p.107, p.113, p.246, p.246, p.246)

     "who chose you to speak?'

   b. *totō e   heigoa? ai  fai  mena ke   totō ai*
     hold ABS what    NEG exist thing SUB hold there

     'hold what? there's nothing there to hold'

   c. *ko    e   heigoa e   mena nā?*
     PRED ABS what  ABS thing that

     'what is that thing?'

(84) a. *kehe <u>tama</u>  [  ka   kai e    tau pateta   ]*
     to   child      FUT eat ABS PL  potato

     'to the child who is going to eat the potatoes'

   b. *mo  e   <u>tagata</u>  [ ne    moto  e    koe  ]*
     with ABS person     N.FUT punch ERG you

     'with the person who you punched'

   c. *e   <u>tama</u>  [ ne    hau  i    Makefu  ]*
     ABS child     N.FUT come LOC Makefu

     'the child who comes from Makefu'

(85)

|  Wh-questions | relative clauses |
| :---: | :---: |
| (Seiter 1980) | |
| pp.106-108 | pp.246-247 |

Note that there are two Wh-strategies available in Niuean; in addition to Wh-*in situ* (83a-83c), Wh-elements may be fronted as well. See Seiter (1980) pp.107-108 for details.

## Paumarí

*family*: Arauan

*source*: Chapman & Derbyshire (1986)

*extraction asymmetry*: none

(86) a. *nahina-a bi-na-abina-hi*      *ida*    *ojoro*
who-ERG 3SG.ERG-CAUS-die-THEME.F DEM.F turtle
'who killed the turtle?'

b. *nahina-ra vi-fini-vini*     *adani*   *isai*
what-ACC 3PL.SUBJ-fear-TR DEM.PL children
'what are the children afraid of?'

c. *nahina ida*    *soko-a-ja*
who    DEM.F wash-DETRANS-IMMED
'who is washing?'

(A. Aikhenvald; p.c. / Chapman & Derbyshire 1986; p.242, p.206, pp.239-240, p.239, p.238)

(87) a. <u>*ida*</u>    [ *akadi-ija'ari-ra*      *na-abini-ki*     ]
DEM.F   1SG.POSS-people-ACC CAUS-die-PART
'those who were killing our people'

b. <u>*kaasi*</u> [ *o-rakhai-vini*　　*o-nofi-ki*　　　*a'ini-ni-a*　　　]
　　beach　1SG.SUBJ-plant-TR 1SG.SUBJ-want-PART upriver-FEM-OBL

'the upriver beach which I wanted to plant'

c. <u>*kidi-abi'i*</u>　　　[ *kahano-ki*　] -*a*　*bi-akadai-'i-hi*
　　3SG.POSS-father　drunk-PART　ERG 3SG.ERG-squash-ASP-THEME.F

'her father who was drunk squashed her'

(88)

| Wh-questions | relative clauses |
|---|---|
| Chapman & Derbyshire (1986) | |
| pp.203-216 | pp.238-240 |

## Q'anjob'al

*family*: Mayan (Q'anjob'alan subgroup)

*source*: CLQ (2005)

*extraction asymmetry*: ergative

(89) a. *maktxel max il-on[-i]*      *naq winaq*
      who      ASP  see-**AF**-INTR CL   man
      'who saw the man?'

   b. *maktxel max y-il-[-a']*        *naq winaq*
      who      ASP  3SG.ERG-see-TR CL   man
      'who did the man see?'

   c. *makxel max way-i*                    (Coon et al 2014; p.225, p.192, p.192
      who    ASP  sleep-INTR                 / CLQ 2005; p.209, p.209, p.209)
      'who slept?'

(90) a. *cham winaq* [ *max-∅*            *watx'-ne-n*   *ko-na*           ]
      CL     man      CMPL-3SG.ABS make-TR-**AF** 1PL.POSS-house
      'the man who built our house'

   b. *naq unin* [ *max-∅*          *ha-tzu'*      ]
      CL   boy   [ CMPL-3SG.ABS 2SG.ERG-nag
      'the boy that you nagged'

530

c. <u>*anima*</u> [ *max-on* *b'et y-etoq* ]
   people CMPL-1PL.ABS go 3SG.POSS-with

   'the people we went with'

(91)

| Wh-questions | relative clauses |
|---|---|
| (CLQ 2005) | |
| p.183, pp.199-202 | pp.209-210 |

**Roviana**

*family*: Austronesian (Northwest Solomonic subgroup)

*source*: Corston-Oliver (2003), Waterhouse (1928)

*extraction asymmetry*: ergative (relative clauses only)

(92) a. *esei hoper-i-a* *sa* *igana hie*
   who spear-TR-3SG.OBJ DEF fish DEM

   'who speared this fish?'

   b. *na* *sa* *si* *kaqu tepa-i-a* *rau*
   DEF what FOC FUT ask.for-TR-3SG.OBJ 1SG.ERG

   'what shall I ask it?'

   c. *esei si* *zukezuke* *hoi*      (Waterhouse 1928; p.39, p.233, p.146 /
   who FOC torch.RED there       Corston-Oliver 2003; p.275, p.275, p.275)

   'who is that with a torch?'

