

CONDITION BIAS IN SPLIT-ALIGNMENT SYSTEMS:
A TYPOLOGICAL STUDY OF NORTH AMERICAN LANGUAGES

Caleb Crandall Hicks

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in
partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department
of Linguistics

Chapel Hill
2015

Approved by:

David Mora-Marín

Misha Becker

Randall Hendrick

Katya Pertsova

J. Michael Terry

©2015
Caleb Crandall Hicks
ALL RIGHTS RESERVED

ABSTRACT

Caleb Crandall Hicks: Condition Bias in Split-Alignment Systems:
A Typological Study of North American Languages
(Under the direction of David Mora-Marín)

This dissertation is a study of the strategies employed in the indigenous languages of North America for distinguishing grammatical subjects from grammatical objects. The degree to which intransitive subjects, transitive subjects, and direct objects are the same or different in their form is a property known as alignment. An interesting feature of alignment is that the sameness or difference of subjects and objects can vary according to certain grammatical conditions. This phenomenon, known as *split-alignment*, has been the focus of linguistic investigation from a range of theoretical perspectives since the 1970s. It is frequently remarked that some grammatical conditions are more likely than others to induce split-alignment.

In this study, I survey fifteen indigenous North American languages in order to (a) determine all the grammatical conditions responsible for splitting alignment systems in these languages; and (b) identify the factors contributing to the preponderance of certain conditions over others. The fifteen languages are selected at random in a manner which controls for genetic and areal affiliation. I determine that at least thirteen split-inducing conditions are present in these languages. The most common are status of a subject or object as a speech act participant, verbal semantics, lexical specification, and grammatical number. I interpret the skewing toward

these conditions by appealing to four generalizations which highlight those factors involved in the common, and several less common, conditions.

Split-inducing conditions which do not immediately fall out from these generalizations are tentatively treated in diachronic or other language-specific terms.

ACKNOWLEDGEMENTS

There are many people and organizations to whom I wish to express my gratitude for their roles in my life as a graduate student at UNC. First, my family. My wife, Pankuri, left a home halfway around the world to start a new one with me, and then followed me halfway across the country so that I could do this. She always knew I'd be able to, and knew it was just a matter of time. Thank you. My kids, Nyasa and Arlo, really had no idea what I was doing, except that "Daddy is sometimes a teacher and sometimes a student" and that eventually he would become a "Doctor of Philosophy like Mommy." Having children while in grad school is a world unto itself, and whether or not changing diapers was preferable studying, it was always with fascination and adoration that I witnessed two human beings (start to) become themselves. My parents, Tom and Carol, and my sisters, Megan, Rachel, and Whitney, did their best to indulge my linguistically oriented soliloquies and accepted with grace the perpetually confusing news that "it looks like it's going to be one more year."

I must acknowledge my dissertation committee, Misha, Randy, Katya, and Mike, and particularly my advisor David. I received from all of them much more than just their comments and insights on this dissertation; the many courses I took from them over these last years has helped me grow from student-student to scholar-student and I will never forget this experience. David has been an excellent mentor, endlessly patient and perpetually willing to read, re-read, and re-re-read. I owe to him my first introduction to the notion "ergativity" and from there the entire world of alignment phenomena. He was always quick to offer myriad examples of alignment-related issues, especially from Mesoamerica, and to pass along relevant literature.

The Linguistics Department at UNC, consisting, of course, of more than are mentioned above, has been the veritable vehicle of my graduate education. I am grateful for the consistent financial support I have received in the form of teaching assistantships, which meant not only that by teaching, my own

knowledge codified, but also that I could manage my daily life without relying (very much) on ramen for basic nutritional sustenance. Along similar lines, I thank the Royster Society of Fellows for awarding me a Sequoyah Dissertation Completion Fellowship in my final year. The financial support and professional camaraderie that came along with it have been lasting and meaningful.

Barring perhaps exclusively online degrees or various extreme case of graduate isolationism, being a graduate student is at least partially a social experience. Relationships, some of which may even come to bear the label “friendships,” do emerge. My cohort in 2008, Alice, Anne, Bunni, Mary, and me, together reveled in—dare I say *endured*—the beginnings of graduate life. Now I count them among my friends and colleagues, as well as Amy, Emily, Jeff, Jen, and Kline. In particular I wish to highlight how glad I am to know Emily and Kline if for no other reason than that we get each other’s jokes or have become very good at pretending to. (I swear I know as little about vulpine phonotactics as the next guy.)

I owe many thanks to Carrie Monette and the other wonderful staff of the UNC Music Library, who, on a weekly basis, retrieved for me a selection of dissertation writing music from the mysterious Closed Stacks, which I imagine must be some kind of enchanted cavern with floors of platinum and walls laced with mithril. If not for their diligent work, I never would have been exposed to a trove of new music, not the least of which was an edifying array of Smithsonian Folkways Recordings, which are sheer delights to anyone with even a modicum of interest in language and culture. In a similar vein, I would like to acknowledge Gioachino Rossini for reminding me how beautiful things can be, and Domenico Scarlatti for reminding me how badass things can be, and both of them for helping me realize the compatibility of beauty and badassness.

Thank you also to the Google Street View car for immortalizing my graduate student self on the interwebs. My bearded mug will now be visible to all at least until our ever-expanding sun finally consumes the earth, or the image is changed by Google, whichever happens first.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS.....	x
FONT CHOICES & GLOSSING CONVENTIONS.....	xvii
1 INTRODUCTION.....	1
1.1 Preamble to Chapter 1.....	1
1.2 Description of the topic.....	1
1.2.1 What split alignment is.....	1
1.2.2 Objectives of this study.....	5
1.2.3 Significance of linguistic biases.....	5
1.2.4 Significance of North American languages.....	6
1.3 Overview of findings and claims.....	8
1.4 Structure of the dissertation.....	10
2 DEFINITIONS AND THEORETICAL ASSUMPTIONS.....	12
2.1 Introduction to Chapter 2.....	12
2.2 Transitivity.....	14
2.2.1 Approaches to transitivity.....	15
2.2.2 Semantic approaches to transitivity.....	15
2.2.3 Syntactic approaches to transitivity.....	17
2.2.4 Valency types in this study.....	20
2.2.5 Summary of transitivity.....	21
2.3 Core arguments.....	21
2.3.1 Core arguments in BLT.....	22
2.3.2 Core arguments as proto-roles.....	24
2.3.3 Summary of core arguments.....	27
2.4 Alignment patterns.....	27
2.4.1 Nominative-accusative alignment.....	28
2.4.2 Ergative-absolutive alignment.....	29
2.4.3 Double-oblique alignment.....	31
2.4.4 Tripartite alignment.....	33
2.4.5 Neutral alignment.....	35
2.4.6 Hierarchical alignment.....	36
2.4.7 Portmanteau alignment.....	38
2.4.8 Summary of alignment patterns.....	39

2.5 Split-alignment and split inducers.....	40
2.5.1 Split inducer 1 — Person splits.....	42
2.5.2 Split inducer 2 — TAM splits.....	43
2.5.3 Split inducer 3 — Object definiteness	44
2.5.4 Combinatorial split induction	45
2.5.5 Split-S alignment — a special case.....	46
2.5.6 Summary of split-alignment	52
2.6 Analytical issues and decisions	52
2.6.1 Locus of alignment realization	52
2.6.2 Locus of alignment as a split inducer	62
2.6.3 Function and form	64
2.6.4 Valency changing operations	65
2.6.5 Alignment-like but not alignment	67
2.7 Summary of Chapter 2	71
3 STUDY DESIGN AND METHODOLOGY.....	73
3.1 Introduction to Chapter 3	73
3.2 Language sample.....	74
3.2.1 Sampling frame.....	74
3.2.2 Stratification of the sampling frame	75
3.2.3 Selection and elimination procedure.....	85
3.2.4 Final sample.....	86
3.2.5 Precision of the sample	89
3.3 Data collection	89
3.3.1 Source materials.....	90
3.3.2 Remaining bias in the sample	90
3.4 Summary of Chapter 3	91
4 DATA FROM SAMPLED LANGUAGES	93
4.1 Introduction to Chapter 4	93
4.2 Split-alignment in the sampled languages	93
4.2.1 Split-alignment in Blackfoot.....	93
4.2.2 Split-alignment in Crow.....	100
4.2.3 Split-alignment in Pawnee.....	103
4.2.4 Split-alignment in Alsea.....	109
4.2.5 Split-alignment in Haida	114
4.2.6 Split-alignment in Tlingit.....	121
4.2.7 Split-alignment in Laguna Keres	128
4.2.8 Split-alignment in Maricopa	136
4.2.9 Split-alignment in Seri	141

4.2.10 Split-alignment in Alabama	145
4.2.11 Split-alignment in Cherokee	150
4.2.12 Split-alignment in Euchee	159
4.2.13 Split-alignment in Lowland Chontal.....	163
4.2.14 Split-alignment in Misantla Totonac.....	168
4.2.15 Split-alignment in Chol.....	172
4.3 Summary of split inducers in the sampled languages	175
5 ANALYSIS OF RESULTS.....	177
5.1 Introduction to Chapter 5	177
5.2 Overall distribution and Dryerian preferences.....	178
5.3 Four generalizations	184
5.3.1 Generalizations 1 and 2.....	184
5.3.2 Generalizations 3 and 4.....	194
5.4 Residue	200
5.4.1 The other 27%.....	200
5.4.2 Nominal status	202
5.4.3 Animacy	203
5.4.4 Aspect & mood	204
5.4.5 Definiteness & Particularness.....	208
5.5 Split-alignment within Optimality Theory.....	209
5.6 Summary of Chapter 5	211
6 CONCLUSIONS AND FUTURE DIRECTIONS	213
6.1 Introduction to Chapter 6	213
6.2 Study and results summary	213
6.3 Summary of main claims	214
6.4 Future directions	215
6.5 Final Conclusion	217
APPENDIX 1: GENETIC GROUPS PER AREA	219
APPENDIX 2: MEMBER LANGUAGES FROM EACH SELECTED GENUS.....	221
APPENDIX 3: ESTIMATED CONTINENTAL SPLIT INDUCERS	224
APPENDIX 4: FULLY SPECIFIED TABLE OF SPLIT INDUCERS.....	225
REFERENCES.....	231

LIST OF ABBREVIATIONS

1	first person
2	second person
3	third person
A	Set A
A*	Set A*
α	alpha set
ABS	absolutive
ACT	actor set
ACV	active
AGT	agent set
AGTV	agentive
ALL	allative
ALT	alternating
AN	animate
AND	andative
ANTIPASS	antipassive
AOR	aorist
APP	dative applicative
APPL	applicative
ART	article

ASP	aspect
AUX	auxiliary
B	Set B
β	beta set
C	Set C
CL	classifier
CMPL	completive
CONT	contingent mode
DAT	dative
DECL	declarative
DEI	deictic
DEF	definite
DEM	demonstrative
DEP	dependent
DET	determiner
DI	distal realis
DIR	direct
DIST	distributive
DL	dual
DO	direct object
DR	bivalent direct
DST	distal

DT	derived transitive
DUR	durative
EM	emphatic realis
EUf	Euchee female
EUM	Euchee male
EMPH	emphatic
EP	epenthetic phoneme
ERG	ergative
EVD	evidential
EXC	exclusive
FIN	final phrase marker
FOC	focus
FUT	future
H	human
HAB	habitual
HSFC	horizontal surface
HUM	human
ICP	incompletive
IMPFV	imperfective
INAN	inanimate
INC	inclusive
INCH	inchoative

IND	independent
INDF	indefinite
INDIC	indicative
INDV	individuative
INF	infinitive
INFO	information
INPERF	independent perfect
INST	instrumental
INTENS	intensifier
INTER	interrogative
INTR	intransitive
INV	inverse
IRR	irrealis
ITER	iterative
LOC	locative
MDST	mesiodistal
MS	men's speech
N	neuter
NAME	proper name
NEG	negative
NEU	neutral
NF	non-finite

NNEU	non-Euchee
NOM	nominative
NONPAST	non-past
OBJ	object
OBV	obviative
PAST	past
PAT	patient set
PATV	patientive
PF	perfect
PFV	perfective
PHYBEN	physical benefactive
PL	plural
PNCT	punctual
PON	proposition/oblique nominalizer
POSR	possessor
POSS	possessive
PP	postposition
PR	proximal realis
PREP	preposition
PRES	present
PRET	preterite tense
PRO	pronoun

PROX	proximal/proximate
PT	passive of transitive
PV	preverb
R	simple realis
R/R	reflexive/reciprocal
REAL	realis
REC	recessive
REF	referential
REP	repetitive
RFLX	reflexive
SAL	salient
SAP	speech act participant
SBPRF	subordinate perfect
SEQ	sequential
SG	singular
SN	subject nominalizer
STAT	stative
SUB	subject
SUB1	subordinating affix for Class 1 verbs
TR	transitive
TI	transitive inanimate
TV	theme vowel

V	suffix or prefix vowel
VAR	verb stem variation suffix
VSFC	vertical surface, face
WS	women's speech
ZCNJ	Ø-conjugation class
○⇒○	a portmanteau morpheme where A is before the arrow and O is after

FONT CHOICES & GLOSSING CONVENTIONS

The font used in the main text of this dissertation is Times New Roman. For examples requiring the International Phonetic Alphabet or other special characters, I have used Doulos SIL because it bears a typographic similarity to Times New Roman. However, Doulos comes with the disadvantage of not supporting boldface or italic type. In cases where boldface or italic type was desired, I used Charis SIL.

The glossing of linguistic forms basically follows the recommendations outlined in the Leipzig Glossing Rules (Comrie et al. 2005) with some minor idiosyncrasies explained as they come up. The linguistic examples contained herein come from a large variety of languages and from many different sources. The sources do not always (or even often) agree on glossing conventions, especially those from earlier periods of linguistic scholarship. I have made an effort to select only one abbreviation per grammatical category to be used across all examples. This sometimes meant editing an original source so that it conformed with my abbreviation choice, but this resulted in far fewer total abbreviations than I would have had otherwise.

Within the structure of interlinear glosses, I have sometimes added dashes in order to segment morphemes if this was not done in an original source wherever it was crucial to the discussion at hand. I tried not to make changes to any glosses if they were irrelevant to points being made in the text, so as to faithfully represent the original authorship.

1 INTRODUCTION

1.1 Preamble to Chapter 1

This dissertation is a cross-linguistic study of the indigenous languages of North America. In it, I investigate the range and distribution of grammatical conditions which are responsible for differences in the strategies languages use to distinguish subjects from objects. This introductory chapter gives a description of the topic and reasons for its theoretical significance in §1.2. In §1.3, I give an overview of the main findings and claims, and then describe the structure of the dissertation with a chapter-by-chapter summary in §1.4.

1.2 Description of the topic

1.2.1 What split alignment is

All languages employ strategies for distinguishing “subjects” from “objects.” Presumably, this fact has a sound practical reason: it is important for speakers and hearers to be able to specify and identify who is doing something and to distinguish *who* is doing it from *whom* (or *what*) it is getting done to. The reason “subject” and “object” are in quotation marks is because these terms are familiar from traditional grammar, but will be of little use for my purposes. Many languages treat the “subject” of an intransitive verb and the “subject” of a transitive verb as different grammatical constructs. So that I can discuss these constructs separately, I follow Dixon (1994) in using the abbreviation **A** to represent the subject of a transitive predicate, **S** to represent

the subject of an intransitive predicate, and **O** to represent a direct object (these abbreviations will be refined in Chapter 2).

A, S, and O can exist in various theoretically defined relationships to one another. In the domain of morphosyntax, for example, A can be the same as S, such as when A and S are represented by the same pronoun. Or, S can be the same as O. Or, the three can allative be the same, or all different. The “sameness” or “differentness” of A, S, and O is **alignment**. The various alignment patterns (that is, the various configurations of their sameness or differentness) are definable on theoretical grounds and will be elaborated on in Chapter 2. It is not usual that a single alignment pattern is present in all grammatical domains within a language. In other words, different alignment patterns tend to co-exist in a given language, and each pattern tends to appear only and always under an identifiable set of grammatical conditions. This state of affairs, where certain alignment patterns are restricted to certain grammatical conditions, is called **split-alignment** because the alignment system as a whole is *split* (i.e. divisible into multiple pattern types). As defined here, split-alignment is recognized as *any* difference in alignment in a single language. An accessible example comes from the English pronouns. In (1), the pronouns for A and S are the same while those for O are different.

(1) English pronouns where A and S are the same and O is different

	<u>A</u>	<u>S</u>	<u>O</u>
1SG	I	I	me
1PL	we	we	us
3SG.MASC	he	he	him
3SG.FEM	she	she	her
3PL	they	they	them

But the forms in (1) are only a subset of the English pronouns. The rest are shown in (2).

There, there is no distinction in the forms of the A, S, or O pronouns.

(2) English pronouns where A, S, and O are identical

	<u>A</u>	<u>S</u>	<u>O</u>
2SG	you	you	you
2PL	y'all	y'all	y'all ¹
3SG.NEUT	it	it	it

The split in the alignment system of English pronouns exists between the set in (1) and the set in (2). For this language, the grammatical conditions which induce the alignment split (*split inducers* as I'll name them in Chapter 2) are person, number, and gender.

Earnest work on split alignment has been ongoing since the 1970s from within several scholarly traditions. Important insights into split-alignment typology have identified the full range of attested alignment patterns, the apparently common split inducers, and the association of certain alignment patterns with certain other grammatical properties. Much of the literature on split-alignment characterizes, explicitly or otherwise, some conditions as being more likely to induce the split than others. A handful of such statements is shown in (3). All italicized and parenthetical portions are original.

(3) Statements alluding to a condition bias

It is quite usual that languages confine the appearance of the ergative/agentive case to a particular aspect. (Butt 2006:157)

Ergative splits are typically triggered by one of two types of factors: (i) aspect and (ii) nominal features such as person and animacy. (Coon & Preminger 2012:311)

Split ergativity refers to alignment variations conditioned by grammatical features of the verb (e.g. tense, aspect, etc.) or of its core arguments (e.g. 1st/2nd person pronoun vs. other nominals), whereas *split intransitivity* refers to the fact that verbal lexemes occurring in intransitive constructions may divide into two (or possibly more) classes differing in their alignment properties. (Creissels 2009:448)

¹ Or some regional variant.

Typically, such splits occur between person categories, between tense or aspect oppositions, or between main and subordinate clauses. (Dahlstrom 1983:39)

The criteria for the split, or in other words, for partial ergativity, are not so numerous. (Bavant 2008:436)

The impression one gets from these claims is that, more than any others, the grammatical properties of verbal aspect or relative position of an argument on what is often called a *nominal hierarchy* play some special role in inducing an alignment split in grammatical systems. This putative scenario—that alignment splits tend to be caused by one, or a small number of, grammatical conditions—is something I am calling a *condition bias*. (In fairness I must highlight that the statements in (3) refer to one particular type of alignment split, split-ergativity, described in Chapter 2; my interest is in the conditions which split alignment systems between all pattern types.)

There are many other grammatical conditions which are known to be involved in alignment splits (and perhaps some which are not yet known). It is furthermore important to rule out the possibility that the tendency for certain grammatical conditions to correlate with alignment splits is due entirely to non-linguistic factors. Possible factors along these lines could be, for instance, various types of bias in the sampling procedure, inadvertently skewing the data in favor of certain genetic or areal groups. Another non-linguistic factor could be heavy citation of a handful of seminal studies on split-alignment, thereby propagating an early depiction of the phenomenon into the bulk of the modern literature.

1.2.2 Objectives of this study

There are two specific objectives of this study. The first is to determine the range of grammatical conditions which split alignment systems in North American languages. The second is to determine whether some of the conditions are indeed more common than others and to explain the asymmetry if it exists. To preview the results, there are thirteen split inducers found in the languages I sampled and certain of them are definitely more common than others. I go into more detail regarding the findings in §1.3 below.

The way I made these determinations was by investigating fifteen indigenous North American languages, sampled in a manner which controlled for genetic and areal affiliation. The continent was divided into five non-overlapping geographic regions, each with a low likelihood of linguistic diffusion between them. From each region, I randomly selected three languages each from a different language family. This ensured that none of the languages was closely related and that language structure has probably not diffused beyond each regional triplet. The split inducers uncovered in the sample were then catalogued and interpreted statistically.

1.2.3 Significance of linguistic biases

That certain linguistic patterns are more common than others is not just an empirical fact. It begs explanation and informs the theoretical apparatuses that linguists use to provide explanations. Linguistic biases of any kind contribute to a conceptualization of linguistic markedness, aid our characterization of the cognitive and social principles which drive language structure and acquisition, and apprise our understanding of linguistic stability and susceptibility to change over time.

Split-alignment is situated at an interesting intersection. On the one hand, the fact that most (*all*, to my knowledge) languages employ at least one strategy for distinguishing the core arguments of a two-argument predicate is perfectly logical. No doubt it serves to avoid what could otherwise amount to considerable ambiguity. On the other hand, there is no a priori reason why there should be splits in alignment systems in the first place. Why doesn't a language just have the same alignment pattern in all sub-parts of its grammar?

The major modern grammatical theories all have some way of dealing with at least certain types of split-alignment, but naturally their specific approaches, and the theoretical substrate which underpins them, are quite varied. Relational Grammar (Perlmutter 1980, 1983), Transformational Grammar (Chomsky 1965, 1982, 1995), Role and Reference Grammar (Foley & Van Valin 1984), and Cognitive Grammar (Langacker 1999), among others, have each made efforts to confront certain linguistic facts regarding split alignment and represent the observed variation using some kind of formalism. Not all of the theories have been used to address split-alignment with the same level of tenacity. In other words, some theories have more elaborate devices than others for dealing with grammatical relations in general (Farrell 2005:43). In the end, any theory of grammar is obligated to offer a framework for explicating split-alignment phenomena of all kinds, as long as its practitioners purport that its principles are universal.

1.2.4 Significance of North American languages

All of the data for this study come from indigenous languages of North America. For several reasons, North America is an ideal geographic area to explore in relation to the topic of split-alignment. First, taken together, North American languages display all of the major types of

split systems described in the alignment literature, so there are not likely to be any significant gaps in the representation of split-alignment even when all instances are drawn from that continent.

Second, North American languages are hugely diverse typologically and genetically (there are over 50 language families). This means that the likelihood of genetic inheritance accounting for shared patterns is lower than on some other continents where there is only a handful of language families.

Third, North America is known to contain a number of strong linguistic areas within which structural diffusion has definitely occurred across unrelated languages. This has two direct implications for the present project: (a) the sampling procedure can systematically reduce the influence of areal diffusion when determining the cross-continental distribution of the various split-alignment patterns; and (b) it leaves open the option of exploring such areas in greater detail as an explanation for why certain patterns may propagate more readily than others.

Beyond the methodological reasons for focusing on North America, this study also tells us something about those languages themselves. Most languages of Europe and many in Africa and Asia are under no threat of extinction. Their large speaker populations, deep literary traditions, and high social prestige also mean that there is no dearth of scholarly work being done on them. The same cannot be said for the indigenous languages of the Americas. Cross-linguistic studies which put emphasis on a particular region do a service to our understanding of that region. In North America particularly, such studies often contribute to claims about genetic relationships in deep time and migration patterns of early peoples, topics about which there is ample controversy both inside and outside linguistics.

On a personal level, indigenous North American languages are special to me. I think every linguist has that one (or that handful of) language(s) to which s/he is inexplicably attracted. Ejective and lateralized consonants, polysynthesis, noun incorporation—these things do it for me. The more I learn about the structure of these languages, the more deeply they appeal to that part of me which is purely curious about the world. Since I began working on this topic, the support I have received from other linguists specializing in North American languages, as well as from members of Native American communities, has been very encouraging.

1.3 Overview of findings and claims

As stated in §1.2, thirteen split-inducing conditions turned up in the languages I sampled. Here I give a very brief overview of the results; in Chapter 2 I specify how I decided what counted as split-alignment and how I measured and interpreted these various grammatical conditions. The complete list of grammatical conditions found in the sampled languages is enumerated in (4). The number listed parenthetically after each grammatical condition represents the total number of languages in which that condition was involved in splitting the alignment system. Because many languages have more than one grammatical condition involved in an alignment split, the sum of these numbers is much higher than the total number of languages sampled.

- (4) Split inducing conditions in fifteen North American languages
1. Speech act participant status (11 languages)
 2. Verbal semantics (7 languages)
 3. Lexical specification (6 languages)
 4. Number (6 languages)
 5. Participant semantics (4 languages)
 6. Nominal status (3 languages)
 7. Person (3 languages)
 8. Animacy (2 languages)
 9. Aspect (2 languages)
 10. Mood (1 language)
 11. Individuation (1 language)
 12. Definiteness (1 language)
 13. Particularness (1 language)

The list in (4) accomplishes the first goal of this study, which is to determine the range of split-inducing conditions. Of course, the list must be qualified as non-exhaustive because these are the results of investigating a sample, and samples are non-exhaustive by definition.

It is obvious that certain conditions, particularly the first four, are the best represented overall. So in partial answer to the second objective, yes, there is a condition bias. I account for the bias by appealing to four generalizations which I argue provide the best deconstructions in general terms of the factors contributing to the common patterns as well as several of the less common ones. The generalizations are summarized in (5).

- (5) Four generalizations
- Generalization 1: Features cumulated with grammatical relation are likely to be split inducers.
- Generalization 2: Low markedness tends to produce zero expression, which tends to neutralize distinctions between grammatical relations.
- Generalization 3: Verbal semantics results in alignment splits due to the assignment of semantic roles to arguments by verbs.
- Generalization 4: Lexical specification is mostly dependent on the verbal semantics criterion and so induces splits by creating exceptions to semantic class membership.

Generalizations 1 and 2 are the main principles involved in the most common split inducer in the sample, speech act participant (SAP) status, as well as a handful of less common inducers. Generalizations 3 and 4 outline the factors behind the second and third best represented split inducers, verbal semantics and lexical specification, respectively. There is residue which is not accounted for directly by the generalizations, but I offer suggestions in Chapter 5 on how these might be treated. It is furthermore interesting to note that aspect and animacy, which are often cited as “common” or “typical” split inducers, are quite poorly represented in my study.

1.4 Structure of the dissertation

In Chapter 2, I lay out my definitions and theoretical assumptions. In particular, I clarify the domain of investigation by outlining the crucial distinction between one-argument and two-argument predicates and the abbreviations A, S, and O. I use Dowty’s (1991) concept of proto-roles to distinguish the two arguments of a two-argument predicate. Then, I define and give examples of each of the alignment pattern types, drawing most examples from native North American languages. I show several examples of split-alignment and have a special discussion of split-intransitive systems. I also discuss several analytical issues, particularly the loci of alignment and phenomena which are superficially similar to alignment, but which do not count as alignment in this study.

In Chapter 3, I detail the methodology employed. Taking inspiration from Dryer’s (1989, 1992) work on basic word order, I explain the criteria I used for dividing North America into non-overlapping sub-regions and the assignment of language family groups to each region. I enumerate the language selection procedure and discuss why some languages were precluded

from consideration or were discarded if selected. In addition, I go into several types of bias which can contaminate a language sample and how my study design avoids these.

In Chapter 4, I describe the alignment systems of each of the fifteen sampled languages. This includes an explicit identification of the locus (or loci) of alignment, which alignment patterns are present in each language, and which grammatical conditions induce each pattern.

In Chapter 5, I offer my analysis of the split-inducing conditions found in the sample. As previewed in §1.3 above, I propose four generalizations which get a good deal of mileage in interpreting the reason that certain split inducers exist in the first place and that some of them are disproportionately more common than others. I also have a discussion of those inducers which do not fall out naturally from my four generalizations and suggest some ways in which these might be treated.

In Chapter 6, I conclude the dissertation and consider some remaining avenues for future study.

2 DEFINITIONS AND THEORETICAL ASSUMPTIONS

2.1 Introduction to Chapter 2

The issues related to alignment typology—which include such many-fingered topics as transitivity, argument structure, semantic roles, and case marking—are represented in an enormous array of linguistic literature spanning the gauntlet of theoretical orientations. It will neither be a good use of space nor be especially insightful for me to completely review the intellectual histories of these topics here. Nonetheless, some theoretical discussion and background information is necessary in order to position my outlook on split-alignment within the field. Couched in this discussion will also be definitions of terms and an explanation of my criteria for deciding whether a given construction is of one type or another.

I do not approach the topic of alignment typology from a single theoretical orientation and I am not characterizing the patterns within a particular unified framework. In my treatment of split-alignment phenomena throughout this dissertation, and my analysis of the patterns revealed in the study results, I place much greater emphasis on description than on theory. It is more important to me to document the empirical facts and cast them in generalist terms than it is to interpret them with the tools of a formal model of language. This does not mean that my approach is atheoretical or theoretically neutral because as soon as definitions are adopted, or notions are assumed, or previous scholarship is built upon, some degree of theory is necessarily involved, even if it is implicit. It does mean that in dealing with these issues, I draw from a

variety of perspectives in order to construct a suitable set of notions and definitions to inform my sample of languages and the particular problems I am trying to solve. Of the theories and frameworks made available by other scholars, all have strengths, but their strengths are sometimes in different areas. Drawing selectively from these approaches affords me the advantage of their insights while sidelining those aspects less useful to me.

A significant approach to this project comes from Basic Linguistic Theory (Dixon 2010a, b), especially in terms of the notions A, S, and O and their utility for describing alignment patterns. However, I part from Dixon in favoring Dowty's (1991) proto-role theory in order more explicitly to distinguish the two arguments of a transitive verb.

Transitivity is an umbrella concept which subsumes the other relevant aspects of alignment, so I begin section §2.2 with a discussion of transitivity, how I decide whether a construction is transitive or intransitive, and a justification for the valency types which are included in this study. In §2.3, I outline the notion of "core argument" and how I identify and distinguish between the relations A, S, and O. Section 2.4 is about how patterns of alignment are defined and identified, and includes some illustrative examples. In §2.5, I give non-exhaustive examples of various types of split-alignment; this section forms the basis for the split patterns found in my sample, described in detail in Chapter 4. Finally, §2.6 deals with certain analytical issues that crop up in cross-linguistic studies of alignment and how I deal with these.

But first, a caveat to the upcoming subsections. The notions and definitions I adopt are meant as springboards, not anchors. In embracing certain ideas over others, I mean only to describe my orientation to the topics at hand with the understanding that they may turn out to be better for some languages than for others. When comparing languages to one another, especially

if they are diachronically and geographically distinct, a rigid adherence to pre-determined definitions is bound to run into problems. At the same time, scientists of all kinds are obligated to rely on definitions in order to ensure that the objects under investigation are indeed comparable; that it is really the same thing being compared. Where phenomena in a given language seem to diverge from the definitions I assume here, I appeal to whatever language-specific evidence is available to justify my inclusion of those cases as data. Necessarily, this is done on a case-by-case basis and described individually for the sampled languages in Chapter 4.

2.2 Transitivity

Of the many terminological problems faced by linguists working on alignment typology, few are more recalcitrant than a valid cross-linguistic definition of transitivity. In individual languages, it is often possible to establish solid criteria for transitivity by appealing to patterns of agreement or the ability of a construction to undergo certain kinds of processes such as passivization. But single-language criteria usually fail when applied cross-linguistically, especially across typologically disparate languages. They fail because where one language has property x , another may not; or where one language allows process y , another may allow a process similar, but not comparable in detail, to process y . In this subsection, I give an overview of two main approaches to transitivity, which I am calling the “semantic camp” and the “syntactic camp.” I opt for a syntactic definition of transitivity which is basically the same as valency, and port the insights of the semantic versions into my definitions of specific argument types.

2.2.1 Approaches to transitivity

There are two basic camps in the alignment literature for defining transitivity with cross-linguistic validity. One camp defines transitivity in terms of the minimum number of arguments a verb requires in order for a clause that contains it to be grammatical. This is a strictly syntactic definition because it eschews any consideration of the semantic nature of the arguments.

Hypothetical types according to this view are dichotomous: one-argument verbs are intransitive while two-argument verbs are transitive.¹ The other camp defines transitivity in semantic terms.

These definitions take into account the semantic nature of the event and the argument(s), usually by appealing to properties such as agency, volition, affectedness, etc. They often invoke some degree of prototypicality, treating transitivity as a gradient property rather than a binary one. The perspective I adopt is that transitivity itself is best defined syntactically for my purposes because it fosters easier cross-linguistic identification, but that verbal arguments themselves are best distinguished from one another by semantic criteria.

2.2.2 Semantic approaches to transitivity

Orientations to a definition of transitivity which can broadly be characterized as “semantic” are the ones which take into account properties of participants and events. Hopper & Thompson (1980) have been strongly influential. They provide a list of parameters which tend to be true of participants (affectedness, individuation) and events (action, telicity, punctuality, etc.); the collection of parameters in a given event thereby leads to a branding of the event as having “high” or “low” transitivity. Such a conception of transitivity places events on a continuum with

¹ I am ignoring for now three-argument verbs, but I address these shortly in §2.2.4.

interesting consequences, for instance that a two-argument construction can be “more transitive” than another, as in (1), or even that a one-argument construction can be more transitive than a two-argument one, as in (2).

- (1) Sentence (b) is “more transitive” than sentence (a) (Hopper & Thompson 1980:253, ex. 3)
 - a. Jerry likes beer.
 - b. Jerry knocked Sam down.
- (2) Sentence (b) is “more transitive” than sentence (a) (Hopper & Thompson 1980:254, ex. 6, 7)
 - a. Jerry likes beer.
 - b. Susan left.

That the (a) and (b) sentences in (1) and (2) above can differ in transitivity is exactly the conclusion that Hopper & Thompson argue for. But it makes comparing constructions across languages and across events in the same language prone to certain problems. Consider, for instance, the construction *Jerry likes beer*, which is argued to be the “less transitive” one in both pairs above. Do all instances of this construction in the same language have the same transitivity? Imagine that Jerry, who just turned 21, is about to try beer for the first time. He’s not sure if he will like it or not. He hesitantly takes his first sip. With surprise and delight, he goes for a second sip. His buddy sitting next to him exclaims, “Jerry likes beer!” In this scenario, the construction *Jerry likes beer* has several of the properties of “high” transitivity, such as punctuality, affirmation, and realis. Although Hopper & Thompson admit a crucial discourse component of transitivity, it is not clear how foregrounding or backgrounding are specifically relevant to the scenario just described. From a comparative point of view, it is desirable that the assignment of a transitivity value to *Jerry likes beer* is consistent for every instance of that construction, and does not rely on calibration with other constructions.

All of the semantic approaches have in common some notion of prototypicality; for some, this is the defining feature (Næss 2007). The less prototypical, or sometimes called “non-canonical,” transitive verbs are the ones which are assigned fewer high transitivity parameters. They are usually psychological predicates or take an experiencer semantic role (Belletti & Rizzi 1988, Scheepers et al. 2000, Bennis 2004) and often take morphologically distinct forms from those of the prototypical transitives. It is common for typologists working on alignment to systematically exclude these kinds of verbs from consideration when the goal is to understand how the major patterns work. It is important to me that I do include the so-called non-canonical transitive clauses in this study. Experiencer and psych verbs reveal a sub-class of the kinds of alignment splits I aim to investigate. Let’s say that in some language, a verb meaning something like *love* takes a dative subject (i.e. a subject with marking whose form is identical to the marking of an indirect object). That is not a reason to exclude that verb from analysis. To the contrary, this kind of scenario actually shows a split because *love* (and whatever other verbs behave similarly) induces a different alignment pattern than “canonical” transitive verbs like *kill* or *kick*. Comrie’s (1989) approach to grammatical relations is a defense of this practice. By extending the assignments of A and O from canonical transitive arguments to non-canonical ones, we are simply asserting that “A and [O] are thus syntactic terms, whose prototypes are defined in semantic terms” (ibid:111).

2.2.3 Syntactic approaches to transitivity

None of the preceding discussion is intended as an argument against semantic approaches to transitivity. I mean only to highlight that in practice, I have to decide whether a construction in

a given language is transitive or not, and thereby how to assign argument roles. The easiest way to do this consistently is to refer to a syntactic definition of transitivity which is based on the maximum number of arguments permitted by the predicate. Intransitive predicates subcategorize for exactly one argument. Transitive predicates subcategorize for two. This conception of transitivity essentially equates it to valency. I recognize that some scholars insist on a formal distinction between transitivity and valency. In the sources I consulted to obtain the data for this study, *intransitive* and *transitive* are nearly ubiquitous terms, much more so than *monovalent* and *bivalent*, or *one-place predicate* and *two-place predicate*.

As Givón (2001:109) notes:

While the two definitions of transitivity, semantic and syntactic, seem independent of each other, there is in fact a strong overlap between the *populations* of events and clauses they predict. That is, in most languages, the vast majority of *simple* clauses that are semantically transitive are also syntactically transitive.

Moving forward, I treat transitivity and valency both as indicating the maximum number of permissible arguments.

A quick example of each transitivity type is shown below in (3) and (4) from Caddo, a Caddoan language spoken in Oklahoma.

- (3) An intransitive clause in Caddo (Melnar 2004:43, ex. 52)

k'apáhci? hák-ku-haka-wa-yúk-sa?
 chicken INDIC-1PAT-INDV-PL-vanish-IMPV
 'my chickens are vanishing'

- (4) A transitive clause in Caddo (Melnar 2004:75, ex. 36)

ci-n-ba-ki = sáw-hah
 1AGT-APP-food-scrape-HAB
 'I'm scraping corn kernels off the cob'

The construction in (3) contains an intransitive verb, *-yúk-* ‘vanish’ with exactly one argument, *k’apáhci?* ‘chicken.’ The construction in (4) contains a transitive verb, *-sáw-* ‘scrape’ and has two arguments, *ci-* ‘I’ and *-ki=* ‘food’ (interpreted contextually as ‘corn kernels’).

Many languages allow certain arguments to be zero.² The zero realization of third-person arguments is especially common, a phenomenon that will be discussed at length in Chapter 5. As brief examples, consider (5) and (6) from Mandan, a nearly extinct Siouan language from North Dakota.

- (5) A transitive clause with a zero subject in Mandan (Kennard 1936:10)
no-hε-soc
us-see-PRET
‘he saw us’
- (6) A transitive clause with a zero object in Mandan (Kennard 1936:9)
wa-i’sεk-ktoc
I-make-FUT
‘I will make it’

In (5), the third-person subject has no overt realization (and might be equally interpretable as *she*, given appropriate context). In (6), it is the third-person object that is zero. In both (5) and (6), the clause and verb are still transitive. The reason is that it is not whether a verb *has* two overt arguments which makes it transitive; it is whether it *permits* two overt arguments: full noun arguments in (5) and (6) would, of course, be overt. Verbs that forbid two arguments are intransitive. Related to this is the optionality of certain arguments, even full nouns (Faulhaber 2011:8-10). I consider *eat* to be transitive whether it appears in *Rahul ate the kantola* or *Rahul ate*; the optionality appears to be lexically specific, as shown by a well known and (generally

² Zero realization is distinct from pro-drop, where pronouns are standardly omitted but permissible for pragmatic or discourse reasons. I will not go into further detail about pro-drop phenomena.

regarded to be) ungrammatical construction like **Rahul devoured*. This is not to claim that every verb has exactly one transitivity value; I grant that some verbs can be genuinely ambitransitive. A verb like *cool* is legitimately transitive in *Tetsu cooled the sōchū* but intransitive with an unaccusative reading in *The sōchū cooled*. Such cases are dealt with individually as they arise in Chapter 4.

2.2.4 Valency types in this study

This study is an analysis of the three relations S, A, and O (discussed below in §2.3). These relations are intended to be interpreted only in reference to one-argument and two-argument constructions. There are other relations (with their own abbreviations) involved in three-argument constructions (ditransitives), but these will not feature in the present study. A good deal of work on alignment typology has brought ditransitives into the conversation (Bickel 2011, Comrie 2005, Dixon 2010b, Haspelmath 2011), inarguably to the benefit of the field. Even the earliest research on such topics can appear incomplete without a treatment of three-argument constructions, as revealed by Butt's (2009:31), critique of Fillmore's (1968) famous work on case and language classification:

Fillmore's original assumption excludes a consideration of indirect objects. But what justifies this assumption? Some languages use case to identify indirect objects (generally with the dative case), some do not — why should this not provide a useful basis for classification into language types?

As ditransitives have come to be almost expected in a thorough study on alignment, I must address their exclusion from my study. My reasons for excluding ditransitive clauses from consideration are not theoretical, but practical. A more complete typology of the conditions

which induce split-alignment would incorporate data from ditransitives, and obviously this study would be better if I did so. But a dissertation can only be so big and a person can only spend so much time working on it. So in the interest of time, I leave ditransitives to future work. This project must therefore be understood to claim relevance only to a restricted set of clausal valency types, wherein ditransitives are simply not addressed.

2.2.5 Summary of transitivity

The approach to transitivity I am adopting is one based on valency rather than semantics. On this view, intransitive verbs permit exactly one argument while transitive verbs permit two. In addition, I am excluding ditransitive constructions from this study in order to make it more manageable.

2.3 Core arguments

The notion of transitivity described above hinges on the notion *core argument*. An *argument* is an entity that participates in the event or state specified by a predicate; it is usually expressed as a noun phrase, pronoun, or pronominal. While transitivity itself is treated here as a syntactic phenomenon, core arguments involve a strong semantic component. Their semantics is essential to being able to tell them apart when they occur in the same construction. The idea of core argument I adopt moving forward is inspired by Dixon (1979, 1994, 2010a, b), described in §2.3.1. But I part from Dixon in favor of Dowty's (1991) proto-role approach for defining the distinction between types of arguments (§2.3.2).

2.3.1 Core arguments in BLT

Basic Linguistic Theory (BLT) is a theoretical framework spelled out most fully by Dixon (2010a, b) in a three-volume book of the same title. BLT is a descriptive theory, not an explanatory one. In other words, it provides tools, nomenclature, and conceptual distinctions adequate for describing linguistic form. It does not offer explanations for why languages are the way they are, but is, in principle, compatible with explanatory theories. Dryer (2006:211) has a useful summary of BLT's position within linguistic theory generally:

[Basic Linguistic Theory] can thus be roughly described as traditional grammar, minus its bad features (such as the tendency to describe all languages in terms of concepts motivated for European languages), plus necessary concepts absent from traditional grammar. It has supplemented traditional grammar with a variety of ideas from structuralism, generative grammar (especially pre-1975 generative grammar and relational grammar), and typology.

As a descriptive theory, the terminology and concepts of BLT will be the primary means by which I identify and talk about the relevant constructions in the languages in my sample (Chapter 4). However, the sources I rely on to provide that data are sometimes peppered, sometimes laden, with other theoretical substrates. This has meant that part of my job in researching these languages has been to translate their various descriptions into terms which foster cross-linguistic comparison. Usually, this was fairly easy to do simply because grammatical relations are a common topic in descriptive grammars and scholars often get right to the point. But not always, and in those cases it required heavier interpretation on my part of language descriptions to put them into a cross-linguistic context.

Dixon (1979, 1994, 2010a, b) makes a case against using the label “subject” to refer both to the single argument of an intransitive verb and to one of the two arguments of a transitive

verb. His motivation comes from languages with ergative constructions, where the “subject” of a transitive verb is marked differently (e.g. takes a different case marker, has a different pronoun, etc.) from the “subject” of an intransitive verb. As long as intransitive subjects, transitive subjects, and direct objects can have different properties, and as long as we are going to discuss those properties, we ought to use terminology which allows us easily to distinguish them without reusing the same term. Dixon’s suggestion is to use the labels S, A, and O.³ The abbreviation **S** represents the single argument of an intransitive verb, **A** represents the “subject” of a transitive verb, and **O** represents the “direct object” of a transitive verb (in §2.3.2, I describe how I distinguish A from O). S, A, and O are core arguments.⁴

A crucial terminological and theoretical point is this: the labels S, A, and O, were selected by Dixon as labels of convenience. They bear obvious resemblance to the fully worded conceptual notions subject, agent, and object, respectively. But S, A, and O do not **stand for** subject, agent, and object. To put it differently, “S” is not to be interpreted as meaning “subject,” etc. “S” is to be interpreted as representing the single argument of an intransitive verb. “A” is to be interpreted as representing the more subject-like of the two arguments of a transitive verb. “O” is to be interpreted as representing the less subject-like (or more object-like) of the two arguments of a transitive verb, as defined by criteria described in §2.3.2.

In transitive constructions, where there are two arguments, there must be some way of distinguishing between them. Dixon’s method is that the A function is assigned to “that role

³ Some authors prefer ‘P’ to ‘O,’ presumably as it invokes “patient” rather than “object.” I will continue to use ‘O,’ as this frees me up to use ‘SAP’ to mean “speech act participant” without the possibility that ‘SAP’ could be misinterpreted as referring to the argument labels.

⁴ Dixon distinguishes the term *core argument* from *peripheral argument* on the basis of its obligatoriness. Peripheral arguments for him include adverbials and certain adpositional phrases. These may be omitted without affecting grammaticality, though the omission will often affect the completeness of an expression or the naturalness of its use. In this sense, what Dixon calls a peripheral argument is what is otherwise referred to as an adjunct.

which is most likely to be related to the success of the activity” (2010b:128), and the O function goes to whichever one is not most likely. However, it is not clear to me how one determines which argument is “more related to the success of the activity” or even if it is possible to do so. For a two-argument predicate, both arguments are fully required for the activity to be successful, or else there would be no activity. Asking which one is “more related to the success of the activity” seems to me like asking, “Which contributes more to the area of a rectangle: length or width?” The definition of area requires length and width to interact; neither one “contributes more.” Likewise, in a garden-variety transitive clause like *Xiaohe julienned the jicama*, without the jicama, the julienning would be totally unsuccessful—in fact, it would not occur at all—even though *Xiaohe* is in A function here. That *Xiaohe* is A and *the jicama* is O is easily determined by the method proposed by Dowty (1991), discussed next.

2.3.2 Core arguments as proto-roles

The identification of the core argument of an intransitive verb is seldom difficult because the core argument is the only argument. The single argument of an intransitive verb is S. Because transitive verbs have two arguments, there must be some procedure for deciding which argument is designated as A and which as O. Theories of argument structure owe a great deal to the notion of semantic roles (also called “semantic relations”) for advancing mechanisms by which this is achieved. Important early work on this topic was Gruber (1965) and Fillmore (1968). Fillmore proposed a handful (six explicitly, plus others suggestively) of universal semantic cases which are variously mapped onto surface forms. Since then the list of semantic roles has been modified and expanded to at least ten or so generally accepted roles, including AGENT, PATIENT, THEME,

EXPERIENCER, BENEFICIARY, INSTRUMENT, LOCATION, GOAL, SOURCE, and STIMULUS (Saeed 2009). The different roles reflect the particular manner in which an argument participates in the event or state specified by the verb, and what its relationship is to any other arguments.

