Differential Case-Marking in Ecuadorian Siona

Justin Thomas Case

A thesis presented to the University of Ottawa
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in Linguistics

Department of Linguistics
Faculty of Arts
University of Ottawa

© Justin Thomas Case, Ottawa, Canada, 2026

This thesis was created in the context of a cotutelle agreement with Leiden University (Netherlands). This version is modified from that published by LOT publishers — LOT719. Although the content is unchanged, due to formatting differences, page numbers may not always align across version. Bibliographical details for the LOT version are provided below:

ISBN: 978-94-6093-504-6
DOI: https://dx.medra.org/10.48273/LOT0719
NUR: 616

Promotores: (Supervisors)	prof. dr. A.P. Salanova prof. dr. J.S. Doetjes	University of Ottawa Leiden University
Co-promotor: (Co-supervisor)	dr. M. Bruil	Leiden University
Overige leden: (Committee members)	prof. dr. É. Mathieu prof. dr. H. de Hoop dr. G.J. van der Wal dr. N.D. Siddiqi	University of Ottawa Radboud University Leiden University Carleton University

The research carried out to form this dissertation is partially self-funded and partially funded by the Ontario Graduate Scholarship (awarded twice: 2020-2022) under the project title "Animacy strewn across the grammar in Ecuadorian Siona".

This dissertation is dedicated to my *airob<u>a</u><u>i</u>* friends in Sototsiaya
and to the *b<u>a</u><u>i</u>coca* language. *Ai deoji si'ahua'i !*

Contents

Acknowledgments ix

Abstract xi

Abbreviations xv

1 Introduction **1**
- 1.1 DCM patterns attested in Siona 2
- 1.2 Research questions and dissertation outline 5
- 1.3 The Siona people and their language 7
  - 1.3.1 Siona in its genealogical context 8
  - 1.3.2 Siona in its areal context 11
- 1.4 The representation of Siona data 12
  - 1.4.1 A taxonomy of Siona example types 12
  - 1.4.2 Glossing conventions: Reporting parsed Siona data . . . 14
  - 1.4.3 Orthography matters: Presenting Siona data 15

2 Differential case-marking in the typological record **18**
- 2.1 DCM: A working definition 18
- 2.2 Ingredients for a holistic description of DCM 19
  - 2.2.1 The formal dimension: A typology of DCM patterns . . . 20
  - 2.2.2 The functional dimension: DCM across grammatical relations 22
  - 2.2.3 The semantico-pragmatics of DCM 25
    - 2.2.3.1 Multi-dimensionality: Ranking DCM triggers . 29
    - 2.2.3.2 Ambiguity avoidance in the broader DCM picture 31

3 Siona case grammar: A preliminary description **34**
- 3.1 Subject-oriented DCM patterns 38
- 3.2 Non-Subject-oriented DCM patterns 41
  - 3.2.1 Direct object-oriented DCM patterns 42
  - 3.2.2 Promotion and noun class mobility patterns 46

3.2.3 Indirect object-oriented DCM patterns 50
3.2.4 Other non-Subject DCM patterns 52
3.3 Spatial DCM patterns 54
3.3.1 Static location-oriented DCM patterns 56
3.3.2 Goal-oriented DCM patterns 59
3.3.3 Extended uses of the Goal-marker, -na 62
3.4 Interim summary: Siona DCM as a paradigm 65

4 A comparativist study: Siona case-marking in typology 67
4.1 Methodology: Case-marking questionnaire 69
4.2 Typologizing Siona case-marking 70
4.2.1 Typologizing Siona case: The Formal domain 70
4.2.1.1 The case-marker inventory (size) 70
4.2.1.2 The case-marker inventory (members) 72
4.2.1.3 Single versus multiple case-marking patterns . 76
4.2.2 Typologizing Siona case: The Functional domain 77
4.2.2.1 Situating Siona case-marker correspondences . 77
4.2.2.2 Functional gaps in Siona case-marking 88
4.2.3 Typologizing Siona DCM properties 91
4.2.3.1 Situating Siona Subject-oriented DCM 91
4.2.3.2 Situating Siona Direct object-oriented DCM . . 94
4.2.3.3 Situating Siona Indirect object-oriented DCM . 100
4.2.3.4 Situating Siona spatial DCM patterns 103
4.3 Conclusion: Siona case-marking in typology 107