(93) a. *hierana sa* <u>*koreo*</u> [ *sapu tupa-na* *e* *zone* ]
   this DEF boy REL punch-3SG.NSUF PERS John

   'this is the boy that punched John'

b. *hierana sa* <u>*koreo*</u> [ *sapu tupa-i-a* *e* *zone* ]
   this　　DEF boy　　REL punch-TR-3SG.OBJ PERS John
   'this is the boy that John punched'

c. *hierana sa* <u>*tie*</u> [ *sapu kote taloa* ]
   this　　DEF man　　REL FUT leave
   'this is the man who is going away'

(94)

| Wh-questions (Waterhouse 1928) | relative clauses (Corston-Oliver 2003) |
|---|---|
| pp.232-233, p.249 | pp.274-275, p.278 |

　　　　According to Corston-Oliver (2003), in subordinate clauses only, all alignment is lost in Roviana. Thus Wh-questions are correctly predicted not to exhibit an extraction asymmetry – since morphological case is still available as a cue (92) – but in addition relative clauses are correctly predicted to exhibit an extraction asymmetry. In transitive subject relative clauses only, the embedded clause must be nominalized and the transitive object is encoded via an idiosyncratic suffix on the verb (93a).

**Seediq**

*family*: Austronesian (Formosan subgroup)

*source*: Holmer (1996)

*extraction asymmetry*: accusative

(95) a. *ima wada qmita Pawan?*　　　　(Holmer 1996, p.86, p.86, p.74, p.74, p.76)
　　　who PRET see.AF Pawan
　　　'who saw Pawan?'

b. *ima wada qtaun na ka Pawan?*
who PRET see.PF NA KA Pawan

'who did Pawan see?'

c. (intransitive Wh-data not available)

(96) a. <u>*seediq*</u> [ *mnapa*     *laqi* ]
person    carry.AF.PRET child

'the person who carried a child'

b. <u>*laqi*</u> [ *napa*     *na ka seediq* ]
child   carry.PF.PRET NA KA person

'the child carried by the person'

c. <u>*rseno*</u> [ *tmnalang* ] *nii*
man    run.AF.PRET   this

'this man who ran'

(97)

| Wh-questions | relative clauses |
|:---:|:---:|
| (Holmer 1996) | |
| pp.85-87 | pp.74-75 |

Like Kapampangan, another classic Austronesian extraction asymmetry which is technically accusative in nature, since in this respect transitive and intransitive subject Ā-extraction pattern together in opposition to transitive object extraction.

**Selayarese**

*family*: Austronesian (South Sulawesi subgroup)

*source*: Finer (1998), Finer (1997), Finer (1994)

*extraction asymmetry*: ergative

(98) a. *inai la-erang-i*      *loka-ñjo*
       who 3.ERG-bring-**3.ABS** banana-DEF

       'who brought the bananas?'

    b. *apa la-taro i Baso ri lamari*
       what 3.ERG-put DET Baso in cupboard

       'what did Baso put in a cupboard?'

    c. *inai ak-kelong*        (Finer 1997; p.689, p.689 / Finer 1998;
       who INTR-sing            p.291, p.291 / D. Finer, p.c.)

       'who sang?'

(99) a. <u>*pakoko*</u> [ *to-la-pallu-ñjo-i*      *jukuʔ-na*    ]
       farmer    REL-3.ERG-cook-DEF-**3.ABS** fish-3.POSS

       'the farmer who cooked his fish'

    b. <u>*jukuʔ*</u> [ *nu-la-pallu-ñjo*     i    *Ali* ]
       fish      REL-3.ERG-cook-DEF DET Ali

       'the fish Ali cooked'

    c. <u>*tedong*</u> [ *nu-ak-kelong-ñjo*    ]
       buffalo    REL-INTR-sing-DEF

       'the buffalo who sang'

(100)

| Wh-questions (Finer 1997) | relative clauses (Finer 1998) |
|---|---|
| pp.687-689 | pp.291-293 |

Note that the nature of the asymmetry here is the <u>omission</u> of absolute morphology of transitive object and intransitive subject Ā-extraction, which is still an ergative extraction asymmetry.