Deciding which semantic role(s) should be assigned to a given argument can be gnarly, particularly as there is sometimes quite a subtle distinction between certain roles. For instance, the PATIENT role canonically relays a change of state, while the THEME role canonically relays a change in location. Consider the sentence *From his clandestine perch in the treehouse, Franzl dropped the water balloon on his sister*. As anyone with experience throwing water balloons knows, the event described will result in both a change of location and (if things go as Franzl surely hopes) a change of state in the water balloon. Is *water balloon* therefore both a PATIENT and a THEME? And is getting wet a significant enough change of state to warrant calling *his sister* a PATIENT, or is she a GOAL (as the entity which the water balloon moves to), or even a LOCATION (as the place where the balloon exploded), or is she all of these things? It is not entirely clear, and some scholars have worked to propose more explicit tests for determining semantic roles and strategies for assigning them to arguments (Jackendoff 1972, 1983, Cruse 1973, Chomsky 1982).

Dowty (1991) offers a helpful perspective on the difficulties in assigning and interpreting semantic roles, and his solution is more-or-less the idea I adopt moving forward. His suggestion is that the fine-grained individual semantic roles can be collapsed into larger units which he calls “proto-roles.” An argument is either a proto-agent or a proto-patient, and these are the only two role types that any argument ever is. The semantic properties attributable to each proto-role are presented by Dowty as a (preliminary and non-exhaustive) list of entailments of the predicate in

question. These are shown below in (7) and (8) (reproduced from Dowty 1991:572, items (27) and (28)).

- (7) Contributing properties for the Agent Proto-Role:
 - a. volitional involvement in the event or state
 - b. sentience (and/or perception)⁵
 - c. causing an event or change of state in another participant
 - d. movement (relative to the position of another participant)
 - (e. exists independently of the event named by the verb)
- (8) Contributing properties for the Patient Proto-Role:
 - a. undergoes change of state
 - b. incremental theme
 - c. causally affected by another participant
 - d. stationary relative to movement of another participant
 - (e. does not exist independently of the event, or not at all)

The assignment of each role is formalized by the Argument Selection Principle, reproduced below in (9).

- (9) Argument Selection Principle (Dowty 1991:576)

In predicates with grammatical subject and object, the argument for which the predicate entails the greatest number of Proto-Agent properties will be lexicalized as the subject of the predicate; the argument having the greatest number of Proto-Patient entailments will be lexicalized as the direct object.⁶

In the terminology I am employing, Dowty's Proto-Agent is my A and his Proto-Patient is my O. Notice that the Argument Selection Principle is restricted to two-argument predicates. It therefore does not contribute at all to the role assignments of intransitive arguments. And with good reason: intransitive subjects (or what I am calling S) may be either Proto-Agents or Proto-

⁵ As suggested by Dowty and elaborated on by Primus (1999), the *sentience* property includes experiencers.

⁶ Dowty elaborates on the Principle with two corollaries and a nondiscreteness qualification. The first corollary allows for either of two arguments to be lexicalized as the subject when they have an equal number of properties from both proto-roles.

Patients. This, in fact, forms the basis for one of the major types of split-alignment, so-called split-S or split-intransitive systems, discussed in detail in §2.5.5.

Recall the sentence *Xiaohe julienned the jicama* from the previous subsection. That *Xiaohe* is A and *the jicama* is O is easily determined by Dowty’s method: properties (a)-(d) for Proto-Agents are true of *Xiaohe* and properties (a)-(d) for Proto-Patients are true of *the jicama*.

2.3.3 Summary of core arguments

The concept of *core argument* I adopt for this dissertation is principally from Dixon (1979, 1994, 2010b), but amended with insights from Dowty (1991). I am employing Dixon’s abbreviations S, A, and O and retaining his general definitions of these terms. S is for the single argument of an intransitive verb, as in Dixon’s analysis. In order to distinguish between transitive arguments, I define A as the Dowtyian proto-agent argument, and O as the Dowtyian proto-patient argument.

2.4 Alignment patterns

Alignment patterns are configurations of the “sameness” or “differentness” of the three relations S, A, and O. Seven alignment types may be defined on purely theoretical grounds, and it turns out that all of them are attested in actual languages. For reference purposes, each of the seven types has its own name. In this subsection, I define each of the types and show a brief example to illustrate the pattern, drawing from Native American languages where possible.

In the upcoming discussion, it is important to keep in mind that each of the types represents a pattern found in some language; no claim is here made that the pattern permeates the entire grammar of that language. In most cases, the opposite is true: a pattern is restricted to a particular subset of constructions within a language. Therefore, it does not make sense to talk about an “ergative language” unless that is taken to mean “a language with ergative alignment in some constructions, and some other alignment pattern(s) elsewhere,” in which case it would make equally little sense to label the entire language with one of its other alignment types. In terms of the actual distribution of these patterns, it is preferable to speak in terms of ergative (or other pattern names) *constructions*, rather than languages.

2.4.1 Nominative-accusative alignment

Nominative-accusative alignment is the pattern in which A and S are marked the same and O is marked differently.⁷ A convenient shorthand for representing alignment patterns is by means of a math-inspired formula, where ‘=’ means ‘is marked the same as’ and ‘≠’ means ‘is marked differently from.’ A formulaic representation of the nominative-accusative pattern is shown in (10).

- (10) Formulaic representation of the nominative-accusative alignment pattern
 $(A = S) \neq O$

Nominative-accusative alignment is the pattern most familiar from case marking in western Indo-European languages. Non-second-person pronouns in English offer a clear example. Consider, for instance, the third-person feminine pronoun, which is *she* when it is in S or A function but *her* when it is in O function, as in (11).

⁷ What I mean by “marked” will be elaborated on in subsection §2.6.

(11) Nominative-accusative alignment in English

She pinches **her**.

She sleeps.

In the transitive clause, the A pronoun is realized as *she*, and so is the S pronoun in the intransitive clause. The O pronoun is realized as *her*. Since S and A are marked the same, and O is marked differently from either S or A, the pattern is nominative-accusative. The name *nominative-accusative* comes from the canonical case functions in traditional grammar, where *nominative* refers to subjects and *accusative* refers to direct objects. The names derive from grammatical conventions inherited from Latin and Greek, but the pattern itself is found in languages all over the world. As will be made clear in Chapter 4, in some languages this pattern is a relatively minor one or does not exist at all.

2.4.2 Ergative-absolutive alignment

Ergative-absolutive alignment is the pattern in which S and O are marked the same and A is marked differently. The etymology is not entirely transparent, but a common interpretation is that *ergative* comes from the Greek *ergátes* ‘worker.’⁸ The A argument takes the label *ergative* while the S and O arguments take *absolutive*. A formulaic representation is shown in (12).

(12) Formulaic representation of the ergative-absolutive alignment pattern

A ≠ (S = O)

Ergative-absolutive alignment, like the nominative-accusative type, is found in all sorts of languages. A fine example of the pattern comes from Coast Tsimshian (Sm’algyax), spoken on the northern coast of British Columbia, in (13)-(15) below.

⁸ This assumed etymology may actually stem from a misinterpretation of scholarship from the early twentieth century on the Eastern Trans-Fly language Miriam Mir, where the term was used to denote spatial relations and was possibly derived from Latin *ergā* ‘against, near’ (see Butt 2006:154-158 for a summary).

- (13) Absolutive **S** in Coast Tsimshian (Mulder 1994:51, ex. 54)⁹

la wila diduuls-**u**
PAST be alive-**1SG.ABS**
'I am still alive'

- (14) Absolutive **O** in Coast Tsimshian (Mulder 1994:52, ex. 57)

ada wil m way-**u**
and then 2SG.ERG find-**1SG.ABS**
'and then you found me'

- (15) Ergative **A** in Coast Tsimshian (Mulder 1994:51, ex. 53)

n=dm man-gad-n
1SG.ERG=FUT up-take-2SG.ABS
'I will take you up'

The clause in (13) contains a morpheme, *-u*, indicating the first-person singular S argument. This same morpheme is present in (14), where it indicates the first-person singular O argument. Therefore S and O are marked the same in this language, at least for the first-person singular marker (I'm using ambiguous language for now, in calling it a "morpheme" and a "marker," but I'll distinguish between pronouns and pronominals in §2.6). In the gloss, S and O both get the abbreviation ABS for *absolutive*. The A argument is indicated differently, as the proclitic *n=*, and it is glossed as ERG for *ergative*. Since S and O are marked the same, and A is marked differently from either S or O, the pattern is ergative-absolutive.

By itself, ergativity is the topic of a great many analyses from different theoretical perspectives. Accounting for ergative constructions, and in particular the co-existence within a single language of both ergative-absolutive and nominative-accusative construction types, has been a rich source of scholarship in theoretical syntax (see among others Johns 1992, Bittner & Hale 1996, Manning 1996). Coon's (2013) formal analysis of split-ergativity in Chol is reviewed

⁹ The interlinear glosses of the Coast Tsimshian constructions are courtesy of Mithun (1999:209-10), who cites earlier work by Mulder.

briefly in §5.4.4. Optimality Theory (OT) is another promising framework in which split-ergativity has been (and is being) explored. OT is uniquely conceived to handle issues of markedness, cross-linguistic variation, and interlinguistic variation—all of which are firmly embedded in the empirical facts surrounding ergativity. In §5.5, I give a brief overview of Aissen’s (2001, 2003) OT-based interpretation of split-alignment. She appeals to harmonic alignment of prominence scales and local constraint conjunction in order to derive grammatical contexts in which various argument-marking patterns may occur in a language.

2.4.3 Double-oblique alignment

Double-oblique alignment is the mirror-image of ergative-absolutive. In a double-oblique pattern, A and O are marked the same while S is the odd one out. This pattern may strike the reader as odd and probably rare; s/he would be right. It is probably the least-discussed of the possible alignment types and the literature on attested examples is thin. The reasons for this are theoretically sound: it doesn’t do much good if the construction doesn’t distinguish between the two arguments in a transitive clause, but still bothers to distinguish a single intransitive argument which never co-occurs with the other two. Nominative-accusative and ergative-absolutive patterns both distinguish the two transitive arguments, and they re-use one of the marking strategies for S; they only differ in which strategy they re-use (nominative-accusative patterns re-use the A marker for S while ergative-absolutive patterns re-use the O marker for S). Presumably, re-using a marker relieves pressure on the cognitive and memory systems of speakers.¹⁰ But if

¹⁰ Not that patterns, or languages for that matter, are obligated to go easy on cognitive and memory systems. There certainly are plenty of instances where they apparently do not.

the pattern does not distinguish between A and O, that could potentially lead to ambiguity and confusion. A formulaic representation is shown in (16).

- (16) Formulaic representation of the double-oblique alignment pattern
(A = O) ≠ S

Two comments on this pattern are in order. First, it is usually not correct to say that a clause with a double-oblique pattern does not distinguish between A and O. In double-oblique alignment, the *morphological* marking of A and O may be identical, but there will probably be a consistent constituent order (or at least a constituent order that is pragmatically licensed in a given context) which makes the interpretation of the clause obvious. I'll talk more about constituent order, and the problems with using it as a metric for determining alignment type, below in §2.6.1.3. There may also be other aspects of the morphology, such as agreement, which could serve to disambiguate. Second, it is important to remember that there are no double-oblique languages, only double-oblique constructions. The constructions are restricted to particular grammatical conditions and do not impose that alignment pattern onto constructions in other parts of the grammar.

The now classic example of double-oblique marking is from Rošani, an Iranian language from Tajikistan and Afghanistan. First- and second-person pronouns, demonstratives, and human interrogatives are inflected according to the case roles of the nouns they specify. In the past tense only, those marking A-function and O-function nouns are taken from the same conjugational set; those for S-function nouns are taken from a different set. The examples in (17)-(19) using first-person pronouns will serve to demonstrate.

- (17) S pronoun from the “absolute” conjugational set in Rošani (Payne 1980:156, ex. 13c)
az-um tar x̄ār viĵ
 I-1SG to town be.PERF
 ‘I’ve been to town’
- (18) A pronoun from the “oblique” conjugational set in Rošani (Payne 1980:156, ex. 13a)
mu t̄ā wunt
 I you see.PAST
 ‘I saw you’
- (19) O pronoun from the “oblique” conjugational set in Rošani (Payne 1980:156, ex. 13b)
 t̄ā **mu** wunt
 you **me** see.PAST
 ‘You saw me’

In (17), the first-person pronoun in S function is taken from what Payne (1980) calls the “absolute” set of pronouns, and is realized as *az*. The first-person pronoun in A function in (18) and the one in O function in (19) both come from the “oblique” set of pronouns and are realized as *mu*. Since both A and O come from the oblique set, the name Payne gives to this alignment pattern is “double oblique.” As noted above, this does not mean that there is no way to distinguish A from O. The constituent order in the Rošani past tense is consistently AOV, so there is never any ambiguity. Still, the morphologies of the pronouns, taken alone, have ambiguous forms.

2.4.4 Tripartite alignment

Tripartite alignment is the pattern in which S, A, and O are all marked differently from one another. A formulaic representation is shown in (20).

- (20) Formulaic representation of the tripartite alignment pattern
 $A \neq S \neq O$

Nez Perce, a Sahaptian language from northern Idaho, has tripartite alignment on third person arguments. Those in A function receive a phonologically conditioned variant of the suffix *-nim*, those in O function are marked with a direct object suffix *-ne* or *-e*, and those in S function are zero-marked. The examples in (21) and (22) show this pattern.

- (21) Zero-marked third-person S in Nez Perce (Rude 1986:126, ex. 3)

hi-páay-na **háama**
 3.NOM-arrive-ASP **man**
 ‘The man arrived.’

- (22) Differentially marked A and O in Nez Perce¹¹ (Rude 1986:126, ex. 6)

háama-nm péé-’wi-ye **wewúkiye-ne**
man-ERG 3TR-shoot-ASP **elk-DO**
 ‘The man shot the elk.’

It is worth mentioning that Rude (1986) does not use the term *tripartite* to refer to this alignment pattern. Rather, she interprets the construction in (22) as ergative (hence ERG in the interlinear gloss), but “atypically ergative” because the direct object is overtly realized and therefore there is no absolutive, under her assumption that absolutives are necessarily zero-marked (ibid:124). This pattern does not count as ergative under my definitions because alignment is contingent on all three roles A, S, and O; it is not enough that there is some special way of indicating A, there must also be a distinct way of indicating S and O. The S- and O-marking strategy need not be via zeros, it just must be identical for both S and O yet distinct from A. The only way to construe the Nez Perce alignment pattern in (21) and (22) according to my definitions is as tripartite because A, S, and O are all marked differently (see also Creissels 2009:454).

¹¹ Unfortunately, Rude (1986) does not provide an example with *háama* ‘man’ in O function, which would round out the examples nicely. Nonetheless, the direct object marker on *wewúkiye* ‘elk’ serves the purpose of showing tripartite marking.

2.4.5 Neutral alignment

Neutral alignment is the pattern in which S, A, and O are all marked alike. Neutral alignment often shows up where a particular grammatical person/number category is morphologically zero; this will be discussed at length in Chapter 5. A formulaic representation is shown in (23).

- (23) Formulaic representation of the neutral alignment pattern
A = S = O

Alignment is neutral for third person pronominals in Tlingit, a language isolate from southeastern Alaska, provided that the A and O are not both definite (alignment in Tlingit will be described in detail in §4.3.6). The examples in (24)-(26) show the zero marker in the expected position for overt person morphemes.

- (24) Zero-marked third-person **S** in Tlingit (Crippen 2012:323, ex. 302b)
ȳu-**Ø**-ȳa-gut-h
PFV-**3.SUB**-CL-go.SG-VAR
'he went'
- (25) Zero-marked third-person **A** in Tlingit (Crippen 2012:324, ex. 305a)
ḡad-ȳu-**Ø**-si-ku-ȳ
1SG.OBJ-PFV-**3.SUB**-CL-know-VAR
'he knows me'
- (26) Zero-marked third-person **O** in Tlingit (Crippen 2012:324, ex. 305b)
Ø-ȳu-ḡa-si-ku-ȳ
3.OBJ-PFV-1SG.SUB-CL-know-VAR
'I know him'

The alignment pattern shown above is neutral because S, A, and O are indicated identically by zero-marking.

2.4.6 Hierarchical alignment

Hierarchical alignment is a pattern which differs substantially from those discussed already. In transitive clauses showing this pattern, there is some means by which the relationship of A to O is expressed which refers to the relative position of the arguments along a grammaticalized hierarchy.

The direct-inverse pattern is a well known version of hierarchical alignment. In a direct-inverse pattern, there is a morpheme indicating the relevant categories (person, number, animacy, etc.) of the A and O arguments, but neither those morphemes themselves nor their position within the clause, relays which one is A and which one is O. There is an additional morpheme which encodes the directionality of action between the arguments (i.e. which is acting on which), thereby specifying the roles of the two arguments. The formulaic representation I have shown for the other patterns is not appropriate for direct-inverse alignment because the strategy of marking A or O is dependent on the relative ranking of A and O with respect to each other (discussed below). In other words, a short-hand expression like $A = O$ or $A \neq O$ simply does not capture the nature of the direct-inverse pattern.

Directionality of action is determined according to a hierarchy of grammatically relevant argument categories. There is a bewildering list of labels which have been given to such hierarchies since the 1970s. I am following Bickel & Nichols in calling it an *indexability hierarchy* because this term captures that “a referent can be identified – or ‘indexed’ – from within the speech-act situation” (2007:224). The indexability hierarchy ranks grammatical categories and semantic features such as speech act participant (SAP), pronoun, noun, [\pm animate], [\pm human], and the list goes on. Not all of the categories are relevant for every

language with phenomena sensitive to such a hierarchy. While there are strong cross-linguistic tendencies for the ranking of categories, there are also language-specific rankings, so an overall ranking cannot be thought of as universally valid. When any A acts on a lower ranking O, then the morpheme which encodes directionality is termed *direct*. When any A acts on a higher ranking O, then that morpheme is termed *inverse*.

Direct-inverse marking thus requires two distinct but interacting morphological components in order to encode argument relations. The first component is the morphemes for the relevant grammatical categories of the A and the O, which does not say which is which. The second is the directionality marker, which specifies whether the higher ranked argument is A or O.

A clean example of the direct-inverse pattern comes from Plains Cree, an Algonquian language from southern central Canada. In Plains Cree, first person outranks third person. The first-person morpheme, *ni-*, is a verbal prefix. The third-person morpheme, *-w*, is a suffix. As the positions of these morphemes are invariant regardless of argument role, only the directionality marker indicates which is A and which is O, as in (27) and (28).

- (27) Direct marking in Plains Cree (Zúñiga 2006:24, ex. 12a)

ni-wāpam-ā-w
1-see-DIR-3
'I see him/her'

- (28) Inverse marking in Plains Cree (Zúñiga 2006:24, ex. 12b)

ni-wāpam-ikw-w
1-see-INV-3
's/he sees me'

In (27), the direct marker is used because the first-person argument is A, the third-person argument is O, and first person outranks third person. In (28), the inverse marker is used because

the roles are reversed: the lower ranking third-person argument is A and the higher ranking first-person argument is O.

The direct and inverse markers are only used in transitive clauses because intransitive clauses don't need to distinguish between two arguments. Therefore the marking of S will necessarily be distinct from the marking of both A and O in that there will never be a directionality marker in intransitive clauses. The morpheme representing the S may, however, be identical to the morpheme representing the person/number/animacy categories of either A or O argument. In the Plains Cree examples above, *ni-* represents a first-person argument in both (27) and (28), and it does the same in the intransitive construction below in (29).

(29) Intransitive construction in Plains Cree (Zúñiga 2006:75, ex.45b)

nit¹²-api-n
1-sit-SAPSG
'I sit'

Other versions of hierarchical alignment may use a specific morpheme to indicate directionality of action, but do not require the person/number information of the arguments to be indicated as well.

2.4.7 Portmanteau alignment

What I am calling portmanteau alignment is a pattern in which both A and O, and their directionality of action, are conflated into a single morpheme. Laguna Keres, a Keresan language from New Mexico, employs a portmanteau alignment pattern when both A and O are speech act participants. There are two morphemes: *sra-* encodes a first-person A with a second-person O;

¹² *Nit* is a phonologically conditioned allomorph of the first-person morpheme.

dy- encodes a second-person A with a first-person O. The examples in (30) and (31) show the pattern. The double-headed arrow ‘ \Rightarrow ’ means “acts upon.”

- (30) First-person A and second-person O in Laguna Keres (Lachler 2006:146, ex. 5g)

sra-ukacha

1 \Rightarrow **2**-see

‘I see you.’

- (31) Second-person A and first-person O in Laguna Keres (Lachler 2006:146, ex. 5h)

dy-ukacha

2 \Rightarrow **1**-see

‘You see me.’

The portmanteau pattern is similar to direct-inverse because the notion directionality is involved, but it is not the same because the A and O arguments are not differentially marked and there is no designated directionality marker. Alignment in Laguna Keres will be discussed at length in §4.2.7.

2.4.8 Summary of alignment patterns

The seven types of alignment are summarized in list form below, referring where possible to their formulaic representations as a convenient short-hand.

1. Nominative-accusative: $(A = S) \neq O$
2. Ergative-absolutive: $A \neq (S = O)$
3. Double-oblique: $(A = O) \neq S$
4. Tripartite: $A \neq S \neq O$
5. Neutral: $A = S = O$
6. Hierarchical: A and O are distinguished by means of a directionality morpheme which indicates the direction of action
7. Portmanteau: A single morpheme encodes both A and O

There are other alignment patterns besides these, particularly involving instances where portmanteau morphemes turn out to be identical to those used in certain intransitive constructions. These will be discussed as they come up in Chapter 4.

2.5 Split-alignment and split inducers

Most (maybe all) languages do not have the same alignment pattern implemented in every construction type. One alignment pattern occurs under certain grammatical conditions, while at least one other pattern is used under other conditions. Said differently, the alignment system as a whole in a particular language is split between different patterns on the basis of certain grammatical conditions. What those conditions can be, and how we can explain their distribution in a subset of human languages, is the topic of this dissertation. In this subsection, I give a few examples of alignment splits occurring under different conditions (§2.5.1-3) and introduce my use of the term “split inducer.” The main objective of this subsection is to give the reader a clearer picture of the kind of phenomenon I am focusing on and a small taste of the different forms it may take. As all of Chapter 4 is devoted to different instances split-alignment in my sample, this section will be brief and non-exhaustive.

Any grammatical condition which coincides with the implementation of a particular alignment pattern in a single language is a *split inducer*. In principle, any grammatical condition could be a split inducer. Apparently, not just any grammatical condition actually is a split inducer, which is interesting because it suggests there may be constraints on the kinds of grammatical categories which can interact with alignment.

The term *split inducer* imparts certain implications about the nature of causation within split-alignment systems. The main implication is that if a given grammatical condition is called a split inducer, then that grammatical condition is actually affecting change; it is responsible for causing the split. It is unclear whether this is always indeed true. It may alternatively be the case that a given grammatical condition just happens to co-occur with an alignment split, perhaps by historical accident or because that condition is dependent of some cognitive (or other) principle which is truly responsible for inducing the split. My use of the term *split inducer* does not distinguish between these possibilities.

From the outset, before it is known whether a given grammatical condition is causative or coincidental of split-alignment, it is desirable not only that I admit the possibility that *all* grammatical conditions associated with split-alignment are indeed split inducers, but also that I assume, as a methodological decision, that they are. The reason is that it is easier to whittle down the set of grammatical conditions post-hoc than it is to throw new conditions onto the pile as they are proved as split inducers. The only way to know in advance is to launch a thorough investigation of each possible split inducer in terms of its usage, history, and context within each language. Such investigations are probably dissertation-worthy in their own right. In this study, I admittedly sacrifice a certain degree of single-language depth for cross-linguistic breadth. The advantage, though, is that the outcome of my study is a set of testable notions which can (and should) be the objects of future research.

Of the known split inducers, some of them are quite robustly attested across languages and/or are oft-discussed in the relevant literature. Others of them do not feature very strongly in general discussions of split-alignment phenomena. In §2.5.1 and §2.5.2, I describe in superficial

detail two of the better known split inducers, respectively person splits and TAM splits. In §2.5.3, I describe object definiteness as a comparatively minor split inducer. In §2.5.5, I have a somewhat fuller exposition of split-S systems, as these are particularly well represented in the alignment literature and deserve special mention.

2.5.1 Split inducer 1 — Person splits

An easy-to-find example of split-alignment is one induced by the grammatical person category. Sometimes such splits are induced by the status of the argument as a speech act participant (SAP) or not. In these cases, all SAPs have one alignment pattern and all non-SAPs have a different pattern. Other times, the split inducer is even more narrow, down to the specific person, where first, second, and third persons all show different alignment. An instance of this latter type comes from Seri, an isolate spoken on and near Isla Tiburón in northwest Mexico. Ignoring for now the plural forms, consider the agreement paradigm in (32).

- (32) Partial paradigm of Seri pronominals (constructed from Marlett 1990:514, 521 & 2014:607)

	<u>A</u>	<u>S</u>	<u>O</u>	
1SG	h-	hp-	him-	
2SG	m-	m-	ma-	
3SG	Ø	Ø	Ø	A⇒O i-

In Seri, singular first-person arguments show tripartite marking because they are marked differently in S, A, and O functions. Singular second-person arguments show nominative-accusative alignment because S and A are the same while O is different. In the third-person, alignment is neutral because none of them is overtly marked. An exception to the neutral pattern occurs when both A and O are in the third person, in which case there is portmanteau alignment

via the single marker *i-*. So here, on the basis of person alone, there are four different alignment patterns.

2.5.2 Split inducer 2 — TAM splits

One of the commonly described split inducers is the megacategory TAM, an abbreviation for *tense, aspect, modality/mood*. Really there are multiple potential split inducers in this category, as any one language can use, say, only aspect to split the alignment system while still distinguishing tense from aspect from modality from mood (or any confluences and combinations thereof). An aspect-induced split occurs in the Iroquoian language Cherokee. For this example, I focus only on the third-person forms (Cherokee split-alignment is discussed in detail in Chapter 4). In four of Cherokee’s eight aspects, alignment is split-S, with one set of markers indicating A and S_a (termed Set A) and another set indicating S_o and O (termed Set B).¹³ (33) shows the third-person markers for the two sets.

(33) Third-person Set A and Set B markers (Scancarelli 1987:55)

	<u>Set A</u>	<u>Set B</u>
3SG	a-, ka-	u:-

An intransitive verb stem such as *aliski:* ‘dance’ is treated as agentive, while the intransitive verb stem *yert* ‘laugh’ is non-agentive. So typically, the third-person marker on ‘dance’ will come from Set A and the third-person marker on ‘laugh’ will come from Set B, as in (34) and (35), which are in the imperfective aspect.

¹³ S_a indicates the S role which patterns (morphologically and usually semantically, as well) with A; S_o patterns with O. These “split-intransitive” systems are described further in §2.5.5.1.

- (34) Set A in the imperfective (Scancarelli 1987:65, ex. 16b.i)

a-aliski: = sko:ʔi

3SG.A-dance=IMPFV

‘he’s always dancing’

- (35) Set B in the imperfective (Scancarelli 1987:65, ex. 16b.ii)

u:-ye:t = sko:ʔi

3SG.B-laugh=IMPFV

‘he’s always laughing’

In other aspects, however, all arguments get the Set B marker. (36) and (37) are examples of the same verbs shown above but in the perfective aspect.

- (36) Set B in the perfective (Scancarelli 1987:65, ex. 16d.i)

u:-aliski: = sv:ʔi

3SG.B-dance=PERF

‘he (has) danced’

- (37) Set B in the perfective (Scancarelli 1987:65, ex. 16d.ii)

u:-ye:t = sv:ʔi

3SG.B-laugh=PERF

‘he (has) laughed’

So in Cherokee, aspect is a split inducer. It turns out that aspect-based splits are not as common in my sample as one might expect from a review of the literature. But I leave it until Chapter 5 to discuss this fact.

2.5.3 Split inducer 3 — Object definiteness

An example of a split inducer which, as far as I can tell, has not yet been so widely recognized is object definiteness. I include it here as an instance of how different from one another split inducers can be. In the Alaskan isolate Tlingit, discussed fully in Chapter 4, there is

an ergative morpheme which appears on A provided O is definite, as in (38). But if O is indefinite rather than definite, the ergative marker on A is ungrammatical, as in (39).

(38) Ergative marking on A (Crippen 2012:33, ex. 5a, modifying Story & Naish 1973:183)

ak éesh-ǂ útl̥xi ʔa-Ø-Ø-sa-.i-:
1SG.POSS father-ERG soup 3.OBJ-ZCNJ-3.SUB-CL-cook-VAR
'my father is cooking soup'

(39) No ergative marking on A when O is indefinite (Crippen 2012:35, ex. 9)

ax éesh (*éesh-ǂ) t'á ʔa-wsit'éx
1SG.POSS father (*father-ERG) king.salmon 3.OBJ-PFV-3.SUB-CL-fish
'my father was fishing for king salmon'

In (38), where O is definite, A gets the ergative marker. The definiteness of O is contextually determined; there is a specific soup that father is cooking. In (39), where O is indefinite, there is no ergative marker on O. The indefiniteness of O in this case comes from the fact that father was not fishing for a particular king salmon; any king salmon would do. (The Tlingit alignment split shown above is not, incidentally, induced by a tense change; the object definiteness split holds across tenses.)

2.5.4 Combinatorial split induction

It is often the case that split induction is achieved only when grammatical conditions operate in concert. For instance, it might be the case that a particular alignment pattern only occurs on first-person arguments, and then only in the perfective aspect. In such an instance, both conditions are considered to be split inducers. So, ignoring any other alignment splits, the split inducers would be (a) first-person and (b) perfective aspect, rather than a single split inducer composed of both conditions, first-person+perfective aspect. In this sense, a split inducer is not

to be interpreted necessarily as the *only* condition relevant to a given split-alignment event. Each split inducer is a condition relevant to splitting the alignment system in a given language.

Of course, the observation that some split inducers occur in combination with others should be ignored. It has the potential to reveal interesting dependencies and relationships between split-inducing categories. But for counting purposes, each distinct category which contributes to split-alignment is treated as a split inducer. The reason is that one of the objectives of this study is to determine what kinds of categories are involved in split induction in the first place, and so compiling an itemized list is the first step toward this. It must be noted, of course, when certain split inducers occur (only) in combination with others. Such issues will be discussed in Chapter 5.

2.5.5 Split-S alignment — a special case

Split-S alignment, also known as split-intransitive, is any pattern in which there is more than one way of marking S. The split-S pattern is one of the major types of split-alignment discussed in the literature, and some scholars treat it as a full-fledged alignment type, on par with nominative-accusative, ergative-absolutive, etc. In §2.5.5.3, I will defend my treatment of split-S marking as a type of split-alignment, where the split is conditioned by verbal semantics or lexical specification. This has consequences for the analysis, as well as general theoretical repercussions. But first, I'll use §2.5.5.1-2 to give an overview of this pattern and highlight some of the different forms it can take.

2.5.5.1 Prototypical split-S alignment

Prototypical split-S alignment involves S being marked like A in some situations, but like O in other situations. For notational purposes, a subscripted ‘a’ or ‘o’ may be placed on S in order to identify how it aligns; though not all authors follow this convention, I find it to be quite useful. A basic formulaic representation is shown in (40).

(40) Formulaic representation of the prototypical split-S alignment pattern

$$(A = S_a) \neq (S_o = O)$$

Whether S is treated like A or like O in a particular instance usually has some semantic motivation.¹⁴ It is often the case that when S is agentive or active, it patterns with A, but when it is patientive or stative, it patterns with O. Differences across languages and theoretical orientations regarding the specific semantic categories relevant for grouping S with A or O have spurred a handful of label names falling under the split-S umbrella, including *agent-patient*, *active-stative*, *active-inactive*, *active*, and *agentive*, among others. Some authors treat these more-or-less as synonyms, while others insist on their distinctiveness. I will not go into these various types further because it will not enlighten my findings or claims to take a stance on the nomenclature. However, in characterizing the motivation for splits in some languages, it will be helpful to appeal to certain of these terminological differences on a language-particular basis. I’ll do this as needed in Chapter 4.

A split-S pattern occurs with first-person singular and second-person arguments in Lakota, an Athabaskan language from South Dakota, as in (41)-(44).

¹⁴ For this reason, Wichmann (2008) argues forcefully for the label “semantic alignment” as opposed to all others.

- (41) **A** in Lakota (Dahlstrom 1983:42, ex. 10d)
 wičha-**wa**-gnayã-pi
 AN.3PL.ACC-**1SG.AGT**-trick-PL
 ‘I tricked them’
- (42) **S_a** in Lakota (Dahlstrom 1983:41, ex. 9a)
wa-lowã
1SG.AGT-sing
 ‘I sing’
- (43) **O** in Lakota (Dahlstrom 1983:41, ex. 9c)
ma-ya-gnayã-pi
1SG.PAT-2AGT-trick-PL
 ‘you (pl.) tricked me’
- (44) **S_o** in Lakota (Dahlstrom 1983:41, ex. 9b)
ma-hãska
1SG.PAT-be.tall
 ‘I am tall’

The marker for the first-person singular A in (41) is *-wa-*, just as it is for S_a in (42). The marker for O in (43) is *ma-*, just as it is for S_o in (44).

2.5.5.2 Other kinds of split-S patterns

A sub-type of split-S alignment is called *fluid-S* or *fluid-intransitive*. In this pattern, certain intransitive verbs have Ss which are not *consistently* S_a or S_o. Rather, the marking of S like A or O is flexible, depending on the particular semantics of the event at hand. In Chickasaw, a Muskogean language originally from the southeastern United States, the first-person A marker is *-li* and the first-person O marker is *sa-*, as shown in (45) and (46).

- (45) First-person **A** in Chickasaw (Munro & Gordon 1982:81, ex. 1a)
 kisili-**li**
 bite-**1SG.A**
 ‘I bite him.’

- (46) First-person **O** in Chickasaw (Munro & Gordon 1982:81, ex. 1b)
sa-kisili
1SG.O-bite
 ‘He bites me.’

Certain intransitive arguments, rather than being marked like A or O on the basis of a verb’s lexicalized agency or activeness, are marked like A or O on the basis of their semantics in a given event. This distinction is shown in (47) and (48), which show different marking of S for the same verb.

- (47) First-person O-like **S** in Chickasaw (Munro & Gordon 1982:81, ex. 2b)
sa-hotolhko
1SG.S_O-cough
 ‘I coughed.’

- (48) First-person A-like **S** in Chickasaw (Munro & Gordon 1982:81, ex. 3b)
 hotolhko-**li**
 cough-**1SG.S_A**
 ‘I coughed (on purpose).’

There are still other kinds of split-S patterns in which the ways of marking the S argument correspond neither to A nor O. An example of this type comes from San Francisco del Mar Huave, a Huavean language from the Laguna Superior region of southeastern Mexico. There are two ways of marking S in this language, based on whether the intransitive root takes a required grammatical affix as a prefix or a suffix. As these two intransitive verb classes are distinguished formally by their affix requirements, the semantic content of their verbs is largely motley.¹⁵ From the perspective of alignment, prefixing S stems and suffixing S stems often receive different marking from each other, according to the aspect that the verb is conjugated for.

¹⁵ Nonetheless, Kim (2008) identifies certain semantic tendencies. The prefixing intransitives tend to be bodily functions, verbs of being, changes of state, and motions. The suffixing intransitives tend to be body postures and inchoatives. Still, there is significant overlap since emotional states are found in both classes and some verbs’ semantic categories are not entirely clear (e.g. ‘die’, ‘burn’, ‘be born’, ‘quiet down’, all of which Kim labels as “other”). She (2008:247) also notes the different classes for ‘ascend’ and ‘rise’.

The marking of each S is sometimes different from, and sometimes identical to, that of A. It is always different from O. So San Francisco del Mar Huave is an example of a language with split-S alignment where the formulaic representation $(A = S_a) \neq (S_o = O)$ is an inappropriate characterization.

2.5.5.3 Status of split-S alignment

The split-S alignment pattern is different in a significant way from the other alignment types discussed so far in that it seems to be inherently split. Nominative-accusative, ergative-absolutive, tripartite, and neutral alignment types are cohesive patterns in and of themselves. An alignment system may be split *between* them, e.g. one alignment pattern is used under certain conditions while a different pattern is used under other conditions. But the way that split-S alignment is often characterized builds “splittedness” into the description of the pattern. It is possible, however, and in my view preferable, to view split-S alignment as decomposable into the nominative-accusative and ergative-absolutive patterns.

Taking the Lakhota constructions as an example, the A in (41) and the S in (42) are marked identically to each other, and differently from how O is marked in (43). This is a nominative-accusative pattern, formulaically $(A = S) \neq O$. This pattern occurs when S is agentive (according to the semantic principles which determine agentivity in Lakhota; see Mithun 1991). The O in (43) and the S in (44) are marked identically to each other, and differently from how A is marked in (41). This is an ergative-absolutive pattern, formulaically $A \neq (S = O)$. This pattern occurs when S is non-agentive (or patientive).

An analysis along these lines treats split-S marking not as an alignment pattern all by itself, but as an alignment *system* which is split between two uniform (not split) alignment patterns, nominative-accusative and ergative-absolutive, on the basis of the agentivity of S. I am aware that this is a controversial position to take as several scholars have expressly argued against this treatment of split-S systems (Wichmann 2008 and others). Their argument is that by treating split-S alignment as decomposable, the analyst is assuming that the category S is a natural unit in these languages; in other words, s/he is making the implicit (and in their opinion incorrect) assumption is that there really is a unitary concept “S” in these languages in the first place.

My justification for treating it as decomposable is based on the utility of S as a label for a specific argument of a verb *with a particular valency*, rather than as a representation for any kind of unified category in a given language. The same is true for A and O arguments: they need not be interpreted as unified categories in order to refer to them when identifying alignment patterns. Precedence for this way of thinking comes from what Haspelmath (2011) calls the “Bickelian approach,”¹⁶ assumed by Bickel (2011) and Bickel & Nichols (2009). As Bickel (ibid:413, fn. 9) puts it:

Split intransitivity is sometimes taken to challenge the universality of the notion ‘S.’ But S is defined here purely by numerical valence, as an argument licensed by an intransitive predicate; and in all languages with split intransitivity that I am aware of, intransitive verbs behave differently from transitive verbs in at least some morphological or syntactic effects, minimally with regard to the number of syntactic argument positions they license.

Furthermore, most (or at least a great many) languages have some constructions in which S, A, and/or O is atypically marked, and so the notion of ‘unified concept’ as it pertains to any

¹⁶ As opposed to the Comrian and Dixonian approaches.

one of these labels must be sought on grounds other than valency if treating them as unified concepts is indeed one's objective.

In order to facilitate ease of reference to the split-S patterns, while at the same time noting their decomposability, I will use the abbreviations 'nom-acc' or 'erg-abs' subscripted onto 'split-S' when discussing them in Chapters 4 and 5.

2.5.6 Summary of split-alignment

In §2.5, I have given several examples of split-alignment. The sameness or differentness configurations of A, S, and O are used to define the alignment pattern itself; where the configuration differs under certain grammatical conditions, the alignment system is said to be "split." The grammatical conditions which induce the split are referred to as "split inducers."

2.6 Analytical issues and decisions

2.6.1 Locus of alignment realization

Up to this point, I have been using terms like "marked the same as" or "indicated differently from" in order to describe the various alignment patterns, without explaining what marking is. "Marking of alignment" means whatever morphosyntactic device(s) is/are used to specify the A, S, or O status of an argument. There are basically four ways in which this is done: case marking, pronouns and pronominals, agreement, and constituent order. These four types are not necessarily discrete; in some languages, it may be difficult to separate them or tell them apart if they can be separated. Even theoretically, it is sometimes not clear exactly what distinguishes certain of these, as their functions may overlap or interact in complex ways in certain situations

(Siewierska & Bakker 2009). They may also be used in combination in a given construction. But discriminating precisely among the various morphosyntactic encoding mechanisms will not advance my objectives in this chapter. Instead, my aim in this subsection is to show, of all the construction types in a language, where I look for the alignment patterns; in other words, to summarize the various loci of alignment realization.

In this subsection, I briefly characterize the four types as they relate to alignment marking, covering case, pronouns and pronominals, and agreement in §2.6.1.1-2. I argue in §2.6.1.3 against using constituent order to determine alignment patterns in this study.

2.6.1.1 Case marking

One way in which grammatical relations are signaled is by what are traditionally called *case markers*. *Case*, as used here, refers to the inflection of nouns and their associates¹⁷ on the basis of their function in a clause. The identification of grammatical relations is only one of the duties case marking may serve, but for the remainder of this subsection, I'll restrict the discussion to case as it relates to alignment. An example of case marking showing alignment comes from Jensen & Johns' (1989) analysis of Baker Lake Inuktitut, an Eskimo-Aleut language from Canada.¹⁸ In (49) ergative case is marked by a nominal suffix and absolutive case in both (49) and (50) is not expressed.

¹⁷ By *associates*, I mean any of the various constituents which may be dependent on, referential with, or modifying of nouns. This may include, for instance, determiners, demonstratives, adjectives, classifiers, pronouns, and the like. Which associates a given noun has, and which ones are inflected for case, varies by language.

¹⁸ I am calling it *Baker Lake Inuktitut* because the authors say in a footnote that their data came from "a native speaker of Inuktitut from Baker Lake, Northwest Territories, Canada" (Jensen & Johns 1989:226), but in their paper, they refer to the language as *Eskimo*, which is really a (sub)family of languages.

- (49) Case marking in Baker Lake Inuktitut (Jensen & Johns 1989:219, ex. 21a)

arna-**up** angut kunik-paa
woman-ERG man.ABS kiss-3SG/3SG
'The woman kissed the man.'

- (50) Case marking in Baker Lake Inuktitut (Jensen & Johns 1989:210, ex. 1)

arnaq ani-vuq
woman.ABS go.out-3SG
'The woman is going out.'

Case marking itself may be realized in a variety of ways. A common strategy is affixation, where a particular affix, with a particular case designation (or set of designations), attaches to the noun and/or its associates.¹⁹ Case markers on nouns reveal an alignment pattern when they are identical or different in the various roles S, A, and O. The especially common case markers in this context are the ones glossed as nominative, ergative, accusative, absolutive, and dative.

There is a very strong cross-linguistic tendency for nominative and absolutive cases to be zero-marked (Handsuh 2014), though as usual there are counterexamples. This zero marking, when translated into a text medium, can raise certain analytical questions. Some authors notate zero-marked cases (and other zero-marked categories) with the symbol for the null set, ' Ø '. Others may omit the case in the gloss, or simply gloss the zero-marked case into the noun itself using a period or some similar convention in order to signal that "a single object-language element is rendered by several metalanguage elements (words or abbreviations)" (Comrie et al. 2005). An interesting case in point comes from two instances of the same construction in Straits Salish, a Salishan language from southern Vancouver Island. Evans (1995) cites the construction

¹⁹ Suffixation is the dominate strategy cross-linguistically, but it can also be done by pre- or circumfixation, cliticization, adpositions, consonant mutation, tone changes, or stem changes (Spencer 2009, Siewierska & Bakker 2009).

from a conference talk by Jelinek (1989), as in (51). Jelinek’s own (1995) version differs in orthography and glossing, shown in (52).²⁰

(51) No glossing of absolutive (Evans 1995:209, ex. 16)

mək’^w=ł w’ na-t tsə sčenx^w

all=1PLU LINK eat-TR DET be.fish

‘We at all the fish/We all ate the fish/We all at all the fish/We ate the fish up completely.’

(52) Glossing of zero-marked absolutive (Jelinek 1995:514, ex. 65)

mək’^w=ł ’əw’ ŋa-t-Ø cə sčeenəx^w

ALL=1PNOM LINK eat-TR-3ABS DET fish

‘We ate all the fish. Or: We all ate the fish. Or: We ate the fish up completely.’

The theoretical issue revolves around the status of zero-marked forms in mental grammar and the extent to which that status is accurately represented in the gloss. Is the absolutive case in Straits Salish an actual morpheme with zero phonetic realization, as implied by the use of the Ø? Or is it a case category that really does not have a morpheme (zero or otherwise), but whose function can be represented as “built in” to the noun? Since authors do not usually state whether they believe in a theoretical distinction between morphologically unmarked categories and zero-marked morphemes, my only recourse is to treat these glossing strategies as equivalent.

2.6.1.2 Pronouns, pronominals, and agreement markers

Personal pronouns represent another vehicle by which grammatical relations may be distinguished.²¹ As defined here, personal pronouns deictically identify arguments. They often encode certain grammatical properties of their nominal referents, such as person, number, gender, proximity, clusivity, definiteness, animacy, and others. Not to be overlooked is that

²⁰ Since I’ve retained verbatim glossing of (51) and (52), the abbreviations are not identical to those in my abbreviations list for this dissertation. Nonetheless, their meanings should be obvious.

²¹ This study does not look at possessive, reflexive, relative, or other non-personal types of pronouns.

pronouns are “targets for agreement” (Corbett 2006) and some of the variation in their forms may be a direct result of case inflection. It may not therefore be possible to separate “pronoun form” from “case” in all instances²². Languages with case systems that affect pronouns vary in the nature of that influence; some case affixes attach directly to pronouns, yet in others, case-inflected pronouns are fully distinct morphologically.

The character of pronoun systems can have interesting consequences for alignment. Broadly speaking, there are two main possibilities. The first and most obvious is that alignment is actually encoded on the pronouns themselves by virtue of their different forms. Such is the case, as in many languages, in the Sahaptian language Umatilla Sahaptin from Oregon. In Umatilla Sahaptin, pronouns reveal a mostly²³ nominative-accusative alignment pattern, as shown in (53).

- (53) Selected personal pronouns in Umatilla Sahaptin (Zúñiga 2006:147, citing Rigsby & Rude 1996)

	<u>A/S pronouns</u>	<u>O pronouns</u>
1SG	ín	ináy
1DL	napiiní	napiinamanáy
1PL	náma	naamanáy
2SG	ím	imanáy
3SG	pín	paanáy

The second way for pronoun systems to have consequences for alignment is when alignment is sensitive to the morphological status of the pronouns themselves, such as whether they are free or bound. Free pronouns are stand-alone words, typically with stress. Bound pronouns, as clitics or affixes, are obligatorily attached to other words, usually verbs, though not

²² Hudson (1995), for example, argues against case inflection on pronouns in English, generally thought to be the only surviving relic of a formerly robust case system. He does not make claims about case influence on pronouns cross-linguistically.

²³ There are other alignment patterns in subparts of the Umatilla Sahaptin pronoun system which I have omitted here. See Zúñiga (2006:146-9) for a concise summary.

always. They are sometimes called *pronominals*, *pronomials*, *pronominal markers*, *pronominal affixes/clitics*, or *incorporated pronouns*. For consistency throughout this dissertation, I will refer to free forms as “pronouns” and bound forms as “pronominals,” and I will avoid using “pronominal” in its adjectival sense unless specifically in reference to the bound pronominals. As an example of this second kind of influence, consider again Umatilla Sahaptin. Whereas the pronouns show mostly nominative-accusative alignment, alignment of the pronominal enclitics is mostly portmanteau, as in (54).