5 A variationist study: DCM in a narrative sample 111
5.1 The Siona narrative dataset 112
5.2 Methodology: The variationist program 113
5.2.1 From tales to tokens: Coding the narrative sample . . . 114
5.2.2 The variable-rule technique: An overview 118
5.3 Results: DCM in the narrative dataset 120
5.3.1 Distribution of marking alternatives in the sample . . . 120
5.3.2 S-marking in the Siona narrative dataset 124
5.3.2.1 Excluding invariable S-tokens 124
5.3.2.2 Exploring the *varbrul* results for S-oriented DCM 126
5.3.3 P-marking in the Siona narrative dataset 129
5.3.3.1 Excluding invariable and outlier P-tokens . . . 130
5.3.3.2 Exploring the *varbrul* results for P-oriented DCM 132
5.3.4 L-marking in the Siona narrative dataset 137
5.3.4.1 Exploring the *varbrul* results for L-oriented DCM 139
5.4 Discussion: Siona DCM in usage 142
5.4.1 Discourse-related Triggers in the dataset 144
5.4.2 Referentiality-related Triggers in the dataset 145
5.4.3 Sentence-level Triggers in the Siona narrative dataset . 146

6 Information structure and Siona DCM **148**
6.1 Methodology: Controlled elicitation 150
6.2 Case-marking as a focus strategy 152
6.3 Focal case-marking: The size question 156
6.3.1 Broad and narrow interpretations for focal case-marking 156
6.3.2 Focal case-marking within complex nominals 159
6.4 Distribution check: Siona DCM and focus 166
6.4.1 Siona DCM patterns across focal construction types . . . 167
6.4.1.1 Case-marking and information-seeking questions 171
6.4.1.2 DCM and opposition constructions 174
6.4.2 Compatibility of focal case- and discourse-marking . . . 176
6.4.3 In summary: Focus-sensitive Siona DCM 182
6.5 Towards a theory of discourse-driven DCM 183

7 Conclusions **186**
7.1 Summary of findings 189
7.1.1 A typological investigation of Siona case-marking 189
7.1.2 A corpus-based investigation of Siona DCM 190
7.1.3 An elicitation-driven investigation: Focus-sensitive case-marking 192
7.2 Future research directions 195
7.3 Towards a formal analysis of Siona DCM 196

Appendix A Northwest Amazonia case-marking dataset **200**
A.1 Methodology matters 200
A.1.1 Presenting the NWA sample 200
A.1.2 Formulating and administering the questionnaire 203
A.1.2.1 Coding the Formal domain of the questionnaire 204
A.1.2.2 Coding the Functional domain of the questionnaire 206
A.1.2.3 Coding the DCM domain of the questionnaire . 209
A.2 Presenting the NWA dataset 211
A.2.1 The dataset: The Formal domain 213
A.2.2 The dataset: The Functional domain 216
A.2.3 The dataset: The DCM domain 231

Appendix B Siona DCM narrative dataset (supplement) **251**
B.1 Token-by-token annotation protocols 252
B.1.1 Annotating sentence-level factors 252
B.1.2 Annotating referential token-level factors 254
B.1.3 Annotating information structural properties 258
B.2 Results: Statistics and illustrative data 261
B.2.1 The results for S-oriented DCM 262
B.2.2 The results for P-oriented DCM 266
B.2.3 The results for L-oriented DCM 278
B.2.4 Miscellaneous examples for discussion 281

Bibliography **289**

Curriculum Vitae 310

Acknowledgments

This dissertation is the culmination of years of work, which would not have been possible without the support of my family, my friends, and my colleagues. I count myself incredibly lucky to have worked with so many talented individuals, and these words are a small token of my appreciation for the time and effort you have all invested in this project and in me.

First and foremost, my deepest gratitude goes to the Siona speakers, the *airobai̯* of Sototsiaya. I will never forget my first visit in 2018. The warmth of the community, children jumping into the stream, and the patience of those who took the time to teach me about their language. Since then, through multiple returns in 2019, 2022, 2023, and 2025, the warm reception has never wavered. To the families who hosted me, shared their meals, and opened their culture to me: my life has been enriched by your generosity and your friendship. I dedicate this work to my Siona friends. I truly hope that this work contributes, in some small way, to the maintenance and to the celebration of the *bai̯coca* language for the generations to come.

I would also like to recognize my linguist friends who shared the field with me through the years: Alex Nastevski, Teresa Rabideau, Olivia Griffin, Jurriaan Wiegertjes, Jasmine Desjardins, Jelle Christiaans, Isaac Eaton, Veronika Voevodina, and Rochelle Valenzuela. (Those of you who are not mentioned here, are not forgotten.) A special mention goes to my dear friend and colleague, Paloma Jeretič. I have fond memories of our long chats on hammocks in Sototsiaya, working out the puzzles of Siona grammar as a syntax-semantics team. You have made me a better linguist, and I look forward to our future collaborations.

My doctoral training was shaped by exceptional minds at both the University of Ottawa and Leiden University. At Ottawa, I am grateful for the guidance and support of Ana Arregui, Dennis Ott, and Éric Mathieu, to name a few, particularly in the earlier stages of my doctoral training. I also learned greatly from Laura Sabourin and Tania Zamuner who taught me the foundations of quantitative analysis. In Leiden, I benefited immensely from the insights and collaborations of Jenneke van der Wal and Rik van Gijn. I have learned a great

deal from our discussions over the past few years, and I am a better linguist for it.