**Shipibo**

*family*: Panoan (Mainline Panoan subgroup)

*source*: Valenzuela (2003)

*extraction asymmetry*: ergative (internally-headed relative clauses only)

(101) a. *tso-n-ki      Tsoma      atsa xeati      xea-ma-ibat-a*
　　　who-ERG-INT Tsoma.ABS yuca drink.ABS drink-CAUS-PAST-PART
　　　'who invited Tsoma to drink yuca beer?'

　　 b. *jawe-ki      Wesna-n      Tsoma      xea-ma-ibat-a*
　　　what.ABS-INT Wesna-ERG Tsoma.ABS drink-CAUS-PAST-PART
　　　'what did Wesna invite Tsoma to drink?'

(Valenzuela 2003; p.378, p.379, p.195, p.482, p.483)

　　 c. *jawe-rin      nato*
　　　what.ABS-INT2 this
　　　'what is this?'

535

(102) a. [ *pitso-n*    <u>*bake*</u>    *natex-a*   ] *-tonin-ra joshin*    *pike*
        parakeet-ERG child.ABS bite-PART   ERG-EV  banana.ABS eat-CMPL

        'the child the parakeet bit ate he banana'

    * 'the parakeet that bit the child ate the banana'

  b. [ *mi-bé*    <u>*ainbo*</u>    *jo-a*       ] *-ra no-n*    *onan-yama-ke*
        2SG-COM woman.ABS come-PART-ABS  EV 1PL-ERG know-NEG-CMPL

        'we don't know the woman who came with you'

(103)

| Wh-questions | relative clauses |
|:---:|:---:|
| Valenzuela (2003) | |
| pp.378-385 | pp.445-486 |

In Wh-questions and externally-headed relative clauses, morphological case provides a sufficient cue such that Ā-extraction is never ambiguous. However, in internally-headed relative clauses, a constraint operates such that only absolutive arguments may serve as heads (102a-102b). See Valenzuela (2003) Chapter 10 for discussion of relativization in Shipibo.

**Sinaugoro**

*family*: Austronesian (Papuan subgroup)

*sources*: Tauberschmidt (1999), Tauberschmidt (1992)

*extraction asymmetry*: none

(104) a. *ḡomi tau-mi    rua-rua   dei   na   b-e            kira-mi       au*
      2PL    man-2PL two-RED who ERG REM-3SG.SUBJ say-2PL.OBJ 1SG

   *tubu-gu              kwaḡwa    bo-ḡo            vaḡi-a*
   grandchild-1SG.POSS bird.name REM-2PL.SUBJ kill-3SG.OBJ

   'who told you to kill my grandchild-bird?'

   b. *kara  o         ura-ni*                    (Tauberschmidt 1999; pp.90-91,
      what 2SG.SUBJ like-IMPF                        p.90, p.90, p.74, p.74)

   'what do you want?'

c. *dei-dei*    *be-ḡe*         *iaḡoma-ni*
   who-RED REM-3PL.SUBJ come-IMPF

'who will come?'

(105) a. [ *ḡarivata e*        *lema-ia-to*          ] *tau-na*          *e*
         banana  3SG.SUBJ steal-3SG.OBJ-PERF    man-3SG.POSS 3SG.SUBJ
      *raga-rekwa-rekwa-to*
      run-loose-RED-PERF

'the man who stole the bananas ran away'

b. [ *ḡe*          *voi-a-to*          ] *dagara-na*        *asi*  *b-o-ni*
     3PL.SUBJ pay-3SG.OBJ-PERF    thing-3SG.POSS NEG REM-2SG.SUBJ-IMP
   *fitoḡa-ia*
   throw.away-3SG.OBJ

'don't throw away the thing they bought'

c. (intransitive relative clause data not available)

(106)

| Wh-questions | relative clauses |
| --- | --- |
| Tauberschmidt (1999) | |
| pp.50-51 | p.74 |

**Sm'algyax**

*family*: Tsimshianic (Coast Tsimshian subgroup)

*sources*: Mulder (1994), Peterson (2015)

*extraction asymmetry*: ergative

(107) a. *'yagay   'wii   gyisiyaasg-at   in-t   deentg  =asga  łgu   alasg-m*
       instead great northwind-3   TOP-3 avenge ABS    little weak-CNN.ADJ
       *yetsisk*
       land-animal