- (54) Selected pronominals in Umatilla Sahaptin (Zúñiga 2006:147, citing Rigsby & Rude 1996)
- | | |
|------------|-----------------|
| <u>A⇒O</u> | <u>enclitic</u> |
| 1SG⇒2SG | =maš |
| 2SG⇒1SG | =nam pá- |
| 1SG⇒2PL | =mataš |

This distinction between free and bound (i.e. between pronoun and pronominal) raises the issue of whether pronominals are really types of pronouns or types of agreement markers. Givón (1975) argues that it is neither possible nor fruitful to distinguish between pronominality and “agreement,” as they form a diachronic continuum and synchronically usually retain their anaphoric interpretations. Cysouw (2003:13-4) casts the issue as rooted in differences in theoretical orientation:

In most of the generative literature on pronominal marking, this continuum [between agreement and pronominalization] is not accepted. Independent pronouns and inflectional person marking are considered to be two completely different aspects of linguistic marking. The reason for the persistence is probably the high status of the *projection principle* as formulated by Chomsky (1981:29). It implies that at every level of syntactic analysis, the arguments of each predicate are to be present (overt or covert). Independent pronouns are possible instantiations of arguments; inflectional person marking is seen as agreement of the predicate with these arguments.

Wichmann (2009:800) adopts a similar outlook:

Rather, what is more generally true of case marking is that it signals the relationship of an argument to a predicate at the clause level. Adherents to the opposite view that case marking is a phenomenon restricted to nouns customarily describe potential candidates for case marking pronominal affixes attached to verbs under the rubric of agreement.

At any rate, the distinction between pronominals and agreement markers only matters for the present study inasmuch as it is necessary to identify alignment patterns of pronominals as different from those of agreement markers. I have not found any instances where this seems warranted.

2.6.1.3 Constituent order

Constituent order is one way that splits in grammatical relations may be signaled. An example comes from Ch'orti', a Mayan language spoken near the center of the Guatemala-Honduras border, in which relative topicality of the arguments determines the position of each with respect to the verb. The transitive construction in (55), and the two intransitives ones in (56) and (57) show the split.

- (55) **A** is topic and preverbal while **O** is non-topic and postverbal in Ch'orti' (Quizar 1994:124, ex. 5)

poréso e **ib'ach** kone'r ma'chi ub'ak're e **jaja'r**
therefore the **armadillo** today not he-fear-it the **rain**
'Therefore, nowadays the armadillo is not afraid of the rain.'

- (56) **S** is topic and preverbal in Ch'orti' (Quizar 1994:129, ex. 19)

e **ixik** yaja' ente'to domíngo e'kmay ta chinam [...]
the **woman** that every Sunday descend-she to town
'Every Sunday that woman went to town [...]

(57) S is non-topic and postverbal in Ch'orti' (Quizar 1994:129, ex. 21)

inte'to a'xin **uwixka'r** maku' chinam
every.time she-go **his-wife** into town

b'an uche kónde ak'otoy e **ixik**
thus he-do-it when she-come the **woman**

'Every time his wife went into town, thus he would act when the woman came back.'

In (55) and (56), A and S are both topics and therefore both preverbal. This is a nominative-accusative pattern. In (55) and (57), O and S are both non-topics and therefore both postverbal. This is an ergative-absolutive pattern. (This is not, according to Quizar 1994, topic fronting, because Ch'orti' does not have a basic verb-initial word order.)

In this subsection I argue against using constituent order as a locus for alignment in this study on the grounds that its fundamentally sequential nature, as opposed to morphological form, precludes it from ever instantiating certain alignment patterns attested elsewhere in language. Recall that alignment is defined in this dissertation as a configuration of all three roles A, S, and O. Constituent order in a minimal intransitive clause, containing only a verb and its sole argument, will necessarily be the order of V and S. There will be exactly two options for the position of S: it will either be pre-verbal or post-verbal.

In a minimal transitive clause, the order will factor in both arguments, for a total of six possible permutations of V, A, and O. Both A and O can be pre-verbal or post-verbal, in principle. If the two arguments are consistently on opposite sides of the verb, then determining alignment is simple because S will align either with A or with O in terms of being on either side of the verb. But if both A and O are on the same side of the verb, then alignment cannot be determined in this way because they will both be either pre-verbal or post-verbal.

Imagine what a nominative-accusative pattern would look like where the locus of alignment is constituent order. In such a pattern, S and A would have to get equivalent placement with respect to the verb because this would be the only criterion applicable to both transitives and intransitives. In other words, both S and A would have to be either pre-verbal or post-verbal. O would have to be the opposite of whatever S and A are. Either of the constituent orders in (58) would satisfy this requirement.

(58) Nominative-accusative constituent orders

	<u>Transitive</u>	<u>Intransitive</u>
(a)	AVO	SV
(b)	OVA	VS

Likewise, an ergative-absolutive pattern with constituent order as the locus would see O and S on the same side of the verb, with A on the opposite side. Both sets of orders in (59) show such a sequence.

(59) Ergative-absolutive constituent orders

	<u>Transitive</u>	<u>Intransitive</u>
(a)	AVO	VS
(b)	OVA	SV

The nominative-accusative and ergative-absolutive patterns laid out in (58) and (59) both share the property of having A and O on opposite sides of the verb. That is precisely the feature allowing those two alignment patterns because they permit S to match up with either transitive argument, but not both.

Then what kind of alignment pattern would it be if both A and O are on the same side of the verb? The logical possibilities are shown in (60).

(60) Logical possibilities of A and O on the same side of the verb

	<u>Transitive</u>	<u>Intransitive</u>
(a)	AOV	SV
(b)	OAV	SV
(c)	VAO	VS
(d)	VOA	VS
(e)	AOV	VS
(f)	OAV	VS
(g)	VAO	SV
(h)	VOA	SV

Consider the first four constituent orders in (60a)-(d) where S, A, and O are all on the same side of the verb. It may be tempting at first to consider this to be neutral alignment because A, S, and O are all preverbal. But such an interpretation only works under the assumption that argument position *with respect to the verb* is the only thing that matters. In transitive clauses, however, the position of A and O with respect to *each other* is a crucial metric. And this metric is unavailable for the intransitive clauses because there is only one argument. The fact that a metric is available only to transitives is not an issue by itself; other patterns already discussed in §2.4 have this property, too, such as direct-inverse and portmanteau alignment. But of course, the pattern shown above is neither of those: it is not direct-inverse because there is no extra morpheme indicating directionality of action and it is not portmanteau because there is no fusion of grammatical relation and directionality.

If anything, the pattern in (60a)-(d) is most like tripartite because there one way of identifying S (by its position relative to the verb), one way of identifying A (by its position relative to the verb and to O), and one way of identifying O (by its position relative to the verb and to A), and each of these is necessarily different.

Now consider the last four orders in (60e)-(h). In those orders, A and O are both on the same side of the verb, but S is on the opposite side. This seems like it could be a double-oblique pattern, but it is not. The reason it is not a double oblique pattern is the same as the reason why the orders in (a)-(d) are not neutral: there is still a way of distinguishing A from O, and that is by their position with respect to each other. So which alignment pattern is expressed by (e)-(h)? Surprisingly, again the answer must be tripartite. Here the identification of S is achieved by its position relative to the verb, that of A by its position relative to the verb and O, and that of O by its position relative to the verb and A.

All of this means that the only possible alignment patterns determined by constituent order are nominative-accusative, ergative-absolutive, and tripartite. As these constitute only a subset of the alignment patterns revealed within other loci, in this project I ignore constituent order when determining alignment.

2.6.2 Locus of alignment as a split inducer

The loci of alignment I have been describing in §2.6.1 above amount to what might be called “subparts” or “subdomains” of a grammar, e.g. case morphology, pronouns, pronominals, etc. In viewing these as loci of alignment, I am, for the most part, looking at each of these subparts to determine whether there is an alignment split within them. It is also essential that I consider whether there is an alignment split *between* them; or to put it differently, whether different alignment patterns on different loci actually count as a *split*. For example, it might be that some language has one alignment pattern on the pronominals, but a different one on the nouns and pronouns. Such a state of affairs, where relational distinctions are not applied

uniformly across all classes of nominals, is what Dixon (1994:95) calls a “meta-split” and what Iggesen (2009) refers to as “case-asymmetry.”

Dixon (1994:94) says that distinct marking on bound versus free forms “is best regarded not as a distinct kind of conditioning, but as a secondary phenomenon, explainable in the same terms as [an] NP-conditioned split.” However, as one of the long-term aims of studies such as this one is to explain the preponderance of certain marking strategies over others, it is useful for me to treat the locus of alignment as a potential split inducer provided that alignment is *non-neutral* on the loci in question. The reason for the non-neutrality stipulation is that certain instantiations of identical form should not be considered “neutral alignment” because they do not actually reveal alignment at all.

To give an example, the Mayan language Sacapultec has a robust ergative-absolutive pattern on the pronominals. But the full pronouns are identical to one another within each person/number category, no matter whether they are A, S, or O. Is it fair to say that Sacapultec shows a split induced by nominal status, since pronominals are ergative-absolutive and pronouns are neutral? No, because the neutralization of a distinction between A, S, and O on the pronouns casts serious doubt on any hypothesis that alignment is even realized on the pronouns in the first place. The pronoun category in Sacapultec represents a locus or grammatical domain in which alignment simply doesn’t show up. In order to count pronouns as a locus for alignment in Sacapultec, there would need to be evidence somewhere in the language that at least some pronouns show distinct morphology for different grammatical relations.

On the other hand, there are genuine instances of a split induced by nominal status because alignment is non-neutral and different for different types of nominals. It occurs, for

instance, in the extinct isolate Alsea, described fully in Chapter 4. In Alsea, pronominals show nominative-accusative alignment, but nouns and pronouns are ergative-absolutive. In this case, nominal status is a split inducer because there are different alignment patterns in the different loci and in none of them is a distinction neutralized.

2.6.3 Function and form

It is well known that case markers (and morphemes in general, for that matter) do not always have a one-to-one correspondence between phonological form and grammatical function. So is it the function of a morpheme or its form which determines alignment? The answer is that it is its function as the identifier of A, S, or O, and so restricted by this function, the various forms it may take.

As an example, consider that the form of the recipient or beneficiary in a three-argument clause might also indicate the S or A of certain verbs. This is especially so of psych verbs which take something like a semantic EXPERIENCER as an argument. Other functions are possible, too, as in these examples from the Muskogean language Koasati, spoken in Louisiana.

- (61) **Dative** marker for beneficiary in Koasati (Kimball 1991:131, ex. 63)

cim-acó:li-t

2SG.DAT-sew-PAST

‘She sewed it for you.’

- (62) **Stative S** in Koasati (Kimball 1991:132, ex. 67)

cim-ho?-pa

2SG.STAT.SUB-hurt-INTER

‘Are you hurt?’

- (63) Idiosyncratic **O** in Koasati (Kimball 1991:132-3, ex. 71)
 s-**cim**-aha:láhl
 INST-2**S.OBJ**-amuse
 ‘He amuses you.’

The bold-faced constituent in (61)-(63) is the same in every instance, *cim*-, but with different functions. Kimball (1991) is careful to relay the functional differences in his glossing choices, but not all authors do this.²⁴ If Kimball had glossed *cim*- in (62) and (63) as 2SG.DAT, it would give the (inaccurate) impression that the morpheme in those constructions is functionally dative, but it would not change the characterization of the alignment system because (62) would still show the marking of S and (63) would still show the marking of O, since these abbreviations are defined in terms of the number of permissible arguments. The takeaway point is that constructions like (62) and (63) should not be excluded from an analysis of alignment just because they contain a morpheme which is syncretic with the dative. The construction in (61), however, should be excluded because *cim*- does not represent an A, S, or O there.

2.6.4 Valency changing operations

Constructions in which a verb’s valency has been derived by a syntactic or morphological process do not contribute to a description of that language’s split-alignment behavior. Of course, valency operations are a crucial aspect to any grammar and often reveal interesting features of the language regarding transitivity or the status of grammatical relations. But it would be remiss to use, say, a passive construction as an example of an alignment split when compared to the corresponding active. Valency changing operations fall mainly into four categories and,

²⁴ As the rock band Rush observes, “Everybody got mixed feelings about the function and the form” (Vital Signs, Lee et al. 1981).

depending on the language, are usually signaled by overt morphology or marked syntax (Dixon 2010a). The categories are passive, antipassive, causative, and applicative.

These four categories are broad generalizations. There is substantial cross-linguistic variation in their actual instantiations, as well as subtypes of and countertypes to these categories. I will not go into a full explication of these processes here; where relevant to the discussion at hand, I will describe the necessary facts on a case-by-case basis. A very brief set of examples, however, will illustrate the basic idea. The pair in (64) compares an active sentence to a passive one from Innu, an Algonquian language from eastern Canada. The O in the active is promoted to S in the passive; the morpheme *-kani-* signals the passivization.

(64) Passive in Innu (Drapeau 2012:181, exx. 1, 2)

cipaym-w išk-wâtêm-ilu
to.close.TI-3 door-OBV
's/he closes the door'

cf. cipay-**kani**-w išk-wâtêm
to.close.TI-PT-3 door
'the door is closed'

In (65), from Sierra Popoluca spoken near the northern coast of the Isthmus of Tehuantepec, the O in the transitive construction is demoted, in this case omitted entirely, in the corresponding antipassive. The antipassive morpheme *-ʔoʔy-* indicates the derivation.

(65) Antipassive in Sierra Popoluca (Boudreault 2009:509, exx. 13.26, 13.27)

nikk-pa ʔi-wiit-W jeʔm yoomo-tam
go-ICP 3ERG-massage-DEP.TR that woman-PL.HUM
'She (the midwife) goes to massage these (pregnant) women.'

cf. ʔagi-Ø-wiit-**ʔoʔy**-pa
INTENS-3ABS-massage-ANTIPASS-ICP
'She massaged a lot.'

Zuni, spoken in the southwestern United States, provides an example of a causative. In (66), the one-argument construction can be compared to the two-argument causative construction, indicated with the causative morpheme *-kʔa-*.

(66) Causative in Zuni (Stout 1973:208, exx. 12, 13)

ʔaʷ-akcekʔi ʔaʷ-aʷ-ka
 PL-boy PL-go-PAST
 ‘The boys went.’

cf. taʔpuʔpuʷ ʔaʷwan kʔakw-an ʔaʷ-aʷ-**kʔa**-kka
 governor their house-LOC PL-go-CAUS-PAST
 ‘The governor sent them home.’

An applicative construction is shown in (67) from Movima, an isolate from central northern Bolivia. The applicative morpheme *-pa* is used to render a one-argument verb into a two-argument one.

(67) Applicative in Movima (Haude 2006:401, exx. 145a, b)

josi:-cheʔ
 laugh-R/R
 ‘to laugh’

cf. josi-**pa**:-na as jankwa = n
 laugh-APPL-DR ART.N say=2
 ‘I laugh at what you are saying.’

2.6.5 Alignment-like but not alignment

There are certain phenomena which behave in alignment-like ways, but which do not qualify as alignment-marking patterns because they do not reveal the semantic roles of participants. An instance of this is noun incorporation in Southern Tiwa, a Kiowa-Tanoan

language spoken in central New Mexico. In Southern Tiwa, animate S and animate A are unincorporable. Animate O, however, may be incorporated, as (68)-(70) show.²⁵

(68) Animate S is unincorporable in Southern Tiwa (Allen et al. 1984:299, exx. 56a, b)

(a) Musade we-seur-mi
cat A.NEG-fall-PRES.NEG
'The cat is not falling.'

(b) * We-musa-seur-mi
A.NEG-cat-fall-PRES.NEG
'The cat is not falling.'

(69) Animate A is unincorporable in Southern Tiwa (Allen et al. 1984:299, exx. 58a, b)

(a) Hliawrade Ø-k'ar-hi yede
lady A:A-eat-FUT that
'The lady will eat that.'

(b) Ø-hliawra-k'ar-hi yede
A:A-lady-eat-FUT that
*'The lady will eat that.'²⁶ ('She will eat that lady' is the required reading.)

(70) Animate O is incorporable in Southern Tiwa (Allen et al. 1984:294-5, exx. 12, 13)

(a) Seuanide ti-mũ-ban
man 1S:A-see-PAST
'I saw the/a man.'

(b) Ti-seuan-mũ-ban
1S:A-man-see-PAST
'I saw the/a man.'

In (70), the unincorporated animate O may be grammatically incorporated. (68) and (69) show that incorporation of an animate S or A is ungrammatical. This is a nominative-accusative

²⁵ The interlinear glosses from the Allen (1984) paper import the theoretical framework of uninetwork relational grammar (Perlmutter 1980) by encoding theory-specific concepts such as "final subject," "final direct object," etc. into the glossing abbreviations. I have not included the meanings for these abbreviations in the abbreviation list for this dissertation. Please see Allen (1984) and sources therein for a discussion of their glossing choices.

²⁶ The gloss provided for this construction is given in Allen et al. (1984:299) as follows:

'She ate that lady.'
*'The lady ate that.'

I can only assume that this is a typo because future tense morphology is in the example itself, even though the gloss is rendered in the past tense.

pattern in the sense that S and A respond the same way to incorporability, while O responds differently. Incorporation in Southern Tiwa gets much more interesting. When arguments are inanimate rather than animate, incorporability is apparently ergative-absolutive. Inanimate S and O are *obligatorily* incorporated. (I say “apparently” ergative-absolutive because Allen et al. say that they do not have examples of inanimate As, so it is not possible to define this pattern as ergative-absolutive with certainty according to how I have defined that term.) The Southern Tiwa data therefore show a kind of split-incorporation system, which has enjoyed (or endured) various theoretical treatments (Franz 1985, Sadock 1985, Rosen 1990, Heck & Richards 2010). But it would not be appropriate to say that Southern Tiwa noun incorporation is an example of split-alignment because noun incorporation is not the means by which grammatical relations are expressed. Grammatical relations in Southern Tiwa are expressed by verbal prefixes which encode person and number (in a complex and theoretically significant manner, but I refer the reader interested in the details to the sources mentioned just above).

Syntactic ergativity, which has been covered in a good deal of alignment-related literature, also does not count as alignment for this project. The now somewhat famous example of syntactic ergativity is from Dixon’s (1979) exposition of conjoined clauses in Dyirbal. When one intransitive and one transitive clause are conjoined, S is permissibly elided if it is co-referential with O, and O is permissibly elided if it is co-referential with S, as in (71) and (72).

- (71) Elided S in Dyirbal (Dixon 1979:62, ex. 6)
- | | | | | | |
|--------|------------|------------|------|-------|---------------|
| ɲuma | yabu-ɲgu | buɾa-n | Ø | _____ | banaga-ɲu |
| father | mother-erg | see-nonfut | conj | (S) | return-nonfut |
- ‘Mother saw father and (he) returned.’ or ‘Father was seen by mother and returned.’

- (72) Elided O in Dyirbal (Dixon 1979:62, ex. 7)
- | | | | | | |
|--------|---------------|------|-------|------------|------------|
| ɲuma | banaga-ɲu | Ø | _____ | yabu-ɲgu | buɾa-n |
| father | return-nonfut | conj | (O) | mother-erg | see-nonfut |
- ‘Father returned and mother saw (him).’ or ‘Father returned and was seen by mother.’

It is not possible in Dyirbal to conjoin clauses meaning *Father returned* and *Father saw mother* and also elide the common element *father* because *father* is S in the intransitive clause and A in the transitive clause. The conjunction-elision operation only works for elements in S and O relation.²⁷

I am not saying that this property of conjunction (or, in some other languages, subordination), which treats S and O alike, is not ergativity; I am saying that it doesn’t qualify as *alignment* because coordination is not the means by which A, S, or O are actually identified in Dyirbal.

There are many other processes and phenomena which behave in alignment-like ways because they treat various subsets of A, S, and O as members of the same category with respect to that process or phenomenon, but which do not count as alignment. They are too numerous to list exhaustively, but as a final example consider the English preposition used “when the head precedes in a nominalized clause” (Donohue 2008:26), as in (73).

- (73) (a) English **A** preceded by *by*
 an inquiry by **the state department**
- (b) English **S** preceded by *of*
 the death of **Anselmo**
- (c) English **O** preceded by *of*
 a transfer of **funds**

²⁷ It would be possible to conjoin these clauses if A is converted to a derived S via an antipassive operation.

The preposition in (73a) is *by* because *the state department* is A. In (b) and (c), the preposition is *of* because *Anselmo* is S and *funds* is O. This is an ergative-absolutive pattern, but it is not alignment because preposition choice in nominalized clauses is not the means by which grammatical relations are identified in English.

Thus the determining factor for me is the mechanism(s) by which grammatical relations are signaled in the language. If a phenomenon operates by treating A and S as members of the same category, or S and O, or any other arrangement which can be labeled using alignment terminology, it only counts as alignment if that is the way that A, S, and O are actually expressed.

2.7 Summary of Chapter 2

The objective of Chapter 2 was to contextualize my approach to alignment typology. I am treating transitivity as an essentially valence-dependent notion, rather than a semantically based one. The abbreviations A, S, and O, taken from Dixon (1994, 2010a, b), represent the different core arguments: S represents the single argument of an intransitive verb, A and O the two arguments of a transitive verb. A and O are distinguished from one another by the semantic criteria proposed by Dowty (1991) such that the proto-agent is A and the proto-patient is O.

I described and showed brief examples of seven alignment patterns, and then discussed how different patterns can be induced under different grammatical conditions. The condition-dependent occurrence of alignment patterns is what is meant by “split-alignment,” and the conditions themselves are termed “split inducers.”

I discussed a handful of analytical issues, in particular the various loci of alignment and the exclusion of constituent order from consideration. I also reviewed several phenomena which

behave in alignment-like ways and to which may be attributed certain alignment labels, but which do not actually qualify as alignment for this study.

3 STUDY DESIGN AND METHODOLOGY

3.1 Introduction to Chapter 3

This study seeks to explain the distribution of split inducers across the indigenous languages of North America. Fortunately, it is not practical to examine each and every one of these languages.¹ Rather, a sample must be generated which can arguably represent the entire population. The objective of Chapter 3 is to describe the procedure I used for generating such a sample and my methodology for sorting out the relevant language data.

In §3.2.1, I explain the geographic and linguistic parameters of my sampling frame (i.e. the set of languages which could potentially go into the final sample). In §3.2.2., I discuss how the frame was stratified in order to reduce the likelihood of areal and genetic bias. The specific procedure for selecting individual languages for the sample, and reasons for eliminating others, is detailed in §3.2.3. Finally, in §3.3, I talk about the source materials I used for the study and some unavoidable bias that remains in the sample.

¹ I say *fortunately* because if it were practical to examine each of them, that would mean that the number of indigenous North American languages would be significantly less than it is even today. Most linguists are aware of the extreme endangerment faced by the majority of these languages. But given this dire and truly unfortunate fact, it would be much worse if the number of available languages were small enough to facilitate an exhaustive investigation.

3.2 Language sample

The point of this subsection is to explain how the 15 languages in my sample got there. Wrapped up in this topic is a precise denotation of the sampling frame, a discussion of how certain biases were attenuated, and an account of the language selection process.

3.2.1 Sampling frame

The universe of investigation consists of all languages indigenous to North America. For several reasons, North America is an ideal geographic area to explore in relation to the topic of split-alignment. First, taken together, North American languages display all of the major types of split-systems described in the alignment literature, so there are not likely to be any significant gaps in the representation of split-alignment even when all instances are drawn from that continent. Second, North American languages are hugely diverse typologically and genetically (there are over 50 language families). This means that the likelihood of genetic inheritance accounting for shared patterns is lower than on some other continents where there are only a handful of language families. Third, North America is known to contain a number of strong linguistic areas within which structural diffusion has definitely occurred across unrelated languages. This has two direct implications for the present project: (a) the sampling procedure can systematically reduce the influence of areal diffusion when determining the cross-continental distribution of the various split-alignment patterns; and (b) it leaves open the option of exploring such areas in greater detail as an explanation for why certain patterns may propagate more readily than others.

For the purposes of this study, *North America* means the geological continent of North America, including Greenland, Newfoundland, and the Aleutian Islands to the southwest of Alaska. The southern border is drawn north-south through central Honduras such that western Honduras, which contains the southernmost contiguous Mesoamerican languages, is treated as part of North America, as is all of El Salvador. Eastern Honduras and the remaining central American isthmus are not considered to be a part of North America as defined here.

The languages from North America considered are those spoken (or whose direct ancestor languages were spoken) prior to the arrival of the Europeans. Mithun (2010) estimates the number of time-of-contact languages to be over 2000. Many have since gone extinct. In principle, the extinct languages are also included in the sampling frame, but practical challenges preclude most of these from actual inclusion.

The sampling frame “is the means of access to the universe [of investigation]” (Bell 1978:126). The frame would ideally contain all languages which have been described in the linguistic literature, perhaps supplemented with field work where possible and appropriate. In my case, constraints on time and money (certainly not on inclination) make field work on these languages unfeasible. Thus the frame is the set of available linguistic resources on North American languages. This mostly means descriptive grammars, but also includes journal articles and other published literature.

3.2.2 Stratification of the sampling frame

Bell (1978) got the ball rolling on language sampling methods for general typological research. His concerns in that paper revolve around establishing a foundation for the appropriate

application of different types of samples (probability and non-probability), and for developing an appreciation of various types of errors and biases. Probability samples offer the greatest degree of error estimation and statistical representation of a given population. The chance of any particular sample being selected is known. Therefore, if a researcher wishes to make claims about the relationship between the sample and the universe of investigation, then a probability sample is the only way to do so with statistical certainty. Still, error and bias can contaminate probability samples, but there are strategies for avoiding these.

One strategy is the use of a stratified sample. To obtain such a sample, the frame is divided into non-overlapping categories. Then *random* languages (to avoid inadvertent bias) are selected from each category. But using a stratified sample means that two issues first need to be resolved: how to go about stratifying the frame and how many languages should be randomly selected from each strata.

Frame stratification cannot be arbitrary; it must take into account sources of bias in language samples. Two major such sources, according to Bell, are genetic and areal bias. Stratifying the frame according to these factors helps to ensure that the sampled items (i.e. languages or language groups) are independent of one another. In other words, if a trait is shared by two closely related languages, then this should not count as two instances of the trait, because the languages may have inherited it from their common ancestor. Likewise, a trait which passed from one language to another by areal diffusion should also count as one instance. By stratifying the frame according to genetic and areal criteria, the researcher can ensure that the sampling procedure does not favor languages or language groups which may show non-independent instances of the trait.

Bell's proposal achieves genetic, but not areal stratification. He separates the world's languages into 16 linguistic stocks and selects a time-depth of 3500 years to group together smaller categories within each stock. He then determines on the basis of "historical knowledge [and] glottochronological estimates" (ibid:147), presumably obtained from published literature and his own expertise, the number of groups in each stock which are separated by 3500 years or more. He arrives at 478 such groups and these make up his stratified frame. The 478 categories are non-overlapping and where there is a genetic relationship between them, it is at least as distant as 3500 years.

Since the publication of Bell's seminal paper, a number of scholars have taken various approaches in refining language sampling procedures. One line of approaches is more or less a direct extension of Bell's method, but aims to improve it by standing up to areal bias, as well as genetic. This avenue is taken up, to various degrees and with nontrivial differences, by Tomlin (1986), Perkins (1989, 1992), Dryer (1989, 1992), and Nichols (1992), among others. The crux of these approaches is to stratify the frame so as to control for both genetic and areal affiliation. This is typically done by first dividing the languages in the frame into genetic groups of a comparable time-depth, as Bell does. Then the frame is divided by certain areal criteria. In Tomlin's case, these are what he calls "non-controversial linguistic or cultural areas" (1986:29), though he does not specify his sources for this determination.² By contrast, Dryer (1989) argues for the (possible) existence of large linguistic areas the size of entire continents. Dryer's method is novel in a number of other ways, and as his approach is influential to my proposal, it is worth describing in some detail.

² Except for one example: South Asia, for which he cites Masica (1976), and which he refers to as South *Africa*, presumably the result of a typographic error.

Dryer has the goal of stratifying languages with a world-wide sample in mind. Since my project is restricted to North America, the number of languages and the specific regions are different from his, but the principle is nearly identical. First, Dryer forms subfamilial language groups which he calls *genera*. Genera are “roughly comparable to the subfamilies of Indo-European, like Germanic and Romance,” (ibid:267) a determination he makes by his own expertise, but which roughly follows the categorization by Ruhlen (1987). The genera are drawn from a large sample of 542 languages. Dryer assumes that within a given genus, typological characteristics will be mostly similar. For Dryer, it is genera, not languages, which are the units counted.

To achieve areal independence, Dryer categorizes the genera according to five continental areas: Africa, Australia-New-Guinea, Eurasia, North America, and South America.³ In justification of the proposal for large linguistic areas, Dryer gives evidence of linguistic properties which are common within such large areas, but uncommon outside them, suggesting that linguistic traits may diffuse over enormous geographic space. Dryer needs only one such property to support the existence of a large linguistic area. In this sense, his linguistic areas are different from the general notion *linguistic area*. The general notion refers to rather contained geographic regions in which sets of languages, some of which are usually unrelated, share a number of diffused linguistic properties.⁴

³ In Dryer (1992), he increases the number of large areas to six: Africa, Eurasia, Southeast Asia & Oceania, Australia-New Guinea, North America, and South America.

⁴ The exact number of properties they must share in order to be considered part of a linguistic area is controversial (see Campbell, Kaufman & Smith-Stark 1986:530-6 and Campbell 2006:1-31 for discussion), but in general, the more properties which are shared, the stronger the case is for the existence of a linguistic area.

Having grouped genera into these five large areas, Dryer can establish the number of genera within each area which contain languages with property X, compared to genera containing languages without X, or languages with a different property Y, etc. Thus he can ask, “How many genera in each large area contain languages with property X?” rather than “How many languages in each large area have property X?” If, for each of the five areas, more genera contain languages with the property than without it, then this is taken to be confirmation of an overall linguistic preference for that property.

To illustrate, Dryer (1989:269-70) provides an analysis supporting a linguistic preference for SOV word order. In each of the five large areas, he counts the number of genera containing SOV languages and compares that to the number of genera containing SVO languages. In all five areas, there are more SOV-language genera than SVO ones, so SOV is taken to be an overall linguistic preference. If there were no preference for either word order (i.e. if the probability of each word order were .5), then the chance of all five areas having more SOV-language genera would be one in thirty-two ($0.5^5 = 0.03125 = 1/32$), assuming genetic and areal independence of the areas. The observation that all five areas do show more SOV-language genera confirms a preference for SOV with less than .05 statistical significance ($0.03 < 0.05$).

Dryer’s strategy is criticized by Dunn et al. (2011), who adopt a Bayesian phylogenetic approach to language sampling and argue that Dryer’s correlations are not as strong as he claims. Their method fosters the ability to follow correlated features as they change over the course of the history of a language family. However, Dunn et al. do not make any effort toward areal stratification and the language families they investigate are selected for their phylogenetic

completeness rather than their genetic diversity. I return to this issue in my discussion of my overall findings in §5.2.

The sampling frame of my study is stratified along genetic and areal lines, in a manner which draws heavily on Dryer's (1992) method, with certain modifications. Recall that Dryer divides his frame into genealogical groups (*genera*) of 3500-4000 years and areal groups (large continental areas). Obviously, areal stratification in units as large as continents would not make sense for me because I am only looking at a single continent. Instead, I stratify North America into non-overlapping geographic areas with a low likelihood of linguistic diffusion between them; this is detailed below in §3.2.2.1. Genetic stratification of my sampling frame is largely the same as that proposed by Dryer, but of course is restricted to North American genera; this is described in §3.2.2.2.

3.2.2.1 Areal stratification

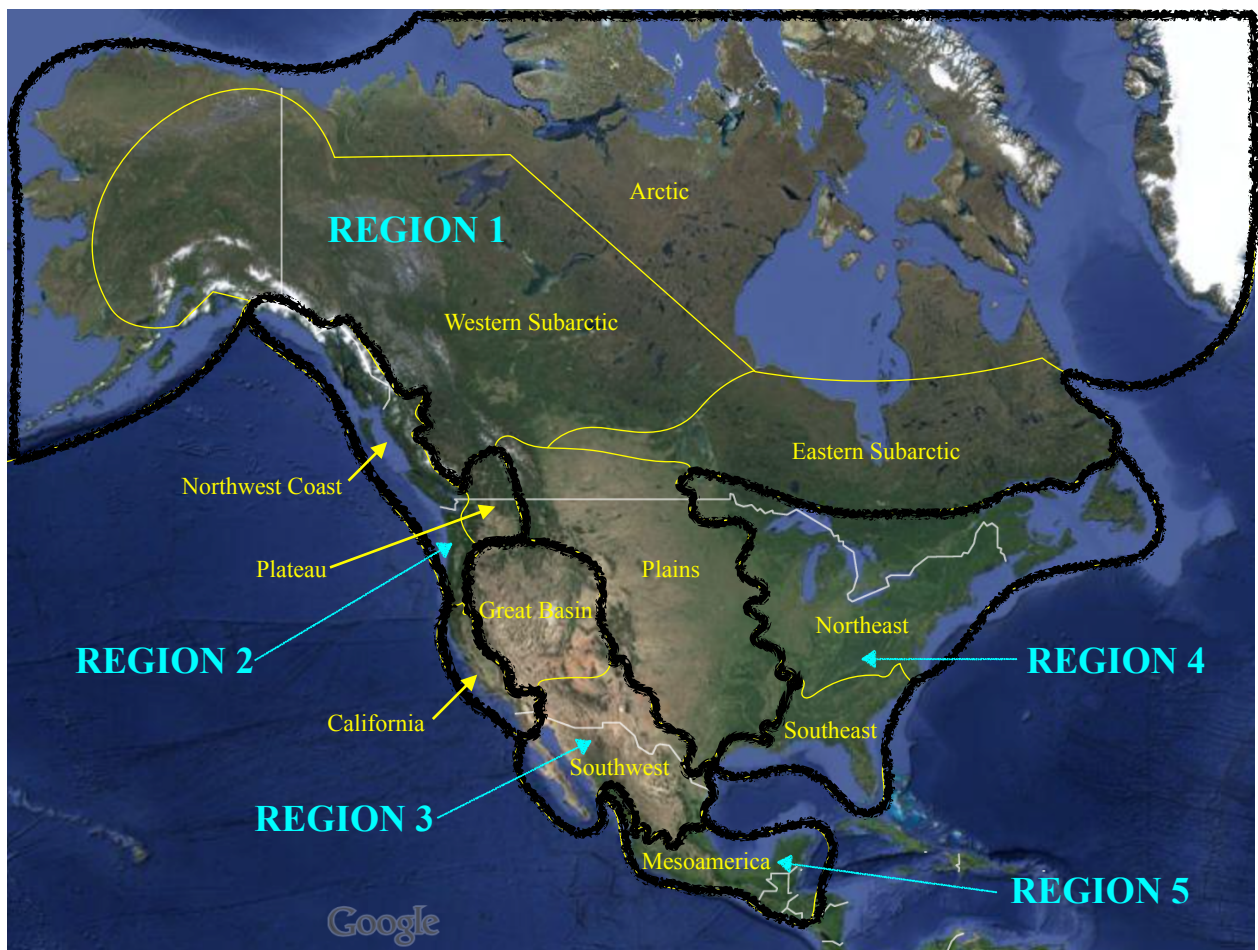
In a *worldwide* sample, the ideal sampling frame must be stratified such that every language group can be assigned to exactly one geographic region so as to avoid areal bias. In North America, however, avoiding areal bias is not as simple as carving up the region into sub-areas. The reason is that North America is home to a sizable number of linguistic areas which are claimed to contain diffused linguistic traits. Sherzer (1976) identifies at least 11 such areas north of Mexico; further south, Mesoamerica is a very well established linguistic area. Indeed, it appears that nearly every part of the continent is a member of some linguistic area or another.

The way I reduce bias in the sample is to cluster certain of the contiguous linguistic areas together so that the new regions themselves cover a bigger geographic area and there are fewer total areas to be sampled from. The five areas I use as divisions of North America are enumerated in (1) below. The figure in (2) shows the geographic boundaries of each region.

(1) Five geographic sampling areas in North America

1. Arctic, Subarctic, & Plains
2. Northwest Coast, Plateau, & California
3. Great Basin & Southwest
4. Northeast & Southeast
5. Mesoamerica

(2) Geographic stratification of North America



The yellow lines and labels represent the cultural and linguistic areas within North America. The black lines show the borders of the five larger regions, divided for areal stratification. Note that although Region 1 is significantly larger than the others, much of it is uninhabited and it actually contains the fewest genera of all regions.

To be clear, these groups are formed in a quasi-arbitrary way. The names in each group (“Arctic,” “Subarctic,” etc.) are the names of already established linguistic or cultural areas and have been adopted as standard geographic units within North American studies (Sherzer 1976, Campbell 1997, Mithun 1999). The grouping of these regions together as I have done achieves three things. First, it reduces the total number of areas to be sampled from, which eases the burden and time demand of data collection. Second, it is necessary to ensure that each area has enough genetic groups inside it. Otherwise, treating “Arctic” as its own area would mean that at most, two genetic groups (Eskimo and Aleut) could be sampled from it because these two are the only groups in the Arctic. Third, it amounts to an odd number of total areas, making it possible to state whether certain split inducers are or are not in the majority across the continent.

It is always the case that some areas within each region are geographically contiguous. They are only arbitrary in that I do not presuppose that they form natural geo-linguistic units. But this is not a problem for the sampling procedure because I only assume that diffusion has not occurred (or is not evident) *between* these areas; in other words, they are independent of one another on areal grounds.⁵ In grouping these areas together, I recognize that there almost certainly are diffused traits *within* them. But this is just the point, because I can determine

⁵ In so doing, I am following Dryer’s (1992) logic for dividing up the world into large continental areas. Many of his groups (e.g. Australia-New Guinea or Oceania-Southeast Asia) are similarly arbitrary, as he freely admits.

whether certain alignment patterns are more common within such areas than outside them, thus appealing to areal explanations for the presence of certain traits.

3.2.2.2 Genetic stratification

From each geographic region, the same number of genetic groups will be selected at random. Dryer (1989, 1992) introduces the concept *genus* as a genetic unit whose time-depth is 3500-4000 years, making it comparable across the world. His 1992 list and the expanded version in WALS (2011) both provide a hefty number of genera from North America. However, it is questionable whether these are indeed commensurate because he does not provide the actual time-depths of the genera, stating only in general terms:

If there is evidence of time depth of groups, the genus would not have a time depth greater than 3500 or 4000 years. A genus may have a time depth much less than this, but if the time of the split of one group of languages from other languages in the family appears to be greater than 4000 years, then this constitutes a reason to say that this group of languages is a separate genus (Dryer 2011).

So for Dryer, the time-depth of a genus is never greater than 4000 years, but it is not clear what the minimum is or if there is one. He goes on to say:

The decision as to which groups to treat as genera here are best described as my own educated guesses. In many instances they are based on conversations [I have] had with specialists. However, in the absence of a tradition within the field of attempting to identify groups of comparable time depth in different parts of the world, they should not be considered more than educated guesses.

In any case, it turns out that the vast majority of Dryer's genera are in line with Campbell's (1997) more detailed assessment of familial relationships in North America. Besides, having genetic groups of a comparable time-depth is not the real issue for this project. The real

issue is selecting genetic groups which reduce the chances that two randomly selected but identical alignment patterns are identical as a consequence of common inheritance.

The total number of genera for all North American regions is 105. Appendix 1 gives an exhaustive list of genera per region. For the sake of illustration, the ten genera in Region 1 are shown below in (3).

- (3) Genera of Region 1: Arctic, Subarctic, Plains
1. Aleut
 2. Algonquian
 3. Aranama-Tamique
 4. Beothuk
 5. Caddoan
 6. Eskimo
 7. Karankawa
 8. Kiowa
 9. Siouan
 10. Tonkawa

Each genus in North America is assigned to exactly one region. In some cases, this meant doing so forcibly, as some genera contain languages or entire subgroups in more than one region. An example is Algonquian. Within Algonquian: Blackfoot, Cheyenne, Cree-Montagnais, and Ojibwa-Potawatomi would be assigned to Region 1, whereas Arapaho would be assigned to Region 3, and Eastern Algonquian, Fox, Menominee, and Miami-Illinois to Region 4. Similar spreads occur for Siouan and Muskogean. In such cases, each genus is assigned to whichever area contains the largest geographic range of it. Therefore all of Algonquian and Siouan is assigned to Region 1, and all of Muskogean to Region 4.

Dryer's method allows for a simple binomial test to determine whether the observed distributions are linguistic preferences, or artifacts of other phenomena. Using this strategy, I can ask for each split inducer x and each area y , how many genera in y contain a language with x ?

3.2.3 Selection and elimination procedure

I select three genetic groups at random from each of the five areas, for a total of 15 genetic groups. Then, I randomly select one language from each genetic group, for 15 languages total. Unlike in Dryer's method, I only sample one language per genus, and not all genera are included in the sample. The reason for this is that my sample size (= 15, discussed below) is less than the total number of genera (= 105) from which languages are drawn. Taking exactly one language from each available genus would necessitate a sample size of 105, and taking more than one language from some genera would increase the size further. In this sense, my sampling method is similar to that of Nichols (1992), who also takes one language per genetic group. However, Nichols is deliberate in selecting languages which are considered typical or representative of the family (*ibid*:27). I agree with Bickel (2008:223) that it is a "thorny question of what is the best representative of a branch or genus," especially if the linguistic variable of interest has not been fully studied for each branch. For this reason, I instead select an individual language randomly from each genus.

The random selection procedure, both for the genera within each area and the language within each selected genus, was facilitated by a free online random sequence generator (random.org 2014). The website uses atmospheric noise to generate random numbers or sequences within specified ranges. The list of genera for each region was obtained from Campbell (1997) and cross-referenced with Dryer (2011).⁶ These were entered into a spreadsheet alphabetically, enumerated, and then selected at random. For each selected genus, the list of individual member languages was obtained from the Ethnologue website (Lewis et al. 2014),

⁶ There are some differences between these sources, mainly in terms of the language groups mentioned and the names they are given. Dryer only lists extant languages, but Campbell lists the known extinct ones.

enumerated, and then selected at random. The randomly selected languages were then investigated for their alignment properties. If a language was determined to have split-alignment, it was entered into the sample. If a language was determined not to have split-alignment, or if determination was inconclusive, a new language was selected from a different genus in the same area.⁷ A more detailed procedure is laid out in (4).

(4) Procedure for creating a probability sample

1. Assign each genus in Dryer's (2011) and Campbell's (1997) compendiums to exactly one of the five areas in North America.
2. Starting with Region 1, randomly select one genus.
3. From the genus selected by (2), randomly select one language.
4. Determine, for the language selected by (3), if it (a) has split-alignment, or (b) does not have split-alignment at all.
 - If it is determined that the language does have split-alignment, then categorize it according to its split inducers.
 - If it is determined that the language does not have split-alignment, enter it into a database and label it *not applicable*. Repeat step (2) until (4a) is achieved.
 - If available information is insufficient for a determination, enter the language into a database and label it *undetermined*. Repeat step (2) until (4a) is achieved.
5. Repeat the process from step (2) starting with Region 2, and so on, until a language has been selected from all three genera.
6. Repeat the steps (2)-(5) until three languages have been selected per region for a total of 15 languages.

See Appendix 2 for a list of member languages from each selected genus.

3.2.4 Final sample

The stratification of the continent, combined with the selection and elimination procedure described above, resulted in a wide geographic and genetic range of languages admitted into the

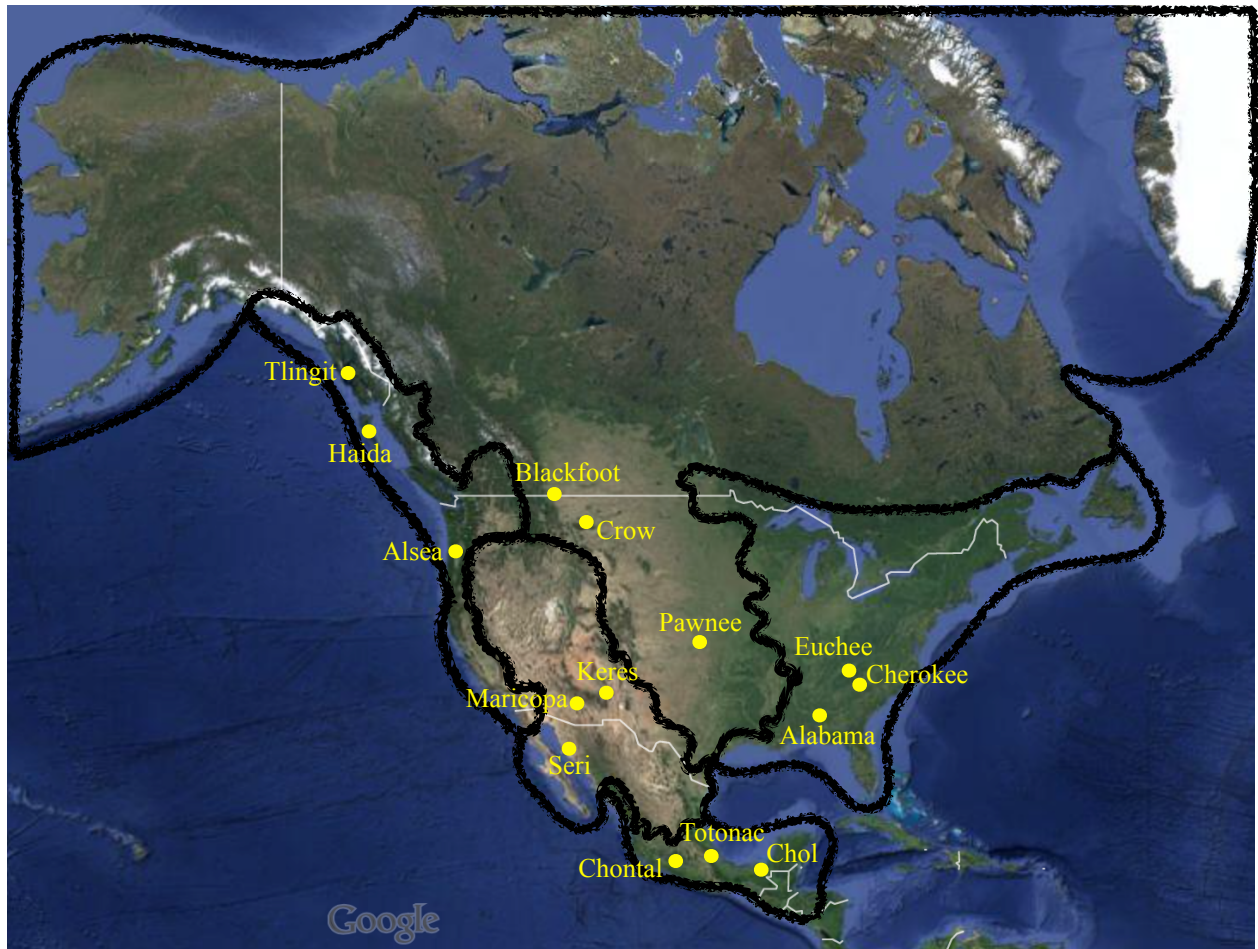
⁷ The new language came from a different genus in the interest of time. I assume, as Dryer (2011) does, that typological properties of languages within a genus are more likely to be similar than dissimilar. Therefore if the first language selected at random from a genus does not have split-alignment, then it is more likely that others in that genus also will not. Finding a split-alignment language is therefore more likely in a new genus. In practice, though, most languages do have split-alignment; the main reason for moving on to a different genus is that the first one selected was either extinct or lacking in descriptive resources. Rather than spending time researching each individual language within such a genus, it was more efficient to just move on to the next one.

sample. The complete list is shown in (5), organized according to region; a map showing the approximate locations of each language is shown in (6).