This project would have been impossible without my extraordinary supervisory team. At Ottawa, my deepest gratitude goes out to Andrés Salanova, who has played a pivotal role in my development throughout this entire process. You pushed me to be a well-rounded linguist, encouraged me to embrace the descriptive dimensions of our field, and are a role model of how to foster meaningful relationships with Indigenous communities. Your mentorship is foundational to my professional development.

To Jenny Doetjes: thank you for your unrelenting kindness and for all of the advice in developing the final version of this doctoral project. Your guidance on precise and concise writing and scientific reporting, and your infectious curiosity, even for phenomena outside your immediate research, have had a considerable impact on me and have elevated this dissertation.

And to Martine Bruil: working with you on Siona changed my life. Thank you for your patience as I "learned the ropes" and for showing me the vital importance of working with naturalistic, textual data. You have inspired me to aspire to high descriptive standards and to look at the puzzles from various angles. I feel privileged to count you not just as a supervisor and colleague, but as a close friend. Thank you for making me feel so at home in Leiden.

Finally, I must thank my wonderful family. To my brother Brian and his wife Gabrielle, my dad Terry, my father Tom and his wife Alicia, and to my in-laws in California, Jon, Donna, Joni and Josh: thank you for your support and encouragement through the years. To my mom, Carlene: I cannot properly put into words how much your tireless support has meant. Whether it was a phone call to calm my anxieties or acting as a sounding board for my ideas, you have been a constant throughout this doctorate.

Last, but certainly not least, to my wife, Julia. You have been my pillar through every stress-filled night and every moment of doubt. I am so grateful for your encouragement and for the work you do that allows me to pursue what I love. Having you join me in Sototsiaya to teach English was a highlight for me, and I look forward to many more journeys together.

Thank you all for being part of this chapter in my academic life.

Abstract

Ecuadorian Siona (henceforth Siona) is spoken in the most northwestern region of the Amazon basin. The language is currently spoken by approximately 300 people, spread across a half-dozen communities in the Napo-Putumayo region. The language belongs to the western branch of the Tukanoan language family and exhibits a number of typical traits of this family and other languages in the area. A particularly intriguing trait concerns the distribution of case-markers to indicate the role of core arguments (e.g., Subject, Direct object) and various spatial arguments in the sentence. The exact usage of case-markers is not random, rather it is conditioned by the interplay of several factors such as animacy status, information structure, and communicative efficiency or ambiguity avoidance. These are collectively referred to as *differential case marking* (DCM). Two DCM patterns are illustrated in the examples below — Subject-oriented DCM in (1), and animate Object-oriented DCM in (2):

(1) a. *jaiye joro aide'oji iño*
hai-je **horo** ai + de'o-hi ihño
big-CL:GEN flower big + become-3S.M.PRS.ASS here
'Many flowers grow here.'

b. *jaiye jorobi aide'oji iño*
hai-je **horo-bi** ai + de'o-hi ihño
big-CL:GEN flower-SBJ big + become-3S.M.PRS.ASS here
'Many FLOwers grow here (i.e., not something else).'
[VOL: 20230622ejabi001.012a-c]

(2) a. *bayë jaibai a'yëbaire*
baa-jɨ **hai-bãi** **a'j-ɨ-bãi-re**
have-N3S.PRS.ASS many-PL.AN older_sibling-CL:M-PL.AN-N.SBJ
'(I/we/they) have many (older) siblings.'
[VOL: 20230622ejabi001.011b]

b. *më'bi yë'ni ñañe bạhuë*
mɨ'ɨ-bi **jɨ'ɨ-ni** ñaa-ñe bãã-wɨ
2SG-SBJ 1SG-N.SBJ2 see-CL:GEN NEG.AUX-N3S.PST.ASS
‘ “YOU didn't see ME (i.e., not vice versa).” ’
[NAT: 20140925salsu001.032]

The aim of this work is to refine the analysis of Siona DCM patterns based on the research questions in (3):

(3) a. Which DCM patterns can be identified in Siona? What properties do these DCM patterns have in common, and how do they differ?

b. Which factors determine the selection of case marking alternatives in a given context — such as animacy status, focus, specificity? A combination thereof? How do the active factors interact to determine the DCM patterns in usage?

c. How does Siona DCM compare to other DCM systems described in the Tukanoan literature? In the northwestern Amazonia area? In the broader DCM typology?

To address the questions in (3), this dissertation assembles original data collected via diverse methodologies — including a comparative typological study in Chapter 4, a corpus study in Chapter 5, and an elicitation-based analysis of focus effects in Siona DCM in Chapter 6.