       'instead the great northwind avenged the little weak animal'

  b. *waab-a           awaan nah   dzab-u*         (Mulder 1994, p.135, p.135,
      house-CNN.NP DET    PAST make-1SG      p.135, p.143, p.142, p.142)
      'that's the house that I built'

  c. *awta      uks-haytg-it      gi-sga       lax małiitg-m       kyoox*
      porcupine toward-stand-3 DEM-CNN.P top green-CNN.ADJ grass
      'it was porcupine who stood at the edge on the green grass'

(108) a. <u>*'yuuta*</u> [ *dm   in-t    baa-'n     boot* ]
       man       FUT REL-3 run-CAUS boat

       'the man who will run the boat'

  b. <u>*hakwdak*</u> [ *dzab-is          niyaa-t            * ]
      bow         make-CNN.PRED grandfather-3.POSS
      'a bow which his grandfather had made'

  c. *łimkdii-t-ga          [ gu   da    dzag-a     di-dsa     aamt*
      sister-3.POSS-CNN.P   REL CNN.P dead-CNN.P DEM-CNN.P good
      *da          sganaktda* ]
      CNN.POSS some.time
      'their sister who had been dead there for some time'

(109)

| Wh-questions | relative clauses |
|:---:|:---:|
| (Mulder 1994) | |
| pp.135-140 | pp.140-143 |

According to Mulder (1994), the particle *in-* characterizes instances of transitive subject Ā-extraction only, but in addition *gu* is optional in cases of transitive object and intransitive subject (108c) extraction.

## Sorani Kurdish

*family*: Indo-European (Iranian subgroup)

*sources*: McCarus (2009), Blau (1980), Thackston (2015)

*extraction asymmetry*: none

(110) a. *dwe  aw mirov-a barz-á=y   ke dît?*
    yesterday DEM man-LINK tall-DEM=3SG.ERG who saw

    'who saw that tall man yesterday?'

  b. *dwe  aw mirov-a barz-á ke=y   dît?*
   yesterday DEM man-LINK tall-DEM who=3SG.ERG saw

   'whom did that tall man see yesterday?'

  c. *dwe  ke mird?*         (W. M. Thackston, p.c.)
   yesterday who died

   'who died yesterday?'

(111)

| Wh-questions |
| :---: |
| McCarus (2009) |
| p.615 |

In Sorani Kurdish, the Wh-element moves to an immediately pre-verbal position, which in principle could be a source of ambiguity in a V-final language. However, there is a morphophonological requirement for the ergative marker to cliticize

onto the direct object, effectively serving the same function that morphological case would in identifying grammatical roles. As predicted, then, there is no extraction asymmetry; see Chapter 5 for more discussion of Wh-questions in Sorani Kurdish.

**Suena**

*family*: Trans-New Guinea (Binandere subgroup)

*source*: Wilson (1974)

*extraction asymmetry*: none

(112) a. *ai  meni weti*          (Wilson 1974; p.16, p.17, p.17)
      who ERG  did.3SG.SUBJ.INT

      'who did it?'

   b. *ni  are tosi*
     2SG who speak.2SG.SUBJ.INT

     'to whom are you speaking?'

   c. *ou  awe ni*
     DEM who INT

     'who is that?'

(113)

| Wh-questions |
| --- |
| Wilson (1974) |
| pp.16-17, p.106 |

**Tiriyó**

*family*: Carib (Taranoan subgroup)

*source*: Meira (1999), Carlin (2004)

*extraction asymmetry*: ergative (relative clauses only)

(114) a. *a-kï      n-i-pono-ja-n*               (Carlin 2004; p.253, p.116, p.235 /
         WH-ANIM 3>3-TR-tell-PRES-NCERT          Meira 1999; p.573, p.573, p.574)

         'who is telling it?'

    b. *a-kï-jamo-hpe         tï-wë-e            i-ja,     mëhparë-ton*
         WH-ANIM-PL-INDEF COREF-shoot-NF 3-GOAL game.animals-PL

         'what on earth kinds of game animals did he shoot?'

    c. *a-kï       n-e:-ja-n*
         WH-ANIM 3>3-come-PRES-NCERT

         'who is coming?'