(5) Sampled languages

<u>Region</u>	<u>Language</u>	<u>Genus</u>
1	Blackfoot	Blackfoot
	Crow	Missouri River Siouan
	Pawnee	Caddoan
2	Alsea	Alsea
	Haida	Haida
	Tlingit	Tlingit
3	Laguna Keres	Keresan
	Maricopa	Cochimí-Yuman
	Seri	Seri
4	Alabama	Muskogean
	Cherokee	Southern Iroquoian
	Euchee	Yuchi
5	Chol	Mayan
	Lowland Chontal	Tequistlatecan
	Misantla Totonac	Totonacan

(6) Map showing approximate locations of sampled languages



The sample design offers two major advantages. First, because of genetic stratification, the languages in the sample are guaranteed either to be unrelated to one another or, if they are related, that their common ancestor was spoken at least 3500 years ago. This means that if the same split inducers occur in different languages, they are unlikely to have been inherited from a common ancestor. Second, because of areal stratification, each regional triplet of languages has a very low likelihood of areal contact with the other twelve languages. This means that if the same split inducers occur in languages in different regions, they are unlikely to have been diffused.

3.2.5 Precision of the sample

The sampling strategy comes with a cost, which is that precision will be low due to the unequal density of genera per region. Since the same number of genera are taken from each region, the regions with many genera are contributing the same number of units as the regions with fewer genera. This means that the selection probability for any individual genus is extremely low in dense regions, but extremely high in sparse regions. The usual way of handling such a situation is to sample the regions proportionally so that the number of genera selected from each region is proportional to the region's population. But doing so automatically voids the entire motivation behind areal stratification: the point is to avoid the chances of areal diffusion contaminating the sample, but if more genera are selected from certain regions, then it increases the likelihood that two randomly selected genera will contain languages with diffused traits. I discuss this issue more in Chapter 5 with respect to the specific results.

3.3 Data collection

Data collection was started in concurrence with the random selection procedure so that variation in the availability of resources could immediately be accommodated in the sample itself. Thus the sample was in constant flux until suitable data could be found on all selected languages. In §3.3.1, I discuss the kinds of resources I used to access the sampled languages and I discuss some remaining bias in §3.3.2.

3.3.1 Source materials

The materials I used to collect data on alignment patterns in each language were linguistic publications (or occasionally unpublished manuscripts). Usually, these were descriptive grammars in the form of doctoral dissertations or books. To supplement the grammars, I also used journal articles and chapters from edited volumes. Personal contact with experts in certain languages was also sometimes necessary. Unfortunately, it was not possible to undertake field work for this project.

It is worth noting that not all grammars are equal. Some are more clearly written, more comprehensive, more logically organized, and less theory dependent than others. Some are very old and written in now antiquated linguistic terminology. In some cases, an alignment property would be described in general terms, but without relevant examples, making it impossible to verify that the property actually exists in the language. At times, the poor quality of a grammar was justification for throwing a language out of the sample if other resources were not available. This brings me to a brief discussion of the penultimate type of bias in the language sample: bibliographic bias, which I take up in the next subsection.

3.3.2 Remaining bias in the sample

The fact that some descriptive grammars are better than others is, of course, not the fault of the described language. So when it affects the contents of a language sample, it is really a consequence of bibliographic bias. Bakker (2011:100) supposes that adequate grammars exist for only about one-third of the existing languages; I do not imagine that North American languages are especially privileged in this regard. The only way to completely avoid bibliographic bias is to

have at hand equally comprehensive descriptions of every language in the sampling frame. I guess this is wishful thinking.

A final type of bias remaining in the sample is what might be called ontological bias, and as far as I can tell, is totally unavoidable at present. The ontological bias has to do with the fact that the languages which exist today make up a tiny fraction of the languages which have ever existed. This is potentially problematic when we extrapolate from available language data to claims about “language.” Linguists generally seem to be unworried by this. Languages of the past are assumed to be not fundamentally different from the languages of today, as described by the General Uniformity Principle (Lass 1997:26). In other words, the fraction of extant languages should still be statistically representative of the historical population. Still, if we consider that the extant languages worldwide must be infinitesimal compared to the extinct ones, it is not hard to imagine that some linguistic phenomena are lost in the sift. In an interesting estimation of worldwide historical language counts, Bakker (2011:101) figures that about 240,000 languages have existed in the last 40,000 years,⁸ which would make a current language population of 6000 a mere 2.5% of the total. But since access to the vast majority of extinct languages, particularly in deep time, is currently impossible, there really is no other choice but to use what we have.

3.4 Summary of Chapter 3

In this chapter, I have described in detail the methodology employed for sampling languages to represent genera within North America. Genera were defined as familial groupings

⁸ This figure is based on the assumption that an average of 6000 languages were spoken at any given moment between 40,000 years ago and the present and that “a language changes every 1000 years to the extent that we may conceive of it as a ‘new’ language” (Bakker 2011:101).

with a time-depth of at least 3500 years and were taken from Dryer (2011) and Campbell (1997). The geographic space of North America was divided into five non-overlapping areas with a low likelihood of areal diffusion occurring between them. Each genus was assigned to exactly one geographic area. Using an online random sequence generator, a three genera were selected from each region, and from each genus one language was selected, for a total of fifteen genera and fifteen languages. This procedure is motivated by an effort to avoid genetic and areal bias.

4 DATA FROM SAMPLED LANGUAGES

4.1 Introduction to Chapter 4

Chapter 4 summarizes the patterns of split-alignment found in the languages sampled. Trudging through a large set of linguistic examples is often a dry and tedious enterprise, but it is crucial that I am able to show attested instances of the patterns claimed to occur in each language so as to facilitate independent verification. This chapter proceeds rather formulaically: starting in §4.2, each sampled language is briefly introduced and then its alignment behavior is summarized. The contents of each subsection are mostly the same; minor differences are due to different aspects of alignment relevant to the particular language being discussed. The order of languages is first by region (1–5) and within each region, alphabetically by language name.

4.2 Split-alignment in the sampled languages

4.2.1 Split-alignment in Blackfoot

4.2.1.1 Blackfoot the language

Blackfoot is an Algonquian language spoken by around 3300 people in southern Alberta and northern west Montana. Within Algonquian, Blackfoot comprises one of the main branches of the Algonquian subfamily.¹ Its relatives include Cheyenne, Arapaho, Ojibwa-Potawatomi, Fox,

¹ The Algonquian languages not in Algonquian are the two Californian languages in the Ritwan subfamily: Wiyot (extinct) and Yurok (revitalizing).

Shawnee, Miami-Illinois, Cree-Montagnais, and several languages in the Eastern Algonquian group (Campbell 1997). It is grouped in Region 1: Arctic, Subarctic, and Plains.

4.2.1.2 Overview of Blackfoot alignment

Like other Algonquian languages, Blackfoot has a robust direct-inverse alignment system in which the hierarchically highest argument is indicated as a verbal affix and then a second morpheme specifies the directionality of action. Intransitive and transitive verbs both have separate paradigms for grammatically animate and inanimate arguments, but these do not change the alignment system. A split occurs when an O argument is non-particular and is therefore not encoded in the agreement morphology; the result is a transitive clause which is marked identically to an intransitive one, a nominative-accusative pattern.

4.2.1.3 Intransitive argument marking in Blackfoot

In Algonquian linguistics, animacy is the basis for the two grammatical gender categories of nominals. Animate nouns include all sentient beings (humans and animals), most plants, spirits, as well as some culturally significant nouns. Inanimate nouns include all others. These are formally grammatical categories.

The affixes for animate intransitive S are shown in (1) and brief examples are given in (2) and (3).

(1) Animate S affixes (modified from Frantz 2009:16-23)

	<u>SG</u>	<u>PL</u>
1	nit-	+ -hpinnaana (-o'pa for inclusive)
2	kit-	+ -hpoaawa
3.PROX	-wa	-yi
3.OBV	-yini	-yi

(2) First-person singular animate S (Frantz 2009:22, ex. h)²

nit-áak-ahkayi
1-fut-go.home
'I'm going home.'

(3) Third-person singular animate S (Frantz 2009:23)

sspítáá-**wa**
be.tall-**3**
'S/he is tall.'

The *inanimate* intransitive paradigm necessarily consists only of third-person forms because first- and second-person arguments must be animate. The pronominals in the inanimate intransitive paradigm are identical to the third-person animate forms, so I will not go into further detail about them here.

4.2.1.4 Direct-inverse alignment in Blackfoot

The direct-inverse system of Blackfoot is quite rich and it will not be useful to detail all of its various permutations here; the differences, as far as I can tell, do not induce an alignment split. For illustrative purposes, a few brief examples showing how the pattern works will suffice. In a transitive construction, there are two arguments. Their marking refers to an indexability hierarchy, shown below in (4).

² Some examples given by Frantz (2009) are not glossed while others are. For consistency, I have tried to gloss all the examples used here based on the descriptions and explanations Frantz provides in his grammar and co-authored (1995) dictionary. Of course, it's possible that I have missed some morphemes where segmentation is possible.

- (4) Blackfoot indexability hierarchy for person marking³ (Bliss & Jesney 2005)
 $2 > 1 > 3\text{PROX} > 3\text{OBV} > 3\text{INAN}$

In constructions involving one third-person argument, both the SAP and non-SAP argument are overtly realized on the verb; this is possible because third-person arguments are marked by a suffix while the other persons are marked by prefixes, so there is no competition for morphological slots. When the directionality of action is down the hierarchy (a higher ranking A acts on a lower ranking O), then the direct morpheme is used, as in (5). When directionality goes the other way, with a lower ranking A acting on a higher ranking O, then the inverse morpheme is used, as in (6).⁴

- (5) **Direct** marking with a non-SAP argument (Frantz 2009:52, ex. a)

nit-sik-ákomimm-**a**-wa nit-ána
 1-INTENS-love-**DIR**-3SG 1-daughter
 ‘I love my daughter.’

- (6) **Inverse** marking with a non-SAP argument (Franz 2009:56, ex. c)

nit-sik-ákomimm-**ok**-a nit-ána
 1-INTENS-love-**INV**-3SG 1-daughter
 ‘My daughter loves me.’

When both arguments are SAPs, only the highest ranking one shows up on the verb.

Direct and inverse morphemes again are used to show directionality of action, but their form is different from those used in non-SAP constructions, as shown below in (7) and (8).

- (7) **Direct** marking with only SAP arguments (Bliss & Jesney 2005:96, ex. 4a)

kit-iino-**o**
 2-saw-**DIR**
 ‘I saw you.’

³ The indexability hierarchy in Algonquian languages is well known for ranking second person higher than first person. In Blackfoot, the $2 > 1$ ranking holds for the choice of argument prefix, but for the actual directionality morpheme, as well as the argument suffixes (which encode number as well as person), the ranking is $1 > 2$ (Macaulay 2005).

⁴ Due to regular phonological processes, the intensifier morpheme in both (5) and (6) is underlyingly *ik-* and the 3SG suffix in (6) has had the glide deleted.

- (8) **Inverse** marking with only SAP arguments (Bliss & Jesney 2005:96, ex. 4b)
kit-iino-**oki**
2-saw-INV
'You saw me.'

Direct-inverse alignment is typically discussed only in reference to transitive constructions, but since alignment is defined here as a relationship between all three grammatical roles, it is necessary to point out that direct-inverse alignment implies that S will always have a distinct marking pattern from transitive arguments. The reason is simply that intransitive predicates do not have two arguments and therefore do not invoke the indexability hierarchy nor require a directionality morpheme. Even though the affix marking the argument itself is the same in the intransitive construction in (1) as in the transitive one in (5), since there is no directionality morpheme on intransitives, S is marked differently by virtue of not involving a direct or inverse morpheme.

4.2.1.5 Nominative-accusative alignment in Blackfoot

The nominative-accusative pattern occurs when an O role is overtly realized as a full noun, but is "non-particular," to use Frantz's term. Non-particular nouns do not refer to a specific entity, but rather identify a class of objects without a unique referent. They are usually translated into English as an indefinite noun or a mass noun. Necessarily, these will be third-person arguments because first- and second-person arguments are always referential. As described above, third-person Os will normally be realized on the verb with the suffix *-wa* and a directionality morpheme. Non-particular Os are not marked on the verb and do not induce use of a directionality morpheme; the verbs are morphologically identical to intransitive ones, but since

the clauses contain a full noun O, the constructions cannot be interpreted as having a single argument.⁵

- (9) Non-particular **O** (Bliss 2009:62, ex. 10b)

nit-ikamo's-i **ihthpomo'pi**
 1-steal-ai **money**
 'I stole money.'

- (10) Non-particular **O** (Frantz 2009:102, ex. j1)

iihpomm-aa-wa⁶ **ónniki-i**
 buy-ai-3sg **milk-nonpart**
 'He bought milk.'

And here is a transparent triad of examples from Blackfeet, the dialect spoken in Montana:

- (11) Particular **O** is marked on the verb (Hardy 1989:35, ex. 11a)

nit-awayaki-a-**w** ama-aiksiniw
 1sg-hit-dir-**3sg** 3sg.this-pig
 'I hit this pig.'

- (12) Non-particular **O** is not marked on the verb (Hardy 1989:35, ex. 11c)

nit-awayaki-aaki **aiksinii**
 1sg-hit-intr **pig**
 'I hit a pig.'

- (13) Marking of a non-particular **O** is ungrammatical (Hardy 1989:35, ex. 11b)

*nit-awayaki-a-**w** aiksinii
 Intended meaning: 'I hit a pig.'

The constructions in (9), (10), and (12) indicate the A relation in the same way that the S relation is indicated in (2) and (3), and the O relation is indicated differently by virtue of not

⁵ Blackfoot is sometimes described as a language with "morphological transitivity," determined by agreement morphology, as well as "syntactic transitivity," determined by full noun argument realization. Under such a view, the constructions with non-particular Os would be morphologically intransitive but syntactically transitive. Since I am adopting a valency-centered definition of transitivity, these are necessarily transitive for me.

⁶ This -wa indexes A, not O.

appearing on the verb. This is a nominative-accusative pattern induced by the particularness of O.

4.2.1.6 Summary of split-alignment in Blackfoot

Alignment in Blackfoot is split between the direct-inverse and nominative-accusative patterns on the basis of the particularness of O. Transitive constructions with particular Os have a direct-inverse alignment pattern. Those with non-particular Os do not mark the O on the verb, and have argument inflection in which A and S are marked the same, a nominative-accusative pattern. The table in (14) summarizes the system.

(14) Tabular summary of split-alignment in Blackfoot (plurals omitted)

<u>Person</u>	<u>A</u>	<u>S</u>	<u>O</u>	<u>A & O</u>	<u>Pattern</u>	<u>Inducer</u>
1		nit-			direct-inverse	particularness (O = particular)
2		kit-				
3.PROX		-wa				
3.OBV		-yini				
1⇒2				kit...oki		
2⇒1				kit...o		
1⇒3				nit...a...wa		
3⇒1				nit...ok...wa		
⋮						
1	nit-	nit-	Ø		nominative-accusative	particularness (O = non-particular)
2	kit-	kit-	Ø			
3.PROX	-wa	-wa	Ø			
3.OBV	-yini	-yini	Ø			

4.2.2 Split-alignment in Crow

4.2.2.1 Crow the language

Crow, known endonymically as B́ilukaailaaú, is a Siouan-Catawban language spoken by several thousand people in southern Montana (Lewis et al. 2014, Graczyk 2007). Together with Hidatsa, it is a member of the Missouri River Siouan branch of Siouan. Crow is somewhat more distantly related to Dakota, Omaha, Osage, and Hock. It is grouped in Region 1: Arctic, Subarctic, and Plains.

4.2.2.2 Overview of Crow alignment

The basic alignment system of Crow is active-stative, with subjects of transitive and intransitive active verbs taking Set A pronominal prefixes and subjects of intransitive stative verbs and objects of transitive verbs taking prefixes from Set B. Third person arguments are not overtly marked, so alignment is neutral in the third person. The split is induced by three conditions: (1) verbal semantics; (2) lexical specification of exceptional verbs (as active or stative); and (3) SAP status.

4.2.2.3 Split-S alignment in Crow

In the first and second persons, there is an overt marker indicating the roles A and S_a, taken from a conjugational set termed “Set A,” and a different marker indicating the roles S_o and O, taken from the conjugational set “Set B.” The markers are verbal prefixes and referred to by Graczyk (2007) as pronominals. The table in (15) shows a simplification of the two pronominal sets. Phono- and morphophonological variants are not shown.

(15) A and B pronominal prefixes (Graczyk 2007:121, Table 6.1)

	<u>A-set</u>	<u>B-set</u>
1SG	baa-	bii-
2SG	dá(a)-	dii-
3SG	Ø	Ø
1PL	baa- + PL	balee
2PL	dá(a)- + PL	dii- + PL
3PL	Ø + PL	Ø + PL

In (16), an active intransitive verb takes the Set A pronominal prefixes.

(16) Active **S** with Set A pronominal prefixes (Graczyk 2007:179, ex. 2)

(a) **baa**-xalússhi-k

1A-run-DECL

‘I was running’

(b) **da**-lée-k

2A-go-DECL

‘you went’

(c) **Ø**-disshí-k

3A-dance-DECL

‘he was dancing’

In (17), a stative intransitive verb takes the Set B pronominal prefixes.

(17) Stative **S** with Set B pronominal prefixes (Graczyk 2007:179, ex. 3)

(a) **bii**-apáa-k

1B-cold-DECL

‘I am cold’

(b) **dii**-háchka-k

2B-tall-DECL

‘you are tall’

(c) **Ø**-baakuhpáa-k

3B-sick-DECL

‘he is sick’

In (18), a transitive verb takes a Set A prefix for indexing the subject and a Set B prefix for indexing the object.

(18) Transitive verb with Set A for **A** and Set B for **O** (Graczyk 2007:178, ex. 1)

(a) **dii-waa**-lichí-k
2B-1A-hit-DECL
'I hit you'

(b) **bii-láa**-lichí-k
1B-2A-hit-DECL
'you hit me'

(c) **bii-Ø**-lichí-k
1B-3A-hit-DECL
'he hit me'

(d) **Ø-baa**-lichí-k
3B-1A-hit-DECL
'I hit him'

Whether a Crow verb falls into the active category or the stative one is based, for the most part, on verbal semantics. Graczyk (2007) does not expressly define the semantic requirements for active or stative categorization, but his examples and discussion line up with the canonical ones as outlined, for instance, by Van Valin (1990) and Mithun (1991). There are, however, a handful of exceptional verbs which fall into the opposite of the expected classes. For instance, *biíshi* 'tell a lie' is stative, while *shée* 'die,' *chilii* 'be afraid,' and *ili* 'be alive' are active. The alignment behavior of these verbs must therefore be lexically specified.

4.2.2.4 Summary of split-alignment in Crow

A and S are marked alike, differently from O, when their verbs are semantically active or lexically specified as active. This is a nominative-accusative pattern. S and O are marked alike, differently from A, when their verbs are semantically stative or lexically specified as stative. This is an ergative-absolutive pattern. These two alignment patterns only occur in the first and second

persons. In the third person, arguments are not overtly marked, and so alignment is neutral. The table in (19) summarizes the various alignment patterns and the conditions which induce them.

(19) Tabular summary of split-alignment in Crow⁷

<u>Person</u>	<u>A</u>	<u>S</u>	<u>O</u>	<u>Pattern</u>	<u>Inducer</u>
1	baa	baa	bii	split-S _{nom-acc}	SAP status (+SAP) & verbal semantics (active), SAP status (+SAP) & lexical specification (active)
2	dá(a)	dá(a)	dii		
1	baa	bii	bii	split-S _{erg-abs}	SAP status (+SAP) & verbal semantics (stative), SAP status (+SAP) & lexical specification (stative)
2	dá(a)	dii	dii		
3	Ø	Ø	Ø	neutral	SAP status (–SAP)

4.2.3 Split-alignment in Pawnee

4.2.3.1 Pawnee the language

Pawnee is a Caddoan language now spoken Oklahoma, though historically the Pawnees are from Nebraska and Kansas (Parks & Pratt 2008). The language is joined in the Caddoan family by Caddo, Wichita, Kitsai (now extinct), and Arikara, the latter being its closest relative. There are two dialects of Pawnee: South Band and Skiri. They show phonetic (and small phonological) differences, but are mutually intelligible (Parks 1976:1). There are no remaining fluent speakers, though heritage speakers remain and teaching materials are promoted. It is grouped in Region 1: Arctic, Subarctic, and Plains.

⁷ Where there are multiple split inducers, this is represented symbolically in the inducer column. The ampersand represents split inducers which must cooccur in order to induce the pattern. The comma represents another (set of) split inducer(s) for the same pattern.

4.2.3.2 Overview of Pawnee alignment

Pawnee has several alignment splits on the basis of SAP status, individuation, and animacy. The most obvious split occurs in the person-marking paradigm, which is nominative-accusative for SAPs and neutral for non-SAPs. A nearly complementary scenario occurs in the number-marking paradigm such that SAPs receive the same plural marker (and thus show neutral alignment), while non-SAPs have different markers for the roles A, S, and O. Depending on whether plural third-person arguments are individuated or unitary,⁸ or whether they are animate or inanimate, third-person plural marking is either ergative-absolutive or tripartite.

4.2.3.3 Person marking in Pawnee

Person marking in Pawnee is achieved via affixation of a person morpheme to a verbal complex which also includes affixes for mode, aspect, and often number. Alignment is nominative-accusative for SAPs and neutral for non-SAPs, as indicated in the table in (20).

(20) Person-marking paradigm in Pawnee (format modified from Parks 1976:164, Table 6)

	<u>A</u>	<u>S</u>	<u>O</u>
1	-t-	-t-	-ku-
2	-s-	-s-	-a-
3	Ø	Ø	Ø

In the first and second persons, A and S are marked the same while O is different, a nominative-accusative pattern. In the third person, all three relations are zero-marked, making alignment neutral. The three examples in (21)-(23) show the pattern with first-person arguments.

⁸ These concepts are elaborated on below, but I have adopted the terms themselves from Parks' work. They appear to be quite similar (maybe identical) to a distributive vs. collective distinction familiar from other languages. It may be worth exploring the extent to which these categories match those proposed for other languages.

- (21) First-person **A** (Parks & Pratt 2008:35)
 ta-**t**-raar-raspii-Ø
 1.INDIC-**1**.AGTV-3.INAN.PL.PATV-look.for-PFV
 ‘I looked for them.’
- (22) First-person **S** (Parks & Pratt 2008:38)
 ta-**t**-uks-at-Ø
 INDIC.1-**1**.AGTV-AOR-go-PFV
 ‘I went.’
- (23) First-person **O** (Parks & Pratt 2008:39)
 raa-ku-ti-Ø-a-**ku**-ri-ut-i-awahc-raawii’at-Ø
 just-INDF-INDIC-3.AGTV-PV-**1**.PATV-PHYBEN-PV-SEQ-at.least-help-PFV
 ‘Oh, you should help me!’

The constructions in (24)-(26) show third-person arguments with the neutral pattern.

- (24) Third-person **A** (Parks & Pratt 2008:43)
 ar-ri-Ø-kiss-raar-takipuh-Ø
 EVD-3.CONT.AGTV-**3**.AGTV-bone-PL-gather-PFV
 ‘He gathered the bones.’
- (25) Third-person **S** (Parks & Pratt 2008:46)
 ti-Ø-kikak-Ø
 INDIC-**3**.AGTV-cry-PFV
 ‘He cried.’
- (26) Third-person **O** (Parks & Pratt 2008:46)
 ra-s-ku-hisaask-a
 INF-2.AGTV-INF-call-SUB1
 ‘for you to call him’

4.2.3.4 Number marking in Pawnee

The Pawnee number-marking system is split between three types (neutral, tripartite, and ergative-absolutive) on the basis of person, subject individuation, and object animacy. Because grammatical role is encoded in certain number markers, this is not just a case of split-number, but is split-alignment as well. The plural morpheme for A, S, and O is *rak-* in the first and second

persons.⁹ In the third person, S and O both take *rar-* if S is individuated and O is inanimate. A takes *si...ir*. This is the ergative-absolutive scenario, where $(S = O) \neq A$. However, if S is unitary, then it takes *ir* as the plural marker; if O is animate then it takes *ak* as the plural marker. Thus if S is unitary and/or if O is animate, then $S \neq O \neq A$ and the pattern is tripartite. The table in (27) shows the number-marking morphemes.

(27) Pawnee plural markers for all argument-person combinations (Parks 1976)

	<u>A</u>	<u>S</u>	<u>O</u>
1	rak	rak	rak
2	rak	rak	rak
3	si...ir	rar, ¹⁰ ir ¹¹	ak, ¹² rar ¹³

Third-person S and O may be pluralized alike if S is individuated (the plural Ss *each* performed some action individually) and O is inanimate. Both arguments, if in the third-person, are pluralized with *rar-*. This is the ergative pattern, exemplified below. (28) and (29) show third-person plural S marked with *rar-*, while (30) and (31) show third-person plural O marked with *rar-*.

(28) Third-person plural individuated S marked with *rar-* (Parks 1976:167)

ti-**rar**-kukstakuk-Ø
INDIC-**PL**-run-INPRF
‘they (individually) ran’

⁹ Based on the definitions established in Chapter 2, this does not count as neutral *alignment* because the SAP plural markers are suffixes on whatever they pluralize and do not themselves encode any information about grammatical relations. The third-person plural markers do indeed encode grammatical relations. For this reason, the first- and second-person plural forms are not included in the summary table in (37) below and do not factor into the alignment description of Pawnee.

¹⁰ individuated plural

¹¹ unitary plural

¹² animate direct object plural

¹³ inanimate direct object plural

- (29) Third-person plural individuable **S** marked with *rar-* (Parks 1976:167)
 ti-**rar**-ke:ac-ar-Ø
 INDIC-**PL**-long-INCH-INPRF
 ‘they (individually) became long’
- (30) Third-person plural inanimate **O** marked with *rar-* (Parks 1976:169)
 ta-t-**rar**-huras-Ø
 INDIC-1.SUB-**PL**-find-INPRF
 ‘I found them (objects)’
- (31) Third-person plural inanimate **O** marked with *rar-* (Parks 1976:169)
 ta-t-**rar**-uh-Ø
 INDIC-1.SUB-**PL**-give-INPRF
 ‘I gave them to them’

The tripartite pattern occurs for third-person arguments in two scenarios: either **S** is individuable and **O** is animate (both must be true), or **S** is unitary (acting as a collective group) and it doesn’t matter whether **O** is animate or inanimate. Examples of individuable **S** and inanimate **O** are given above in (28)-(31); in these cases, the plural is *rar-*. If **S** is unitary, it is pluralized with *ir-*, as in (32) and (33).

- (32) Third-person plural unitary **S** marked with *ir-* (Parks 1976:167)
 ti-**ir**-kukstakuk-Ø
 INDIC-**PL**-run-INPRF
 ‘they (as a body or group) ran’
- (33) Third-person plural unitary **S** marked with *ir-* (Parks 1976:167)
 ti-**ir**-ke:ac-ar-Ø
 INDIC-**PL**-long-INCH-INPRF
 ‘they (group) became long’

If **O** is animate, it is pluralized with *ak-*, as in (34).

- (34) Third-person plural animate **O** marked with *ak-* (Parks 1976:169)
 ta-t-**ak**-huras-Ø
 INDIC-1.SUB-**PL**-find-INPRF
 ‘I found them (people)’

In the third person, A is pluralized by means of the dual number prefix *si-* and the indefinite subject/switch reference marker *ir-*, which are combined to form the discontinuous “compound” plural *si...ir*, as in (35).

- (35) Third-person plural A marked with *si...ir* (Parks 1976:168)
si-ti-ir-kusk-Ø
PL-INDIC-PL-put-INPRF
 ‘they put it’

Parks (1976:168) goes on to describe a further restriction to the plural third-person A-marking pattern: “If however, there is a first or second person object, *ir-* is not used, only *si-*.” An example is below in (36).

- (36) Third-person plural A and first-person object (Parks p.c. 1/2/14)
si-ti-ku-is-Ø
PL-INDIC-1.OBJ-hit-INPRF
 ‘they hit me’

Thus in constructions with a plural third-person A, that A is always marked with either *si...ir* (if the object is also third-person) or simply with *si* (if the object is first- or second-person).

4.2.3.5 Summary of split-alignment in Pawnee

There are four alignment patterns in Pawnee: nominative-accusative, ergative-absolutive, neutral, and tripartite. The nominative-accusative pattern pervades the singular SAP markers, while the neutral pattern occurs with singular non-SAPs, so number and SAP status are both factors. Alignment is also encoded in the non-SAP plural markers because different roles take different plural morphemes. Ergative-absolutive alignment shows up when S is individuated but O is inanimate. There are three different tripartite patterns, depending on different combinations

of the individuation of S and the animacy of O. The table in (37) summarizes the various Pawnee alignment patterns.

(37) Tabular summary of split-alignment in Pawnee

<u>Person/ number</u>	<u>A</u>	<u>S</u>	<u>O</u>	<u>Pattern</u>	<u>Inducer</u>
1SG	-t-	-t-	-ku-	nominative-accusative	SAP status (+SAP) & number (singular)
2SG	-s-	-s-	-a-		
3SG	Ø	Ø	Ø	neutral	SAP status (–SAP) & number (singular)
3PL	si...ir/si	rar	rar	ergative-absolutive	SAP status (–SAP) & number (plural) & individuation (S = individuating) & animacy (O = inanimate)
	si...ir/si	rar	ak	tripartite	SAP status (–SAP) & number (plural) & individuation (S = individuating) & animacy (O = animate)
	si...ir/si	ir	ak	tripartite	SAP status (–SAP) & number (plural) & individuation (S = unitary) & animacy (O = animate)
	si...ir/si	ir	rar	tripartite	SAP status (–SAP) & number (plural) & individuation (S = unitary) & animacy (O = inanimate)

4.2.4 Split-alignment in Alsea

4.2.4.1 Alsea the language

Alsea is an extinct language which also contained the Yaquina dialect; a relationship to other languages or families has not been confirmed. It was spoken on the Oregon coast until its last speaker died in 1942. From the audio recordings and language descriptions produced prior to extinction, modern linguists have continued to describe and analyze the language. It is grouped in Region 2: Northwest Coast, Plateau, and California.

4.2.4.2 Overview of Alsea alignment

Alignment in Alsea is split between nominative-accusative and ergative-absolutive, depending on nominal status. Verbal arguments are indicated by full nouns, free pronouns, and bound person agreement markers. The A relation is marked with an ergative morpheme prefixed to a determiner preceding a full noun phrase or cliticized to a pronoun. The pronominal agreement markers, however, are nominative-accusative.

4.2.4.3 Ergative-absolutive full (pro)nouns in Alsea

In most cases, full nouns are preceded by a morphologically complex determiner. The first morpheme is often a spatial deictic, and the second is often a referential marker; this basic template is exemplified in (38).

(38) Noun phrase preceded by **determiner** (Buckley 1988a:28, ex. 3)

k̲ = in	nu'ns-áa	m̲úhu'	ku-s	tsu'dái's
FUT=1SG.SUB	eat-TR	now	PROX-REF	salmon
'I will eat these salmon now.'				

The ergative marker is *qa-/xa-*.¹⁴ It occupies the first position of a determiner in which it is placed, either co-occurring with or replacing the deictic morpheme. The ergative marker is obligatory on determiners before full nouns in the A relation. The nominal status of the object (whether it is a noun, pronoun, or pronominal) is irrelevant. An example is shown in (39).

(39) Ergative **A** with nominal A and nominal O (Buckley 1988a:29, ex. 13)

tem̲úhu'	ya's-au-ʔyái'-nx	qa-s	mó'luptsiní'sla	a-s	qatsí'li'ʔ
and.now	say-DUR-ITER-3SG.OBJ	ERG-REF	coyote	DST-REF	wolf
'Then Coyote spoke to Wolf.'					

¹⁴ The choice between these allomorphs is not fully understood. Buckley (1988a:29) states that "there is no clear phonological pattern to the alternation," though Mithun (2010a) suggests a contact-induced borrowing scenario between Alsea and a neighboring Coosan language which might account for it.

The example in (39) shows the ergative marker on a determiner preceding a full noun subject of a transitive verb. The full noun *object* in (39) has a determiner without *qa-*, and the ergative marker is expectedly also absent from full noun subjects of intransitive verbs, as in (40) and (41).

- (40) Absolutive and nominal S (Buckley 1988b:11, ex. 6)

temúhu' kexk-ái' = slo' ts-hítə-k
and.then assemble-INCH=all DET-body-POSS
'Thereupon all the people assembled.'

- (41) Absolutive and nominal S (Buckley 1988b:17, ex. 35)

temúhu' qaúwi's ats-sáa-k tp-ái'-xa kwí'-ks = auk
and.then first DET-sister-POSS jump-INCH-CMPL canoe-ALL=inside
'First his elder sister jumped into the canoe.'

Pronouns in the A relation receive ergative marking and like nouns, are free morphemes.

The personal pronouns are usually only used for emphasis.¹⁵ Unlike full nouns, pronouns receive the ergative marker directly, rather than on a determiner, as in (42) and (43).

- (42) A pronoun with ergative clitic (Buckley 1988a:30, ex. 22)

xa = kts-áa qa = ní'x
2SG.OBJ=wear-TR ERG=2SG
'You put it on!'

¹⁵ The "emphatic" nature of the pronoun is probably not at the heart of licensing the ergative marker; rather the key feature is the presence of a free pronoun as opposed to a bound pronominal. Evidence for this comes from the observation that the interrogative morpheme, *ú'k*, which Gene Buckley (p.c. 1/26/14) agrees could be considered a pronoun, receives the ergative marker in transitive constructions, as in (i) and (ii).

- (i) *ú'k* in S function without ergative marking (Buckley 1988a:31, ex. 25)

k = uk^u ú'k = ən qauwí's ʒo'h-ái'-m
FUT=up who=Q first climb-INCH-INTR
'Who will climb up first?'

- (ii) *ú'k* in A function with ergative marking (Buckley 1988a:31, ex. 26)

k = qa-ú'k = ən múhu' pé'x-ai' a-s məhayát-au
FUT=ERG-who=Q now visit-INCH DIST-REF old-DIM
'Who will go now to the little old man?'

In any case, it is difficult to discern from the old Alsea texts and Frachtenberg's (1920) English translations whether there are any instances of free personal pronouns whose pragmatic context is not emphatic.

- (43) A pronoun with ergative prefix (Buckley 1988a:31, ex. 23)

qáltə hí'kə **xa**-qátsə ʒə-mk'é'n-au-x
 always just **ERG**-3SG TR-hit-DUR-CMPL
 'He always hit it.'

An S pronoun does not receive ergative marking, as in (44).

- (44) S pronoun with no ergative clitic (Buckley 1989:21, ex. 5)

k̥ = in qʷám-aw-tx-am tə-**qʷún**
 FUT=1SG.SUB behind-DUR-HAB-INTR DEI-1SG
 'I'll walk behind.'

So the full nouns and free pronouns have a distinct A form by virtue of the ergative morpheme on a preceding determiner (for nouns) or directly on the pronoun.

4.2.4.4 Nominative-accusative pronominals in Alsea

Pronominal agreement markers align in a nominative-accusative fashion. Pronominal subjects are marked by clitics on the first word in the clause, no matter its lexical category; these clitics encode person and number. (38) above, repeated below as (45), shows a transitive construction in which A is a pronominal and does not receive ergative marking.

- (45) Pronominal A with no ergative marker (Buckley 1988a:30, ex. 16)

k̥ = **in** nu'ns-áa m̥u'hu' ku-s tsu'dái's
 FUT=**1SG.SUB** eat-TR now PROX-REF salmon
 'I will eat these salmon now.'

In a more revealing example, the third-person pronominal A in (46) has the same shape as the third-person pronominal S in (47). The third-person pronominal O in (46) is different from either of these.

- (46) Pronominal **A** and pronominal **O** (Buckley 1988a:28, ex. 2)
 tem = **iŋx** pxe'ltsu's-ái'-**nx** a-ts-ŋí'a-**k̥** = iŋx
 and = **3PL.SUB** ask-INCH-**3SG.OBJ** DST-3POSS-mother-3POSS=3PL
 'They asked their mother.'
- (47) Pronominal **S** (Buckley 1988a:28, ex.1)
 tem = **iŋx** múhu' ay-ái'
 and = **3PL.SUB** finally go-INCH
 'Finally they started out.'

The examples in (45)-(47) reveal the nominative-accusative pattern. The pronominal agreement markers are the same for S in (47) as for A in (46). The marker for a third-person object in (46) is *-nx*, and though it has a singular referent, number is not encoded in the object markers and so it would be the same for a plural object.¹⁶ The chart in (48) shows the nominative-accusative pattern for all three persons.

- (48) Pronominals (modified from Buckley 1988, Kinkade 2005:63, Mithun 2010a:92).
- | | <u>A</u> | <u>S</u> | <u>O</u> |
|-----|----------|----------|----------|
| 1SG | =an, =in | =an, =in | 1 -mts |
| 2SG | =aɣ | =aɣ | 2 -u: |
| 3SG | Ø | Ø | 3 -nx |

Since the A and S columns are identical, with a different O column, these morphemes show nominative-accusative alignment.

4.2.4.5 Summary of split-alignment in Alsea

All free nominal forms (nouns and pronouns) show ergative-absolutive alignment. The bound pronominals show nominative-accusative alignment. Thus the split is conditioned by

¹⁶ I have retained the gloss used by Buckley (1988). Although he states that number is not realized in the object markers, he has nonetheless included a number distinction in the gloss.

nominal status, with nouns and pronouns having one alignment pattern and pronominals having another. The table in (49) summarizes the system.

(49) Tabular summary of split-alignment in Alsea

<u>Person</u>	<u>A</u>	<u>S</u>	<u>O</u>	<u>Pattern</u>	<u>Inducer</u>
(all)	qa=/xa=	Ø	Ø	ergative-absolutive	nominal status (non-pronominal)
1	=an/=in	=an/=in	-mts	nominative-accusative	nominal status (pronominal)
2	=aɣ	=aɣ	-u:		
3	Ø	Ø	-nx		

4.2.5 Split-alignment in Haida

4.2.5.1 Haida the language

Haida is a language consisting of two dialects (Masset and Skidegate) spoken respectively on the northern and southern portions of Haida Gwaii, also known as the Queen Charlotte Islands, which lie in the North Pacific Ocean off the western coast of British Columbia. It is moribund, with fewer than 60 speakers, and an isolate (Lewis et al. 2014). It is grouped in Region 2: Northwest Coast, Plateau, and California.

4.2.5.2 Overview of Haida alignment

Haida has a split-S system for first-person pronouns (both singular and plural) and second-person singular pronouns. Second-person plural and third person pronouns show neutral alignment. The semantic features which determine the pronoun choice for any intransitive verb are [AGENCY] and [CONTROL]. This characterization comes almost exclusively from Hori's (2008) analysis. The split-S patterns are conditioned both by verbal semantics (some verbs are

always of one type or another, based on their [AGENCY] and [CONTROL] values) and by participant semantics (some verbs may be of either type; this is a fluid-S subsystem).

4.2.5.3 Split-S alignment in Haida

In an effort to avoid implication, Hori (2008) calls the two categories of pronoun α and β (rather than, say, *agent* and *patient*, respectively). The α -set corresponds roughly (with certain peculiarities) to agents and the β -set to patients. The table below shows the personal pronouns in Haida.

(50) Personal pronouns in Haida (Hori 2008:27, modified from Table 1)¹⁷

	α	β
1SG	łaa (łə=)	dii
1PL	t'alaan	?iitl'ə
2SG	daa	dən
2PL	dalaan	
3	łaa (łə=)	

Hori (2008) treats agency and control as the formal semantic features [AGENCY] and [CONTROL], respectively. Under his definition, “a verb has the feature [AGENCY] if it requires a participant that performs the activity or instigates the situation denoted by the verb” (ibid:37) and gives three tests which can be used to determine the agency of a particular verb; the verb must pass at least two of the three tests in order to qualify as having [AGENCY].¹⁸ Hori does not explicitly define [CONTROL], but does give two tests, at least one of which must be passed by a

¹⁷ Parenthetical forms indicate bound (proclitic) versions of the pronouns.

¹⁸ **For [AGENCY]:** TEST 1: “The verb can be used as a predicate in replying to the question ‘What was/is/will be S doing?’, or ‘What did S do?’” TEST 2: “The verb can be used as a predicate in an imperative, as a request.” TEST 3: “The verb can be used as a predicate in replying to the question ‘What is going on?’ or ‘What happened?’” (Hori 2008:38).

verb in order for it qualify as having [CONTROL].¹⁹ In featural terms, verbs which always take pronouns from the α -set are designated as [+AGENCY, +CONTROL]; verbs which always take pronouns from the β -set are [-AGENCY, -CONTROL]. A handful of verbs, which Hori (2008) calls *intermediate verbs* can take either set of pronouns, which he accounts for formally by assigning to them either + or – values for [AGENCY] and gradient feature values for [CONTROL]. He suggests that for the intermediate verbs, the choice may be due to the speaker’s perception of these semantic features in the context of a particular action (2008:46-47). Examples of intermediate verbs include those meaning ‘vomit,’ ‘dream,’ ‘stagger,’ and ‘be angry’ (Hori 2008:34). The sentences in (51)-(56) show the use of the α pronouns for both A and S relations.

- (51) First-person singular **A** (Hori 2008:28, example 5a)

1aa = ?uu tɬəway tɬəguɬga-gən
1SG.α=FOC boat[DEF] build-PAST
 ‘I built the boat.’

- (52) First-person singular **S** with the α set (Hori 2008:30, example 11a)

1aa = ?uu ɬaw-gəŋ-giin-’i
1SG.α=FOC fish-HAB-PAST-INFO
 ‘I used to fish.’

- (53) First-person plural **A** (Hori 2008:29, example 5c)

ciina **t’alaan** dlən-gən
 fish **1PL.α** wash-PAST
 ‘We cleaned the fish.’

- (54) First-person plural **S** with the α set (Hori 2008:30, example 11c)

Vancouver = gu **t’alaan** ɬgangulɬa-gən
 Vancouver=PP **1PL.α** work-PAST
 ‘We worked in Vancouver.’

¹⁹ **For [CONTROL]:** TEST 1: “The verb can take the suffix -t’ajəŋ ‘to try to V’.” TEST 2: “The sentence in which the verb is used as a predicate can be modified by a manner adverb *k’udɬan* ‘on purpose’” (Hori 2008:39).

- (55) Second-person singular **A** (Hori 2008:29, example 6a)
daa = gwaa dəwjay qyaan̩a
2SG.α=INTER cat[DEF] see[EVD]
 ‘Did you see the cat?’
- (56) Second-person singular **S** with the α set (Hori 2008:31, example 12a)
 daGaɬ = gwaa **daa** xyaala
 yesterday=INTER **2SG.α** dance[EVD]
 ‘Did you dance yesterday?’

In all sentences above, the choice of pronoun comes from the α -set, no matter whether the subject is A (the odd-numbered examples) or S (the even-numbered examples). The sentences below in (57)-(62) show the use of β pronouns for both O and S relations.

- (57) First-person singular **O** (Hori 2008:29, example 8a)
dii ‘laa qin̩-gən
1SG.β 3 see-PAST
 ‘He/she saw me.’
- (58) First-person singular **S** with the β set (Hori 2008:31, example 14a)
 gəm **dii** q’ud-Gən̩-gən
 NEG **1SG.β** hungry-NEG-PAST
 ‘I was not hungry.’
- (59) First-person plural **O** (Hori 2008:30, example 8b)
?iitl’ə nəŋ kun-gən
1PL.β somebody hit-PAST
 ‘Somebody hit us (on vehicle).’
- (60) First-person plural **S** with the β set (Hori 2008:31, example 14b)
?iitl’ə hit’agan ?ina-gən
1PL.β a.little.while.ago grow-PAST
 ‘We were young.’
- (61) Second person singular **O** (Hori 2008:30, example 9a)
dəŋ xaagay xidxiidən
2SG.β dog[DEF] chase[PAST]
 ‘The dog chased you.’

- (62) Second-person singular **S** with the β set (Hori 2008:31, example 15a)

dəŋ = gwaa **giitgi**
2SG.β=inter be.ready
 ‘Are you ready?’

In all sentences above, the choice of pronoun comes from the β -set, no matter whether the relation is O (the odd-numbered examples) or S (the even-numbered examples).

4.2.5.4 Neutral alignment in Haida

The second-person plural pronouns are all the same whether in A, O, or S relation, and the semantics of intransitive verbs is irrelevant, as in (63)-(66).

- (63) Second-person plural **A** (Hori 2008:29, example 6b)

daGalayGa = gwaa **dalaan** ?axada-?yaana
 the.next.day=INTER **2PL** seine-outward[EVD]
 ‘Did you (pl) go seining the next day?’

- (64) Second person plural **O** (Hori 2008:30, example 9b)

xaagaay = ?uu **dalaan** xidxiidən
 dog[DEF]=FOC **2PL** chase[PAST]
 ‘The dog chased you (pl).’

- (65) Second-person plural **S** with α semantics (Hori 2008:31, example 12b)

daGał = gwaa **dalaan** xyaala
 yesterday=INTER **2PL** dance[EVD]
 ‘Did you (pl) dance yesterday?’

- (66) Second-person plural **S** with β semantics (Hori 2008:31, example 15b)

dalaan = gwaa **giitgi**
2PL=INTER be.ready
 ‘Are you (pl) ready?’