Chapter 2 reviews the DCM literature and lays the groundwork for the analysis of Siona DCM. Chapter 3, entitled ‘Siona case grammar: A preliminary description’, describes the diverse set of Siona DCM patterns, expanding considerably upon the generalizations outlined in Bruil (2014, §4.4). On the basis of diverse data points, this chapter provides new insights regarding the principles of several DCM patterns as pairs of PLAIN versus PROMINENT marking-alternatives, where the former is often zero-marking, as shown in (1a). A distinction between Subjects and non-Subjects is described, where only non-Subjects exhibit an obligatory case-marking requirement for animate nouns. This chapter also introduces a productive pattern where a class of inanimate nominalizations display the obligatory case-marking requirement in the singular. These are called *promoted noun phrases* in this dissertation. Patterns of DCM are described for both argumental and spatial noun phrases, where promoted spatial nouns also adhere to the non-Subject obligatory case-marking generalization. These observations form the foundation for the analysis of Siona via various methodologies.

Chapter 4 reports a comparative typological analysis of (differential) case-marking features in a sample of languages from the northwestern Amazon region, presented in full in Appendix A. This chapter is designed as a comparison of various case-marking properties based on a structural questionnaire. In particular, traits such as inventory size, case syncretisms, and the semantic-pragmatic factors underpinning DCM are scrutinized across the sample. The questionnaire is conceived with Siona as the standard of comparison so that

the findings position the case-marking properties attested in the target language of this dissertation in its typological context. The sample comprises a balanced set of Tukanoan and non-Tukanoan languages spoken within the delimited area — i.e., the putative northwestern Amazonia area. The chapter provides insight into which features may have been inherited from an ancestral language, Proto-(Western-)Tukanoan, and/or where contact-induced grammatical shifts may have arisen. Ultimately Siona and its closest relatives display an admixture of typical Tukanoan traits — e.g., a significant overlap between argumentative and locative DCM, and traits shared with local non-Tukanoan languages — e.g., the presence of DCM in the case of the subject and the crucial role of information structure in DCM patterns.

Chapter 5 presents a corpus analysis of case-marking patterns in a sample of Siona narrative texts (8078 words, 1914 nominal tokens). All eligible Subject, Object, and Location tokens are manually coded as marked or unmarked — as shown on the Subject noun phrases in (1). This forms the dependent variable for a statistical analysis of the narrative corpus via the implementation of the variable-rule technique, i.e., *variable-rule* analysis (Labov 1966, 1970, *et seq.*; Sankoff and Labov 1979; Sankoff 1988). This chapter tests the statistical effects of a handful of plausible factors mentioned in the typological literature, including some already mentioned in Chapter 3: word order, affectedness, animacy, specificity, topicality, and contrastivity. Besides an anticipated degree of variability in usage, statistical support is yielded for Siona DCM as being primarily driven by information structure. Contrastivity shows the largest effect across all three tested grammatical relations (Subjects, Direct objects, and Locations). In other words, a noun phrase is far more likely to receive overt (i.e, PROMINENT) case-marking when its referent is contrasted with salient alternative referents at that moment in the narrative. Topicality shows the second largest effect, although this effect is not observed for Subject-oriented DCM in the sample. A handful of smaller statistical effects regarding word order and specificity status are found, but these effects differ in nature from the effects of information structure. This study also demonstrates that spatial noun phrases receive overt case-marking significantly more often than do argumental noun phrases, despite a clear overlap regarding the statistically significant factors favouring case-marking.

Chapter 6 characterizes the distribution of case-marking in contexts traditionally associated with focus, reporting data collected via controlled elicitations. The results indicate that the relevant overt (PROMINENT) case-marking arises strictly in a subset of these focus contexts — i.e., in certain question-answer pairs such as in (4), an in opposition constructions as in (5). This same case-marking is generally associated with the *emphatic contrast* reading in out-of-the-blue contexts, as was shown above in (1b):

(4) a. QUESTION: *quebi iño aide'oquë ?*
ke-e-bi ihño ai + de'o-kɨ
WH-CL:GEN-SBJ here big + become-2/3S.M.PRS.N.ASS
'<u>What</u> grows here?'

b. ANSWER: *jaiye jorobi aide'oji iño*
hai-je **horo #(-bi)** ai + de'o-hi ihño
big-CL:GEN flower-SBJ big + become-3S.M.PRS.ASS here
'Many FLOWers grow here.' [VOL/SUG: 20230622ejabi002.003a-b]

(5) [*bahuë yohuë hueroye* NEG] [*te'e guënarore huerohuë* AFF]
bãã-wɨ **jo-wɨ** wero-je **te'-e**
NEG.AUX-N3S.PST.ASS canoe-CL:CONT buy-CL:GEN one-CL:GEN
gĩna-ro #(-re) wero-wɨ
metal-CL:RECIP-N.SBJ buy-N3S.PST.ASS
'[I didn't buy a BOAT NEG], [(I) bought a COOKing pot AFF].'
[VOL: 20230623ejepa001.013]