(115) a. *an-po=n-ai          <u>wïtoto</u>, [ kaikui i-tuuka-<span style="color:red">ne</span>-npë        ]*
         WH-LOC=3S$_A$-COP person   dog     3-hit-**A.NMLZR**-PAST

         'where is the person who hit the dog?'

    b. *<u>kaikui</u> ë-waarë, [ pahko i-n-tuuka-hpë        ]*
         dog    2-COGN [ father 3-O.NMLZR-hit-PAST

         'do you know the dog that my father hit?'

    c. *<u>mëërë</u> [ ëturu-ketï-npë      ] ji-tuuka*
         3AN    talk-S.NMLZR-PAST   1O-hit

         'that one, who had been talking, hit me'

(116)

| Wh-questions | relative clauses |
|:---:|:---:|
| Meira (1999) | |
| pp.560-561 | pp.573-574 |

Like in Paumarí, Tiriyó has several clausal frames available, and it uses of them to distinguish different types of Wh-extraction, thus heading off any potential ambiguity. These do not constitute a completely unexpected change in morphology – rather the language is simply making use of existing morphosyntax from plain, declarative sentences and thus this is not considered an ergative extraction asymmetry *per se*.

In relative clauses, however, it appears that the combination of free word order, multiple clausal frames and appositive relativization strategy warrants extra disambiguating morphology.

**Trumai**

*family*: language isolate

*sources*: Guirardello (1999), Guirardello (2010)

*extraction asymmetry*: ergative (relative clauses only)

(117) a. *te  yi-k  in  ha  midoxos*
    who YI-ERG FOC 1SG call
    'who is calling me?'

   b. *te  yi in  Kumaru-k  padi ke*
    who YI FOC Kumaru-ERG wait KE
    'whom is Kumaru waiting for?'

(Guirardello 2010; p.219, p.219, p.218 / R. Guirardello, p.c.)

c. *te  yi in  sa  ke*
   who YI FOC dance KE

   'who is dancing?'


(118) a. *ha  hu'tsa chï in  <u>axos</u>-a-tl*  [ *ha  aton mud husa-t'  chï-k* ]
   1SG see  FOC/TENSE child-EV-DAT  1SG pet  neck tie-NMLZR REL

   'I saw the boy who tied my pet'

b. *ha  hu'tsa chï in  <u>axos</u>-a-tl*  [ *hai-ts  midoxos-t'a ke* ]
   1SG see  FOC/TENSE child-EV-DAT  1SG-ERG call-NMLZR REL

   'I saw the boy whom I called'

c. *ha  hu'tsa chï in  <u>axos</u>-a-tl*  [ *esa-t'  ke* ]
   1SG see  FOC/TENSE child-EV-DAT  dance-NMLZR REL

   'I saw the boy who danced'


(119)

| Wh-questions | relative clauses |
|---|---|
| Guirardello (1999) | |
| pp.231-232 | pp.410-422 |


Relative clauses in Trumai are complicated by the fact that there exist two transitive clausal frames (ERG-ABS and ABS-DAT), and that they can occur in various permutations of word order (pre-nominal, post-nominal, headless or appositive). In Chapter 5, it is argued that these factors produce a 'compounding effect' which in turn requires additional disambiguation morphology (the transitive subject relativizer *chï-k*).

**Tseltal**

*family*: Mayan (Tseltalan subgroup)

*source*: Polian (2013)

*extraction asymmetry*: none

(120) a. *mach'a la    y-il         te   Xun=e*             (Polian 2013; p.262, p.218,
      who    CMPL 3SG.ERG-see DET Juan=DET            p.789, p.789, p.789)

    'who did Juan see?' or 'who saw Juan?'

   b. *mach'a tal*
      who     come.CMPL

    'who came?'

(121) a. *te    <u>ts'i'</u> [ te    la      s-ti'-at=e                    ]*
      DET dog    DET CMPL 3SG.ERG-bite-2SG.ABS=DET

    'the dog that bit you'

   b. *te    <u>ts'i'</u> [ te    la      a-man=e              ]*
      DET dog    DET CMPL 2SG.ERG-bite-=DET

    'the dog that you bought'

   c. *te    <u>ts'i'</u> [ te    tal=e              ]*
      DET dog    DET come.CMPL=DET

    'the dog that came'

(122)

| Wh-questions | relative clauses |
|:---:|:---:|
| (Polian 2013) | |
| pp.217-234 | pp.783-794 |

Ā-extraction of the transitive subject in Tseltal is not characterized by any special change in morphology, which means that potential ambiguity is tolerated (120a) and thus this constitutes an incorrect prediction for the disambiguation hypothesis. Note however that passive morphology can help to disambiguate; see discussion in Polian (2013) pp.261-262.