And the same is true for third persons:

- (67) Third-person singular **A** (Hori 2008:29, example 7a)
 k'aay dii 'laa dəw-χalgən
 apple 1SG.β 3 get-tell[PAST]
 'He/she told me to get an apple.'
- (68) Third-person singular **O** (Hori 2008:30, example 10c)
 ʔaa = ʔuu 'lə = xidxiidən
 1.SG.α=FOC 3=chase[PAST]
 'I chased him/her.'
- (69) Third-person singular **S** with α semantics (Hori 2008:31, example 13c)
 yaan = ʔuu 'lə = xudsk'aju + ʔiwʔan-di-gən
 truly=FOC 3=whistle+big-DUR-PAST
 'He/she was whistling really loud.'
- (70) Third-person singular **S** with β semantics (Hori 2008:32, example 16c)
 gəm 'lə = 'laa-gəŋ-ga
 NEG 3=good-NEG-NONPAST
 'He/she is not fine.'
- (71) Third-person plural **A** (Hori 2008:29, example 7b)
 k'iway = ga ʔiitl'e 'laa xisgaləŋ-gu-gən
 street[DEF]=PP 1PL.β 3 wave-PL-PAST
 'They waved us to the street.'
- (72) Third-person plural **O** (Hori 2008:30, example 10d)
 daa = gwaa 'lə = qin-gaawa
 2SG.α=INTER 3=see-PL[EVD]
 'Did you see them?'
- (73) Third-person plural **S** with α semantics (Hori 2008:31, example 13d)
 'le = q'əw-sklə-gu-gən
 3=sit-completely-PL-PAST
 'They sat down.'
- (74) Third-person plural **S** with β semantics (Hori 2008:32, example 16d)
 'lə = χaldan̄aa-gu-ga
 3=be.slave-PL-NONPAST
 'They are slaves.'

Full nominals in Haida neither inflect for case nor induce verb agreement (Enrico 2003:74). The order of arguments is affected in transitive clauses on the basis of what Enrico (2003) calls “potency,” which is a Haida-specific version of a general indexability hierarchy.²⁰ However, since alignment is defined in this dissertation as a correspondence between all three relations A, S, and O, and since Haida word order is verb-final, word order cannot be used to make claims about alignment in the language and therefore full nominals cannot enter into this discussion.

4.2.5.5 Summary of split-alignment in Haida

The alignment system in Haida is split between the split-S patterns and the neutral pattern on the basis of person and number. All first-person, but only singular second-person, pronouns have split-S alignment. It cannot therefore be said that the split is based on SAP status, since plural second persons align together with the third persons but differently from singular second persons. For the same reason, number must also be considered a split inducer. Within the split-S system, the choice between the two pronoun sets is based on semantic control and semantic agency. Second-person plural and all third-person pronouns have the neutral pattern because they take identical forms whether in A, S, or O function. A summary in table form is shown in (75).

²⁰ The hierarchy is as follows: known single adult free humans > non-adult and/or enslaved and/or unknown and/or grouped humans > non human higher animals > inanimates and lower organisms (lower than fish) (Enrico 2003:76).

(75) Tabular summary of split-alignment in Haida

<u>Person/ number</u>	<u>A</u>	<u>S</u>	<u>O</u>	<u>Pattern</u>	<u>Inducer</u>
1SG	ɬaa	ɬaa	dii	split-S _{nom-acc}	person (1) & verbal semantics (agentive), person (1) & participant semantics (agentive), person (2) & number (singular) & verbal semantics (agentive), person (2) & number (singular) & participant semantics (agentive)
1PL	t'alaan̩	t'alaan̩	ʔiitl'ə		
2SG	daa	daa	dən̩		
1SG	ɬaa	dii	dii	split-S _{erg-abs}	person (1) & verbal semantics (patientive), person (1) & participant semantics (patientive), person (2) & number (singular) & verbal semantics (patientive), person (2) & number (singular) & participant semantics (patientive)
1PL	t'alaan̩	ʔiitl'ə	ʔiitl'ə		
2SG	daa	dən̩	dən̩		
2PL	dalaan̩	dalaan̩	dalaan̩	neutral	person (2) & number (plural), person (3)
3	'laa	'laa	'laa		

4.2.6 Split-alignment in Tlingit

4.2.6.1 Tlingit the language

Tlingit has around 600 speakers (Lewis et al. 2014) and is spoken in and around the Alexander Archipelago of extreme southeastern Alaska. It is sometimes considered to be a member of the Eyak-Athabaskan family, which would make Kutchin, Beaver, Chipewyan, Dogrib, Navajo, and Apache its distant relatives. However, due to likely areal diffusion in the Northwest Coast linguistic area, its relationship status is not at all clear (Campbell 1997:114). A conservative approach would classify it as an isolate. It is grouped in Region 2: Northwest Coast, Plateau, and California.

4.2.6.2 Overview of Tlingit alignment

Tlingit has three alignment patterns: nominative-accusative, ergative-absolutive, and neutral. The split between nominative-accusative and ergative-absolutive alignment occurs in two different subsystems of argument marking. First, it occurs with full nouns and independent pronouns, where arguments in A function receive an ergative marker only when the verb's object is definite. The second subsystem involves the pronominal prefixes, which operate for the most part on a split-S system. Here, alignment is based on lexically fixed agentivity, where verbs taking agents receive a “subjective” prefix whether they are transitive or intransitive, and verbs taking patients receive an “objective” prefix. There is an additional split in the pronominals, where third-person marking is neutral in most situations, but nominative-accusative if both the subject and the object are definite. So here, it is SAP status and definiteness at play in conditioning the alignment split.

4.2.6.3 Alignment of nouns in Tlingit

Most transitive verbs take an ergative subject, which is marked by the clitic =ǵ. Some examples are shown in (76) and (77).²¹ Both sentences have a direct object without this clitic, and (78) shows an intransitive construction whose subject also lacks it.

²¹ Throughout this dissertation, I have made certain orthographic modifications to the Tlingit examples cited. Since source materials did not always converge on a single orthographic convention (or even a single phonetic description of their orthographic choices), I have defaulted to the orthography used in Leer (1991). This was necessary to maintain consistency and comparability of the examples. I have only modified the orthography of the specific forms under discussion, rather than doing a complete orthographic overhaul of every example in full.

- (76) Ergative marking on **A** (Crippen 2012:33, ex. 5a, modifying Story & Naish 1973:183)

ak̥ éesh = ǵ útl̥xi ʔa-Ø-Øsa-.i-:
 1SG.POSS father=ERG soup 3.OBJ-ZCNJ-3.SUB-CL-cook-VAR
 ‘my father is cooking soup’

- (77) Ergative marking on **A** (Crippen 2012:33, ex. 5b, modifying Story & Naish 1973:183)

haa s’aatí = ǵ ch’a du yéet ʔa-ka-ŷu-Ø-ŷa-ka-h
 1PL.POSS master=ERG just 3H.POSS son 3.OBJ-HSFC-PFV-3.SUB-CL-say-VAR
 ‘our master sent his own son’

- (78) Absolutive marking on **S** (Crippen 2012:164, ex. 50b, citing Edwards 2009:199)

kee-x’é shuká-t á-wé sha-u-Ø-da-nuk-h-ch
 dawn-mouth ahead-PNCT FOC-MDST head-IRR-3.SUB-CL-mv.vert-VAR-HAB

ax léelk’w
 1SG.POSS grandparent
 ‘my grandfather gets up before dawn’

If O is indefinite, then A will not receive ergative marking, as (79) shows. In this context, the indefiniteness of O is pragmatic and without overt morphology; *my father* is not fishing any specific king salmon.

- (79) No ergative marking on **A** when O is indefinite (Crippen 2012:35, ex. 9)

ax éesh t’á ʔa-wsit’éx
 1SG.POSS father king.salmon 3.OBJ-PFV-3.SUB-CL-fish
 ‘my father was fishing for king salmon’

* ax éesh = ǵ t’á ʔa-wsit’éx
 1SG.POSS father-ERG king.salmon 3.OBJ-PFV-3.SUB-CL-fish
 ‘my father was fishing for king salmon’

4.2.6.4 Alignment of pronominals in Tlingit

The pronominals are verbal prefixes which inflect for person and number (but, in the case of third-person arguments, number is not encoded but definiteness is). These pronominals

participate in a split-intransitive alignment system, so they may be grouped into two categories, shown in (80).²²

(80) Pronominals in Tlingit (modified from Leer 1991a:58, Figure 10 and Crippen 2012:315, Table 19.1)

	<u>“Subject”</u>	<u>“Object”</u>
1SG	xa-	xad ²³ -
1PL	tur-	ha’-
2SG	i-	?i-
2PL	yi-	yi’-
3	Ø-	Ø- ~ ?a-
3.INDF	du-	qu-

The decision between the “subject” and “object” pronominal is determined by lexical specification of the verb. The general tendency is that verbs taking semantic agents as subjects require the “subject” pronominal and those taking semantic patients require the “object” pronominal. There is no fluidity in a verb’s selection of pronominal (Leer 1991a:50, fn. 3). Some intransitive verbs which would be expected to take “object” pronominals on the basis of their semantics in fact take “subject” pronominals, such as ‘be lazy,’ ‘be afraid,’ ‘be wise,’ and ‘be born.’ The split here is conditioned by verbal semantics and lexical specification. A few examples of the first-person pronominals are shown in (81)-(84).

²² Literature on Tlingit pronominals (Story 1972, Story & Naish 1973, Leer 1991a, Crippen 2012) has apparently adopted nomenclature from Boas in calling the categories *subject* and *object*, even though the system is split-intransitive and the object pronominals are also used for the sole arguments of some one-argument verbs. Leer (1991:48, fn. 2) comments that “the terms ‘agent’ and ‘patient’ might be on the whole more accurate.” I don’t want to add my own interpretive layer to this issue by renaming these categories for this dissertation. But in light of the definitions used here and the conceptual apparatus I am employing, it is worthwhile keeping in mind that the labels *subject* and *object* are probably not representative of the structure of the language. For this reason, I will continue to use these terms as they are readily understood by scholars of Tlingit, but I will place them in double quotation marks, “subject” and “object,” respectively, to maintain the theoretical distinction relevant to split-intransitive systems.

²³ Note that the lack of tone on this form distinguishes it from the first-person singular pronoun *xád*.

- (81) First-person singular **A** marked with “subject” pronominal (Crippen 2012:316, ex. 290)
 útłxi Ø-Ø-**ʔa**-sa-.i-ʔ
 soup 3.OBJ-ZCNJ-**1SG.SUB**-CL-cook-REP
 ‘I cook soup (frequently).’
- (82) First-person singular **S_a** marked with “subject” pronominal (Crippen 2012:317, ex. 292b)
 i ʔán-t ʔu-**ʔa**-ʔa-gut-ʔ
 2SG.POSS near-PNCT PFV-**1SG.SUB**-CL-go.SG-VAR
 ‘I came by you.’
- (83) First-person singular **S_o** marked with “object” pronominal (Leer 1991:49, ex. 2b)
ʔad-ʔak’é
1SG.OBJ-be.fine
 ‘I am fine.’
- (84) First-person singular **O** marked with “object” pronominal (Crippen 2012:318, ex. 295b)
ʔad-ʔu-Ø-si-tin-h
1SG.OBJ-PFV-3.SUB-CL-see-VAR
 ‘he saw me’

In the third person, there are two sub-categories of pronominal: the first is the standard one, used in most situations; the second is for indefinite arguments. For the standard third-person pronominals, alignment is split between neutral, with A = S = O since they are all zero-marked, and nominative-accusative, with A and S being zero-marked and O marked with *ʔa-*. Marking of O with *ʔa-* requires that the subject of that verb is third-person definite as well as the object. The sentences in (85)-(88) show the neutral pattern. The sentence in (89) shows the marking of O with *ʔa-*.

- (85) Third-person **A** marked with “subject” pronominal (Crippen 2012:324, ex. 305a)
 ʔad-ʔu-Ø-si-ku-ʔ
 1SG.OBJ-PFV-**3.SUB**-CL-know-VAR
 ‘he knows me’

- (86) Third-person **S_a** marked with “subject” pronominal (Crippen 2012:323, ex. 302b)
 yü-**Ø**-yä-gut-h
 PFV-**3.SUB**-CL-go.SG-VAR
 ‘he went’
- (87) Third-person **S_o** marked with “object” pronominal (Crippen 2012:323, ex. 302d)
Ø-yä-na-h
3.OBJ-PFV-CL-die-VAR
 ‘he died’
- (88) Third-person **O** marked with “object” pronominal (Crippen 2012:324, ex. 305b)
Ø-yä-ṣa-si-ku-ÿ
3.OBJ-PFV-1SG.SUB-CL-know-VAR
 ‘I know him’
- (89) Third-person **O** marked with the “object” pronominal *?a-* (Crippen 2012:325, ex. 305c)
?a-yä-**Ø**-si-ku-ÿ
3.OBJ-PFV-3.SUB-CL-know-VAR
 ‘he knows it’

The reason (86) and (87) are both shown above is to illustrate that for the third-person standard pronominals, there is no distinction between **S_a** and **S_o**, as there is for the other pronominals. The difference between (88) and (89) shows the differential marking of **O** depending on the status of the subject. As stated earlier, marking of **O** with *?a-* requires that the subject is a third-person definite pronominal. In (88), the subject is first-person, not third-person, so that **O** takes the usual marking of zero. In (89), the subject is third-person (and is definite), so that **O** takes marking with *?a-*.

The second sub-category of pronominal involves indefinite arguments. These follow the same split-intransitive pattern as the first- and second-person pronominals, with **A** = **S_a** as *du-*

and $S_o = O$ as *qu-*; thus, alignment of the third-person indefinite pronominals is split between nominative-accusative and ergative-absolutive. The sentences in (90)-(92) serve as examples.²⁴

- (90) Third-person A marked with indefinite “subject” pronominal (Crippen 2012:324, ex. 304a)

ÿaaÿ Ø-ÿu-**du**-ÿa-jak-ÿ
 whale 3.OBJ-PFV-**3.IND.H.SUB**-CL-kill-VAR
 ‘somebody killed a whale’

- (91) Third-person S_a marked with indefinite “subject” pronominal (Crippen 2012:174, ex. 166a)

i ÿa-t’-e-’ yuh = ɣ’a-Ø-**du**-ÿa-tan-k
 2SG.POSS VSFC-behind-LOC ALT=mouth-ZCNJ-**IND.H.SUB**-CL-handle-REP
 ‘people are talking behind your back’

- (92) Third-person S_o marked with indefinite “object” pronominal (Leer 1991:44, ex. 64)

qu-šäyadihé’n
3.IND.H.OBJ-are.many
 ‘there are a lot of people; people are many’

4.3.6.6 Summary of split-alignment in Tlingit

A summary in table form of the alignment patterns in Tlingit is shown in (93).

²⁴ Unfortunately, I was unable to find a sentence with the indefinite “object” pronominal *qu-* marking a third-person O , which would have rounded out the examples provided. It could presumably be something meaning ‘I saw someone’ or ‘I saw people.’

(93) Tabular summary of split-alignment in Tlingit

<u>Person/ number/ definiteness</u>	<u>A</u>	<u>S</u>	<u>O</u>	<u>Pattern</u>	<u>Inducer</u>
3.REF	-ž	-Ø	-Ø	ergative- absolutive	nominal status (non-pronominal) & definiteness (O = definite)
1SG	ᵿa-	ᵿa-	ᵿad-	split-S _{nom-acc}	nominal status (pronominal) & SAP status (+SAP) & verbal semantics (active), nominal status (pronominal) & SAP status (+SAP) & lexical specification (active), nominal status (pronominal) & SAP status (-SAP) & definiteness (A = definite & O = definite)
1PL	tu'-	tu'-	ha'-		
2SG	i'-	i'-	ʔi-		
2PL	yi-	yi-	yi'-		
3.IND	du-	du-	qu-		
3.DEF	Ø	Ø	ʔa-		
1SG	ᵿa-	ᵿad-	ᵿad-	split-S _{erg-abs}	nominal status (pronominal) & SAP status (+SAP) & verbal semantics (patientive), nominal status (pronominal) & SAP status (+SAP) & lexical specification (patientive), nominal status (pronominal) & SAP status (-SAP) & definiteness (indefinite) & verbal semantics (patientive), nominal status (pronominal) & SAP status (-SAP) & definiteness (indefinite) & lexical specification (patientive)
1PL	tu'-	ha'-	ha'-		
2SG	i'-	ʔi-	ʔi-		
2PL	yi-	yi'-	yi'-		
3.IND	du-	qu-	qu-		
3.DEF	Ø	Ø	Ø	neutral	nominal status (pronominal) & SAP status (-SAP) & definiteness (A = definite & O = indefinite)

4.2.7 Split-alignment in Laguna Keres

4.2.7.1 Laguna Keres the language

Laguna Keres is part of a dialect chain belonging to the Keresan family of New Mexico.

Along with Acoma and several other dialects, the Laguna variety is a member of the Western

Keresan branch, whose speakers report high intelligibility with one another, but marginal

intelligibility with Eastern Keresan (Lachler 2006). It is threatened, with most fully fluent speakers being elderly and population estimates in the few thousands. Its endonym is K'awaigame Dzeenyi. It is grouped in Region 3: Great Basin and Southwest.

4.2.7.2 Overview of alignment in Laguna Keres

Laguna Keres has four alignment patterns. A split-S system results in a split between nominative-accusative and ergative-absolutive alignment patterns based on verbal semantics, lexical specification, and SAP status of the arguments. The third pattern is portmanteau alignment, which is split from the rest on the basis of both A and O being SAPs. In addition, there is a minor fourth alignment pattern, neutral, which occurs with third-person forms in the imperative mood.

4.2.7.3 Split-S alignment in Laguna Keres

Case marking in Laguna Keres is achieved by prefixation of a person marker to the verb. There are three sets of person markers, labeled by Lachler (2006) as A, B, and C. Sets A and B are shown below in (94). Since Set C is used only in the portmanteau alignment pattern, I'll discuss those in §4.2.7.5.

(94) Set A and Set B prefixes²⁵ (Lachler 2006:143, Table 7.1)

	<u>Set A</u>	<u>Set B</u>
1	si-	srg-
2	sr-	kidr-
3	g-	dzi-

²⁵ Shown here is the paradigm for the direct mood. There are other markers in the paradigms of the other four moods. For each of the other moods, the relations between the markers is the same as it is for the direct mood, with the exception of the third-person forms in the imperative mood, discussed below.

The choice between Set A and Set B prefixes on intransitive verbs is achieved according to broad semantic tendencies: Set B prefixes mainly occur on verbs describing mental states, physical states, or involuntary actions; Set A prefixes tend to occur with verbs denoting activities (Lachler 2006:143). As usual, there are exceptions. So verbal semantics and lexical specification are both split inducers here. The sentences in (95)-(97) below show Set A prefixes and those in (98)-(100) show Set B prefixes; these all contain only intransitive verbs.

- (95) First-person Set A prefix for S (Lachler 2006:140, ex. 1a)
si-usrp'etrutsa
1A-limp
 'I am limping.'

- (96) Second-person Set A prefix for S (Lachler 2006:140, ex. 1b)
sr-usrp'etrutsa
2A-limp
 'You are limping.'

- (97) Third-person Set A prefix for S (Lachler 2006:141, ex. 1c)
g-usrp'etrutsa
3A-limp
 'He/She is limping.'

- (98) First-person Set B prefix for S (Lachler 2006:141, ex. 3a)
srg-utyishu
1B-be.afraid
 'I am afraid.'

- (99) Second-person Set B prefix for S (Lachler 2006:142, ex. 3b)
kidr-utyishu
2B-be.afraid
 'You are afraid.'

- (100) Third-person Set B prefix for S (Lachler 2006:142, ex. 3c)
dzi-utyishu
3B-be.afraid
 'He is afraid.'

The reason for including lexical specification as a split inducer, rather than only verbal semantics, is due to a handful of apparent exceptions to the semantic tendencies for verbs to take either Set A or Set B prefixes. As an example, consider a few of the verbs in each category, listed below in (101) and (102).

(101) Some intransitive verb stems requiring Set A prefixes (Lachler 2006:141, exx. 2a-e)

- ushch'i 'to have diarrhea'
- ausrgitsi 'to be brave'
- ausinyitsa 'to hurry along'
- umi 'to leave'
- uusrbiitsa 'to whistle'

(102) Some intransitive verb stems requiring Set B prefixes (Lachler 2006:142, exx. 4a-f)

- unu 'to be selfish'
- unashiya 'to be in a hurry'
- ubayatsa 'to burst out laughing'
- uhima 'to believe'
- udyumidruwi 'to forget'
- uunawats'i 'to have a secret'

The first verb listed in (101), *-ushch'i* 'to have diarrhea' is a body function and not a transparent example of an activity or a state achieved voluntarily. Similarly, the second item, *-ausrgitsi* 'to be brave' might be thought to fall into the 'mental state' category. Nonetheless, these both require Set A prefixes. Notice, also, that two nearly identical verbs require different sets of prefixes: *-ausinyitsa* 'to hurry along' takes Set A, while *-unashiya* 'to be in a hurry' takes Set B, though there might be a state/activity contrast going on there. Lachler (2006:144) furthermore

states that speakers do not “have the option to code an argument as either agent or patient, depending on the perceived volitionality of the action.”²⁶

4.2.7.4 Transitive constructions in Laguna Keres

With transitive verbs, there are, of course, two arguments. However, the morphological frame of Laguna Keres verb complexes provides only one structural position for the placement of a person marker. This is resolved in one of three ways, depending on the hierarchical status of the arguments, and the result is another split in the alignment system.

If either argument is in the third person, then speakers must decide whether to affix the marker for the agent or the one for the patient. As is expected from split-intransitive systems, Set A markers indicate semantic agents and Set B markers indicate semantic patients of transitive verbs. The three marking options for transitive verbs are as follows: (1) In the instance where one, but not both, of these roles is in the third person, the decision is straightforward: mark the SAP. (2) If both roles are in the third person, mark the agent. (3) If neither argument is in the third person, then use a third set of person markers, Set C.

The first option, where exactly one argument is in the third person, is shown in (103) and (104). The third-person argument is a non-SAP, and so the SAP argument will be either first or second person. It is the SAP argument, whether agent or patient, that is marked in this case.

²⁶ Lachler’s analysis of Laguna Keres verbal agreement treats *all* verbs as being lexically specified for its prefix set. His argument is motivated by the exceptions to verbal semantics listed above and detailed more fully in his dissertation. But exceptions can be found for almost any pattern in a language; to argue forcefully that “the choice of a Set A prefix or Set B prefix [...] must be learned on a verb-by-verb basis” (2006:143) would require something like a novel verb task to see how speakers assign verbs they have never encountered before, particularly whether they appeal to semantics in order to inflect the verb. Such a study would be very interesting. In lieu of that data, though, I have opted to assume that the exceptions really are exceptions, and that verbal semantics and lexical specification therefore both play a role.

- (103) First-person Set A prefix marking **A** on a transitive verb (Lachler 2006:144, ex. 5a)
si-ukacha
1A-see
 'I see him.'
- (104) Second-person Set A prefix marking **A** on a transitive verb (Lachler 2006:144, ex. 5b)
sr-ukacha
2A-see
 'You see him.'

In (103) and (104), the patient role is a non-SAP, and so it is not marked. The opposite case is in (105) and (106), where the agent role is a non-SAP, and is likewise not marked.

- (105) First-person Set B prefix marking **O** on a transitive verb (Lachler 2006:145, ex. 5d)
srg-ukacha
1B-see
 'He sees me.'
- (106) Second-person Set B prefix marking **O** on a transitive verb (Lachler 2006:145, ex. 5e)
kidr-ukacha
2B-see
 'He sees you.'

As (103)-(106) show, if one argument is an SAP but the other is not, then the SAP will be expressed, regardless of its relational role. The second marking scenario on transitive verbs is if both arguments are in the third person (are non-SAPs). In this event, it is the agent which is marked, as in (107).

- (107) Third-person Set A prefix marking **A** on a transitive verb (Lachler 2006:145, ex. 5c)
g-ukacha
3A-see
 'He sees him.'

Considering the data provided so far for the intransitive verbs in (95)-(100) and the transitive verbs in (103)-(107), we are now in a position to begin identifying alignment types in Laguna Keres. The nominative-accusative alignment pattern, where $(A = S) \neq O$, is conditioned

by verbal semantics and the lexical specification of intransitive verbs as requiring a Set A marker and the +SAP status of A of transitive verbs with both an SAP and a non-SAP argument.

The ergative-absolutive alignment pattern, where $A \neq (S = O)$, is conditioned by the lexical specification of intransitive verbs as requiring a Set B marker and the +SAP status of O of transitive verbs with both an SAP and a non-SAP argument.

4.2.7.5 Portmanteau alignment in Laguna Keres

Finally there is the third option for the marking of transitive arguments, in which both arguments are SAPs. In this case, a single marker encodes both A and O. One marker encodes a first-person A with second-person O, the other encodes the reverse. Lachler (2006) calls these the Set C markers; they are shown in (108).

- (108) Set C prefixes (Lachler 2006:146)
1 \Rightarrow 2 sra-
2 \Rightarrow 1 dy-

The Set C prefixes are therefore only used on transitive verbs when both A and O are SAPs. Some examples are below in (109) and (110).

- (109) Set C marking first-person A and second-person O (Lachler 2006:146, ex. 5g)
sra-ukacha
1 \Rightarrow 2C-see
'I see you.'
- (109) Set C marking second-person A and first-person O (Lachler 2006:146, ex. 5h)
dy-ukacha
2 \Rightarrow 1C-see
'You see me.'

Thus the third alignment type is portmanteau, conditioned by A and O both being SAPs.

4.2.7.6 Neutral alignment and summary of split-alignment in Laguna Keres

The fourth alignment type is a comparatively minor one, and occurs only in the imperative mood. There are five grammatical moods in Laguna Keres: direct, negative direct, indirect, negative indirect, and imperative. Allomorphs of the person prefixes are varied in each of the mood paradigms, but in all cases except one, the relationships between prefixes are identical to what has been described above. The single exception is in the imperative mood, where the third-person A prefix is the same as the third-person O prefix: *pi-*. Since split-intransitivity operates equally in all moods, the fact that S is sometimes marked like the A-marking *pi-* and sometimes marked like the O-marking *pi-* means all three argument relations are identical under these conditions. Therefore, the fourth alignment type is neutral, conditioned by the third person in the imperative mood.

Laguna Keres has four alignment types: nominative-accusative, ergative-absolutive, portmanteau, and neutral. The first two are conditioned by verbal semantics, lexical specification, and SAP status. Portmanteau alignment is induced by status of A and O as SAPs. Neutral alignment is induced by non-SAP status and mood. A summary table is shown in (110).

(110) Tabular summary of split-alignment in Laguna Keres

<u>Person</u>	<u>A</u>	<u>S</u>	<u>O</u>	<u>A & O</u>	<u>Pattern</u>	<u>Inducer</u>
1	si-	si-	srg-		split-S _{nom-acc}	SAP status (A = +SAP & O = –SAP & verbal semantics (agentive), SAP status (A = –SAP & O = +SAP & verbal semantics (agentive), SAP status (A = +SAP & O = –SAP & lexical specification (agentive), SAP status (A = –SAP & O = +SAP & lexical specification (agentive)
2	sr-	sr-	kidr-			
3	g-	g-	dzi-			
1	si-	srg-	srg-		split-S _{erg-abs}	SAP status (A = +SAP & O = –SAP & verbal semantics (patientive), SAP status (A = –SAP & O = +SAP & verbal semantics (patientive), SAP status (A = +SAP & O = –SAP & lexical specification (patientive), SAP status (A = –SAP & O = +SAP & lexical specification (patientive)
2	sr-	kidr-	kidr-			
3	g-	dzi-	dzi-			
3	pi-	pi-	pi-		neutral	SAP status (–SAP) & mood (imperative)
1⇒2				sra-	portmanteau	SAP status (A = +SAP & O = +SAP)
2⇒1				dy-		

4.2.8 Split-alignment in Maricopa

4.2.8.1 Maricopa the language

Maricopa is a Cochimí-Yuman language spoken in Arizona. It is joined in the Yuman branch by languages including Quechan, Mojave, Yavapai, and several others sometimes collectively known as Diegueño. The Cochimí branch of the family is extinct. Lynn Gordon declines to estimate the number of speakers at the time of writing her 1986 grammar of the language, but cites others with a community member approximation around 500. According to the Ethnologue, there were 100 speakers in 2007 (Lewis et al. 2014). It is grouped in Region 3: Great Basin and Southwest.

4.2.8.2 Overview of alignment in Maricopa

Maricopa has three alignment patterns: nominative-accusative, neutral, and portmanteau.

The splits are conditioned by status of an argument as a pronominal, pronoun, or full noun phrase, as well as its status as an SAP.

4.2.8.3 Alignment of pronominals in Maricopa

Pronominal markers in Maricopa are verbal prefixes and inflect for person. Their alignment behavior depends entirely on SAP status. The table in (111) shows the non-portmanteau pronominal markers. Note that the cells for first- and second-person object are missing because they are always involved in a portmanteau pattern (explained at the end of this subsection).

(111) Non-portmanteau pronominals in Maricopa (Gordon 1986)

	<u>A</u>	<u>S</u>	<u>O</u>
1	‘-	‘-	
2	m-	m-	
3	Ø	Ø	Ø

If A or S is an SAP, but O is not, then alignment is nominative-accusative, as is clear from the identical forms in the A and S columns and the zero expression of third-person O. Examples (112) and (113) show the marking of S, while (114) and (115) show A and O.

(112) First-person S (Gordon 1986:17, ex. 9)

‘-ashvar-k
1-sing-REAL
‘I sang/am singing.’

(113) Second-person S (Gordon 1986:18, ex. 14)

m-ashvar-k
2-sing-REAL
‘You sang/are singing.’

- (114) First-person **A** with third-person **O** (Gordon 1986:19, ex. 24)

‘-aaham-m²⁷

1-hit-REAL

‘I hit him.’

- (115) Second-person **A** with third-person **O** (Gordon 1986:19, ex. 25)

m-aham-m

2-hit-REAL

‘You hit him.’

When **A** is an **SAP** and **O** is not, the pronominal marking **A** is the same as that marking **S** and different from that marking **O**, which is a nominative-accusative pattern. When all relations **A**, **S**, and **O** are non-SAPs, then alignment is neutral because, as seen in the table in (111), none of the third-person forms is overtly marked. The two sentences in (116) and (117) show this pattern.

- (116) Third-person **S** (Gordon 1986:16, ex. 5)

ashvar-k

sing-REAL

‘He sang/is singing.’

- (117) Third-person **A** with third-person **O** (Gordon 1986:19, ex. 26)

aham-m

hit-REAL

‘He hit him.’

Where the alignment of pronominals becomes richer is when **O** is an **SAP**. In these cases, alignment is portmanteau because there is a single marker which encodes both the **A** and the **O** arguments. The examples in (118)-(121) show the four instances of this pattern.

- (118) First-person **A** with second-person **O** (Gordon 1986:19, ex. 28)

ny-wik-k

1⇒**2**-help-REAL

‘I helped you.’

²⁷ Whether the realis marker is *-k*, as in (112) and (113), or *-m*, as in (114) and (115), depends on lexically specified properties of the element to which that marker directly attaches (Gordon 1986:24).

- (119) Second-person **A** with second-person **O** (Gordon 1986:19, ex. 29)
'nym-wik-k
2⇒1-help-REAL
 'You helped me.'
- (120) Third-person **A** with first-person **O** (Gordon 1986:20, ex. 30)
ny-wik-k
3⇒1-help-REAL
 'He helped me.'
- (121) Third-person **A** with second-person **O** (Gordon 1986:20, ex. 31)
m-wik-k
3⇒2-help-REAL
 'He helped you.'

4.2.8.4 Alignment of personal pronouns in Maricopa

The personal pronouns align similarly to the pronominals, minus the portmanteau pattern.

There is a nominative-accusative pattern for the SAP arguments and a neutral pattern for the non-SAP ones. The table in (122) shows the pronoun forms.

- (122) Personal pronouns in Maricopa (Gordon 1986:59)

	<u>A</u>	<u>S</u>	<u>O</u>
1	'nyaash	'nyaash	'nyip
2	mansh	mansh	many
3	Ø	Ø	Ø

The examples in (123) and (124) show the SAP pronouns.

- (123) First-person **A** pronoun with second-person **O** pronoun (Gordon 1986:59, ex. 236)
nyaa-sh **many** ny-yuu-k
 I-SUB you 1⇒2-see-REAL
 'I saw you.'
- (124) Second-person **A** pronoun with first-person **O** pronoun (Gordon 1986:59, ex. 237)
man-sh **nyip** 'nym-aham-m
 you-SUB me 2⇒1-hit-REAL
 'You saw me.'

Since first- and second-person forms have identical A and S, but different O, pronouns, the pattern is nominative-accusative for SAPs. With no overt third-person pronouns, alignment is neutral for non-SAPs.

4.3.8.5 Alignment of full nouns in Maricopa

Full nouns align in a nominative-accusative pattern because there is a suffix which appears on A and S, but not on O, arguments. The suffix is *-sh*, which also appears on A and S personal pronouns.²⁸ However, alignment of the full nouns is different from that of the pronouns because the neutral pattern is absent on full nouns. The pattern is shown in (125) and (126).

- (125) Full noun **S** marked with *-sh* (Gordon 1986:37, ex. 149)

sny'ak-sh ashvar-k
 woman-SUB sing-REAL
 'The/a woman sang.'

- (126) Full noun **A** marked with *-sh* and O without *-sh* (Gordon 1986:41, ex. 171)

'iipaa-ny-sh qwaaq kyaa-m
 man-DEM-SUB deer shoot-REAL
 'The man shot a/the deer.'

Since the S in (125) and the A in (126) are suffixed with *-sh*, but the O in (126) is not, the pattern is nominative-accusative.

4.2.8.6 Summary of split-alignment in Maricopa

The three alignment patterns, nominative-accusative, neutral, and portmanteau, are conditioned in Maricopa by specific combinations of SAP status and nominal status (noun,

²⁸ Note that the O pronouns are not just the A/S versions without the suffix; they have a different morphological form. So for the pronouns, it is not simply the presence or absence of the suffix which makes the pattern nominative-accusative, but it is for the full nouns.

pronoun, or pronominal). Full nouns show nominative-accusative alignment across the board.

Pronouns show nominative-accusative alignment for SAPs and neutral alignment for non-SAPs.

Pronominals show nominative-accusative alignment when A or S is an SAP but O is not, neutral alignment if A, S, and O are all non-SAPs, and portmanteau alignment if only O is an SAP. A

summary table is shown in (127).

(127) Tabular summary of split-alignment in Maricopa

<u>Person</u>	<u>A</u>	<u>S</u>	<u>O</u>	<u>A & O</u>	<u>Pattern</u>	<u>Inducer</u>
1	‘-	‘-			nominative-accusative	SAP status (O = -SAP & A = +SAP & S = +SAP) & nominal status (pronominal)
2	m-	m-				
3	Ø	Ø	Ø		neutral	SAP status (-SAP) & nominal status (pronominal)
1	'nyaash	'nyaash	'nyip		nominative-accusative	SAP status (+SAP) & nominal status (pronoun)
2	mansh	mansh	many			
3	Ø	Ø	Ø		neutral	SAP status (-SAP) & nominal status (pronoun)
	-sh	-sh	Ø		nominative-accusative	nominal status (noun)
1⇒2				ny-	portmanteau	SAP status (O = +SAP) & nominal status (pronominal)
2⇒1				'nym		
3⇒1				ny-		
3⇒2				m-		

4.2.9 Split-alignment in Seri

4.2.9.1 Seri the language

Seri is an isolate spoken on and near Isla Tiburón in northwest Mexico. Known endonymically as Cmiiq̣ue Iitom, there is minimal dialectal variation. The Ethnologue lists a

population of 900 in 2007 (Lewis et al. 2014), though it is not clear how many fully fluent speakers there are. It is grouped in Region 3: Great Basin and Southwest.

4.2.9.2 Overview of alignment in Seri

Core participants in Seri are signaled by the use of pronominal agreement markers prefixed to the verb stem. The alignment system is split between tripartite, nominative-accusative, neutral, and portmanteau alignment types, all on the basis of person and number for the agreement markers. First-person singular agreement markers show a tripartite pattern with different forms for A, S, and O. First-person plural and second-person markers of both numbers show a nominative-accusative pattern with A and S marked the same and O marked differently. Third person agreement markers of both numbers show neutral alignment because none of them is overtly marked. If both A and O are third person, then a portmanteau morpheme is used.

4.2.9.3 Alignment of agreement markers in Seri

The subject and object agreement markers are shown in (128).²⁹

²⁹ There are some orthographic differences between Marlett's publications in the 1980s and '90s and his more recent work from the 2000s. What is represented by a glottal stop in his 1990 paper is represented by *h* in later publications. I am using the *h*-versions here. For consistency, I have changed *ʔ* to *h* in the examples taken from his earlier works. Similarly, *š* in earlier work is *z* in the more recent literature. I have likewise changed *š* to *z* throughout.

(128) Agreement markers (constructed from Marlett 1990:514, 521 & 2014:607)

	<u>A</u>	<u>S</u>	<u>O</u>	
1SG	h-	hp-	him-	
1PL	ha-	ha-	hizi-	
2SG	m-	m-	ma-	
2PL	ma-	ma-	mazi-	
3SG	Ø	Ø	Ø	A/O i-
3PL	Ø	Ø	Ø	A/O i-

The tripartite pattern occurs only on first-person singular arguments. As shown in (129)-

(131), different agreement markers are used for each of the argument types A, S, and O.

(129) First-person singular **A** (Marlett 1990:515, ex. 28a)

h-χo-tis

1SG.SUB-EM-point.at

‘I pointed at it/him/her!’

(130) First-person singular **S** (Marlett 1990:515, ex. 28e)

hp-mi-panšχ

1SG.SUB-PR-run

‘I am running’

(131) First-person singular **O** (Marlett 1990:522, ex. 37a)

him-mi-kášni

1SG.OBJ-PR-bite

‘it bit me’

From the examples in (129)-(131), one can see that the first-person singular agreement marking on the verb is different for arguments in A, S, or O function. First-person plural, second-person singular, and second-person plural agreement markers show a nominative-accusative pattern. The second-person singular forms serve as an example in (132)-(134).

(132) Second-person singular **A** (Marlett 1990:515, ex. 28b)

m-t-i:

2SG.SUB-R-hear

‘did you hear it?’

- (133) Second-person singular **S** (Marlett 1981:31, ex. 36m)

m-im-hak

2SG.SUB-PROX-blind

‘you are blind’

- (134) Second-person singular **O** (Marlett 1990:523, ex. 37f)

ma-h-yo-aho

2SG.OBJ-1SG.SUB-DI-see

‘I saw you’

Third-person arguments take on a neutral pattern because they are not overtly marked on verbs, as shown in (135)-(137).

- (135) Third-person singular **A** (Marlett 1990:523, ex. 37h)

hizi-yo-aho

1PL.OBJ-DI-see

‘s/he saw us’

- (136) Third-person singular **S** (Marlett 1990:515, ex. 28g)

t-afp

R-arrive

‘did s/he/it arrive?’

- (137) Third-person **O** (Marlett 1981:33, ex. 40g)

ma-s-o:ktam-χo

2PL.SUB-IRR-look.at-EMPH

‘you (pl) should look at it/them.’

When both subject and object of a transitive verb are in the third person, then the agreement marker is *-i*, as in (138).

- (138) Third-person **A** and third-person **O** (Marlett 2014:607, ex. 55)

z i-taho³⁰

one 3:3-see

‘s/he saw one

Thus when both arguments are non-SAPs, agreement is portmanteau.

³⁰ The verb stem is *aho*; *t* is epenthetic for syllabic reasons.

4.2.9.4 Summary of split-alignment in Seri

Within the pronominals alone, Seri has four alignment patterns. Tripartite alignment shows up with first-person singular pronominals. Nominative-accusative alignment occurs on first-person plural and all second person pronominals. Neutral alignment occurs on all third-person pronominals, as these are zero-marked, unless both A and O are third person, in which case alignment is portmanteau. Thus the alignment system is split by person, number, and SAP status. A summary table is shown in (139).

(139) Tabular summary of split-alignment in Seri

<u>Person/ number</u>	<u>A</u>	<u>S</u>	<u>O</u>	<u>A & O</u>	<u>Pattern</u>	<u>Inducer</u>
1SG	h-	hp-	him-		tripartite	person (1) & number (singular)
1PL	ha-	ha-	hizi-		nominative-accusative	person (1) & number (plural), person (2)
2SG	m-	m-	ma-			
2PL	ma-	ma-	mazi-			
3SG	Ø	Ø	Ø		neutral	SAP status (A = +SAP & O = -SAP), SAP status (A = -SAP & O = -SAP)
3PL	Ø	Ø	Ø			
3⇒3				i-	portmanteau	SAP status (A = -SAP & O = -SAP)

4.2.10 Split-alignment in Alabama

4.2.10.1 Alabama the language

Alabama is a Muskogean language from the southeast United States. It is joined by Koasati, Mikasuki, Creek, and Seminole in the Eastern branch of the family; Choctaw and

Chickasaw form Western Muskogean.³¹ Population estimates are in the low hundreds. The Alabama probably resided in north central Mississippi in the sixteenth century (Campbell 1997:148), though nowadays the language is spoken in Alabama and Texas. It is grouped in Region 4: Northeast and Southeast.

4.2.10.2 Overview of alignment in Alabama

Alignment in Alabama is split between the nominative-accusative,³² ergative-absolutive, and neutral alignment types. SAP pronominal markers are fluid-S on the basis of whether or not the argument has control over the action. Thus, alignment of SAP pronominals is split between nominative-accusative and ergative-absolutive. Non-SAP pronominals are zero-marked, and so show neutral alignment.

4.2.10.3 Fluid-S alignment in Alabama

There are two sets of pronominals in Alabama, shown below in (140).

³¹ This internal grouping is based on classifications by Mary Haas and Karen Booker; an alternative analysis by Pamela Munro puts Creek and Seminole together into Northern Muskogean and the rest into nested subgroups in Southern Muskogean; see Campbell (1997:147-8) for a summary.

³² I am treating full nouns in Alabama as not participating in alignment behavior. Several analyses of case marking on Alabama nouns (Munro & Gordon 1982, Lupardus 1982, and others) have characterized a basically nominative-accusative alignment system in which *-k* is suffixed to subjects and *-n* to non-subjects (direct and indirect objects alike, a so-called “oblique” category of noun). However, Davis & Hardy (1988) argue, convincingly in my view, that these suffixes are really not case markers at all, but rather markers of pragmatic functions such as topicality, expectation, and prominence, and may occur on any of the relations S, A, or O. Thus these markers should not be interpreted as indicating grammatical relations.

(140) Pronominals in Alabama (modified from Lupardus 1982:67, Table 2 and :70, Table 4)³³

	<u>SET I</u>	<u>SET II</u>
1SG	-aa/-li ³⁴	ca-
2SG	is-/-ci	ci-
3SG	Ø	Ø
1PL	(h)il-/-(hi)li	po-/ko-
2PL	has-/-haci	haci-
3PL	Ø	Ø

The choice between the two sets depends on the degree of control an argument exerts in an event. Arguments with control are indexed with the Set I marker; those without control, or with comparatively less control, are indexed with a marker from Set II. Before showing examples, it is necessary to point out that the two sets show apparently identical morphemes for some of the person/number categories (compare the second-person singular and the second-person plural in both sets). But these morphemes are only identical in form, not in structure: where there is overlap, the Set II markers are only prefixes and the Set I markers are never prefixes; their distribution is therefore complementary.

Choice of marker is not lexically fixed, but is responsive to changing degrees of control an argument may have in a given event. Thus this split-intransitive system is also fluid-S. Sentences with first-person arguments serve as examples in (141)-(144).

- (141) First-person Set I marker on a controlling **A** (Hardy & Davis 1993:456, ex. 3a)
 achi-lkomòo-**li**-ti
 2SG.II-embrace-**1SG.I**-ASP
 ‘I hugged you.’

³³ I am following Hardy & Davis (1993) in naming these sets according to the theoretically neutral Roman numerals I and II, rather than “agentive” and “patient” as Lupardus (1982) does.

³⁴ The allomorphy in the first person singular is morphologically conditioned by tense. -aa is restricted to future constructions while -li occurs in non-future ones. Since neither allomorph recurs in Set II, the difference does not bear on the alignment system. The other instances of allomorphy are phonologically conditioned.

- (142) First-person Set I marker on a controlling **S** (Hardy & Davis 1993:456, ex. 1)
 waliika-**li**-ti
 run-**1SG.I**-ASP
 ‘I ran.’
- (143) First-person Set II marker on a non-controlling **S** (Hardy & Davis 1993:456, ex. 4)
cha-chaaha-hchi
1SG.II-be.tall-ASP
 ‘I’m tall.’
- (144) First-person Set II marker on a non-controlling **O** (Hardy & Davis 1993:456, ex. 3b)
acha-lkomòo-chi-ti
1SG.II-embrace-2SG.I-ASP
 ‘You hugged me.’

The use of different affix sets on the same verb reveals why the control feature is relevant in Alabama alignment. Intransitive subjects of the same verb can be marked with either set of pronominal affixes depending on its control over the action, as in the contrast between (145) and (146) and between (147) and (148).

- (145) First-person Set I marker on a controlling **S** of verb *toḥohka* ‘cough’ (Hardy & Davis 1993:458, ex. 10b)
 toḥohka-**li**-ti
 cough-**1SG.I**-ASP
 ‘I coughed (on purpose, meaningfully).’
- (146) First-person Set II marker on a non-controlling **S** of verb *toḥohka* ‘cough’ (Hardy & Davis 1993:458, ex. 10a)
cha-toḥohka-ti
1SG.II-cough-ASP
 ‘I coughed.’
- (147) Second-person Set I marker on a controlling **S** of verb *nakaata* ‘leave’ (Hardy & Davis 1993:458, ex. 9b)
 nakaṭ = **chi**-ya-ahi-o-ⁿ?
 leave=**2SG.I**=leave-FUT-ASP-Q
 ‘Are you going to leave?’

- (148) Second-person Set II marker on a non-controlling **S** of verb *nakaala* ‘leave’ (Hardy & Davis 1993:458, ex. 9a)

chi-nakaala-lah-ool-o

2SG.II-leave-FUT-EVID-ASP

‘You will die.’ (idiomatic)

Hardy & Davis (1993) defer to Klaimain’s (1991) formal definition of *control*, in which it is viewed as a universal construct which distinguishes such properties as animacy and intentionality from those such as agency and undergoing. The control construct offers an explanation for the use of Set I and Set II markers with active and stative verbs, respectively, because stative subjects are typically not in control, while active ones are. It also explains why Set I markers correspond to arguments in A relation and Set II markers correspond to those in O relation: they state “[...A]n undergoer of a transitive event [...] is by definition the participant that ‘does not perform, instigate or control any situation’ (Foley and Van Valin 1984:29)” (1993:458). No O arguments take the Set I markers in Alabama.

Third-person pronominals likewise show neutral alignment because they are not overtly marked. Three sentences with a third-person singular argument serve as examples in (149)-(151).

- (149) Third-person singular **A** in Alabama (Lupardus 1982:27)

ifa-n Ø-ibi-ti-~

dog-OBL **3SG**-kill-PROX-Q

‘Did he kill the dog?’

- (150) Third-person singular **S** in Alabama (Lupardus 1982:68)

Ø-nooco

3SG-slept

‘He slept.’

- (151) Third-person singular **O** in Alabama (Lupardus 1982:71)

il-hiica-ti

1PL-see-ASP

‘We saw him.’

4.2.10.4 Summary of split-alignment in Alabama

Alignment in Alabama is pervasively split-intransitive with fluid-S behavior for SAPs of most verbs. Neutral alignment occurs with non-SAPs of all verbs. A summary table is shown in (152).