Besides the (often predictable) distribution of overt case-marking in (4) and (5), judged as odd where the case-marking is absent, it is shown that case-marking is not used in certain focus contexts where it might be expected. For example, speakers avoid case-marking in contexts of correction, with focus-sensitive morphemes such as *-hẽ* (also) and *-se'e* (only), and in certain interrogative constructions such as alternative questions (X or Y), or complex WH questions (which student did you see?). Ultimately, this chapter refines the description of the distribution of Siona DCM alternatives on the basis of focus and/or contrastivity, distinguishing rigid from more subtle effects.

Considered together, the various forms of evidence assembled in this dissertation paint a picture of Siona as primarily conditioned by the interplay of an animacy class system and information structure. The former determines whether case-marking is obligatory in certain contexts, and which case-markers are available to a particular noun phrase, whereas the latter conditions the use of overt PROMINENT case-marking. Various insights are presented regarding the subtler role of other, less important factors (e.g., word order, specificity) contributing to the variable nature of Siona DCM. Part of this variation is due to ambiguity avoidance and communicative efficiency, which merits greater attention in future research. These factors differ fundamentally from the animacy and discourse factors scrutinized extensively here.

Abbreviations

1	first-person
2	second-person
3	third-person
ACC	accusative
ABL	ablative
ACC	accusative
ADD	additive
ANIM	animate
APPL	applicative
ASS	assertive
AUX	auxiliary
BEN	benefactive
CL	classifier
CNT.EXP	counter-expectative
COM	comitative
COMP	companion
CONT	container (classifier)
COP	copula
DAT	dative
DCM	differential case-marking
DEF	definite
DEM	demonstrative
DEP	dependent
DERIV	derivation
DES	desiderative
DIST	distal
DS	different-subject
EXCL	exclusive
F	feminine
FOC	focus
FUT	future
GEN	generic (classifier)
GOAL	goal
INAN	inanimate
INCL	inclusive
INTENS	intensive
LOC	locative
M	masculine
MED	medial
N.ASS	non-assertive
N.SBJ	non-subject
N3S	non-third-person singular
NEG	negative
NMLZ	nominalizer
OPP	opposition
PL	plural
POL	polar
POSS	possessor
PRO	pronoun
PROG	progressive
PROX	proximate
PRP	prospective
PRS	present
PST	past
Q	question (interrogative)
RCT.PST	recent past
REDUP	reduplication
REP	reportative

RND	round (classifier)	SS	same subject
SEL	selection/selective	TOP	topic/topical
SG	singular	TEMP	temporal
SOURCE	source	TRS	transitive
SPEC	specific	WH	*wh* question item

The following symbols are utilized to represent complex items in glosses throughout this dissertation. Further discussion of the glossing conventions is found in Section 1.4.2.

< - > indicates affixation; e.g., *jɨ'ɨ-bi* (1SG-SBJ)
< = > indicates cliticization; e.g., *jɨ'ɨ=kato* (1SG=TOP)
< + > indicates compounding; e.g., *je'ja* + *wɨ'e* (school, lit. teaching + house)
< . > extends the grammatical gloss of a morpheme; e.g., *wɨ'e-mahka* (house-DIM.SG)
< _ > extends a lexical gloss where multiple English elements correspond to a single Siona element in the translation; e.g., *ko'e-ñu'ũ* (look_for-HORT)

CHAPTER 1

Introduction

Case-marking morphology is commonly recruited in natural language to indicate the role of a noun in the sentence (Nichols 1986, 1992; Haspelmath 2010; Nichols and Bickel 2013). The classical Latin examples in (1) are a case in point, where the form of each nominal is determined by its role in relation to the predicate. Following terminology in the Latin descriptive tradition, the SUBJECT argument takes the *nominative* form, whereas the DIRECT OBJECT takes the *accusative* form:

(1) Latin (*lat* — INDO-EUROPEAN, ITALIC) [Haegeman 1991, 143]

a. Caeser$^{(\text{SUBJECT})}$ Belgas$^{(\text{OBJECT})}$ vinc-it
C.NOM Belgians.ACC beat-3SG
'Caesar beats the Belgians.'

b. Belgae$^{(\text{SUBJECT})}$ Caeserem$^{(\text{OBJECT})}$ tem-ent
Belgians.NOM C.ACC fear-3PL
'The Belgians fear Caesar.'