**Tukang Besi**

*family*: Austronesian (Celebic subgroup)

*source*: Donohue (1999)

*extraction asymmetries*: ergative (*in situ* Wh-questions), accusative (externally-headed relative clauses), ergative (internally-headed relative clauses)

(123) a. *no-manga-'e      na   kaujawa te    emai*          (Donohue 1999; p.452, p.128,
           3.SUBJ-eat-3.OBJ ABS cassava  ERG who                 p.128, p.373, p.380, p.373)
          'who is eating the cassava?'

   b. * *u-'ita-'e              na   emai*
          2SG.SUBJ-see-3SG.OBJ ABS who
          'who did you see?'

   c. * *no-mai        na   emai i    aba*
          3.SUBJ-come ABS who   OBL previous
          'who arrived just then?'

(124) a. <u>*mia*</u>    [ *t[um]opa te    La Udi* ]
          person    slap.SI    CORE La Udi
          'the person who was slapping Udi'

   b. <u>*mia*</u>    [ *i-'ita-su*          ]
          person    PFX-see-1SG.POSS
          'the person who I am looking at'

546

c. <u>*mia*</u>　　[ *w*[*um*]*ila kua Lia* ]
　　person　go.SI　　ALL Lia

'people who left for Lia'

(125)

| Wh-questions | relative clauses |
|:---:|:---:|
| (Donohue 1999) | |
| pp.123-129 | pp.367-388 |

In Tukang Besi's default VOS frame, *na*-marked transitive objects and intransitive subject Wh-elements cannot remain *in situ* (123b-123c). However, in the alternate VSO frame, it is the *na*-marked transitive subject which cannot remain *in situ*; see discussion in Donohue (1999) pp.128-129. In clefted Wh-questions, there is no restriction.

In terms of relative clauses, externally-headed transitive and intransitive subjects are relativized via the subject infix -*um*- (124a, 124c), while transitive objects are relativized with the prefix *i*- and possessive morphology to encode the transitive subject (124b). Finally, internally-headed relative clauses exhibit the classic ergative constraint in the sense that that strategy is only available to transitive objects and intransitive subjects. See Donohue (1999) Chapter 15 for more details.

**Warlpiri**

*family*: Pama-Nyungan (Ngarrkic subgroup)

*sources*: Legate (2002), Nash (1985), Simpson (1991)

*extraction asymmetry*: none

(126) a. *ngana-ngku kurdu nyanungu-nyangu paka-rnu*
      who-ERG    child  3SG-GEN        hit-N.PAST

    'who$_i$ hit his$_i$ child?'

  b. *ngana ka          nyanungu-nyangu maliki-rli wajili-pi-nyi*
     who    PRES.IMPF 3SG-GEN        dog-ERG  chase-N.PAST

    'who$_i$ is his$_i$ dog chasing?'

  c. *ngana-patu ka-lu        wangka-mi*       (Legate 2002; p.43, p.43,
     who-PL    PRES.IMPF-3PL speak-N.PAST   p.215, p.252, p.252, p.213)

    'which ones are speaking?'


(127) a. <u>*jarntu-ngku*</u> [ *kuja   ngarrka yarlku-rnu* ] *kapu paka-rni*
      dog-ERG      COMP man     bite-PAST    FUT  strike-N.PAST

    'the dog that bit the man, he will belt it'

  b. <u>*ngarrka*</u> [ *kuja   jarntu-ngku yarlku-rnu* ] *ngula-ngku kapu paka-rni*
     man       COMP dog-ERG     bite-PAST    that-ERG   FUT  strike-N.PAST

    'the man whom the dog bit, he is going to belt it'

  c. *yalypilyi ngula-ju* <u>*pama*</u>    [ *kuja-ka           nguna*
     yalypilyi that-TOP delicacy   COMP-PRES.IMPF lie.N.PAST
     *manja-ngawurrpa*   ]
     mulga-belonging.to

    '*yalipilyi* is a sweet scale found on mulga trees'

(128)

| Wh-questions | relative clauses |
|:---:|:---:|
| Nash (1985) ||
| pp.235-237 | pp.240-241 |

## West Greenlandic

*family*: Eskimo-Aleut

*sources*: Bok-Bennema (1991), Sadock (2003), Fortescue (1984) ,Yuan (2013)

*extraction asymmetry*: ergative (relative clauses only)

(129) a. *kia      uqaatig-aa*                      (Fortescue 1984; p.23, p.16, p.24 /
     who.ERG talk.about-3SG>3SG.INT      Bok-Bennema 1991; p.69, p.65, p.239)

     'who talked about it?'

   b. *suna=ana  Jaaku-p    siulittaasu-u-vvigi-gaa*
     what=that Jaaku-ERG chairman-be-have.as.place.of-3SG>3SG.PART

     'what is Jaaku chairman of?'

   c. *kina tikis-sinnaa-va*
     who  come-can-3SG.SUBJ.INT

     'who can come?'