(152) Tabular summary of split-alignment in Alabama

<u>Person/ number</u>	<u>A</u>	<u>S</u>	<u>O</u>	<u>Pattern</u>	<u>Inducer</u>
1SG	-aa/-li	-aa/-li	ca-	fluid-S _{nom-acc}	SAP status (+SAP) & participant semantics (+control)
2SG	is/-ci	is/-ci	ci-		
1PL	(h)il/- (hi)li	(h)il/- (hi)li	po-/ ko-		
2PL	has/- haci	has/- haci	haci-		
1SG	-aa/-li	ca-	ca-	fluid-S _{erg-abs}	SAP status (+SAP) & participant semantics (–control)
2SG	is/-ci	ci-	ci-		
1PL	(h)il/- (hi)li	po-/ko-	po-/ ko-		
2PL	has/- haci	haci-	haci-		
3SG	Ø	Ø	Ø	neutral	SAP status (–SAP)
3PL	Ø	Ø	Ø		

4.2.11 Split-alignment in Cherokee

4.2.11.1 Cherokee the language

Cherokee is the sole representative of the southern branch of Iroquoian. The northern branch contains Mohawk, Seneca, Tuscarora, and several others. The Cherokee Nation in Oklahoma is by far the most populous in terms of tribal membership and reported speakers. The Eastern Band in western North Carolina is considerably smaller. Estimates of the total number of

Cherokee speakers are often around the 10,000-mark, though degree of proficiency is not clear. The Ethnologue reports 130 monolinguals (Lewis et al. 2014). Cherokee courses for adults are offered at several universities and language learning materials are available. It is grouped in Region 4: Northeast and Southeast.

4.2.11.2 Overview of Cherokee alignment

The alignment system of Cherokee is unique among the languages in my sample. It has a standard split-S system for intransitives, though it is unusual for North America in that there are sub-splits conditioned by aspect. Most transitive arguments are indicated by portmanteau morphemes encoding both A and O. Some of these are systematically identical to the intransitive argument markers; others are not. There is also a hierarchical system in Cherokee for third-person arguments only. The conditions SAP status, person, number, and animacy influence whether or not the conflated morphemes align with the intransitive ones.

4.2.11.3 Split-S alignment in Cherokee

Intransitive verbs in Cherokee show a split-S system, though it is different in a major way from split-S systems in the other North American languages discussed elsewhere in this dissertation. As described in §2.5.5, and again for individual languages in this chapter, split-S systems usually have two sets of markers, one for A and S_a, the other for S_o and O. In Cherokee, there are also two sets of markers, but they are not used for marking A or O. This is because, as discussed in the next subsection, transitive verbs have a mostly portmanteau alignment pattern

and therefore cannot be said to have a distinctly marked A argument and O argument. The two sets of intransitive argument pronominals are given in (153).

(153) Two sets of S pronominals in Cherokee (Scancarelli 1987:55, Table 2)

	<u>Set A</u>	<u>Set B</u>
1SG	ci-, k-	aki-, akw-
1+3DL	o:sti:-	o:kini:-
1+3PL	o:ci:-	o:ki:-
1+2DL	ì:ni:-	kini:-
1+2PL	ì:ti:-	ì:ki:-
2SG	hi-	ca-
2DL	sti:-	sti:-
2PL	ì:ci:-	ì:ci:-
3SG	a-, ka-	u:-
3PL	ani:-	u:ni:-

The choice between the two sets is based on verbal semantics and lexical specification. Most verbs requiring B prefixes have stative, positional, or body function meanings. Examples include ‘be angry,’ ‘stand in liquid,’ and ‘sneeze.’ There are a handful of exceptions which, at least synchronically, do not fall into even a broad “patientive” category, such as ‘work’ and ‘take revenge.’ Most verbs taking A prefixes are semantically “active,” but again exceptions such as ‘stink,’ ‘be hanging,’ and ‘breathe’ preclude absolute categorization. A brief example of each type is shown in (154) and (155).

(154) Third-person singular S marked with Set A prefix (Scancarelli 1987:66, ex. 16a.i)

a-aliski:ʔa

3SG.A-dance.PRES

‘he’s dancing’

- (155) Third-person singular S marked with Set B prefix (Scancarelli 1987:66, ex. 16a.ii)

u:-ye:tska

3SG.B-laugh.PRES

‘he’s laughing’

The intransitive marking system itself is split on the basis of aspect. There are eight aspects in Cherokee. Verbs taking B prefixes will take B prefixes in all eight aspects. Verbs taking A prefixes will take A prefixes in four aspects (present, imperfective, punctual, and future), but B prefixes in the other four (perfective, infinitive, pre-inceptive, and propensitative). This is shown in (156) and (157) with examples of the same verbs above but in the perfective aspect.

- (156) Set A verb with S marked with Set B in the perfective (Scancarelli 1987:66, ex. 16d.i)

u:-aliski:sv:ʔi³⁵

3SG.B-dance.PFV

‘he (has) danced’ (c.f. (154) above)

- (157) Set B verb with S marked with Set B in the perfective (Scancarelli 1989:67, ex. 16d.ii)

u:-ye:tsv:ʔi

3SG.B-laugh.PFV

‘he (has) laughed’

In (154), the verb meaning ‘dance’ takes a Set A prefix in the present aspect, but a Set B prefix in the perfective in (156). The verb meaning ‘laugh’ takes a Set B prefix in both aspects. So the marking of S is determined by verbal semantics, lexical specification, and aspect. But the marking of S by itself is not yet “alignment” as defined here; we will also need to consider the marking of transitive arguments.

³⁵ In the orthography adopted by Scancarelli (1987), /v/ represents a nasalized central vowel.

4.2.11.4 Alignment of transitive arguments in Cherokee

Transitive marking is split between portmanteau and hierarchical alignment patterns on the basis of person, number, and animacy. The pervasive pattern for transitive verbs is portmanteau alignment, in which a single morpheme indexes both the A and the O arguments. The prefix specifies person and number for A, and person, number, and animacy for O. I will not show all the possible markers here, but I refer the reader to the sizable tables in Scancarelli (1987:71) and King (1975:56-7). Two brief examples will suffice to show portmanteau marking of the A and O arguments.

- (158) First-person singular **A** and third-person singular animate **O** (Scancarelli 1987:74, ex. 29)

ci:y-e:loha

1SG.A⇒**3SG.AN.O**-feed.PRES

‘I’m feeding him.’

- (159) First-person singular **A** and second-person singular **O** (Scancarelli 1987:295, ex. 13b)

te:-**kv:y**-e:ʔyo:hvski-o:ʔi

DIST³⁶-**1SG.A**⇒**2SG.O**-teach.IMPFV-HAB

‘I teach you.’

It is interesting that certain of the morphemes in the transitive paradigm are identical to the A and B sets of the intransitive paradigm. Specifically, any transitive construction with a third-person singular A will take a pronominal identical to those in Set B, inflected according to the relevant features of O. Any transitive construction with a third-person inanimate O will take a pronominal identical to those in Set A, inflected according to the relevant features of A. It would be possible to construe this situation as representing a unique type of alignment pattern, one with formulaic representations something like $A \Rightarrow O = S_A$ or $A \Rightarrow O = S_B$. To my knowledge, this

³⁶ The distributive here “can be thought of as representing the plurality of things taught, or the extension of the action over time” (Scancarelli 1987:295).

pattern does not have a name of its own; it would be inappropriate to call it neutral alignment because the marker for A and O is already portmanteau, which does not happen in a neutral pattern. For ease of reference, I will just use the formulaic representations I suggested above. Nomenclature aside, this pattern would be triggered by A and O combinations of person, number, and animacy.³⁷

The hierarchical³⁸ pattern is induced when both arguments are third person . The choice of a direct or inverse prefix is determined by the relative ranking of A and O on the Cherokee indexability hierarchy, shown below in (160).

- (160) Cherokee indexability hierarchy (Scancarelli 1987:126, item 12)
 1, 2 > 3 human > 3 non-human animate > inanimate

First and second persons are the highest ranked arguments, but are equal with respect to each other. These are followed by subtypes of third persons, predictably with humans above non-human animates, and with inanimates ranked lowest. When A outranks O, the active prefix is used. When O outranks A, the inverse prefix is used. The forms of the active and inverse prefixes are shown in (161) and (162), with examples sentences in (163) and (164).

- (161) Active prefixes (Scancarelli 1987:80, Tables 5 & 6)

<u>A-stems</u>	<u>B-stems</u>
SG a-, ka-	u:-
PL ani:-	u:ni:-

³⁷ There are theoretical treatments of the transitive paradigm which could result in a slightly different characterization of the alignment patterns. Note that the re-use of some of the intransitive forms for certain transitive ones is obviously not arbitrary. Scancarelli operates in the framework of Extended Word and Paradigm Theory (Anderson 1986) to analyze the apparently A prefixes on transitives “as marking two arguments, but one of those may be a dummy. The B prefixes are one-argument prefixes, and are used on certain transitive verbs when one argument is not marked on the verb” (1987:78).

³⁸ Scancarelli (1987) calls this pattern *active-inverse* because she interprets the alternation to be somewhat comparable (though only somewhat) to active/passive alternations in other languages. The system itself, however, is more broadly described as hierarchical.

(162) Inverse prefixes (Scancarelli 1987:81, Table 7)

	<u>SG O</u>	<u>PL O</u>
SG A	u:-	u:ni:-
PL A	kv:wa-	kv:wani:-

(163) Active marking of human A and non-human animate O (Scancarelli 1987:128, ex. 13b)

ani:ke:hy so:kwil t-**a:n**-ahyathe:-ʔa
 women horse DIST-**ACV**-kick-PRES
 ‘The women are kicking the horses.’

(164) Inverse marking of non-human animate A and human O (Scancarelli 1987:129, ex. 13d)

so:kwil **kv:wan**-ahyvthe:-ʔa ani:ke:hya
 horse **INV**-kick-PRES women
 ‘The horses are kicking the women.’

When A and O are both third-person but otherwise of equal rank, the choice of prefix depends on which is proximate and obviative. Scancarelli (1987) borrows these terms from Algonquian linguistics and uses them almost identically: proximate arguments are foregrounded in actual space or discourse prominence; obviative ones are backgrounded. The A role is usually proximate and O usually obviative, but when O is given or assumed from context, or foregrounded for discourse reasons, then it is proximate.³⁹ Since proximate is higher ranked than obviative, constructions with a proximate O receive the inverse prefix. The active-inverse pattern in Cherokee is induced by combinations of person for A and O, as well as by animacy.

4.3.11.5 Summary of split-alignment in Cherokee

The conditions verbal semantics and lexical specification motivate the choice between the Set A and Set B markers on intransitive verbs, with an additional split induced by aspect.

³⁹ These are translated into English using the passive, but they are not passive in Cherokee in the sense of involving a change in valency.

Various combinations of person, number, and animacy condition alignment for transitives.

Because there are 120 cells in the transitive paradigm chart alone, the summary table below in (165) shows only a few forms.

(165) Tabular summary of split-alignment in Cherokee

<u>Person/ number</u>	<u>S</u>	<u>A & O</u>	<u>Pattern</u>	<u>Inducer</u>
1SG	ci-, k-		split-S _{agt}	verbal semantics (agentive) & aspect (present), verbal semantics (agentive) & aspect (imperfective), verbal semantics (agentive) & aspect (punctual), verbal semantics (agentive) & aspect (future), lexical specification (agentive) & aspect (present), lexical specification (agentive) & aspect (imperfective), lexical specification (agentive) & aspect (punctual), lexical specification (agentive) & aspect (future)
2SG	hi-			
⋮				
1SG	aki-, akw-		split-S _{pat}	verbal semantics (patientive), lexical specification (patientive), verbal semantics (agentive) & aspect (perfective), verbal semantics (agentive) & aspect (infinitive), verbal semantics (agentive) & aspect (pre-inceptive), verbal semantics (agentive) & aspect (propensitative), lexical specification (agentive) & aspect (perfective), lexical specification (agentive) & aspect (infinitive), lexical specification (agentive) & aspect (pre-inceptive), lexical specification (agentive) & aspect (propensitative)
2SG	ca-			
⋮				
1SG⇒2SG		kv:y	portmanteau	person (A = 1 & O = 2), person (A = 2 & O = 1), SAP status (–SAP) & number (plural)
1SG⇒2DL		stv:y		
⋮				
1SG⇒3AN		ci-, k-	portmanteau (A/O = S _A)	animacy (O = –animate)
2SG⇒3AN		hi-		
⋮				
3SG⇒1SG		aki-, akw-	portmanteau (A/O = S _B)	SAP status (A = –SAP) & number (A = singular)
3SG⇒2SG		ca-		
⋮				
3SG⇒3SG (ACV)		a-, ka-	hierarchical	SAP status (A = –SAP & O = –SAP)
3SG⇒3SG (INV)		u:		
⋮				

4.2.12 Split-alignment in Euchee

4.2.12.1 Euchee the language

Euchee is an isolate originally from the southeastern United States, though most speakers now live in Oklahoma. It is variously spelled Yuchi or Uchee (or other versions), but Linn (2011:4) observes that “today, most, though not all, Euchee identify with the ‘eu’ spelling” and I have opted to follow her precedent. The language is moribund with around a dozen fluent speakers, all elderly. It is grouped in Region 4: Northeast and Southeast.

4.2.12.2 Overview of Euchee alignment

Euchee has a split-S alignment system in the first and second persons, determined by verbal semantics and lexical specification, so this is a nominative-accusative/ergative-absolutive split. There is a neutral pattern in the third person. All verbs are either eventive or stative, with distinct pronominal sets required to co-index verbal arguments.

4.2.12.3 Split-S alignment in Euchee

The two sets of pronominals are shown in (166).

(166) Actor and patient pronominal sets⁴⁰ (Linn 2001:131, 135)

	<u>Actor</u>	<u>Patient</u>
1SG	di-	dze-
1PL.INC	‘ō-	‘ōdze-
1PL.EXC	nō-	nōdze-
2SG	ne-	nedze-
2PL	‘ā-	‘ādze-
3SG.EUM.MS	hē-/hō-	hē-/hō-
3PL.EUM.MS	hō-	hō-
3SG.EUM.WS	s’e-	s’e
3PL.EUM.WS	‘o-	‘i-/’o-
3SG.EUF	se-	se-
3SG.NNEU	we-	we-
3PL.NNEU	we-	we-

Eventive verbs can refer “to an activity, process, or motion” (Linn 2001:130), and require a pronominal prefix drawn from the actor set. The examples in (167)-(170)⁴¹ show pronominals from the actor set used on intransitive verbs.

(167) First-person **S** marked with the actor set (Linn 2001:133, ex. 18)

di-k’æ
1SG.ACT-laugh
 ‘I laugh’

(168) First-person **A** marked with the actor set (Linn 2001:143, ex. 29a’)

nedze-**di-**‘nē
2SG.PAT-1SG.ACT-see
 ‘I see you’

(169) Second-person **S** marked with the actor set (Linn 2001:133, ex. 18)

ne-k’æ
2SG.ACT-laugh
 ‘you laugh’

⁴⁰ The SAP patient forms appear to be easily segmentable by adding *dze* to the actor forms. Whether one treats the two sets as containing distinct overt morphemes, as the table in (166) does, or as fully decomposable where the patient form is indicated by a suffix or enclitic *-dze* and the actor form is indicated by Ø, is not at issue here. Either analysis would result in the same “sameness” or “differentness” of the two sets, and the alignment pattern would not be affected.

⁴¹ Interlinear glosses of Euchee sentences are my own, extrapolated from explanations and translations provided by Linn (2001) and personal communication (8/5/14).

- (170) Second-person **A** marked with the actor set (Linn 2001:143, ex. 29b')
 dze-**ne**-'nẽ
 1SG.PAT-**2SG.ACT**-see
 'you see me'

Other examples of intransitive eventive verbs are those meaning 'dance,' 'swim,' 'urinate,' 'sleep,' 'crawl,' and 'stand up.' Certain emotional verbs, which might be treated as states in other languages, are lexically specified as events in Euchee: 'be dizzy,' 'be scared,' 'be angry,' 'be hungry,' and others. Thus verbal semantics and lexical specification are both at play here.

Subjects of intransitive statives and objects of transitives are marked with the patient set of pronominals, as the first-person examples in (171) and (172) show.

- (171) First-person **S** marked with the patient set (Linn 2001:138, ex. 24)
 dze-sh'o
 1SG.PAT-be.tired
 'I am tired'
- (172) First-person **O** marked with the patient set (Linn 2001:143, ex. 29b')
 dze-ne-'nẽ
 1SG.PAT-2SG.ACT-see
 'you see me'

Other examples of intransitive stative verbs are those meaning 'regret,' 'tremble,' 'be ready,' and 'be in pain.'

The table in (166) above reveals that no distinction is made between the actor and patient sets for any of the third-person forms, even though these are distinguished along the lines of number, gender, sex-exclusivity, and in-group status. The third-person pronominals therefore align in a neutral pattern, as in (173)-(176).

- (176) Third-person “active” S (Euchee male, men’s speech) (Linn 2001:133, ex. 18)
hẽ-k’æ
3SG.EM.MS-laugh
 ‘he laughs’
- (177) Third-person “patientive” S (Euchee male, men’s speech) (Linn 2001:138, ex. 24)
hẽ-sh’o
3SG.EM.MS-be.tired
 ‘he is tired’
- (178) Third-person A (Euchee male, men’s speech) (Linn 2011:129, ex. 14b)
 s’æ **hẽ-dze-ti**
 down **3SG.EM.MS-1SG.PAT-kick**
 ‘he kicked me’
- (179) Third-person O (Euchee male, men’s speech) (Linn 2001:144, ex. 31)
hẽ-di-’nẽ
3SG.EM.MS-1SG.ACT-see
 ‘I see him’

4.2.12.4 Summary of split-alignment in Euchee

Euchee has a straightforward split-S system for SAP arguments and a neutral pattern for non-SAP ones. Alignment is based on verbal semantics, lexical specification, and SAP status. The summary table in (180) shows the system.

(180) Tabular summary of split-alignment in Euchee

<u>Person/ number</u>	<u>A</u>	<u>S</u>	<u>O</u>	<u>Pattern</u>	<u>Inducer</u>
1SG	di-	di-	dze-	split-S _{nom-acc}	SAP status (+SAP) & verbal semantics (active) SAP status (+SAP) & lexical specification (active)
1INC	‘õ-	‘õ-	‘õdze-		
1EXC	nõ-	nõ-	nõdze -		
2SG	ne-	ne-	nedze-		
2PL	‘ã-	‘ã-	‘ãdze-		
1SG	di-	dze-	dze-	split-S _{erg-abs}	SAP status (+SAP) & verbal semantics (inactive) SAP status (+SAP) & lexical specification (inactive)
1INC	‘õ-	‘õdze-	‘õdze-		
1EXC	nõ-	nõdze-	nõdze- -		
2SG	ne-	nedze-	nedze- -		
2PL	‘ã-	‘ãdze-	‘ãdze-		
3SG.EUM.MS	hẽ-/hõ-	hẽ-/hõ-	hẽ-/hõ-	neutral	SAP status (–SAP)
3PL.EUM.MS	hõ-	hõ-	hõ-		
3SG.EUM.WS	s’e-	s’e-	s’e-		
3PL.EUM.WS	‘o-	‘o-	‘o-		
3SG.EUF	se-	se-	se-		
3SG.NNEU	we-	we-	we-		
3PL.NNEU	we-	we-	we-		

4.2.13 Split-alignment in Lowland Chontal

4.2.13.1 Lowland Chontal the language

Lowland Chontal is one of two Tequistlatecan languages spoken in southern Oaxaca, Mexico. The other is Highland Chontal. Neither language is to be confused with Tabasco

Chontal, which is in the Mayan family and spoken on the north side of the isthmus (Campbell 1997). Lowland Chontal has about 250 fluent, and around a thousand less-than-fluent, speakers (O'Connor 2010). It is also known as Huamelulteco, after one of the principle towns. It is grouped in Region 5: Mesoamerica.

4.2.13.2 Overview of alignment in Lowland Chontal

Lowland Chontal has a basic agentive-patient (or “agentive/non-agentive” in O'Connor's 2010 terms) split-S system for all persons/numbers except third-person singular. Third-person singular alignment is neutral. Within the split-S system, agent-marking is achieved by a set of free pronouns and patient marking by a set of bound affixes (all except first-person singular are suffixes).

4.2.13.3 Split-S alignment in Lowland Chontal

Agent arguments of both transitive and intransitive verbs are marked with an independent pronoun distinguishing person and number. Patient arguments of both transitive and intransitive verbs are marked with person/number affixes; the first-person singular is a prefix while the rest are suffixes. The morphological forms are shown in the table in (181).

(181) Agent argument markers (modified from O'Connor 2010:110, Figure 1)

	<u>Agent</u>	<u>Patient</u>
1SG	iya'	jl-
1PL	iyank'	-onga'
2SG	ima'	-o'
2PL	imank'	-olwa'
3SG	Ø	Ø
3PL	Ø	-ola'

The agent marker is used for arguments which are semantic agents. O'Connor (2010:111) describes agency "as reflecting the volition or intention of the participant. Agency can be inherent in the lexical semantics, as with *ma-* 'die' or *toj-* 'grow', or it can be attributed or perceived, as with *xux-* 'be late' *jak-* 'disappear.'" This suggests that the system may also be fluid-S and that verbal semantics and participant semantics are both at play, in addition to lexical specification. As shown in the table in (181), agent markers are free morphemes while patient markers are affixes.

The sentences in (182) and (183) illustrate the first-person plural agent marker.

- (182) First-person plural **A** marked with agent marker (O'Connor 2010:110, ex. 2)
 pijl-pa' **iyank'** lantranay'
 kill-PFV.PL **1PL.AGT** chickens
 'We killed the chickens.'
- (183) First-person plural agentive **S** marked with agent marker (O'Connor 2010:110, ex. 3)
 may-pa' **iyank'**
 go-PFV.PL **1PL.AGT**
 'We went.'

The examples in (184) and (185) show the first-person plural patient marker.

- (184) First-person plural patientive **S** marked with patient marker (O'Connor 1999:3, ex. 6)
 toj-p-**onga'**
 grow-PFV-**1PL.PAT**
 'We grew (up).'
- (185) First-person plural **O** marked with patient marker (Waterhouse 1962:21)⁴²
 pénk'ipo-**onga'** máxʔ-éex
 tied-**1PL.PAT** LOC-tree
 'They tied us to a tree'

⁴² In examples taken from Waterhouse (1962), I have had to infer the interlinear glosses from available grammatical and lexical information. In addition, I regularized the relevant orthography to conform with that in O'Connor's more recent work.

4.2.13.4 Alignment of third-person arguments in Lowland Chontal

Third-person markers show neutral alignment in the singular but again split-S in the plural. The singular arguments are not overtly marked, as in (186)-(189).

- (186) Third-person singular A is zero-marked (O'Connor 1999:2, ex. 2)
pijl-pa lantranay'
kill-PFV chickens
's/he killed the chickens'
- (187) Third-person singular agentive S is zero-marked (O'Connor 2010:110, ex. 4)
may-pa
go-PFV.SG
'S/he went.'
- (188) Third-person singular patientive S is zero-marked (O'Connor 2010:111, ex. 8)
tyoj-pa
grow-PFV.SG
'S/he grew (up).'
- (189) Third-person singular O is zero-marked (O'Connor 1999:3, ex. 5)⁴³
mi-pa
tell.someone-PFV
's/he told her/him'

The constructions in (187) and (188), both involving S arguments, but with (187) as agentive and (188) as patientive, show that the split-S system is not operative in the third-person singular. These arguments are neutrally marked across the board. For plural third persons, only the patient is overtly realized, as in (190)-(193).

- (190) Third-person plural A is zero-marked (O'Connor 1999:2, ex. 2)
pijl-pa' lantranay'
kill-PFV.PL chickens
'they killed the chickens'

⁴³ The verb *mi-* 'to tell someone' is described as transitive, not ditransitive in O'Connor (1999).

- (191) Third-person plural agentive S is zero-marked (O'Connor 2010:110, ex. 4)
 may-pa'
 go-PFV.PL
 'They went.'
- (192) Third-person plural patientive S is marked with the patient marker (O'Connor 2010:113, ex. 10)
 xo-gix-p-**ola**'
 be.tired-AND-PFV-**3PL.PAT**
 'they got tired'
- (193) Third-person plural O is marked with the patient marker (O'Connor 1999:3, ex. 5)
 mi-p-**ola**' iyank'
 tell.someone-PFV-**3PL.PAT** 1PL.AGT
 'we told them'

Third-person plural constructions are zero-marked for the A and S_a functions, but overtly marked for the S_o and O functions.

4.2.13.5 Summary of split-alignment in Lowland Chontal

All person/number combinations except the third-person singular show a split between nominative-accusative and ergative-absolutive; neutral alignment exists for the third-person singular. A summary table is shown in (194).

(194) Tabular summary of split-alignment in Lowland Chontal

<u>Person/ number</u>	<u>A</u>	<u>S</u>	<u>O</u>	<u>Pattern</u>	<u>Inducer</u>
1SG	iya'	iya'	jl-	split-S _{nom-acc}	SAP status (+SAP) & verbal semantics (agentive), SAP status (+SAP) & participant semantics (agentive), SAP status (+SAP) & lexical specification (agentive), SAP status (–SAP) & number (plural) & verbal semantics (agentive), SAP status (–SAP) & number (plural) & participant semantics (agentive), SAP status (–SAP) & number (plural) & lexical specification (agentive)
1PL	iyank'	iyank'	-onga'		
2SG	ima'	ima'	-o'		
2PL	imank'	imank'	-olwa'		
3PL	Ø	Ø	-ola'		
1SG	iya'	jl-	jl-	split-S _{erg-abs}	SAP status (+SAP) & verbal semantics (patientive), SAP status (+SAP) & participant semantics (patientive), SAP status (+SAP) & lexical specification (patientive), SAP status (–SAP) & number (plural) & verbal semantics (patientive), SAP status (–SAP) & number (plural) & participant semantics (patientive), SAP status (–SAP) & number (plural) & lexical specification (patientive)
1PL	iyank'	-onga'	-onga'		
2SG	ima'	-o'	-o'		
2PL	imank'	-olwa'	-olwa'		
3PL	Ø	-ola'	-ola'		
3SG	Ø	Ø	Ø	neutral	SAP status (–SAP) & number (singular)

4.2.14 Split-alignment in Misantra Totonac

4.2.14.1 Misantra Totonac the language

Misantra Totonac is a Totonacan language spoken in the Mexican state of Veracruz. There is scholarly agreement that the Totonacan family is divisible into two branches, Totonac and Tepehua, but there is disagreement regarding the internal structure of Totonac itself. Campbell (1997:161) lists two Totonac languages, Puebla and Veracruz, noting parenthetically that there

are “several dialects.” MacKay (1999), on whose grammar this section is based, says that there are four Totonac languages. The Ethnologue, admittedly splittist in its language counting, lists nine Totonac languages. It is not clear the extent to which some of these are mutually intelligible with the others or to which the linguonym *Totonac* represents a macrolanguage.⁴⁴ It is grouped in Region 5: Mesoamerica.

4.2.14.2 Overview of alignment in Misantla Totonac

Misantla Totonac has nominative-accusative and neutral alignment patterns. The pronominals are split between nominative-accusative, which occurs with most person-number combinations, and neutral alignment, which occurs with the zero-marked third-person singular.

4.2.14.3 Alignment of pronominals in Misantla Totonac

All of the pronominals except the third-person singular show nominative-accusative alignment. The third-person singular ones are zero-marked and therefore have neutral alignment. The table in (195) shows the pronominal forms, with certain details given in the footnotes.

⁴⁴ Because my sampling procedure draws randomly from the list in the Ethnologue, the matter is not moot for me. The first language selected by my procedure is given in the Ethnologue as *Yecuatla Totonac* [tlc]. Yecuatla is a town in the state of Veracruz, and while neither Campbell (1997) nor MacKay (1999) designate a distinct variety of Totonac with this name, MacKay (:12) includes this town under the heading *Misantla Totonac* and not under any of the other three variety names. For this reason, I am assuming that the random selection *Yecuatla Totonac* is collateral of *Misantla Totonac*.

(195) Pronominals in Misantla Totonac⁴⁵

	<u>A</u>	<u>S</u>	<u>O</u>
1SG	ik-	ik-	kin- / kinlaa-
2SG	~ / ~-ʔ / supp. ⁴⁶	~ / ~-ʔ / supp.	-na / taa-...-na ⁴⁷
3SG	Ø	Ø	Ø
1PL	(ik-)...-wa ⁴⁸	(ik-)...-wa	kin-...-na / kinlaa
2PL	~-tat	~-tat	taa-...-na
3PL	ta-	ta-	laa-

A few brief examples will serve to illustrate the two patterns. (196)-(198) show first-person singular forms.⁴⁹

(196) First-person singular **A** pronominal (MacKay 1999:169, ex. 202)

kit **ik**-paastak-yaa-na
 I **1SG.SUB**-remember-IMPV-2SG.OBJ
 ‘I remember you.’

(197) First-person singular **S** pronominal (MacKay 1999:129, ex. 45)

kit **ik**-iš-paš-cl
 I **1SG.SUB**-PAST-bathe-CL
 ‘I was bathing.’

⁴⁵ As is clear from the table, some of the pronominals show interesting allomorphy, which is sometimes phonological, sometimes morphological, and sometimes optional. In most cases, the effect of allomorphy on alignment is null. But there are two scenarios in which it may bear on the alignment system, depending on how one analyzes the degree to which person information is encoded in the pronominals. The first scenario is the optional O pronominal *kinlaa*, for first-person singulars and plurals. The presence of this morpheme prevents inflection of A if A is in the second person. The second scenario is the pronominal *taa-...na* for second-person plural Os. The presence of this morpheme prevents inflection of A if A is first-person plural. MacKay uses the term “suspends” to refer to the prevention of inflection of the A argument in these scenarios. The reason that different analyses may influence the characterization of the alignment system is because if one assumes that under the conditions just described, the A argument is rendered by a phonetically null morpheme, then alignment for those person/number combinations would be tripartite. On the other hand, if one assumes that the A argument is encoded into the overt pronominal along with the O argument, then alignment would be portmanteau for those person/number combinations. This is an interesting theoretical issue, but one which I leave to future work to investigate in detail.

⁴⁶ The second-person singular A or S form either glottalizes the preceding vowel, or glottalizes the preceding vowel and suffixes a glottal stop, based on phonological factors. The form is marked by suppletion of the verb stem only for verbs based on ‘go’ and ‘come.’

⁴⁷ The form *taa-...-na* is used when A is first-person plural.

⁴⁸ The *ik* component is optional.

⁴⁹ As is clear from the examples, Misantla Totonac also uses free pronouns. These do not contribute to a description of the alignment pattern because there are no object pronouns.

- (198) First-person singular **O** pronominal (MacKay 1999:165, ex. 177)
 utun **kin**-ta-páasték
 they **1SG.OBJ-3PL.SUB-remember**
 ‘They remember me.’

Since the pronominal marking for A and S is the same, but different for O, alignment is nominative-accusative. The neutral pattern, which occurs in the third-person singular, is shown in (199)-(201).

- (199) Third-person singular **A** pronominal (MacKay 1999:161, ex. 161)
 ut **Ø**-laa-waa
 s/he **3SG.SUB-3PL.OBJ-eat**
 ‘S/he eats them.’

- (200) Third-person singular **S** pronominal (MacKay 1999:201, ex. 397)
 ut **Ø**-tata
 s/he **3SG.SUB-sleep**
 ‘S/he sleeps.’

- (201) Third-person singular **O** pronominal (MacKay 1999:172, ex. 222)
 kit **Ø**-ik-tihwan
 I **3SG.OBJ-1SG.SUB-look.for**
 ‘I look for him/her/it.’

4.2.14.4 Summary of split-alignment in Misantla Totonac

Misantla Totonac shows two alignment patterns, nominative-accusative and neutral, with the split induced by SAP status and number (singular non-SAPs are neutral, all others are nominative-accusative). A summary table is shown below in (202).

(202) Tabular summary of split-alignment in Misantla Totonac

<u>Person/ number</u>	<u>A</u>	<u>S</u>	<u>O</u>	<u>Pattern</u>	<u>Inducer</u>
1SG	ik-	ik-	kin-/kinlaa-	nominative-accusative	SAP status (+SAP), SAP status (–SAP) & number (plural)
2SG	/~ʔ/supp.	/~ʔ/supp.	-na/taa-...-na		
1PL	(ik-)...-wa	(ik-)...-wa	kin-...-na/kinlaa		
2PL	~tat	~tat	taa-...-na		
3PL	ta-	ta-	laa-		
3SG	Ø	Ø	Ø	neutral	SAP status (–SAP) & number (singular)

4.2.15 Split-alignment in Chol

4.2.15.1 Chol the language

Chol, also spelled Ch’ol, is a Mayan language spoken by between 150,000 and 200,000 people mostly in the northeastern municipalities of Chiapas, Mexico. Within Mayan, Chol is most closely related to Chontal,⁵⁰ Ch’orti’, Tzeltal, and Tzotzil. Together these languages make up the Cholan-Tzeltalan subgroup of Mayan (Campbell 1997). Chol is thus more distantly related to the Q’anjob’alan, K’ichean, Mamean, and Yucatecan languages. Its endonym is *lak ty’añ*.

4.2.15.2 Overview of alignment in Chol

Alignment in Chol is split-S, with a split between ergative-absolutive and nominative-accusative types conditioned by aspect. There are two sets of argument markers, Set A and Set B, shown below.

⁵⁰ This is Tabasco Chontal, distinct from the Chontals in the Tequistlatecan family, described in §4.2.13.

- (203) Argument affixes in Chol (modified from Vázquez Álvarez 2011:77, Table 7 and :79, Table 8)

	<u>Set A</u>	<u>Set B</u>
1	k-, j- ⁵¹	-oñ
2	a-, aw- ⁵²	-ety
3	i-, (i)y- ⁵³	-Ø

4.2.15.3 Split-S alignment in Chol

When used in the perfective aspect, Set A indicates A arguments, while Set B indicates S and O. Thus the pattern is ergative-absolutive.

- (204) First-person **A** in the perfective (Vázquez Álvarez 2011:73, ex. 25a)

tyi k-päk'-ä-Ø repolloj
 PFV A1-plant-TR-B3 cabbage
 'I planted cabbage.'

- (205) First-person **S** in the perfective (Vázquez Álvarez 2011: 25, ex. 26b)

tyi wäy-i-y-oñ
 PFV sleep-INTR-EP-B1
 'I slept.'

- (206) First-person **O** in the perfective (Vázquez Álvarez 2011:65, ex. 11a)

tyi i-päy-ä-oñ
 PFV A3-call-TR-EP-B1
 'S/he called me.'

The morpheme indicating the first-person A argument in (204) comes from Set A. The morphemes indicating first-person S in (205) and first-person O in (206) both come from Set B.

The ergative-absolutive pattern only occurs in the perfective aspect.

⁵¹ /__k

⁵² /__V

⁵³ /__V

When used in the non-perfective aspects (imperfective and progressive), Set A indicates A and S arguments, while Set B indicates only O.

- (207) Second-person **A** in the imperfective (Vázquez Álvarez 2011:167, ex. 214b)

tyi a-paty mi a-kuch-Ø tyäl-e
 PREP A2-back IMPFV A2-carry-B3 toward-NF
 ‘You carry it on your back.’

- (208) Second-person **S** in the imperfective (Vázquez Álvarez 2011:26, ex. 27b)

mi a-wäy-el
 IMPFV A2-sleep-NF
 ‘You sleep.’

- (209) Second-person **O** in the imperfective (Vázquez Álvarez 2011:153, ex. 17b)

joñoñ mi j-kāñ-ety
 I IMPFV A1-know-B2
 ‘I know you.’

The morphemes indicating the second-person A argument in (207) and the second-person S argument in (208) are both from Set A. The morpheme indicating the second-person O argument in (209) comes from Set B. So this pattern is nominative-accusative. It requires that the aspect be non-perfective.⁵⁴

4.2.15.4 Summary of split-alignment in Chol

Chol shows an ergative-absolutive and nominative-accusative split on the basis of aspect.

The table in (210) summarizes the system.

⁵⁴ There is another potential layer to the alignment system in Chol, depending on how one analyzes certain constructions which are usually translated into English using intransitive verbs. There is a class of intransitive-like constructions which differ from the others in their morphology. This second class cannot receive person inflection directly on the verb. Instead, they require a light verb *cha'l* ‘do’ as a host for the person affixes; the bare intransitive verb is then positioned after the light verb complex. In addition, the person affix will always come from Set A, never Set B. If one views these constructions as genuinely intransitive, then these would necessarily constitute another split in the alignment system. However, they are very similar to transitive constructions with (zero-marked) third-person objects, which is, in fact, how Vázquez Álvarez glosses them. Rather than treating these as intransitives, I opt to follow Coon (2013) in interpreting them as transitive constructions with verbal complements. Since these take Set A markers in all aspects, like other transitive verbs do, there is no need to amend the characterization of split alignment in order to account for them.

(210) Tabular summary of split-alignment in Chol

<u>Person</u>	<u>A</u>	<u>S</u>	<u>O</u>	<u>Pattern</u>	<u>Inducer</u>
1	k-, j-	k-, j-	-oñ	split-S _{nom-acc}	aspect (non-perfective)
2	a-, aw-	a-, aw-	-ety		
3	i-, (i)y-	i-, (i)y-	Ø		
1	k-, j-	-oñ	-oñ	split-S _{erg-abs}	aspect (perfective)
2	a-, aw-	-ety	-ety		
3	i-, (i)y-	Ø	Ø		

4.3 Summary of split inducers in the sampled languages

The subsections in 4.2 detailed the various split inducers in each of the sampled languages and gave relevant examples to show the consequences on the alignment systems. The tabular summaries at the end of the each of those subsections show (a) the split inducer category, e.g. SAP status; (b) the split inducer value, e.g. –SAP; (c) the alignment pattern associated with each split inducer, e.g. ergative-absolutive; and (d) the morphological form of the relevant locus of alignment for A, S, and O.

Since the emphasis of this project is the split inducers themselves, the table in (211) shows those split inducers occurring in each sampled language.

(211) Split inducers by language

Blackfoot	Particularness
Crow	Lexical specification, SAP status, verbal semantics
Pawnee	Animacy, individuation, number, SAP status
Alsea	Nominal status
Haida	Number, participant semantics, person, verbal semantics
Tlingit	Definiteness, Lexical specification, nominal status, SAP status, verbal semantics
Laguna Keres	Lexical specification, SAP status, verbal semantics
Maricopa	Nominal status, SAP status
Seri	Number, person, SAP status
Alabama	Participant semantics, SAP status
Cherokee	Animacy, aspect, lexical specification, number, person, SAP status, verbal semantics
Euchee	Lexical specification, SAP status, verbal semantics
Lowland Chontal	Number, participant semantics, SAP status, verbal semantics
Mistantla Totonac	Number, SAP status
Chol	Aspect

5 ANALYSIS OF RESULTS

5.1 Introduction to Chapter 5

In this chapter, I present my analysis of the distribution of split-inducing conditions found in the 15 sampled genera, as detailed in Chapter 4. My analysis takes the form of four generalizations which together converge on the empirical observation that certain split inducers exist in the first place and the fact that certain of them are common in the sample while others are rare. These generalizations take the form of statements which appeal to typological, functional, and historical observations in order to boil alignment patterns down to a handful of interrelated factors which contribute to split-alignment behavior. The four generalizations are summarized below in (1).

- (1) Four generalizations
- Generalization 1: Features cumulated with grammatical relation are likely to be split inducers.
 - Generalization 2: Low markedness tends to produce zero expression, which tends to neutralize distinctions between grammatical relations.
 - Generalization 3: Verbal semantics results in alignment splits due to the assignment of semantic roles to arguments by verbs.
 - Generalization 4: Lexical specification is dependent on the verbal semantics criterion and so induces splits by creating exceptions to semantic class membership.

The four generalizations do not operate equally across all genera or across all alignment patterns. Generalizations 1 and 2 are the main principles in factoring in the most common split inducer in the sample, SAP status, as well as number, and a handful of less common inducers. Generalizations 3 and 4 contribute to my description of the second and tied-for-third best

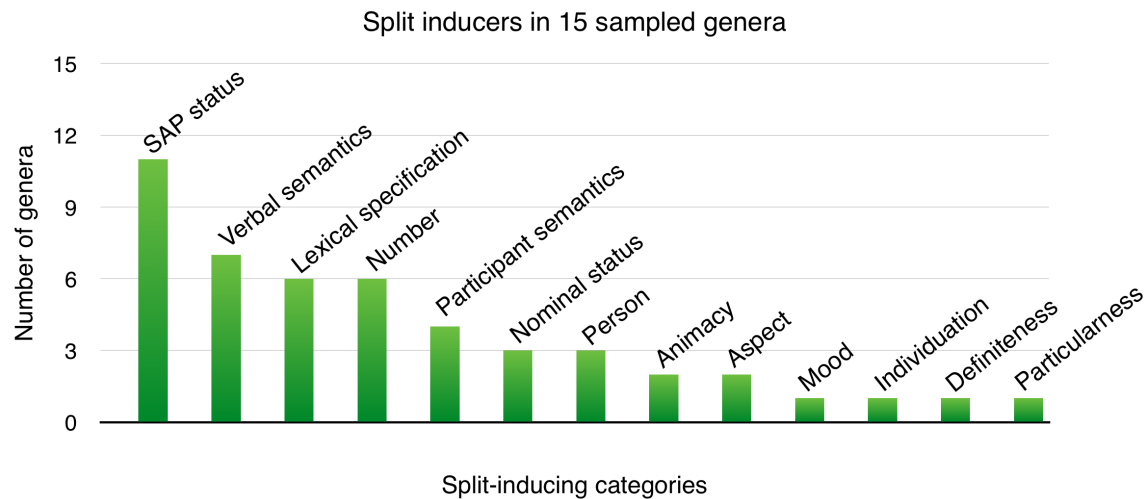
represented split inducers, verbal semantics and lexical specification, respectively. These labels will be defined in §5.3. Of the thirteen split inducers uncovered in the sampled languages, seven of them are attributable to certain of the four generalizations. There is residue which is not accounted for directly by the generalizations, but I offer suggestions on how these might be treated.

In 5.2, I describe the statistical representation of the various split inducers and put this against the backdrop of “Dryerian preferences,” which motivated my study design and sampling methodology. In 5.3, I describe in precise terms the four generalizations and provide evidence from the fifteen sampled languages in support of my claim that they are forces contributing to commonness and rarity of split inducers. In 5.4, I discuss the residue not attributable to the four generalizations. I conclude the chapter in 5.5.

5.2 Overall distribution and Dryerian preferences

The chart in (2) shows the overall distribution of the split inducers found in the fifteen sampled genera.

(2) Representation of all split inducers in the sample



The study methodology laid out in Chapter 3 was largely inspired by Dryer's (1989, 1992) strategy for determining cross-linguistic preferences for basic word order. Recall that Dryer divided his geographic space into non-overlapping areas and sampled genera from each area. This is what I did, too, but to reiterate the differences: in Dryer's case, the geographic space was the entire world whereas in mine it is North America. Dryer furthermore looked at more than one language from many genera, but I sampled exactly one language from each randomly selected genus.

Dryer defines his notion of *linguistic preference* in simple statistical terms. If, in all geographic areas, there are more genera with languages displaying a particular variable (in his case SOV word order) than genera with languages not displaying it, then there is an overall linguistic preference for that variable with less than five percent statistical significance. If most, but not all, areas contain genera with languages displaying the variable, then one can describe that variable in weaker terms, using a label like *trend*, but such a scenario does not count as a

preference for Dryer. So that I can draw on Dryer's definition specifically, without necessarily adopting it myself, I will refer to his notion as a *Dryerian preference*.

In my study, the variables under investigation are the various split inducers. While it is obvious from Chart (2) that some split inducers are more common than others, none of them is a Dryerian preference because none is involved in split-induction in at least two-thirds of the genera from every region. The best represented split inducer, SAP status, is involved in split alignment in over seventy percent of sampled genera. There is one region, however, in which only a third of the genera have it (Region 2). Verbal semantics is a split inducer in almost fifty percent of the genera, but it is not a Dryerian preference because it is absent from two-thirds of the genera in Region 1, two-thirds in Region 3, and two-thirds in Region 5. Lexical specification and number are tied for third place, but again, neither is a Dryerian preference.

The fact that none of the split inducers found in the sample is a Dryerian preference is telling in and of itself. It means that there are no split inducers common enough to permeate at least two-thirds of genera selected at random from each region. Such a finding should temper any inclinations to describe split inducers as being Dryerian preferences, and linguists should exercise caution when characterizing the distribution of split inducers anywhere in the world.

Of course, in a *non*-Dryerian sense, there is quite a lot more to be said about the distribution because it is undeniable, as mentioned already, that the distribution is heavily skewed. It may therefore be worth rethinking the utility of Dryer's methodology. For instance, Dryer's overall conclusions about the correlations of certain word order features are challenged by Dunn et al. (2011), who employ a wholly different methodology for assessing the correlations. They use a Bayesian phylogenetic approach which assesses correlations between

traits within and across four language families whose internal reconstructions are well-established and whose subfamilial time-depths are available. This allows them to “track correlated changes [...] as languages split and diversify” (ibid:79). They argue that the word order correlations predicted by Dryer’s analysis are not strong and that the few genuine dependencies are specific to individual lineages, rather than true across families.