Many languages utilize a case-marking strategy like (1) to encode grammatical relations, and various semantic relations (Simpson 2023). In fact, Iggesen (2013) finds that just over half of the world's languages exhibit some form of case-marking.[1] Case-marking languages vary widely as to the morphosyntactic expression of case-markers, and the number of case-marking categories: ranging from 'two-term systems' to languages with rich inventories exceeding a dozen categories (Blake 2001; Arkadiev 2008). Setting aside discrepancies in inventory size, languages map these categories onto a relatively stable set of relations (cf. Haspelmath 2010 and references therein).

[1] According to Iggesen (2013, *WALS*, chapter 49), 164 languages in his diverse 264-specimen sample (62.12%) have at least two case categories. The remaining 100 languages use another strategy to encode grammatical relations.

This dissertation is concerned with a particular case-marking phenomenon whereby not all nouns with a given relation are marked in the same way. Instead, factors beyond the grammatical or semantic relation itself determine the distribution of certain morphological case-markers without affecting the grammaticality of the sentence: e.g., the animacy status or the specificity status of the referent. This phenomenon is generally called DIFFERENTIAL CASE MARKING (henceforth DCM) (Moravcsik 1978; Bossong 1985; Aissen 1999, 2003). An illustrative instance of a DIRECT OBJECT-oriented DCM pattern is demonstrated in the minimal pair from the Turkic language Sakha in (2). Overt case-marking is found on the specific OBJECT in (2b), but not on its non-specific counterpart in (2a):

(2) Sakha (*sah* — TURKIC)
[Vinokurova 2005, cited in Baker 2015, 1, ex. (1c); 4, ex. (5)]

a. Erel **kinige** atyylas-ta
E.NOM book buy-PST.3SS
'Erel bought a book/books.'

b. Erel **kinige-ni** atyylas-ta
E.NOM book-ACC buy-PST.3SS
'Erel bought the book.'

In a recent survey, Sinnemäki (2014) finds that case-marking languages are more likely to exhibit DCM than obligatory case-marking like that shown for Latin in (1).

DCM patterns are not restricted to DIRECT OBJECTS, as in the Sakha pattern shown in (2). Such patterns are also attested on other core arguments, on spatial arguments, and even on certain oblique arguments (e.g., Chappell and Verstraete 2019; Haspelmath 2021). In fact, a single language may display several DCM patterns, each with their own characteristics. DCM has garnered significant scholarly interest both in typological studies and in detailed, language-particular descriptions (Blake 2001; Klein and de Swart 2011; Baker 2015). A more detailed review and discussion of each of these aspects is undertaken in Chapter 2.

1.1 DCM patterns attested in Siona

This dissertation provides an in-depth description and analysis of the various DCM patterns found in the Ecuadorian variety of the Siona language [*snn* – Western Tukanoan, henceforth Siona]. The analysis developed in these pages complements and expands significantly upon the initial descriptive generalizations regarding the distribution of Siona case-markers sketched in Bruil (2014, §4.4).[2] The present dissertation identifies five primary DCM patterns in the language as shown in Table 1.1 — i.e., labeled the *principled* DCM alternations in Chapter 3. Alongside a brief description of each grammatical relation displaying DCM, the table below provides a relation label for each DCM pattern, used as a short-hand convention throughout this dissertation:

[2]Specific dimensions of Siona DCM patterns are expanded upon in subsequent work: e.g., Case and Jeretič 2021.

Grammatical relation label (used in this dissertation)	Traditional Latin label	Brief prose description
Subject (S)	*nominative*	The single argument of an intransitive clause and the Subject-argument of the (di)transitive clause (i.e., S=A).
Direct object (P)	*accusative*	The direct Object argument in the (di)transitive clause (i.e., P=T).
Indirect object (R)	*dative*	The indirect Object argument in the ditransitive clause. Other dative-like arguments are identified separately.
Static location (L)	*locative*	The static location implicated in the basic locative construction (X is (at) Y), and with other predicates which do not implicate motion.
Spatial goal (G)	*allative*	The dynamic location implicated in directed motion predicates (and change-of-posture predicates), i.e., GOAL.

Table 1.1: Short-hand labels for Siona DCM RELATIONS with brief description

Siona displays straightforward nominative-accusative alignment, like all Tukanoan languages do (Barnes 1999, 2006; Gomez-Imbert 2011; Stenzel 2013d). However, Siona is rather unusual in that it displays DCM on *all* core argumental grammatical relations (i.e., S, P, and R), and on certain relations in the spatial domain (i.e., L and G). A subset of these DCM patterns formally resemble the Sakha pattern shown in (2), where overt case-marking is not strictly required to encode a particular grammatical relation. For instance, the elicited minimal pair in (3) illustrates what might be called the optional *nominative*-marking pattern in Siona:

(3) a. *jaiye joro aide'oji iño*
hai-je **horo** ai + de'o-hi ihño
big-CL:GEN flower big + become-3S.M.PRS.ASS here
'Many flowers grow here.'