(130) a. <u>*piniartuq*</u>      [ *nannu-mik       tuqut-si-suq*      ]
     hunter.ABS        polar.bear-MOD kill-**ANTI**-PART

     'the hunter who killed the polar bear'

   b. <u>*nanuq*</u>        [ *Piita-p    tuqu-ta-a*      ]
     polar.bear.ABS      Piita-ERG kill-PART-3SG

     'the polar bear that Piita killed'

   c. [ *ipis-suq*        ]
     be.sharp-PART

     'the thing that is sharp'

(131)

| Wh-questions | relative clauses |
|---|---|
| Fortescue (1984) | |
| pp.10-18 | pp.49-55 |

In Wh-questions, morphological case on both extracted and non-extracted elements serves to disambiguate. In relative clauses, morphological case on the non-extracted element <u>should</u> provide a sufficient cue, and yet there is still an ergative extraction asymmetry. This is an example of a bad prediction made by the disambiguation hypothesis. See discussion in Chapter 5 for speculation as to the independent constraints which may be responsible.

**Yakima Sahaptin**

*family*: Penutian

*source*: Jansen (2010)

*extraction asymmetry*: none

(132) a. *shí-nɨm   íkush  i-kú-ya*                                  (Jansen 2010; p.203, p.203,
     who-ERG thus  3SG.SUBJ-do-PAST              p.207, p.400, p.399, p.403)

     'who did this to you?'

   b. *tun   nam  á-tkix-sha*
     what 2SG  3SG.OBJ-want-IMPF

     'what do you want?'

   c. *shin-k'a   aw*
     who-next now

     'who's next?'

(133) a. *i-tɬ'yáwi-ya*      *áyat,*    [ *a=sh*     *kwɨ́-nɨm íyax-na*    ]
       <span style="font-variant:small-caps">3SG.SUBJ-die-PAST</span> woman    <span style="font-variant:small-caps">REL=1SG</span> that-<span style="font-variant:small-caps">ERG</span> find-<span style="font-variant:small-caps">PAST</span>

       'that woman died, the one who found me'

    b. *ku*    *kuuk awkú á-sɨ́-ya*         [ *ana-túun sapsikw'a-t*
       and now   then   <span style="font-variant:small-caps">3SG.OBJ-say-PAST</span>    <span style="font-variant:small-caps">REL-ACC</span> teach-<span style="font-variant:small-caps">NMLZR</span>
       *i-náktux-ɨnm-a*               ]
       <span style="font-variant:small-caps">3SG.SUBJ-carry.back-LOC-PAST</span>

       'that's when he spoke the teaching he brought back'

    c. *pá-ni-m*      *íwk'ak <u>itɨ́t</u>*   [ *a=sh*     *wá-ta*     *nimí*     ]
       <span style="font-variant:small-caps">INV-give-LOC</span> that    teeth    <span style="font-variant:small-caps">REL=1SG</span> <span style="font-variant:small-caps">COP-FUT</span> <span style="font-variant:small-caps">1SG.GEN</span>

       'give me those teeth which will be mine'

(134)

| Wh-questions | relative clauses |
|:---:|:---:|
| Jansen (2010) ||
| pp.200-207 | pp.396-404 |

**Yingkarta**

*family*: Pama-Nyungan (Kartu subgroup)

*source*: Dench (1998)

*extraction asymmetry*: none

(135) a. *ngantu*    *wiyarrpa-nma mayu*             (Dench 1998; p.39, p.55, p.22, p.65)
       who.<span style="font-variant:small-caps">ERG</span> hit-<span style="font-variant:small-caps">IMPF</span>     child

       'who hit the child?'

b. *nhaa-rna-nta*        *wankga*
    what-1SG.NOM-2SG.ACC say

    'what will I say to you?'

c. (intransitive Wh-data not available)

(136) a. *ngatha-rna*     *nhanya-purru* <u>*majunpa*</u> [ *nyina-tha* ] [ *pulaya-wu*
      1SG-1SG.NOM see-PAST      turtle     sit-REL.DS    hole-DAT