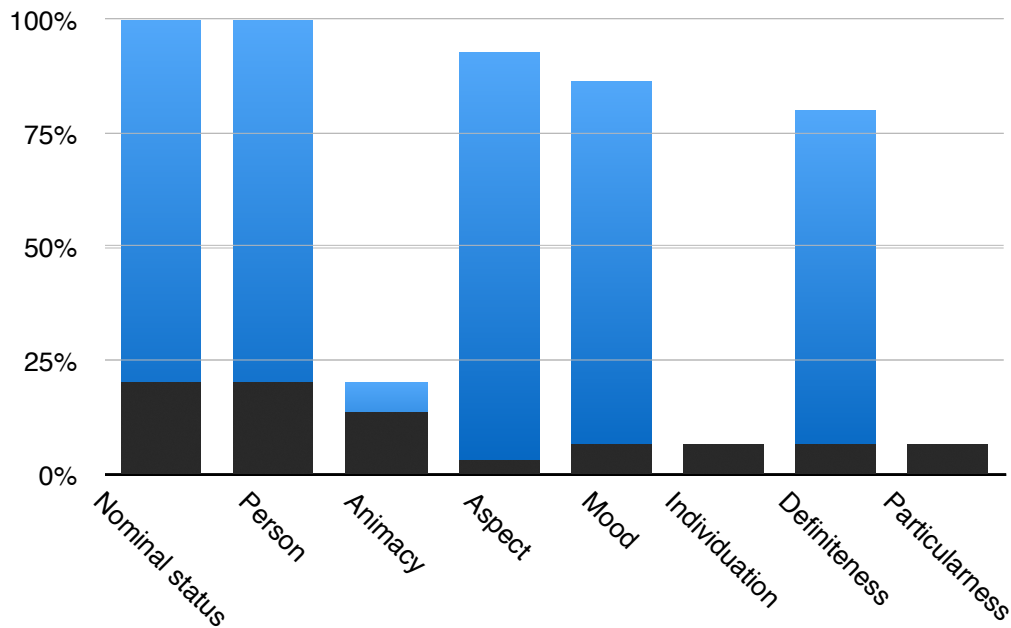
Still, a major strength of Dryer’s method is that it severely reduces the likelihood of genealogical bias simply by virtue of the large number of genera it draws from (and in fact, the large number of totally unrelated families); Dunn et al. only compare four families. My results, though targeting a different phenomenon than word order, largely support Dryer’s methodology because many unrelated languages are taken into account and cross-linguistic trends are still apparent, even if they don’t qualify as Dryerian preferences. In my view, the insight gleaned from Dunn et al.’s method relates to the ability to trace correlations of traits within the histories of individual language families; this would be a fascinating area to explore for alignment behavior. Dunn et al. are able to do it because the families they select already have quantitative phylogenies available, as they are perhaps (among) the best studied languages families. The same cannot be said for all of the families I sampled from (nor that Dryer sampled from). Comparatively less is known about the internal groupings and dates of subfamilial branching for many North American families. Proposals often compete and general scholarly agreement seems to be more the exception than the norm.

The remainder of this section will characterize the study results in terms of commonness and not in terms of Dryerian preferences. Before proceeding, though, it is necessary to qualify the overall statistical results by making two points about the chart in (2). First, the chart shows

the number of genera containing languages with each split inducer *in the sample*. These results are not estimated for all of North America. Applying weights allows for estimates of the number of genera in the entire population expected to have each split inducer. However, there is a very high error of estimation when this is done due to the fact that the areal stratification procedure precluded taking genera from each region proportionally to the region's population. Obviously, a bigger sample would help to reduce the estimation of error, but the data are extremely expensive to collect and considerable time would be required to sample additional genera. Interested readers can see a weighted chart in Appendix 3, which differs somewhat from the one in (2) above.

The second point is that the values shown in (2) are the numbers of languages which have a particular split inducer. Taken at face value, particularness is a less common split inducer than say, aspect because particularness is a split inducer in only one language (Blackfoot), while aspect is a split inducer in two (Cherokee and Chol). In certain ways, this is misleading because not all of the grammatical properties which induce alignment splits across these languages are equally available in all of them. In other words, to say that particularness is "less common" than aspect is numerically true, but not especially informative without knowing in how many languages particularness and aspect are actually realized in the first place. It is therefore more informative to interpret these split inducers not as pure values, but as proportional to the occurrence of each grammatical category within the languages sampled. Recast in this way, a more informative chart is shown in (3), for those split inducers occurring in one-fourth or fewer sampled languages.

(3) Split inducers proportional to the encoding of those properties



The blue columns represent the percentage of sampled languages which encode each grammatical property; all languages encode person, for instance, but only twenty percent encode animacy. The overlapping black columns represent the percentage of each *encoded* property which actually shows up as a split inducer. For example, twenty percent of the languages which encode person have person as a split inducer; two-thirds of the languages which encode animacy have animacy as a split inducer. It is crucial to interpret the data in this way in order to give meaning to the notion “rare” split inducer. The larger the blue column is in proportion to the black, the more significantly rare that property becomes as a split inducer. In this sense, animacy, individuation, and particularness do not come off as very good candidates for the label “rare.” Even though very few languages involve these properties in alignment splits, very few languages in the sample encode these properties in the first place, so it is impossible to tell whether their

low representation as split inducers is due to the fact that they aren't encoded in very many languages or is due to some reason particular to grammatical alignment.¹

The other properties—nominal status, person, aspect, mood, and definiteness—are genuinely rare split inducers because many (in fact most, sometimes all) languages encode these properties and so in theory, these properties *could* be split inducers in many more languages than they actually are. But in fact, they are split inducers in only a handful of these languages.

5.3 Four generalizations

In this subsection, I detail the four generalizations outlined in 5.1 and give evidence supporting my argument that these generalizations are foundational to the grammatical conditions that tend to be split inducers and how common they will tend to be.

5.3.1 Generalizations 1 and 2

Generalizations 1 and 2 are repeated below in (4).

- (4) Generalizations 1 and 2
Generalization 1: Features cumulated with grammatical relation are likely to be split inducers.
Generalization 2: Low markedness tends to produce zero expression, which tends to neutralize distinctions between grammatical relations.

The first generalization refers to cumulative exponence, where multiple pieces of grammatical information are “cumulated into a single set of formatives” (Bickel & Nichols 2007:188). For any morpheme *m*, if *m* codes the feature grammatical relation as well as any other

¹ It is worth mentioning that the chart in (3) only counts the properties encoded with overt and signaling morphology. Many languages do have phenomena sensitive to animacy considerations even in the absence of animacy-signaling morphology (for instance, when only animate nouns can be agents).

feature f , then f is likely to be a split inducer. In other words, features that are cumulated along with the feature grammatical relation will tend to induce changes in morphological form, and will thereby tend to induce different alignment patterns across different values of those features. In the languages I sampled, the most common cumulation was of the features grammatical relation and person. A typical morphological structure is the one schematized in (5a), where the affix is usually a pronominal.

(5a) *affix*-verb, where *affix*: <form, [person, relation, maybe others]>

Another instance of cumulation, this time in pronouns, is schematized in (5b).

(5b) *pronoun* verb, where *pronoun*: <form, [person, relation, maybe others]>

Of course, (5a) and (b) are not the only structures possible. The schematization in (6) is also common; there, one morpheme encodes the person property (and maybe others), while another morpheme encodes grammatical relation. This is what is traditionally called “case marking.” In contrast to cumulative exponence, this scenario is *separative exponence* because the formatives for different grammatical categories are formally distinct.

(6) nominal-*affix*₁ *affix*₂-verb, where *affix*₁: <form, [relation, maybe others]>, *affix*₂: <form, [person, maybe others]>

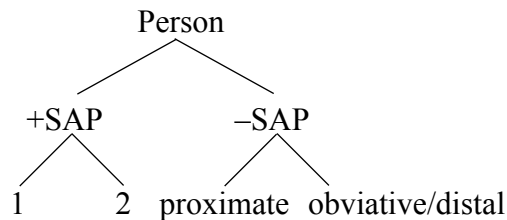
In comparing the two schematized types in (5) and (6), it would be expected on theoretical grounds to find that languages of type (5) tend to have person and grammatical relation encoded on the same morpheme, assuming that grammatical relation is encoded at all, while languages of type (6) would tend to have grammatical relation coded separately from person because coding grammatical relations is one of the primary functions of traditional case

markers. It turns out that this is exactly what is revealed in the languages sampled for this study. Of the languages for which person is involved in split induction, all are of type (5).²

Notice that person is not the most common split inducer. In fact from chart (3), it is clear that it is not even especially common—nonetheless, in those language in which it is a split inducer, it is always cumulated with grammatical relation. From (2), the most common split inducer is SAP status. The status of an argument as – or +SAP (in other words, third or non-third person) is treated here as a split inducer distinct from *individual* person, in which first- versus second-person would have different alignment patterns. The cumulation of person and relation does not, by itself, predict that SAP status will be a common split inducer because the distinction between +SAP and –SAP arguments is a sub-node within the larger person category. In a hierarchical view of the internal structure of the person category, such as the basic and unelaborated one schematized in (7), the first way of splitting the category is according to SAP status. SAPs, comprised of first and second persons, form one subcategory, while non-SAPs, i.e. third persons, make up the other.

² With respect to the observation that all SAP status-split inducing languages are of type (5), two qualifications need to be made. These relate to Maricopa and Tlingit. Maricopa has verbal prefixes encoding person and grammatical relation as well as full pronouns doing the same, and also traditional case markers on full nouns to indicate the A/S relation only. In Tlingit, verbal prefixes are used for person and grammatical relation (and number), but a case marker is used to indicate the A relation. In a language which only uses the case marker on full nouns, the person category is automatically restricted to third because full nouns almost always have third-person referents (except for certain “composite phrasal expressions like ‘the speaker of the present utterance’” [Bresnan 2011:115], which is grammatically but not referentially third person). Maricopa and Tlingit are exactly such languages: the case marker is only used on full noun, and therefore third-person, arguments.

(7) Internal structure of the person category



This fundamental distinction shows up in many areas of language, such as the ability to use lexical expression for reference, the behavior and distribution of demonstratives, constituent order, and cumulation with other grammatical categories. Why should this distinction exist? A leading theory is that denotations of SAPs are situation-sensitive, whereas those of non-SAPs are not. Siewierska (2004:7) nicely summarizes the grammaticalization of the situation-sensitivity of person forms:

All of the above differences are typically seen to be a consequence of the fact that first- and second-person forms are inherently deictic expressions, that is their interpretation is dependent on the properties of the extralinguistic context of the utterance in which they occur. Although the first person is always the speaker of the utterance and the second is the hearer, the actual identity of each depends on who utters the utterance that contains them to whom, when and where. They belong to a class of expressions often referred to as shifters (Jakobson 1971). Third-person forms, on the other hand, are essentially anaphoric expressions. Their interpretation depends not on the extralinguistic but on the linguistic context of the utterance.

The prediction here is that SAP status is more likely than deeper subcategories of person to correspond to differences in the behavior of person-marking generally, and because of the common cumulation of person and grammatical relation, this includes alignment.

In its most straightforward form, SAP status acts as a split inducer when +SAP arguments get one alignment pattern and –SAP arguments get another. The usual scenario is that the –SAP arguments show neutral alignment. In Crow, for instance, +SAP arguments are split-S, while –

SAP arguments are neutral. The reason third-person arguments show neutral alignment is that third-person arguments are all zero-marked. In Euchee, the situation is similar in that +SAP arguments are split-S and –SAP ones are neutral. But in Euchee, unlike in Crow, third-person arguments are overtly marked. In fact, their marking is prolific because the pronominals not only identify an argument as being in the third person, they also distinguish Euchees from non-Euchees, men from women, and men’s speech from women’s. But they do not distinguish A, S, and O.

There are two main ways in which SAP status as a split inducer is not as straightforward as the Crow or Euchee cases. One way is when other split inducers interact with SAP status to produce an alignment pattern. In Pawnee, Lowland Chontal, and Misantla Totonac, for instance, it is only third-person *singular* arguments which are neutrally aligned (like in Crow, these are zero-marked), so number is also a factor. In Tlingit, it is only third-person *definite* arguments which are neutral (again, by virtue of being zeros). In Laguna Keres, neutral alignment only shows up on third-person arguments in the imperative mood.

The other way in which SAP status is involved in less-than-straightforward splits is when specific combinations of the + or – value of SAP status are required of different arguments at the same time in order to induce a particular alignment pattern. This is especially common in Region 3. In Laguna Keres, for example, portmanteau alignment only occurs when *both* A and O are +SAP; if A and O have different values (no matter which has which), alignment is split-S. In Maricopa, the portmanteau pattern requires that O is +SAP, but A can be either + or –. The portmanteau pattern in Seri only occurs when both A and O are –SAP. The languages in which SAP status is involved in split induction are listed in (8).

- (8) Languages in which SAP is involved in split induction
Alabama, Cherokee, Crow, Euchee, Laguna Keres, Lowland Chontal, Maricopa, Misantra
Totonac, Pawnee, Seri, Tlingit

The observation that third-person arguments are points of departure from alignment patterns seen in the first and second persons is also predicted by various interpretations of an indexability hierarchy, in particular that of Silverstein (1976). His analysis is that the higher an argument is on the hierarchy (where $1, 2 > 3$), the more expected it is to be a semantic agent and the less motivation there is to identify it as an agent using explicit morphology. The lower an argument is on the hierarchy, the less expected it is to be a semantic agent, and so explicit morphology (such as an ergative marker) is required to override the low expectation of agency.

The second generalization is that low markedness tends to correlate with zero expression. This is a well known typological observation. With respect to alignment and the high representation of SAP status, the low markedness of third person tends to create neutral alignment by zero expression. Third person is commonly (though of course not universally) unmarked cross-linguistically. Cysouw (2003:61) notes:

Zero marking for speaker or addressee is rare among the languages of the world. In contrast, a zero for the marking of the ‘other’ participant is found commonly, in all parts of the world. This predominance is easily explained by basic markedness. The third person is the unmarked person, and the unmarked person has the highest probability of being zero.

This is true in several senses of *unmarked*. First, it is true in the early sense of Jakobson, whose decades’ long work on markedness starting with the Russian verb led him to propose a *personal* feature for the person category; the personal value (consisting of first and second persons) is the marked one and the lack of this value (corresponding to third persons) is unmarked (Battistella 1996, Bybee 2011). This markedness corresponds literally to overtness: the

marked members have positive realization while the unmarked members are realized as zeros. It is also true in the Greenbergian sense, where markedness is viewed as a scalar, rather than strictly binary, notion. For Greenberg (1966:44-5), third person is less marked than first or second because of its tendency towards contextual neutralization, other-person surrogacy, and high text frequency³ in addition to zero expression.

It is certainly true that many languages in my sample show zero-marking in the third person (or in a subcategory of third person). Nearly 73% of the languages for which SAP is a split-inducer have third person expressed as zero for A, S, and O, which is a neutral alignment pattern. Provided that some other alignment types are found for the SAP arguments, which they always are in the languages I sampled, the neutral pattern for non-SAPs will tend to represent a split in the system on the basis of SAP status.

It is furthermore possible that the high frequency of person marking in general contributes to the maintenance of alignment splits in that domain over time. Since person marking is usually obligatory in some form or another, it is likely to be ubiquitous in all discourses. This frequency may contribute to a resistance to regularization or leveling, so that any splits within the person marking system will tend to remain splits and not be smoothed into a unified alignment pattern. This hypothesis would be interesting to explore in future work.

The interaction of Generalizations 1 and 2 is what makes SAP status so common as a split inducer. Generalization 1 highlights why any feature within the person category would likely be

³ Though Greenberg admits that the “high text frequency” criterion may be susceptible to differences in text genre, for instance in non-conversational texts which tend to have more third-person forms or in “odes addressed to divinities” which favor the second person.

involved in split-alignment, and Generalization 2 delineates why it is SAP status, rather than any other person feature, which is so well represented.

The higher markedness of first and second persons predicts that they will be involved in alignment splits less often, and this is indeed the case. Person as a split inducer means that different values of *individual* person, not just SAP status, show different alignment patterns.

The three languages with person as a split inducer are Haida, Seri, and Cherokee. Seri shows the clearest and most obvious instance of this kind of split. In that language, first person shows tripartite marking, second person shows nominative-accusative, and third person shows neutral. That is quite an interesting situation where each of the three persons gets its own alignment pattern and none of the persons is grouped together. This contrasts with Haida, where second-person plural and both numbers of third person all show neutral alignment. The second-person singular forms pattern along with both numbers of first person. Interestingly, Haida is not a case of neutral alignment by zero expression. The second-person plural pronoun is *dalaan* for A, S, and O, and the third-person pronoun is *'laa* for A, S, and O. So Haida second-person pronouns are not realized identically to the third-person ones, but they do show the same alignment pattern (neutral). But the reason this is not considered a split based on SAP status is that there are SAP arguments (namely, the second-person plural ones) which have the same alignment pattern as the non-SAP arguments. Rather, it seems that Haida is cutting up its person pie a little differently by grouping second-person plural arguments along with the third-person ones for the sake of alignment.

Number is another split inducer which falls out from Generalization 1. Number as a split inducer means that there are different alignment patterns for different values of the number

category for a given argument. In most languages in my sample, there are only two numbers, singular and plural. 40% of sampled languages show number as a split inducer; the list is shown in (9).

- (9) Sampled languages with number as a split inducer
Cherokee, Haida, Lowland Chontal, Misantla Totonac, Pawnee, Seri

The reason for number being fairly well represented as a split inducer is that number and grammatical relation are cumulated in 100% of the languages in (9). In Seri, for example, there is a single morpheme encoding both grammatical relation and number. First-person singular arguments show tripartite marking, as A, S, and O are all different, but first-person plural arguments are nominative-accusative because A and S are the same while O is different. Number and relation are not cumulated equally in all languages. In Pawnee, these categories are only cumulated in the third person. In the other persons, there are separate plural markers. So number as a split inducer in Pawnee is necessarily restricted by person. In Haida, second-person singular arguments are split-S, but second-person plural ones are neutral. Number is only relevant in the second person; first-person arguments, whether singular or plural, are all split-S, and third-person arguments are all neutral.

To be clear, the cumulation of number and grammatical relation does not guarantee that number will be a split inducer, only that it has a higher potential to be one than it does in languages where it is not cumulated with grammatical relation. Of the nine languages where number is not a split inducer, it is nonetheless cumulated with grammatical relation in four. The significant point is that number is *never* a split inducer unless it is cumulated with grammatical relation.

There are two ways in which number cannot be a split inducer. The first is if there is actually no distinction in relational marking between singular and plural forms, even if relation and number are cumulated. If singular and plural versions of a given person morpheme show the same alignment pattern, then there is no split on the basis of number. The second way is if, as alluded to in the previous paragraph, number and relation are not cumulated. In such a case, relational marking does its own thing and number marking does a different thing, but they do not interact and therefore do not result in a number-based split.

A related concept, and one which should be distinguished from alignment, is *split number*. Splits can occur in many grammatical systems; there only needs to be different yet systematic ways of coding a particular property under different conditions. In a split number system, number is coded differently under different conditions, or not coded at all for certain categories of noun. A split-number system can occur with or without split-alignment, depending on whether or not number and grammatical relation are cumulative. If number is always morphologically distinct from grammatical relation, the number system could still be split, but the conditions inducing its splittedness would not necessarily correspond to those splitting in the alignment system (assuming the language does indeed have split alignment). In Pawnee, since number and grammatical relation are cumulated, it could reasonably be analyzed as having both split number and split alignment whose inducers are identical due to the cumulation of those grammatical properties.

The cumulation of number with grammatical relation in Pawnee is made finer grained with the property of individuation, which is cumulated there as well. Individuation as a split inducer means that different alignment patterns occur depending on whether an argument is

“individuated” or “unitary.” These are categories specific (though probably not unique) to Pawnee, and that is the only language in my sample showing a split induced by individuation. Individuation is not the same as number. A plural S (“they”) can do something all together, acting a group, in which case S is considered unitary, or they can each do something individually, but still are *all* doing it, in which case S is individuated⁴. Cumulation of individuation, number, and relation is probably the factor allowing individuation to act as a split inducer in Pawnee. Its rarity likely comes from the relative rarity of individuation as a relevant grammatical property cross-linguistically.

5.3.2 Generalizations 3 and 4

Generalizations 3 and 4 are repeated below in (10).

(10) Generalizations 3 and 4

Generalization 3: Verbal semantics results in alignment splits due to the assignment of semantic roles to arguments by verbs.

Generalization 4: Lexical specification is dependent on the verbal semantics criterion and so induces splits by creating exceptions to semantic class membership.

Verbal semantics is the second best represented split inducer, boasting relevance in 46.7% of sampled languages, which are listed in alphabetical order in (11). In every case, the verbal semantics condition is involved, either alone or in combination with other factors, in the split-S system. This is not surprising because split-S alignment is sometimes even defined as a pattern in which the semantic nature of the verb determines whether S gets marked like A or like O. But

⁴ Examples are easy to think of in English even though individuation is not grammaticalized in that language. Consider these two situations: (i) *All of my friends are playing disc golf this weekend. Jason is going to Blue Ribbon Pines, Matt is going to Bryant Lake Park, and Tim is going to Robbin’s Island. I wish there was a course near my house.* (ii) *All of my friends are playing disc golf this weekend. Too bad I can’t join them.* The first sentence has to have an individuated interpretation, while the second is most naturally unitary.

recall from Chapter 2 (§2.5.5.2) that there are versions of split-S alignment which are not based on verbal semantics, notably the fluid-S subtype, so it is not a matter of necessity or definition that verbal semantics is involved.

- (11) Sampled languages with verbal semantics as a split inducer
Cherokee, Crow, Eucree, Haida, Laguna Keres, Lowland Chontal, Tlingit

Why is verbal semantics a common split inducer rather than a rare one? I'll give two answers, one superficial and one more substantive. The superficial answer is that, as discussed in §2.5.5.3, I am viewing split-intransitivity as a decomposable alignment system rather than an alignment pattern. Those who take issue with this decision will naturally see the healthy representation of verbal semantics in my sample as a consequence of the definitions assumed, rather than of typological principles operating across languages. However, the problem remains how to explain the differential marking of a single intransitive argument, even if one doesn't want to count this as split-alignment as I do.

The more substantive answer refers to Generalization 3: it is expected for verbal semantics to be involved in splits assuming any view of semantics in which predicates assign semantic roles to arguments. Significant insight into this property comes from the Unaccusativity Hypothesis (Perlmutter 1978, and developed within GB by Burzio 1986). Though I am not tying myself to a particular theoretical framework, I assume a general version of the semantics-syntax interface in which inherent semantic properties of verbs restrict (or determine) the kinds of arguments compatible with those verbs. In generative terms, this is realized in the argument structure so that unergative verbs have no internal argument and can only assign a theta role to

their subjects, while unaccusative verbs have no external argument and cannot assign a theta role to their subjects.

The European languages often given as evidence for the unergative/unaccusative distinction usually do not have obligatory morphology which signals class membership (evidence for the distinction comes from more complex constructions involving auxiliary choice, or other diagnostics). But all of the languages sampled for this study do have obligatory and overt agreement morphology which categorizes the two classes of intransitive predicates. The reason verbal semantics is the crucial component in determining class membership follows from Perlmutter's original hypothesis and the series of arguments put forth by Levin & Rappaport Hovav that the distinction "is syntactically represented but semantically determined" (ibid: 1995:30).

It is no wonder that so many languages (not just in my sample, but around the world) show S-marking which is sensitive to the inherent semantics of predicates because those semantics actually determine what kinds of arguments are compatible with the verbs. It remains to be explored why some languages code the semantics explicitly, while others (such as English) reveal the differences somewhat more implicitly. Donohue (2008:74-5) notes:

[...T]he phenomenon known as split intransitivity is a special case of the kind of semantically explicit subsystems that are found in most languages, although generally confined to marking one or other argument of a bivalent predicate. The fact that there is in some cases more than one possible coding choice for the sole argument of a monovalent verb is a natural consequence of the fact that most languages have more than one, or even two, ways to code the arguments of bivalent predicates.

Participant semantics as a split inducer may result from the same mechanism. The difference is that participant semantics as described here pertains exclusively to fluid-S systems, where the semantic properties of an argument in a given event determine alignment.

Generalization 4 is tied directly to Generalization 3. Lexical specification as a split inducer means that certain lexical items trigger an alignment split. This is entirely a consequence of the existence of exceptions in predicate class membership. All of the languages showing a lexically specified split also show a split based on verbal semantics. This is not accidental. The fact that a grammar appeals to verbal semantics for alignment means that there are semantic features or properties of verbs on the basis of which they can be grouped into classes. For most languages with split-S alignment, there are two such classes: an “agentive” class, corresponding to verbs whose S roles are typically agentive, active, volitional, and/or in control; and a “patientive” class, corresponding to verbs whose S roles are typically patientive, stative, non-volitional, and/or not in control. But there are almost always exceptions. The exceptions are verbs whose semantics would predict them to behave like other members of that semantic class, but which instead behave like members of the opposite class with respect to alignment. Note that Generalization 4 is fully dependent on Generalization 3. Lexically specified exceptions to alignment behavior do not, to my knowledge, appear for other split inducers. For instance, I uncovered no languages in which only certain lexical items do *not* show a split in certain aspects, with certain persons, or in the other split inducers, while the majority do.⁵

The exceptions must be lexically specified as behaving in the opposite way expected; they must be learned on a case-by-case basis. As an example, consider the description of split-S

⁵ Counterexamples to this claim would be very interesting. If they exist, then lexical specification as a split inducer and its relationship to verbal semantics would have to be reinterpreted.

alignment in Cherokee from §4.3.11.3. The Set A markers are found on the “active/agentive” class, but also on verbs meaning ‘stink,’ and ‘be hanging,’ which are semantically stative. The Set B markers are placed on verbs in the “stative/body function” class, but they are also used on ‘work’ and ‘take revenge.’ A list of languages with lexical specification as a split inducer is given in (12).

- (12) Sampled languages with lexical specification as a split inducer
 Cherokee, Crow, Euchee, Laguna Keres, Lowland Chontal, Tlingit

Haida is the only sampled language with verbal semantics as a split inducer which does not also have lexical specification as split inducer. Does this mean that the split-S pattern in Haida is exceptionless? Probably not. More likely, it means either that the sources I consulted to discern the alignment patterns in Haida did not happen to show the exceptions or, more likely, I did not notice whatever exceptions were mentioned. It would be very surprising to find a language with no lexical exceptions to semantic class behavior. Perhaps such languages exist, and maybe Haida is one of them, but in my opinion, this is most probably an instance of bibliographic bias or observational failure on my part.

It is not usual for linguists working on alignment to treat lexical specification as a split inducer. More often such cases are simply set aside as exceptions and the emphasis is placed on the non-exceptional members of each category. I think this is a mistake. “Exceptions” in alignment patterns can reveal interesting things about the stability and makeup of semantic classes. When linguists observe that exceptions exist in a paradigm, a pattern, a language, or whatever, that observation is usually a synchronic one. Diachronic processes are frequently fruitful as a means of explanation. For certain exceptions, particularly paradigmatic ones, sound

changes may be invoked to explain why some inflected forms appear anomalous compared to others or why analogical leveling has occurred. But the kinds of exceptions relevant to split induction in this study are probably not based on sound changes because lexical specification refers to exceptions in semantic class membership rather than in paradigm structure. Still, there is a good chance that diachronic explanations can shed light on individual cases. I see two main ways in which this could happen (though there are probably others): lexical fusion and semantic drift.

In lexical fusion, different predicates get combined into a single paradigm. For example, this happened when Middle English *gon* ‘go’ and *wenden* ‘turn’ were fused into the same paradigm resulting in the present tense *go* and the suppletive past tense *went* (Nübling 2010). In regards to split-alignment, if two predicates from different semantic classes had, at some point in the history of the language, become fused, then the agreement properties of only one of them would likely be used on all inflections of the new verb. In other words, one predicate would adopt the inflectional traits of the other.

Semantic drift (or semantic change) is an umbrella phenomenon by which words change meaning over time. It can have several specific vehicles, such as homonymy, member-group relationships, and various types of metaphor (Hock 1991). This could produce split-alignment if, say, a verb originally meaning ‘to smell (something)’ over time acquired a specialized usage meaning ‘to smell (something awful),’ and then eventually came to mean ‘to be smelling awful,’ in other words, ‘to stink.’ In order to result in a lexically specified alignment split, the semantic change would necessarily have to precede an inflectional change. The new predicate would have shifted in its semantic class membership, but would retain the person marker of the original.

Of course, proving that diachronic events have resulted in lexical exceptions to alignment patterns is another matter altogether, one that would require a detailed historical investigation of each of the languages in (12). Such a project would be very interesting indeed, but unfortunately, is outside the scope of this dissertation.

It does not, incidentally, seem especially likely that interparadigmatic analogy would be a common pathway for this kind of lexical specification because if the language has an agreement morphology which is already strongly sensitive to the semantic classes of predicates, it would be more likely that argument marking would be reinforced by the semantics already inherent in each predicate, rather than coaxed into change by analogy to members of the opposite semantic class.

5.4 Residue

There are several instances of “residue” — in other words, patterns which either seem to contradict the generalizations or those which require additional explanation.

5.4.1 The other 27%

In 5.3.1, I stated that nearly 73% percent of languages with SAP status as a split inducer have third person expressed as zero for A, S, and O. But then, what about the other 27% of languages which have SAP status as a split inducer but do not have zero-expressed third persons? These languages are Euchee, Cherokee, and Laguna Keres.

In Euchee, third person arguments do show neutral alignment, but are non-zeros, so the unmarked status of third person cannot be said to contribute to the neutralization in this case. It is interesting that although the third-person pronominals in Euchee do not distinguish between A,

S, and O, they still distinguish singular from plural, Euchee from non-Euchee, men from women, and men's speech from women's speech. The third-person pronominals are the only ones that encode all these properties. Perhaps there are constraints on the number of features cumulable on Euchee pronominals, and the third-person forms reach the upper limit and so speakers leave it to context to disambiguate grammatical relation.

In Laguna Keres, the reason for third person inducing an alignment split is due to the morphological structure of the language. There is exactly one argument slot on Laguna Keres verbs. For intransitive verbs, there is only one argument and so no conflict arises. For transitive verbs, one of the two arguments will appear at the expense of the other. If one argument is an SAP but the other is not, the SAP is the one that shows up. If both arguments are non-SAPs, then the A role is the one that shows up. If both arguments are SAPs, then a third set of person markers is used, and this is where the alignment split occurs. So in Laguna Keres, the SAP split inducer is really a consequence of structural argument inflection and conflict resolution, rather than low markedness and zero expression.

In Cherokee, an alignment split occurs when both arguments are third person; this is when hierarchical alignment comes into play. There are directionality morphemes which indicate whether the action goes up or down the Cherokee indexability hierarchy, but these are not cumulative with person. They coincide with third person because third-person arguments may fall in various places along the hierarchy (human, non-human animate, inanimate), and the directionality of action is grammaticalized in Cherokee.

5.4.2 Nominal status

Nominal status as a split inducer means that different alignment patterns appear on different categories of nominals. The categories are full noun, free pronoun, and bound pronominal. Three languages have nominal status-induced alignment splits: Alsea, Maricopa, and Tlingit. In Alsea and Tlingit, the distinction is between pronominals and non-pronominals. In other words, pronominals behave one way while pronouns and nouns behave another way. This makes sense considering that case morphemes will typically affix to pronouns and nouns, but not to pronominals. In keeping with this prediction, Alsea and Tlingit both have ergative-absolutive alignment on the non-pronominals. On the pronominals, Alsea has nominative-accusative alignment and Tlingit has both split-S and neutral. In Maricopa, all three types of nominals have different behavior.

Why is nominal status a rare split inducer? I suspect the reason is that languages so often encode grammatical relations only on one type of nominal. This “case asymmetry,” occurs in over a quarter of 260 languages sampled by Iggeson (2013) — clearly not the majority pattern, but certainly not difficult to find cross-linguistically. In the languages I sampled, it was very common for relational information to be encoded on pronominals, but full nouns (and often pronouns as well) were left bare or were only inflected for properties other than grammatical relation. These other nominals, then, are not targets for relational marking and so cannot logically be involved in alignment splits.

5.4.3 Animacy

Animacy as a split inducer means that different alignment patterns occur when arguments are specified as being animate or inanimate. Only Pawnee and Cherokee show animacy-based splits. The animacy condition in Pawnee is only relevant in conjunction with the individuation condition. If S is individuating (acting individually, not as a collective group) and O is animate, then there is a tripartite marking pattern. Unitary S induces tripartite alignment regardless of the animacy of O. In Cherokee, most of the argument markers on transitive verbs are portmanteau, but some of them happen to correspond to certain intransitive argument markers. This is what I abbreviated the $A \Rightarrow O = S_A$ pattern in 4.3.11.4. It is induced when a third-person O argument is inanimate. What could be happening here is that inanimate arguments, being lowest ranked on the Cherokee indexability hierarchy, are marginalized so severely that they never acquire an argument marker which identifies them specifically. One could conceivably analyze inanimate Os as not being realized at all, not even in a portmanteau pattern, though since the other $A \Rightarrow O$ arguments are portmanteau, I am treating these as portmanteau as well, pending further investigation.⁶

Given that animacy has been shown to affect language in a wide variety of ways (for instance, in number marking, genitive marking, word order, identification of ‘subject,’ etc.; see Yamamoto 1999), it is surprising that it is not better represented in my sample. Algonquian languages are well known to have agreement patterns sensitive to animacy, but in Blackfoot (the

⁶ Such an outlook would be sketched as follows: The apparently “portmanteau” $A \Rightarrow O$ marker is identical to the agentive S marker if O is inanimate. Therefore, the $A \Rightarrow O$ marker is not really portmanteau in the first place. Rather, it is only marking A, which would be a nominative-accusative pattern since A and S would be the same in that instance. I do not take a stand on this issue at this time. In any case, animacy of O would still be a relevant factor in inducing an alignment split.

only Algonquian language I sampled), changes in animacy affect the forms of certain morphs, but do not fundamentally alter the alignment system itself.

In both Pawnee and Cherokee, it is only the animacy of the O argument which is relevant for alignment. The degree to which this contributes to the relative rarity of animacy as a split inducer is a ripe area for future research; it is difficult to arrive at substantive generalizations of this category on the basis of only two languages, especially when the relevant animacy value is different in each one (in Pawnee O must be animate, in Cherokee it must be inanimate).

5.4.4 Aspect & mood

Aspect as a split inducer means that different alignment patterns occur in different aspects. Given the strong representation of aspect-based splits in the alignment literature (usually in reference to Indic, Basque, Austronesian, or Mayan), I was surprised that only Cherokee and Chol show aspect-induced splits in my sample. In Cherokee, certain intransitive arguments take the Set A marker in four of the eight aspects, but take the Set B marker in the other four. In Chol, ergative-absolutive alignment occurs in the perfective while nominative-accusative occurs in the imperfective. Aspect-induced splits are potentially non-attributable to Generalization 1 because there no is obvious cumulation of aspect and grammatical relation in any language I am aware of. Both Chol and Cherokee have morphemes to indicate aspect which are separate from those to indicate person and relation.

The connection between aspect and split-alignment, particularly split-ergativity, is treated in some functional literature as a consequence of the relative definiteness of past tense or perfective propositions as compared to those in other tenses or aspects (see Dixon 1994). In other

words, the perfective aspect, which describes an event which has already occurred and has been completed, is more likely to contain an A with overt marking distinct from S or O because the agency of that A can readily be ascertained by virtue of the event having been completed. The same certainty is unavailable for future, imperfective, or irrealis events. However, given that only two languages in my sample show aspect-based splits, and that the aspect influence in Cherokee is apparently non-canonical, it is difficult to use my results to weigh in on this issue. Formal treatments of aspect-based splits, however, are insightful for these two languages.

Coon (2013) has an interesting account of the aspect-based split in Chol. She argues that the non-perfective aspect markers (*mi* for imperfective and *choñkol* for progressive) are actually predicates. Specifically, they “are intransitive (unaccusative) stative verbs that take nominalized clauses as their internal arguments” (ibid:110). The perfective morphemes, by contrast, are not predicates. The ergative marking seen in the perfective clauses indicates the external argument, which she believes is itself possessive since ergative is syncretic with the possessive. Under the view that non-perfective morphemes are unaccusative intransitive verbs, they do not have external arguments, and so cannot receive ergative morphology. For Coon, the split in Chol results from facts about Chol clause structure.⁷

If Coon’s story is along the right lines, then perhaps aspect is a relatively rare split inducer in my sample because aspect morphemes are not verbs in most languages I sampled. It is interesting, though, that part of Coon’s analysis also purports to explain why ergative

⁷ Coon’s formal analysis of Chol has precedent in the Mayan comparative literature. Kaufmann (1989) notes that nominalized verbs as complements, with consequences for the ergative or absolutive marking of subjects, is reconstructed for Late Proto-Mayan. He traces the split-ergative pattern in Cholan (of which Chol is a member) through an interesting series of borrowings and reanalyses from Pocom by way of Yucatecan. See also Robertson (1992) for details of the historical development of aspect and related categories within Mayan. Syncretism of ergative and possessive is also well known.

morphology is (apparently) universally in the perfective rather than the non-perfective aspects. Certain language-specific differences notwithstanding, she argues that in every case, the association of ergative morphology with perfectivity is the result of more complex clausal structures in the non-perfective aspects. She supports this view with data from a nice variety of genetically and geographically disparate languages. However Cherokee, the other language I sampled with an aspect-based split, has exactly the opposite pattern. Cherokee uses morphology from the *patient* set on verbs in the perfective, a pattern not expected from Coon's analysis.

Johnson (2012), also working in a generative syntax framework, argues that the Cherokee pattern is a consequence of ergative morphology being licensed by T(ense) and that T is missing from the perfective clauses because they are tenseless. Without T, external arguments are assigned genitive case, which is syncretic with the absolutive. The syncretism of ergative with possessive in Chol and of absolutive with genitive in Cherokee is very interesting and suggests a promising domain for future investigation. For now, it is worth emphasizing that there are several different kinds of alignment splits in Cherokee and it is not assumed here that they are all attributable to a single source synchronically. Still, Scancarelli (1987) comments that the Set A and Set B prefixes common to Iroquoian languages are likely in the past to have been subject and object markers with more regularity than the modern Cherokee system reveals. She (ibid:346) suggests:

Presumably, various factors—pragmatic, semantic, or syntactic, or some combination—operated on the proto-language to yield the present system: but those factors may have operated on different parts of the system at different times.

Mood as a split inducer is even rarer, occurring only in Laguna Keres. In that language, imperative arguments in the third person all take the same marker, *pi-*, regardless of role. Some comments on this claim are in order, since the imperative mood is customarily thought to refer to second-person arguments and it is not intuitive for English speakers to imagine an imperative form of O. This does not seem to be the case in Laguna Keres as the imperative prefixes are specified for all three persons for both the actor and undergoer semantic roles. When these are translated into English, it usually comes out something like ‘let me ...’ and ‘let him/her ...,’ respectively. They are distinct for the SAPs, but identical for the non-SAPs.

Lachler (2006) comments that earlier scholars working on Keresan called the mood “hortative” or “imperative/hortative.” He states “I choose to call it the Imperative because its use is broader than what one would normally expect of a simple hortative” (ibid:180). I think this issue is worth reconsidering because a diachronic pathway for this split might be more forthcoming under the assumption that the Laguna Keres mood system really includes a genuine hortative function. The Laguna Keres “imperative” would qualify as an imperative-hortative system according to the typology outlined by van der Auwera et al. (2013). Suppose, for example, that Proto-Keresan had an imperative system in which only second-person arguments were inflected on imperative verbs, and then this was gradually reanalyzed to permit hortative functions. There would be motivation within the developing imperative-hortative paradigm to distinguish first-person grammatical roles from one another for three reasons: first by analogy to the already distinguished second-person ones; second by analogy to the distinguished first-person non-imperative person markers; and third because first-person hortatives are the rarest cross-linguistically and presumably highly marked, increasing the likelihood of overt

morphology and distinction. However, there would be comparatively less motivation to distinguish the third-person grammatical roles from one another because the low markedness of the third person does not impel distinct morphology beyond simply indicating the “imperative-hortativeness” of the argument. I emphasize that this is speculative. A detailed study of historical Keresan issues would be very interesting in this regard.

5.4.5 Definiteness & Particularness

Definiteness as a split inducer means that different alignment patterns occur under different configurations of the definiteness of arguments. This only occurs in Tlingit. Full nouns in the A role normally receive an ergative suffix, but if the noun is indefinite, then ergative marking is ungrammatical. In Tlingit, definiteness is a semantic notion without overt morphology. It seems to operate according to whether or not there is a specific referent in mind when a full noun is used. For instance, in ‘My father is cooking soup,’ there is a specific soup being made; one could presumably walk into the kitchen and find it simmering on the stovetop. But in ‘My father was fishing king salmon,’ there is no identifiable king salmon being sought; any king salmon which happened to bite would do just as well as any other.

A very similar split inducer is the one in Blackfoot, to which I have given the label *particularness*, based on the term “non-particular” which Frantz (2009) gives to arguments which induce this pattern. Particularness as a split inducer means that alignment is determined by whether or not an overt argument refers to a particular or specific entity in the real world. The class of arguments falling under the rubric of particularness include non-particular mass nouns

(e.g. money, milk) and unknown (but presumed real) entities (as in, e.g. ‘Some boys are playing;’ I don’t know which boys, but I can hear laughter behind the shed).

The definiteness condition in Tlingit and the particularness one in Blackfoot have obvious semantic overlap; it might be possible to group these two inducers together into a single category, something like *referral*, since it seems as though in both cases the crucial factor is whether or not the argument refers to an identifiable entity. However, without a more careful treatment of noun phrases in these two languages, I am not prepared to collapse these inducers into one.

5.5 Split-alignment within Optimality Theory

Important theoretical work with relevance to my findings has been done by Aissen (2001, 2003) in the framework of Optimality Theory (OT).⁸ She exploits two principles of OT, harmonic alignment of prominence scales and local conjunction, which together allow her to account for a wide range of alignment-related phenomena. The scales relevant to her analyses are those for person, semantic role, and relation, with basic versions laid out in (13).

- (13) Some prominence scales (Aissen 2001:64, list 4)
- | | |
|-------------------|----------------------|
| Person scale: | Local > 3rd |
| Role scale: | Agent > Patient |
| Relational scale: | Subject > Nonsubject |

⁸ I assume the reader has a basic understanding of the devices and principles of OT, so I will not go into the definitions of the various technical terms here. The seminal source is Prince & Smolensky (circulated manuscript 1993, finalized version 2008), which deals mainly with its application to phonology. A compendium edited by Legendre, Grimshaw & Vikner (2001) gives a good overview of OT in morphology and syntax.

Aissen's use of "local" corresponds to my use of "SAP": local persons are first and second person. The angled bracket ' > ' means 'ranked higher than' and the relative rankings are claimed to be universal. Alignment of the scales is achieved by associating points or values along one scale to those on another. For example, Su/Local (Su = Subject) refers to the Subject position on the relational scale aligned with the Local position on the person scale; in other words, local-person subject. Because Local outranks 3rd on the person scale, harmonic alignment of the person and relational scales produces $Su/Local > Su/3$ and $Non-Su/3 > Non-Su/Local$. The depressed angled bracket ' > ' means 'more harmonic than.' Thus, local subjects are more harmonic than third-person subjects and third-person non-subjects are more harmonic than local non-subjects. Markedness relations can then be expressed in the forms of standard OT constraint rankings: $*Su/3 \gg *Su/Local$ and $*Non-Su/Local \gg *Non-Su/3$.⁹

Alignment splits, particularly the versions involving split-ergativity and differential object marking,¹⁰ are then treated by combining the harmonic alignment mechanism just described with another operation, local constraint conjunction. In local conjunction, two constraints are joined together (symbolically with an ampersand), creating a new constraint. The new constraint is universally ranked higher than either of the individual component constraints. For instance, the constraint $*\emptyset_C$ (read "star zero case") penalizes outputs with zero morphological expression of case. $*Su/3$, as discussed above, penalizes third-person subjects. Conjoining these constraints produces $*\emptyset_C \& *Su/3$, which penalizes third-person subjects which are not case-marked, but

⁹ Note that because these are avoid constraints, the order of the elements in each ranking is reversed from that in the harmonic alignment.

¹⁰ Differential object marking simply refers to instances where direct objects show formal variation under certain grammatical conditions. See Aissen (2003) for discussion and examples.

does not penalize local-person subjects which are not case-marked nor third-person subjects which are case-marked. The conjoined constraint dominates both $*\emptyset_c$ and $*Su/3$, and is rankable with respect to other conjoined (and individual) constraints.

Using the devices harmonic alignment and local conjunction, Aissen can derive constraints and rankings which permit case marking in some contexts but forbid it in others; it is this feature which facilitates the exposition of (certain types of) split-alignment because split-alignment is, by definition, correspondences among argument realizations which differ by grammatical context. It would be very interesting to explore the implications of my findings against such OT analyses, though differences in certain theoretical assumptions would have to be resolved, and the descriptive terminology I have employed would have to be translated into the OT framework, before this could be done rigorously.

5.6 Summary of Chapter 5

In this chapter, I have argued that the four generalizations below concisely summarize the major factors in determining whether a grammatical property is likely to be a split inducer, and if it is, how common it is likely to be.

(14) Four generalizations

Generalization 1: Features cumulated with grammatical relation are likely to be split inducers.

Generalization 2: Low markedness tends to produce zero expression, which tends to neutralize distinctions between grammatical relations.

Generalization 3: Verbal semantics results in alignment splits due to the assignment of semantic roles to arguments by verbs.

Generalization 4: Lexical specification is mostly dependent on the verbal semantics criterion and so induces splits by creating exceptions to semantic class membership.

Generalizations 1 and 2 together relate to the preponderance of SAP status-induced splits, as well as less represented split inducers which are cumulated with grammatical relation in the same morpheme. Generalizations 3 and 4 identify the preponderance of verbal semantics and lexical specification as split inducers, particularly in reference to the split-S family of alignment systems.

The generalizations do not give substance to all of the alignment facts found in the sampled languages. There are some split inducers which are not attributable to the generalizations; those inducers, however, are always relatively rare. In most cases, the existence of the minor split inducers does not contradict the validity of the four generalizations; it simply requires appealing to additional, often language-specific or historical, factors in order to shed light on their occurrence.

6 CONCLUSIONS AND FUTURE DIRECTIONS

6.1 Introduction to Chapter 6

In this final chapter, I review the study and its results (§6.2) and summarize my main claims (§6.3). In §6.4, I comment on directions for future research. A general conclusion is in §6.5.

6.2 Study and results summary

The objectives of this study were to determine the range of split-inducing conditions in a sample of indigenous genera, and thereby languages, from across the North American continent and to provide an explanation for their relative commonness or rarity. This was done by investigating alignment data from 15 languages selected at random from a sample stratified to control for genetic and areal affiliation. Genetic stratification was achieved by first randomly selecting *genera*, which are sub-familial units with a time-depth of at least 3500 years, and then randomly selecting individual languages from within each genus. This methodology strengthens my claim that the alignment patterns are unlikely to result from shared inheritance. Areal stratification was achieved by dividing the continent into five non-overlapping geographic regions on pre-established linguistic and cultural area grounds. Then each genus was assigned to exactly one area and exactly three languages were selected from each genus. This methodology makes it unlikely that alignment patterns have diffused between any of the five areas.

The sampling of fifteen genera in this way and the subsequent investigation of a language in each of them turned up thirteen split inducers. They are not equally represented. SAP status, verbal semantics, lexical specification, and number are the best represented overall. Interestingly, none of the split inducers qualifies as a Dryerian preference, which would require that it is present in at least two-thirds of the languages sampled from each region. Still, the heavy skewing toward certain split inducers begs explanation.

6.3 Summary of main claims

I account for the skewing by appealing to four generalizations. The first is that when grammatical relations are cumulated with (encoded on the same morpheme as) other grammatical properties, then those properties are likely to split inducers. The reason cumulation predicts split induction is that the properties which tend to be cumulated with grammatical relation, namely person and number, cause morphemic alternations in order to satisfy inflectional requirements. When a morpheme changes form to properly inflect for person and number, the grammatical relation property is carried along with it. This is why person and number are split inducers and part of the reason why SAP status is so common.