b. [**Context:** The speaker conveys that many kinds of flowers grow in this particular location, although other plants (trees, grass) do not.]
jaiye jorobi aide'oji iño
hai-je **horo #(-bi)** ai + de'o-hi ihño
big-CL:GEN flower-SBJ big + become-3S.M.PRS.ASS here
'Many FLOwers grow here (i.e., not something else).'
[VOL: 20230622ejabi001.012a-c]

Unlike the Sakha pattern, where case-marking is associated with a specific reading, overt case-marking in (3b) evokes an interpretative effect whereby the referent is contrasted with salient alternative referents in the context. This effect is related to information structure, where the speaker indicates which referents merit the listener's attention as the discussion or the narrative unfolds. This dissertation labels this the *emphatic contrast* reading, following similar effects described in the Tibeto-Burman optional ergative literature (e.g., Tournadre 1995; LaPolla 1995). A comparable system of case-marking is described for the Yalaku language of Papua New Guinea by Aikhenvald (2015), under the label 'highlighted participant case'. The emphatic contrast effect is observed across each of the DCM patterns enumerated in Table 1.1. Consider the spatial DCM pattern shown in (4) — zero-marking on the locative (L)

argument in (4a) produces a plain reading, whereas the overt *-re* marking in (4b) produces the emphatic contrast reading:

(4) a. *yë' yo'jei Canada ba'iji*
ji̵' jo'he-i̵ **Canada** ba'i-hi
1SG younger.sibling-CL:M C-N.SBJ live-3S.M.PRS.ASS
'My younger brother lives <u>in Canada</u>.'

b. [**Context:** If you pull out a map of Ecuador to show me where your brother lives, except he does not live in Ecuador.]
yë' yo'jei Canadare ba'iji
ji̵' jo'he-i̵ **Canada #(-re)** ba'i-hi
1SG younger.sibling-CL:M C-N.SBJ live-3S.M.PRS.ASS
'My younger brother lives <u>in CAnada</u> (i.e., not some other place).'
[VOL/SUG: 20230619elupa002.011a-b]

The minimal pairs in (3) and (4) demonstrate the information structure-driven usage of overt case-marking across various grammatical relations. This particular function of case-marking is not acknowledged in the previous literature, but is shown to be one of the primary drivers of Siona DCM alternations in this dissertation. This is not a particularly common trigger for DCM in the typological literature — see the discussions in Chapter 2 and Chapter 4.

In addition to the information structure-driven patterns of DCM demonstrated above, this dissertation argues for their interaction with an animacy-based SUBJECT vs. non-SUBJECT divide in the grammar. Put simply, animate non-SUBJECT noun phrases require overt case-marking, whereas inanimates do not. The result is demonstrated for DIRECT OBJECT-oriented DCM in (5) and (6). In the former case, concerning an animate DIRECT OBJECT, a different type of formal DCM is attested where two overt case-markers compete, producing comparable interpretative effects to those noted above. Zero-marking on animate nouns is judged as ungrammatical. Conversely, inanimate DIRECT OBJECTS, like that shown in (6), display the familiar zero- vs. overt case-marking alternation:

(5) a. *huajë gajeire co'esi' cayë*
wahi̵ + gahe-i̵ *(-re) ko'e-si'i kaa-ji̵
new + friend-CL:M-N.SBJ look_for-FUT.N3S say-N3S.PRS.ASS
'I want to look for <u>a new friend</u>.'

b. [**Context:** In a discussion about buying new animals for the house, the speaker would rather find a new friend.]
huajë gajeini co'esi' cayë
wahi̵ + gahe-i̵ *(-ni)/(#-re) ko'e-si'i kaa-ji̵
new + friend-CL:M-N.SBJ2 look_for-FUT.N3S say-N3S.PRS.ASS
'I want to look for <u>a new FRIEND</u>.' [VOL/SUG: 20230619elupa002.018a-c]

(6) a. *huajë jachohuë huerosi' cayë*
wahi̵ + hahcho-wi̵ wero-si'i kaa-ji̵
new + shoot-CL:CONT buy-FUT.N3S say-N3S.PRS.ASS
'I want to buy <u>a new rifle</u>.'

b. [**Context:** At the store, the salesperson tries to sell you an ax.]
huajë jachohuëre huerosi' cayë
wahɨ + hahcho-wɨ #(-re) wero-si'i kaa-jɨ
new + shoot-CL:CONT buy-FUT.N3S say-N3S.PRS.ASS
'I want to buy a new RIfle (i.e., not something else).'
[VOL/SUG: 20230619elupa002.020a-b]