      *yaya-rnuru* ]
      dig-REL.SS

      'I saw a turtle sitting digging a hole'

b. *kutharra mayu patha-lkurra* <u>*nguurru*</u> *ngampu-ngku* [ *thuthu-ngku*
      two      child   hit-PURP     horse     stick-ERG      dog-ERG

      *nguli-rarnu*    ]
      chase-REL.DS

      'two kids want to hit with a stick the horse which the dog is chasing'

c. (intransitive relative clause data not available)

(137)

| Wh-questions | relative clauses |
|:---:|:---:|
| Dench (1998) | |
| p.63 | pp.61-67 |

Relative clauses in Yingkarta are either marked with *-tha* 'different subject' or *-rnuru* 'same subject' subordinating morphology; see Dench (1998) Section 3.8 for discussion of switch-reference. In addition, Yukulta transitive clauses may occur in 'nominative-dative' alignment; in these instances, it is dative-marked transitive objects which provide a disambiguating cue in $\bar{\text{A}}$-extraction (136a).

552

**Yucatec**

*family*: Mayan (Yucatecan subgroup)

*source*: Bricker (1979), Gutiérrez-Bravo (2009)

*extraction asymmetries*: ergative (completive aspect), accusative (durative aspect)

(138) a. *máax* ___ *il-ik-∅* *María*
who         see-IND-3SG.ABS Maria

'who sees Maria?'

   b. *ba'ax t-u* *maan-aj-∅* *le* *ko'olelo'*
what  COMP-3SG.ERG buy-MOD-3SG.ABS DEM woman

'what did the woman buy?'

   c. *máax b'in-∅* *k'íiwik* (Gutiérrez-Bravo & Monforte 2009; p.84, p.86
who   go-3SG.ABS square   / Bricker 1979; p.133, p.126, p.126, p.134)

'who went to the square?'

(139) a. <u>*le*</u> [ *máaš táan u* *hac'-ik-∅* *hwàan-oʔ* ]
DET  who  DUR 3SG.SUBJ hit-NOM-3SG.ACC John-DEIC

'that person who is hitting John'

   b. <u>*le*</u> [ *máaš táan u* *háʔac'-al* *t-u* *mèen hwàan-oʔ* ]
DET  who  DUR 3SG.SUBJ be.hit-NOM DUR-3SG.SUBJ do   John-DEIC

'that person who is being hit by John'

   c. <u>*le*</u> [ *máaš táan u* *b'in k'íiwik-oʔ* ]
DET  who  DUR 3SG.SUBJ go  square-DEIC

'that person who is going to the square'

(140)

| Wh-questions | relative clauses |
|:---:|:---:|
| (Bricker 1979) | |
| pp.111-113 | |

In completive aspect, the disambiguation strategy is to drop TAM morphology and ergative agreement (138), which is an ergative extraction asymmetry. In durative aspect, however – where alignment is nominative-accusative – the disambiguation strategy is to employ passivization in transitive object extraction, which is an accusative extraction asymmetry (139). See Bricker (1979) for more details.

**Yukulta**

*family*: Pama-Nyungan (Tangkic subgroup)

*source*: Keen (1972)

*extraction asymmetry*: none

(141) a. ŋaka-ya =ṯukanta kurka ŋitjinta milyaḻṯa
      who-ERG =3SG>1SG.PAST take 1SG.POSS spear

    'who took my spear?'

b. ŋaka =yikanti yanma ṯiya
   what =2SG.FUT today.INT eat

   'what will you eat today?'

(Keen 1972; p.174, p.175, p.226, p.227, p.268)

c. ŋaka-yana =ɲiŋki puri
   who-DUB 3SG.FUT come

   'I don't know who will come'

(142) a. *tjina-ŋka*        *ṭaṭinta taŋka* [ *ŋala-pakari*       *miyalṭa*
      where-3SG.NOM.PRES that    man    REL-2SG.ACC.3SG.NOM.PRES spear

      *yi:tja* ]
      give

      'where's the man who gave you the spear?'

b. *kuya-yikanta*      *kuri ṭaṭinta maku* [ *ṭiri-ŋala-kanta*
      INT-2SG.NOM.TR.PAST see   that    woman   snake.ERG-REL-3SG.NOM.TR.PAST

      *pa:tja* ]
      bite

      'did you see that woman who the snake bit?'

c. (intransitive relative clause data not available)

(143)

| Wh-questions | relative clauses |
| | Keen (1972) |
|---|---|
| pp.253-254 | pp.227-228, pp.266-268 |