The other part of the reason why SAP status is so common is due to the second generalization, which is that unmarked values of linguistic categories tend to be zero-expressed. Since third-person arguments are the unmarked pronouns, these are frequently zero-expressed whereas other persons are overt. This produces neutral alignment in the third persons and results in an alignment split.

The third and fourth generalizations point to the high representation of verbal semantics and lexical specification as split inducers. On the assumption that verbs select for arguments based on their own semantic properties, it is expected to see the split-intransitive systems show internal splits on the basis of verbal semantics. The fourth generalization, which is dependent on the third, is an admission that there are (almost) always outliers to the semantic categorization of verbs, that is, verbs whose morphology is mismatched to their semantics, and that diachronic developments can often account for individual cases.

6.4 Future directions

As the topic of alignment typology subsumes a huge variety of narrower topics, there are many avenues for future research. Four such avenues stand out to me.

The first avenue for future work involves improving the precision estimate of the results. As explained in Chapter 3, selecting genera proportionally would dissolve the avoidance of areal bias. But there may be some ways in which areal stratification could be achieved without sacrificing (much) precision in the estimation of error. The best way to do this, I think, would be to increase the total number of regions to be sampled from. This has to be done with care so as to maintain the low likelihood of diffusion between regions. Most of the regions contain more than one established cultural-linguistic area, so it is possible to treat some of these as their own regions rather than group them together into a larger region. Naturally, this will increase the sample size and expense of data collection. It wouldn't be possible with all regions, though, because the Arctic, as mentioned in Chapter 3, only contains two genera. And some regions will

definitely contain more genera than others no matter how small they are simply because the genetic diversity in some areas is exceptionally dense.

A second avenue is that the project could be expanded to include other continents, the ultimate goal being a worldwide sample. The best next choice would be South America because the Americas together form a contiguous geological unit. Problems of bibliographic bias are likely to be sizable for South America due to the smaller number of available grammars. Genetic relationships are also not as well worked out for the South American languages as they are for the North American ones. Africa would be another interesting area to explore, as certain alignment patterns, namely ergative-absolutive, are known (or thought?) to be exceptionally rare there (for some examples, see Frajzyngier 1984, Andersen 1988, and Miller & Gilly 2001). Eurasia is undoubtedly the area where alignment phenomena have been the most studied. To my knowledge, there has not yet been a study of Eurasian split alignment which controls for genetic and areal affiliation in the same manner as this one.

A third avenue for further work would be a detailed investigation of the diachronic factors at play in split alignment. In Chapter 5, I alluded to some possible pathways for alignment splits in individual languages, but I have not come close to an actual analysis along these lines. The reason diachrony is important is that historical processes often provide the actual mechanisms by which the structure of languages change. Viewing alignment patterns on a cyclic continuum has been an active area of research for some time (Givón 1980, Finney 2007, Meakins 2015 to name a few). The double-oblique pattern discussed in Chapter 2 is usually interpreted as an intermediate stage between other alignment patterns. Such intermediate stages could help explain the rarity of certain patterns or split inducers.

The study design for this project used Dryer's (1989, 1992) arbitrary time depth of 3500-4000 years to define a genus. But this depth is open to question if it turns out that alignment patterns on the whole are particularly, or particularly not, stable over time. If they are extremely stable over time, then 4000 years might not be enough to sufficiently reduce the likelihood of shared inheritance among distantly related languages. If they are quite susceptible to rapid change, then a much smaller time depth would be possible and in some cases preferable because it would increase the number of genetic groups to be sampled from.

An area related to diachrony is language acquisition. As far as I know, there is not a robust tradition of exploring split alignment from an acquisitionist perspective, though Pye (1990) and Ochs' (1982) study of child Samoan are good foundations.

A fourth avenue ripe for further development is the extension of the principles and methods described here to other valency types. I have not looked at all at three-argument clauses, but the realization of an indirect object argument is also an important way of assessing alignment and splits within alignment systems. Related to this are the several other construction types which were sacrificed in the interest of time, for instance non-personal pronouns.

6.5 Final Conclusion

In the end, this project has contributed to our understanding of grammatical alignment in North American languages by discovering the (probably non-exhaustive) range of split-inducing conditions found in these languages. Taken as a set, this is new empirical data. It has also shed light on the reasons why certain of these split inducers are quite common while others are rare. It is essential, I think, to use caution when making any claims about commonness or rarity of a

given pattern or construction type unless the claim is motivated by a methodologically sound cross-linguistic study.

APPENDIX 1: GENETIC GROUPS PER AREA

Bolded genera represent those selected for inclusion in this study.

Region 1: Arctic, Subarctic, Plains

- | | |
|----------------------|------------------|
| 1. Aleut | 6. Eskimo |
| 2. Algonquian | 7. Karankawa |
| 3. Aranama-Tamique | 8. Kiowa |
| 4. Beothuk | 9. Siouan |
| 5. Caddoan | 10. Tonkawa |

Region 2: Northwest Coast, Plateau, California

- | | |
|-----------------------|-----------------------|
| 1. Alsea | 22. Palaihnihan |
| 2. Bella Coola | 23. Pomoan |
| 3. Cayuse | 24. Sahaptian |
| 4. Central Salish | 25. Salinan |
| 5. Chimakuan | 26. Shasta |
| 6. Chimariko | 27. Siuslaw |
| 7. Chinookan | 28. Southern Wakashan |
| 8. Chumashan | 29. Takelman |
| 9. Costanoan | 30. Takic |
| 10. Coosan | 31. Tillamook |
| 11. Esselen | 32. Tlingit |
| 12. Eyak | 33. Tsamosan |
| 13. Haida | 34. Tsimshian |
| 14. Interior Salish | 35. Tübatulabal |
| 15. Karuk | 36. Washo |
| 16. Klamath-Modoc | 37. Wintuan |
| 17. Kutenai | 38. Wiyot |
| 18. Maiduan | 39. Yana |
| 19. Miwok | 40. Yokutsan |
| 20. Molala | 41. Yukian |
| 21. Northern Wakashan | 42. Yurok |

Region 3: Great Basin and Southwest

- | | |
|-------------------------|-------------------------------|
| 1. Cahita | 12. Mamulique |
| 2. Chichimeco | 13. Maratino |
| 3. Coahuilteco | 14. Naolan |
| 4. Cochimí-Yuman | 15. Numic |
| 5. Comecrudan | 16. Pimic (Tepiman) |
| 6. Corachol | 17. Seri |
| 7. Cotoname | 18. Solano |
| 8. Garza | 19. Tanoan |
| 9. Guaicurian | 20. Taracihitic (Tarahumaran) |
| 10. Hopi | 21. Tubar |
| 11. Keresan | 22. Zuni |

Region 4: Northeast and Southeast

- | | |
|---------------------|------------------------------|
| 1. Adai | 7. Natchez |
| 2. Alagüilac | 8. Northern Iroquoian |
| 3. Atakapan | 9. Southern Iroquoian |
| 4. Catawban | 10. Timucua |
| 5. Chitimacha | 11. Tunica |
| 6. Muskogean | 12. Yuchi |

Region 5: Mesoamerica

- | | |
|-----------------|---------------------------|
| 1. Amuzgo | 11. Otomian |
| 2. Aztecan | 12. Pamean |
| 3. Chinantecan | 13. Popolucan |
| 4. Cuitlatec | 14. Subtiaba-Tlapanec |
| 5. Huavean | 15. Tarascan |
| 6. Jicaquean | 16. Tequistlatecan |
| 7. Lencan | 17. Totonacan |
| 8. Mayan | 18. Xican |
| 9. Mixe-Zoquean | 19. Zapotecan |
| 10. Mixtecan | |

APPENDIX 2: MEMBER LANGUAGES FROM EACH SELECTED GENUS¹

Bolded languages represent those selected for inclusion in this study.

Algonquian (Region 1)

- | | |
|----------------------------|---------------------|
| 1. Abenaki | 13. Menominee |
| 2. Algonquin | 14. Meskwaki |
| 3. Arapaho | 15. Micmac |
| 4. Atikamkwkw | 16. Mohegan-Pequot |
| 5. Blackfoot | 17. Montagnais |
| 6. Cheyenne | 18. Nanticoke-Conoy |
| 7. Cree | 19. Naskapi |
| 8. Delaware | 20. Narragansett |
| 9. Gros Ventre | 21. Ojibwe |
| 10. Kickapoo | 22. Powhatan |
| 11. Lumbee | 23. Shawnee |
| 12. Malecite-Passamaquoddy | 24. Wampanoag |

Caddoan (Region 1)

1. Arikara
2. Caddo
3. Kitsai
- 4. Pawnee**
5. Wichita

Siouan (Region 1)

- | | |
|----------------|-----------------|
| 1. Assiniboine | 8. Lakota |
| 2. Crow | 9. Mandan |
| 3. Dakota | 10. Omaha-Ponca |
| 4. Hidatsa | 11. Osage |
| 5. Ho-Chunk | 12. Quapaw |
| 6. Iowa-Oto | 13. Stoney |
| 7. Kansa | |

¹ Certain variants which the Ethnologue lists as separate languages with regional or group modifications of a general label are omitted here. For example, Moose Cree, Plains Cree, Swampy Cree, etc. are not listed separately here.

Alsea (Region 2)

1. Alsea

Haida (Region 2)

1. Haida

Tlingit (Region 2)

1. Tlingit

Cochimí-Yuman (Region 3)

- | | |
|------------|--------------------|
| 1. Cocopa | 5. Maricopa |
| 2. Cochimi | 6. Mohave |
| 3. Kiliwa | 7. Quechan |
| 4. Kumiai | |

Keresan (Region 3)

- 1. Western Keres**
2. Eastern Keres

Seri (Region 3)

1. Seri

Muskogean (Region 4)

- | | |
|-------------------|-------------|
| 1. Alabama | 4. Koasati |
| 2. Chickasaw | 5. Mikasuki |
| 3. Choctaw | 6. Muskogee |

Southern Iroquoian (Region 4)

1. Cherokee

Yuchi (Region 4)

1. Euchee

Mayan (Region 5)

- | | |
|------------------|---------------------|
| 1. Achi | 16. Mopan Maya |
| 2. Akateko | 17. Poqomchi' |
| 3. Awaketeko | 18. Q'anjob'al |
| 4. Chicomuceltec | 19. Q'eqchi |
| 5. Chol | 20. Sakapulteko |
| 6. Ch'orti | 21. Sipakapense |
| 7. Chuj | 22. Tabasco Chontal |
| 8. Huastec | 23. Tektiteko |
| 9. Itza' | 24. Tojolabal |
| 10. Jakalteko | 25. Tzeltal |
| 11. Kaqchikel | 26. Tzotzil |
| 12. K'iche' | 27. Tz'utujil |
| 13. Lacandon | 28. Uspanteko |
| 14. Mam | 29. Yucatec Maya |
| 15. Mocho | |

Tequistlatecan (Region 5)

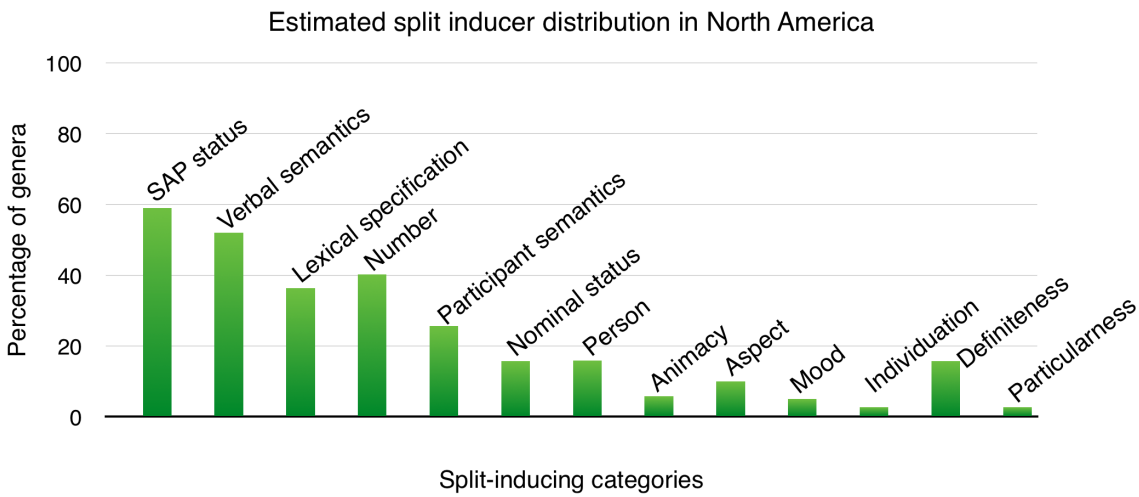
1. Highland Chontal
- 2. Lowland Chontal**

Totonacan (Region 5)

1. Tepehua
- 2. Totonac**

APPENDIX 3: ESTIMATED CONTINENTAL SPLIT INDUCERS

Below is a chart showing the *estimated* distribution of split inducers for the entire North American continent, based on weighted values. The weighting procedure changes the expected commonness of some split inducers because it takes into account the total number of genera in each region. Whichever split inducers happen to show up in the genus-dense regions get scaled up. As mentioned in the main discussion, the standard errors of the weighted values are quite high. Nonetheless, these results are the best we have if we intend to project the obtained data to rest of the continent.



APPENDIX 4: FULLY SPECIFIED TABLE OF SPLIT INDUCERS

Below is a chart showing all of the alignment patterns from each language investigated.

Each pattern is labeled along with the split inducing categories associated with it and the values of each category.

Language and affiliations	Alignment pattern	Inducer 1 category	Inducer 1 value	Inducer 2 category	Inducer 2 value	Inducer 3 category	Inducer 3 value	Inducer 4 category	Inducer 4 value
Blackfoot	direct-inverse	particularness	O = particular						
Region 1	neutral	particularness	O = non-particular						
Algonquian genus									
Algic family									
Crow	split-S _{nom-acc}	SAP status	+SAP	verbal semantics	active				
Region 1									
Missouri River Siouan genus	split-S _{nom-acc}	SAP status	+SAP	lexical specification	active				
Siouan-Catawban family	split-S _{erg-abs}	SAP status	+SAP	verbal semantics	stative				
	split-S _{erg-abs}	SAP status	+SAP	lexical specification	stative				
	neutral	SAP status	–SAP						
Pawnee	nominative - accusative	SAP status	+SAP	number	singular				
Region 1	neutral	SAP status	–SAP	number	singular				
Caddoan genus	ergative-absolutive	SAP status	–SAP	number	plural	individuation	S=individuation	animacy	O=inanimate
Caddoan family	tripartite	SAP status	–SAP	number	plural	individuation	S=individuation	animacy	O=animate
	tripartite	SAP status	–SAP	number	plural	individuation	S=unitary	animacy	O=animate
	tripartite	SAP status	–SAP	number	plural	individuation	S=unitary	animacy	O=inanimate

Alsea	ergative-absolutive	nominal status	– pronominal						
Region 2	nominative - accusative	nominal status	+pronominal						
Alsea genus									
Alsea family									
Haida	split-S _{nom-acc}	person	1	verbal semantics	agentive				
Region 2	split-S _{nom-acc}	person	1	event semantics	agentive				
Haida genus	split-S _{nom-acc}	person	2	number	singular	verbal semantics	agentive		
Haida family	split-S _{nom-acc}	person	2	number	singular	participant semantics	agentive		
	split-S _{erg-abs}	person	1	verbal semantics	patientive				
	split-S _{erg-abs}	person	1	event semantics	patientive				
	split-S _{erg-abs}	person	2	number	singular	verbal semantics	patientive		
	split-S _{erg-abs}	person	2	number	singular	event semantics	patientive		
	neutral	person	2	number	plural				
	neutral	person	3						
Tlingit	ergative-absolutive	nominal status	– pronominal	definiteness	O=+def				
Region 2	split-S _{nom-acc}	nominal status	+pronominal	SAP status	+SAP	verbal semantics	active		
Tlingit genus	split-S _{nom-acc}	nominal status	+pronominal	SAP status	+SAP	lexical specification	active		
Eyak-Athabaskan family	split-S _{nom-acc}	nominal status	+pronominal	SAP status	–SAP	definiteness	A=+def	definiteness	O=+def
	split-S _{erg-abs}	nominal status	+pronominal	SAP status	+SAP	verbal semantics	patientive		
	split-S _{erg-abs}	nominal status	+pronominal	SAP status	+SAP	lexical specification	patientive		
	split-S _{erg-abs}	nominal status	+pronominal	SAP status	–SAP	definiteness	–def	verbal semantics	patientive
	split-S _{erg-abs}	nominal status	+pronominal	SAP status	–SAP	definiteness	–def	lexical specification	patientive

	neutral	nominal status	+pronominal	SAP status	–SAP	definiteness	A=+def	definiteness	O=–def
Laguna Keres	conflated	SAP status	A=+SAP	SAP status	O=+SAP				
Region 3	split-S _{nom-acc}	SAP status	A=+SAP	SAP status	O=–SAP	verbal semantics	agentive		
Keresan genus	split-S _{nom-acc}	SAP status	A=–SAP	SAP status	O=+SAP	verbal semantics	agentive		
Keresan family	split-S _{nom-acc}	SAP status	A=+SAP	SAP status	O=–SAP	lexical specification	agentive		
	split-S _{nom-acc}	SAP status	A=–SAP	SAP status	O=+SAP	lexical specification	agentive		
	split-S _{erg-abs}	SAP status	A=+SAP	SAP status	O=–SAP	verbal semantics	patientive		
	split-S _{erg-abs}	SAP status	A=–SAP	SAP status	O=+SAP	verbal semantics	patientive		
	split-S _{erg-abs}	SAP status	A=+SAP	SAP status	O=–SAP	lexical specification	patientive		
	split-S _{erg-abs}	SAP status	A=–SAP	SAP status	O=+SAP	lexical specification	patientive		
	neutral	SAP status	–SAP	mood	imperative				
Maricopa	conflated	SAP status	O=+SAP	nominal status	+pronominal				
Region 3	nominative - accusative	SAP status	O=–SAP	SAP status	A=+SAP	SAP status	S=+SAP	nominal status	+pronominal
Cochimí-Yuman genus	neutral	SAP status	–SAP	nominal status	+pronominal				
Cochimí-Yuman family	nominative - accusative	SAP status	+SAP	nominal status	pronoun				
	neutral	SAP status	–SAP	nominal status	pronoun				
	nominative - accusative	nominal status	noun						
Seri	conflated	SAP status	A=–SAP	SAP status	O=–SAP				
Region 3	tripartite	person	1	number	singular				
Seri genus	nominative - accusative	person	1	number	singular				

Seri family	nominative - accusative	person	2						
	neutral	SAP status	–SAP						
Alabama	split-S _{nom-acc}	SAP status	+SAP	participant semantics	+control				
Region 4	split-S _{erg-abs}	SAP status	+SAP	participant semantics	–control				
Muskogean genus	neutral	SAP status	–SAP						
Muskogean family									
Cherokee	conflated	person	A=1	person	O=2				
Region 4	conflated	person	A=2	person	O=1				
Southern Iroquoian genus	conflated	SAP status	–SAP	number	plural				
Iroquoian family	conflated (A/O = S _A)	animacy	O=—animate						
	conflated (A/O = S _B)	SAP status	A=—SAP	number	A=singular				
	direct-inverse	SAP status	A=—SAP	SAP status	O=—SAP				
	split-S _{agt}	verbal semantics	agentive	aspect	present				
	split-S _{agt}	verbal semantics	agentive	aspect	imperfective				
	split-S _{agt}	verbal semantics	agentive	aspect	punctual				
	split-S _{agt}	verbal semantics	agentive	aspect	future				
	split-S _{agt}	lexical specification	agentive	aspect	present				
	split-S _{agt}	lexical specification	agentive	aspect	imperfective				
	split-S _{agt}	lexical specification	agentive	aspect	punctual				
	split-S _{agt}	lexical specification	agentive	aspect	future				
	split-S _{pat}	verbal semantics	patientive						
	split-S _{pat}	lexical specification	patientive						

	split-S _{pat}	verbal semantics	agentive	aspect	perfective				
	split-S _{pat}	verbal semantics	agentive	aspect	infinitive				
	split-S _{pat}	verbal semantics	agentive	aspect	pre-inceptive				
	split-S _{pat}	verbal semantics	agentive	aspect	propensitative				
	split-S _{pat}	lexical specification	agentive	aspect	perfective				
	split-S _{pat}	lexical specification	agentive	aspect	infinitive				
	split-S _{pat}	lexical specification	agentive	aspect	pre-inceptive				
	split-S _{pat}	lexical specification	agentive	aspect	propensitative				
Euchee	split-S _{nom-acc}	SAP status	+SAP	verbal semantics	active				
Region 4	split-S _{nom-acc}	SAP status	+SAP	lexical specification	active				
Yuchi genus	split-S _{erg-abs}	SAP status	+SAP	verbal semantics	inactive				
Yuchi family	split-S _{erg-abs}	SAP status	+SAP	lexical specification	inactive				
	neutral	SAP status	–SAP						
Lowland Chontal	split-S _{nom-acc}	SAP status	+SAP	verbal semantics	agentive				
Region 5	split-S _{nom-acc}	SAP status	+SAP	event semantics	agentive				
Tequistlatecan genus	split-S _{nom-acc}	SAP status	+SAP	lexical specification	agentive				
Tequistlatecan family	split-S _{nom-acc}	SAP status	–SAP	number	plural	verbal semantics	agentive		
	split-S _{nom-acc}	SAP status	–SAP	number	plural	participant semantics	agentive		
	split-S _{nom-acc}	SAP status	–SAP	number	plural	lexical specification	agentive		
	split-S _{erg-abs}	SAP status	+SAP	verbal semantics	patientive				

	split-S _{erg-abs}	SAP status	+SAP	event semantics	patientive				
	split-S _{erg-abs}	SAP status	+SAP	lexical specification	patientive				
	split-S _{erg-abs}	SAP status	–SAP	number	plural	verbal semantics	patientive		
	split-S _{erg-abs}	SAP status	–SAP	number	plural	event semantics	patientive		
	split-S _{erg-abs}	SAP status	–SAP	number	plural	lexical specification	patientive		
	neutral	SAP status	–SAP	number	singular				
Misantla Totonac	nominative - accusative	SAP status	+SAP						
Region 5	nominative - accusative	SAP status	–SAP	number	plural				
Totonacan genus	neutral	SAP status	–SAP	number	singular				
Totonacan family									
Chol	split-S _{erg-abs}	aspect	perfective						
Region 5	split-S _{nom-acc}	aspect	imperfective						
Mayan genus									
Mayan family									

REFERENCES

- Aissen, Judith L. 1987. *Tzotzil clause structure*. Dordrecht: D. Reidel Publishing Company.
- Aissen, Judith L. 2001. Markedness and subject choice in Optimality Theory. In: Géraldine Legendre, Jane Grimshaw & Sten Vikner (eds.). *Optimality-Theoretic Syntax*. Cambridge: The MIT Press, 61-96.
- Aissen, Judith L. 2003. Differential object marking: iconicity vs. economy. *Natural Language & Linguistic Theory* 21(3), 435-483.
- Allen, Barbara J., Donna B. Gardiner, & Donald G. Frantz. 1984. Noun incorporation in Southern Tiwa. *International Journal of American Linguistics* 50(3), 292-311.
- Andersen, Torbet. 1988. Ergativity in Pari, a Nilotic OVS language. *Lingua* 75, 289-324.
- Bakker, Dik. 2011. Language sampling. In Jae Jung Song (ed.). *The Oxford Handbook of Linguistic Typology*, 100-127. Oxford University Press, Oxford.
- Battistella, Edwin L. 1996. *The Logic of Markedness*. New York: Oxford University Press.
- Bavant, Marc. Proto-Indo-European ergativity...still to be discussed. *Pozcan Studies in Contemporary Linguistics* 44(4), 433-447.
- Bell, Alan. 1978. Language samples. In Joseph H. Greenberg (ed.). *Universals of Human Language Volume 1: Method and theory*, 123-156. Stanford: Stanford University Press.
- Belletti, Adriana & Luigi Rizzi. 1988. Psych-verbs and T-theory. *Natural Language & Linguistic Theory* 6.3, 291-352.
- Bennis, H.J. 2004. Unergative adjectives and psych verbs. In: Artemis Alexiadou, Elena Anagnostopoulou & Martin Everaert (eds). *The unaccusativity puzzle: Explorations of the syntax-lexicon interface*. Oxford: Oxford University Press, 40-113.
- Bickel, Balthasar. 2008. A refined sampling procedure for genealogical control. *Sprachtypologie und Universalienforschung* 61(3), 221-233.
- Bickel, Balthasar. 2011. Grammatical relations typology. In Jae Jung Song (ed.). *The Oxford Handbook of Linguistic Typology*. Oxford: Oxford University Press, 399-444.
- Bickel, Balthasar and Johanna Nichols. 2007. Inflectional morphology. In Timothy Shopen (ed.). *Language Typology and Syntactic Description Volume 3: Grammatical categories and the lexicon*. Cambridge: Cambridge University Press, 169-240.

- Bickel, Balthasar and Johanna Nichols. 2009. Case marking and alignment. In Andrej Malchukov & Andrew Spencer (eds.). *The Oxford Handbook of Case*. Oxford: Oxford University Press.
- Bittner, Maria & Ken Hale. 1996. Ergativity: toward a theory of a heterogenous class. *Linguistic Inquiry* 27(4), 531-604.
- Bliss, Heather. 2009. Argument structure, applicatives, and animacy in Blackfoot. *Proceedings of the WSCLA* 14, 58-69.
- Bliss, Heather & Karen Jesney. 2005. Resolving hierarchy conflict: local obviation in Blackfoot. *Calgary Papers in Linguistics* 26, 92-116.
- Bresnan, Joan. 2011. Emergence of the unmarked pronoun. In: Géraldine Legendre, Jane Grimshaw & Sten Vikner (eds.). *Optimality-Theoretic Syntax*. Cambridge: The MIT Press, 113-142.
- Bouldreault, Lynda J. de Jong. 2009. *A Grammar of Sierra Popoluca*. Ph.D. dissertation, University of Texas at Austin.
- Buckley, Eugene. 1988a. Ergativity in Alsea. *Occasional Papers on Linguistics: Papers from the 1987 Hoka-Penutian Languages Workshop and Friends of Uto-Aztecan Workshop, held at University of Utah, Salt Lake City, June 18-21, 1987*, 28-38.
- Buckley, Eugene. 1988b. Temporal boundaries in Alsea. *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, 10-22.
- Buckley, Eugene. 1989. The Alsea noun phrase. *Kansas Working Papers in Linguistics: Studies in Native American Languages* V 14(2), 20-36.
- Burzio, Luigi. 1986. *Italian Syntax: A Government-Binding approach*. Dordrecht: Reidel.
- Butt, Miriam. 2006. *Theories of Case*. Cambridge: Cambridge University Press.
- Campbell, Lyle. 1997. *American Indian Languages: The historical linguistics of native America*. Oxford: Oxford University Press.
- Campbell, Lyle. 2006. Areal linguistics: A closer scrutiny. In Yaron Matras, April McMahon & Nigel Vincent (eds.). *Linguistic Areas: Convergence in historical and typological perspective*, 1-31. Basingstoke and New York: Palgrave Macmillan.

- Campbell, Lyle, Terrence Kaufman, and Thomas C. Smith-Stark. 1986. Meso-America as a linguistic area. *Language* 62(3), 530-570.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge: MIT Press.
- Chomsky, Noam. 1982. *Lectures on Government and Binding: The Pisa lectures*. Dordrecht: Foris Publications.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge: MIT Press.
- Comrie, Bernard. 1989. *Language Universals and Linguistic Typology: Syntax and morphology*. Oxford: Blackwell.
- Comrie, Bernard. 2005. Alignment of case marking. In: M. Haspelmath, H.J. Bibiko & C. Schmidt (eds.). *The World Atlas of Language Structures*. Oxford: Oxford University Press, 398-405.
- Comrie, Bernard, Martin Haspelmath, & Balthasar Bickel. 2005. The Leipzig glossing rules: conventions for interlinear morpheme-by-morpheme glosses. Online. <<http://www.eva.mpg.de/lingua/resources/glossing-rules.php>>
- Coon, Jessica. 2013. *Aspects of Split Ergativity*. Oxford: Oxford University Press.
- Coon, Jessica & Omer Preminger. 2012. Towards a unified account of person splits. *Proceedings of the 29th West Coast Conference on Formal Linguistics*, 310-318.
- Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.
- Creissels, Denis. 2009. Uncommon patterns of core term marking and case terminology. *Lingua* 119, 445-459.
- Crippen, James A. 2012. The basics of Tlingit verbal structure. Manuscript.
- Cruse, D.A. (ed.). 1973. *Lexical Semantics*. Cambridge: Cambridge University Press.
- Cysouw, Michael. 2003. *The paradigmatic structure of person marking*. Oxford: Oxford University Press.
- Dahlstrom, Amy. 1983. Agent-patient languages and split case marking systems. *Proceedings of the Ninth Annual Meeting of the Berkeley Linguistics Society*, 37-46.
- Davis, Philip W. & Heather K. Hardy. 1988. The absence of noun marking in Alabama. *International Journal of American Linguistics* 54(3), 279-308.

- Dixon, R.M.W. 1979. Ergativity. *Language* 55(1), 59-138.
- Dixon, R.M.W. 1994. *Ergativity*. Cambridge: Cambridge University Press.
- Dixon, R.M.W. 2010a. *Basic Linguistic Theory Volume 1*. Oxford: Oxford University Press.
- Dixon, R.M.W. 2010b. *Basic Linguistic Theory Volume 2*. Oxford: Oxford University Press.
- Donohue, Mark. 2008. Semantic alignment systems: what's what, and what's not. In Mark Donohue & Søren Wichmann (eds.). *The Typology of Semantic Alignment*. Oxford: Oxford University Press, 24-75.
- Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67(3), 547-619.
- Drapeau, Lynn. 2012. Passives in Innu. *International Journal of American Linguistics* 78(2), 175-201.
- Dryer, Matthew S. 1989. Large linguistic areas and language sampling. *Studies in Language* 13(2), 257-292.
- Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68, 81-138.
- Dryer, Matthew S. 2006. Descriptive theories, explanatory theories, and basic linguistic theory. In Felix K. Ameka, Alan Charles Dench & Nicholas Evans (Eds.). *Catching Language: The standing challenge of grammar writing*, 207-234. Berlin: Mouton de Gruyter.
- Dryer, Matthew S. 2011. Genealogical language list. In Matthew S. Dryer & Martin Haspelmath (eds.). *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library. Accessed online (3/1/12): <http://wals.info/languoid/genealogy>.
- Du Bois, John W. 1981. *The Sacapultec Language*. Ph.D. dissertation, University of California Berkeley.
- Dunn, Michael, Simon J. Greenhill, Stephen C. Levinson & Russell D. Gray. 2011. Evolved structure of language shows lineage-specific trends in word-order universals. *Nature* 473, 79-82.
- Enrico, John. 2003. *Haida Syntax, Volume 1*. Lincoln and London: University of Nebraska Press.
- Evans, Nick. 1995. A-quantifiers and scope in Mayali. In Emmon Bach, Eloise Jelinek, Angelika Kratzer, and Barbara H. Partee (eds.). *Quantification in Natural Languages, Volume 1*, 207-270.
- Farrel, Patrick. 2005. *Grammatical Relations*. Oxford: Oxford University Press.

Faulhaber, Susan. 2011. *Verb Valency Patterns: A challenge for semantics-based accounts*. Berlin: Walter de Gruyter GmbH & Co.

Fillmore, Charles J. 1968. The case for case. In: Emmon Bach and Robert T. Harms (eds.). *Universals in Linguistic Theory*. New York: Holt, Reinhart and Winston, Inc.

Finney, Joseph C. 2007. Austronesian ergativity traced through two cycles. In Ratree Wayland, John Hartmann, and Paul Sidwell (eds.). *SEALS XII: Papers from the 12th annual meeting of the Southeast Asian Linguistics Society 2002*. Canberra: Pacific Linguistics, 37-78.

Foley, W.A. & R.D. Van Valin, Jr. 1984. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.

Frachtenberg, Leo J. 1920. *Alsea Texts and Myths*. Smithsonian Institution, Bureau of American Ethnology Bulletin 67. Washington: Government Printing Office.

Frajzyngier, Zygmunt. 1984. Ergative and nominative-accusative features in Mandara. *Journal of African Languages and Linguistics* 6, 35-45.

Frantz, Donald G. 1985. Syntactic constraints on noun incorporation in Southern Tiwa. *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*, 107-116.

Frantz, Donald G. 2009. *Blackfoot Grammar, Second Edition*. Toronto: University of Toronto Press.

Frantz, Donald G. & Norma Jean Russell. 1995. *Blackfoot Dictionary of Stems, Roots, and Affixes, Second Edition*. Toronto: University of Toronto Press.

Givón, Talmy. 1975. Topic, pronoun, and grammatical agreement. In Charles N. Li (ed.). *Subject and Topic*. New York: Academic Press, 149-188.

Givón, Talmy. 1980. The drift away from ergativity: diachronic potentials in Sherpa. *Folia Linguistica*, 41-60.

Givón, Talmy. 2001. *Syntax: An introduction, volume 1*. Amsterdam: John Benjamins Publishing Company.

Gordon, Lynn. 1986. *Maricopa Morphology and Syntax*. Berkeley & Los Angeles: University of California Press.

Graczyk, Randolph. 2007. *A Grammar of Crow*. Lincoln: University of Nebraska Press.

Greenberg, Joseph. 1966. *Language Universals with Special Reference to Feature Hierarchies*. The Hague: Mouton & Co.

Gruber, Jeffrey S. 1965. *Studies in lexical relations*. Ph.D. dissertation, Massachusetts Institute of Technology.

Hardy, Donald. 1989. The semantic basis of the primary/secondary object distinction in Blackfeet. *Southwest Journal of Linguistics* 9(1), 29-44.

Hardy, Heather K. & Philip W. Davis. 1993. The semantics of agreement in Alabama. *International Journal of American Linguistics* 59(4), 453-472.

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

Haude, Katharina. 2006. *A Grammar of Movima*. Ph.D. dissertation, Radboud Universiteit Nijmegen.

Handschuh, Corinna. 2014. *A Typology of Marked-S Languages*. Berlin: Language Science Press.

Heck, Fabian & Marc Richards. 2010. A probe-goal approach to agreement and non-incorporation restrictions in Southern Tiwa. *Natural Language & Linguistic Theory* 28, 681-721.

Hock, Hans Henrich. 1991. *Principles of Historical Linguistics, Second edition*. Berlin: Mouton de Gruyter.

Hopper, Paul J. & Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56(2), 251-299.

Hori, Hirofumi. 2008. Semantic motivations for split intransitivity in Haida. *Gengo Kenkyu* 134: 23-55.

Hudson, Richard. 1995. Does English really have case? *Journal of Linguistics* 31(2), 375-392.

Iggesen, Oliver A. 2009. Asymmetry in case marking. In In Andrej Malchukov & Andrew Spencer (eds.). *The Oxford Handbook of Case*. Oxford: Oxford University Press, 246-257.

Iggesen, Oliver A. 2013. Asymmetrical case-marking. In: In: Matthew S. Dryer & Martin Haspelmath (eds.). *The World Atlas of Languages Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/50> accessed 4/20/15.

Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge: MIT Press.

- Jackendoff, Ray. 1983. *Semantics and Cognition*. Cambridge: MIT Press.
- Jelinek, Eloise. 1995. Quantification in Straits Salish. In Emmon Bach, Eloise Jelinek, Angelika Kratzer, and Barbara H. Partee (eds.). *Quantification in Natural Languages, Volume 2*, 487-540.
- Jenson, John T. & Alana Johns. 1989. The morphosyntax of Eskimo causatives. In Donna B. Gerdts & Karin Michelson (eds.). *Theoretical Perspectives on Native American Languages*. Albany: State University of New York Press, 209-229.
- Johns, Alana. 1992. Deriving ergativity. *Linguistic Inquiry* 23(1), 57-87.
- Johnson, Meredith. 2012. On the nature of split ergativity in Cherokee. Conference talk: Workshop on American Indigenous Languages, April 27-28.
- Kaufman, Terrence. 1989. Mayan comparative studies. Manuscript.
- Kennard, Edward. 1936. Mandan grammar. *International Journal of American Linguistics* 9(1), 1-43.
- Kim, Yuni. 2008. *Topics in the phonology and morphology of San Francisco del Mar Huave* [PhD Dissertation], University of California, Berkeley.
- Kimball, Geoffrey D. 1991. *Koasati Grammar*. Lincoln: University of Nebraska Press.
- King, Duane Harold. 1975. *A Grammar and Dictionary of the Cherokee Language*. Ph.D. dissertation, University of Georgia.
- Kinkade, M. Dale. 2005. Alsea pronouns. *Anthropological Linguistics* 47(1), 61-76.
- Klaiman, M. H. 1991. Control and grammar. *Linguistics* 29, 623-651.
- Lachler, Jordan. 2006. *A Grammar of Laguna Keres*. Ph.D. dissertation, University of New Mexico.
- Langacker, R. W. 1999. *Grammar and Conceptualization*. Berlin: Moutin de Gruyter.
- Lass, Roger. 1997. *Historical Linguistics and Language Change*. Cambridge University Press, Cambridge.
- Lee, Geddy, Neil Peart & Alex Lifeson. 1981. Vital signs. *Moving Pictures*. Mercury Records.

- Leer, Jeffry A. 1991a. *The Schetic Categories of the Tlingit Verb*. Ph.D. dissertation, University of Chicago.
- Leer, Jeff. 1991b. Evidence for a Northern Northwest Coast language area: promiscuous number marking and periphrastic possessive constructions in Haida, Eyak, and Aleut. *International Journal of American Linguistics* 57(2), 158-193.
- Legendre, Géraldine, Jane Grimshaw & Sten Vikner (eds.). *Optimality-Theoretic Syntax*. Cambridge: The MIT Press.
- Levin, Beth & Malka Rappaport Hovav. 1995. *Unaccusativity: At the syntax-lexical semantics interface*. Cambridge: The MIT Press.
- Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig (eds.). 2014. *Ethnologue: Languages of the World* 17th edition. Dallas: SIL International. Online version: <http://www.ethnologue.com>.
- Linn, Mary Sarah. 2001. *A Grammar of Euchee*. Ph.D. Dissertation, University of Kansas.
- Lupardus, Karen Jacque. 1982. *The Language of the Alabama Indians*. Ph.D. dissertation: University of Kansas.
- Macaulay, Monica. 2005. On the 2 > 1 prominence hierarchy of Algonquian. *LSO Working Papers in Linguistics 5: Proceedings of WIGL 2005*, 1-24.
- MacKay, Carolyn J. 1999. *A Grammar of Misanla Totonac*. Salt Lake City: University of Utah Press.
- Manning, Christopher. 1996. *Ergativity: Argument structure and grammatical relations*. Stanford: CSLI Publications.
- Masica, C.P. 1976. *Defining a linguistic area: South Asia*. University of Chicago Press, Chicago.
- Marlett, Stephen A. 1981. *The structure of Seri*. Ph.D. dissertation, University of California, San Diego.
- Marlett, Stephen A. 1990. Person and number inflection in Seri. *International Journal of American Linguistics* 56(4), 503-541.
- Marlett, Stephen A. 2009. A Grammar of Seri, chapter 16: Pronouns [working draft]. Accessed online 1/22/14 at http://www.und.nodak.edu/instruct/smarlett/Stephen_Marlett/GrammarDraft.html.

- Meakins, Felicity. 2015. From absolutely optional to only nominally ergative: the life cycle of the Gurindji ergative suffix. In: Francesco Gardani, Peter Arkadiev, Nino Amiridze (eds.). *Borrowed Morphology*. Berlin: Walter de Gruyter, Inc., 189-218.
- Melnar, Lynette R. 2004. *Caddo verb morphology*. Lincoln: University of Nebraska Press.
- Miller, Cynthia L. and Leoma G. Gilley. 2001. Evidence for ergativity in Shilluk. *Journal of African Languages and Linguistics* 22, 33-68.
- Mithun, Marianne. 1991. Active/agentive case marking and its motivations. *Language* 67(3), 510-546.
- Mithun, Marianne. 1999. *The Native Languages of North America*. Cambridge: Cambridge University Press.
- Mithun, Marianne. 2010a. Ergativity and language contact on the Oregon Coast: Alsea, Siuslaw and Coos. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 26(2), 77-95.
- Mithun, Marianne. 2010b. Contact and North American languages. In: Raymond Hickey (ed.). *The Handbook of Language Contact*. Malden: Wiley-Blackwell, 673-693.
- Mulder, Jean Gail. 1994. *Ergativity in Coast Tsimshian (Sm'algyax)*. Berkeley: University of California Press.
- Munro, Pamela & Lynn Gordon. 1982. Syntactic relations in Western Muskogean: a typological perspective. *Language* 58(1), 81-115.
- Næss, Åshild. 2007. *Prototypical transitivity*. Amsterdam: John Benjamins Publishing Company.
- Nichols, Johanna. 1992. *Linguistic Diversity in Space and Time*. Chicago: University of Chicago Press.
- Nübling, Damaris. 2010. How do exceptions arise? On different paths to morphological irregularity. In: Horst J. Simon & Heike Wiese (eds.). *Expecting the Unexpected: Exceptions in grammar*. Berlin: De Gruyter Mouton, 139-162.
- Ochs, Elinor. 1982. Ergativity and word order in Samoan child language. *Language* 58(3), 646-671.
- O'Connor, Loretta. 1999. Agentive system of Lowland Chontal. *Santa Barbara Papers in Linguistics 9: First Annual Workshop on American Indigenous Languages*, 144-57.

O'Connor, Loretta. 2010. Aspectual classes and non-agentive morphosyntax in Lowland Chontal. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 26(2), 109-120.

Parks, Douglas R. 1976. *A Grammar of Pawnee*. New York & London: Garland Publishing, Inc.

Parks, Douglas R. & Lula Nora Pratt. 2008. *A Dictionary of Skiri Pawnee*. Lincoln: University of Nebraska Press.

Payne, John R. 1980. The decay of ergativity in Pamir languages. *Lingua* 51, 147-186.

Perkins, Revere D. 1989. Statistical techniques for determining language sample size. *Studies in Language* 13(2), 293-315.

Perkins, Revere D. 1992. *Deixis, Grammar, and Culture*. John Benjamins Publishing Company, Amsterdam/Philadelphia.

Perlmutter, David. 1978. Impersonal passives and the Unaccusative Hypothesis. *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society*, 157-189.

Perlmutter, David. 1980. Relational grammar in current approaches to syntax. *Syntax and Semantics, Ann Arbor, Michigan* 13, 195-229.

Perlmutter, D.M. (Ed.). 1983. *Studies in Relational Grammar I*. Chicago: University of Chicago Press.

Primus, Beatrice. 1999. *Cases and Thematic Roles: Ergative, accusative and active*. Tübingen: Max Niemeyer Verlag.

Prince, Alan & Paul Smolensky. 2008. *Optimality Theory: Constraint interaction in generative grammar*. Malden: Blackwell.

Pye, C. 1990. The acquisition of ergative languages. *Language* 28, 1291-1330.

Quizar, Robin. 1994. Split ergativity and word order in Ch'orti'. *International Journal of American Linguistics* 60(2), 120-138.

Robertson, John S. 1992. *The History of Tense/Aspect/Mood/Voice in the Mayan Verbal Complex*. Austin: University of Texas Press.

RANDOM.ORG. 1998-2014. www.random.org.

Rosen, Carol. 1990. Rethinking Southern Tiwa: the geometry of a triple-agreement language. *Language* 66(4), 669-713.

- Rude, Noel. 1986. Topicality, transitivity, and the direct object in Nez Perce. *International Journal of American Linguistics* 52(2), 124-153.
- Ruhlen, Merritt. 1987. *A Guide to the World's Languages: Volume 1: Classification*. Stanford: Stanford University Press.
- Saeed, John I. 2009. *Semantics: Third edition*. Malden: Wiley-Blackwell.
- Sadock, Jerrold M. 1985. The Southern Tiwa Incorporability Hierarchy. *International Journal of American Linguistics* 51(4), 568-572.
- Scancarelli, Janine. 1987. *Grammatical Relations and Verb Agreement in Cherokee*. PhD dissertation, UCLA.
- Scheepers, Christoph, Barbara Hemforth, Lars Konieczny. 2000. Linking syntactic functions with thematic roles: Psych-verbs and the resolution of subject-object neutrality. In: Barbara Hemforth & Lars Konieczny (eds.). *German sentence processing*. Springer Netherlands, 95-135.
- Sherzer, Joel. 1976. *An Areal-typological Study of American Indian Languages North of Mexico*. Amsterdam: North-Holland Publishing Company.
- Siewierska, Anna. 2003. Person agreement and the determination of alignment. *Transactions of the Philological Society* 101:2, 339-370.
- Siewierska, Anna & Dik Bakker. 2009. Case and alternative strategies: word order and agreement marking. In Andrej Malchukov & Andrew Spencer (eds.). *The Oxford Handbook of Case*. Oxford: Oxford University Press, 290-303.
- Spencer, Andrew. 2009. Case as a morphological phenomenon. In Andrej Malchukov & Andrew Spencer (eds.). *The Oxford Handbook of Case*. Oxford: Oxford University Press, 185-199.
- Story, Gillian. 1972. *A morphological study of Tlingit*. Dallas: SIL.
- Stout, Carol. 1973. Problems of a Chomskyan analysis of Zuni transitivity. *International Journal of American Linguistics* 39(4), 207-223.
- Story, Gillian L. & Constance M. Naish. 1973. *Tlingit Verb Dictionary*. College: University of Alaska.
- Tomlin, Russell S. 1986. *Basic Word Order: Functional principles*. Croom Helm, London.

- van der Auwera, Johan, Nina Dobrushina & Valentin Goussev. 2013. Imperative-hortative systems. In: Matthew S. Dryer & Martin Haspelmath (eds.). *The World Atlas of Languages Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter72> accessed 4/16/15.
- van Valin, Robert D. 1990. Semantic parameters of split intransitivity. *Language* 66(2), 221-260.
- Vázquez Álvarez, Juan Jesús. 2011. *A Grammar of Chol, A Mayan Language*. PhD dissertation, University of Texas at Austin.
- Waterhouse, Viola Grace. 1962. The grammatical structure of Oaxaca Chontal. *International Journal of American Linguistics* 28(2), 1-121.
- Wichmann, Søren. 2008. The study of semantic alignment: retrospect and state of the art. In: Mark Donohue & Søren Wichmann (eds.). *The Typology of Semantic Alignment*. Oxford: Oxford University Press.
- Wichmann, Søren. 2009. Case relations in Tlapanec, a head-marking language. In Andrej Malchukov & Andrew Spencer (eds.). *The Oxford Handbook of Case*. Oxford: Oxford University Press, 797-807.
- Yamamoto, Mutsumi. 1999. *Animacy and Reference*. Amsterdam: John Benjamins Publishing Company.
- Zúñiga, Fernando. 2006. *Deixis and Alignment: Inverse systems in indigenous languages of the Americas*. Amsterdam: John Benjamins.