The same *-re* vs. *-ni* alternation shown in (5) is found on Recipients, Experiencers, and Benefactee arguments. For this reason, this dissertation opts for the neutral labeling convention for these case-markers: *-re* (N.SBJ, as NON-SUBJECT) and *-ni* (N.SBJ2), following several Amazonianists (e.g., Aikhenvald 1999, 2007; Farmer 2015; or the related OBJECTIVE label in Zúñiga 2007, *et seq.*).[3] In fact, this dissertation recognizes instances of spatial *-ni*, which are not recognized in earlier descriptions. In the bulk of instances, *-ni* is associated with the emphatic contrast reading as anticipated. This is shown the naturalistic sentences in (7):

(7) a. *më'bi yë'ni ñañe bahuë*
mɨ'ɨ-bi **jɨ'ɨ-ni**$^{(P)}$ ñaa-ñe bãã-wɨ
2SG-SBJ 1SG-N.SBJ2 see-CL:GEN NEG.AUX-N3S.PST.ASS
'YOU didn't see ME (i.e., not vice versa).'
[NAT: 20140925salsu001.032]

b. *. . . si'ahua'i jaoni bateña*
si'a-wa'i **hã-o-ni**$^{(L)}$ ba-re-ña
all-PL.AN DEM.MED-CL:F-N.SBJ2 be-N2/3S.PST.N.ASS-REP
'. . . everybody was THERE (they say).' [NAT*: 20151023orocr001.112]

The present work addresses the rich set of DCM patterns found in Siona grammar by establishing the formal patterns at play, and the interpretative effects which are evoked by particular case-markers (or the lack thereof) in a given context. In previous work (e.g., Bruil 2014, §4.4), generalizations are put forth regarding the effect of other possible factors for certain DCM patterns — for example, the specificity-driven usage of *-re* on DIRECT OBJECTS, or ambiguity avoidance-based usage of SUBJECT *-bi* marking. On the assumption that several factors may conspire to account for the actual distribution of case-markers in Siona speech, these hypotheses are scrutinized in this dissertation.

1.2 Research questions and dissertation outline

With the aim of adequately characterizing Siona DCM patterns, and of fitting these patterns into the broader typology, this dissertation addresses the research questions laid out in (8):

(8) a. Which DCM patterns can be identified in Siona? What properties do these DCM patterns have in common, and how do they differ?

[3]Other Tukanoanists, including Bruil 2014, *et seq.*; Johnson and Levinsohn 1990; Stenzel 2008, 2013c; etc. utilize the label OBJ (similar to SBJ for the S-marker), which indicates the dominant marking function of this case-marker. See Stenzel 2008, p. 175, for a brief discussion concerning *-re* and the 'labeling issue'.

b. Which factors determine the selection of case marking alternatives in a given context — such as animacy status, focus, specificity? A combination thereof? How do the active factors interact to determine the DCM patterns in usage?

c. How does Siona DCM compare to other DCM systems described in the Tukanoan literature? In the northwestern Amazonia area? In the broader DCM typology?

In order to adequately address the questions in (8), a diverse methodological program is followed. This dissertation espouses the view that linguistic research is strengthened where diverse, corroborating evidence is made available (e.g., Polinsky 2010; Aihenvald 2018; Davidson 2020; Matthewson 2004, 2022 for recent discussions). I wear several linguistic hats, so to speak, in the different chapters of this dissertation, with the aim of avoiding 'methodological monotheism', following Matthewson (2022, 21), who acknowledges Farrell Ackerman for the term. A general outline of this dissertation is provided below.

Chapter 2 establishes the working definitions and terminology utilized in this work, and navigates certain dimensions of the robust DCM literature in the form of a state-of-the-art. Chapter 3 provides the *preliminary description* of Siona DCM, establishing the empirical facts and framing the various targeted studies presented in the subsequent chapters of this dissertation. Several novel contributions are made here — such as the abstract PLAIN-PROMINENT dichotomy, permitting a paradigmatic analysis, and the description of the noun class-shifting procedure of *promotion*, which forms the basis of the analysis of grammatical animacy classes. The baseline analysis of Siona DCM is established in the *preliminary description* is tested and expanded in the following three content chapters, conceived as independent studies.

For Chapter 4, I put on my *comparativist* hat. This chapter develops a typological study, which serves to situate the baseline Siona DCM patterns laid out in Chapter 3 among other (non-)Tukanoan languages spoken in the putative NWA area. The discussion in this chapter is built around a typological dataset − i.e., the NWA case-marking dataset, assembled in Appendix A. This study gathers various case-marking and DCM facts via the administration of a structural questionnaire (e.g., as found in Krasnoukhova 2012; Birchall 2014), against a modest sample languages spoken in the NWA area. As such, this chapter is not a component of the language-particular analysis developed later in this dissertation per se; instead it contextualizes these language-specific facts, and lays the groundwork for a diachronic analysis of Siona DCM. To date these dimensions of Siona DCM have not received adequate attention.

For Chapter 5, I put on my *variationist* hat. This chapter outlines a quantitative, corpus-driven investigation of DCM patterns in a modest manually-coded Siona narrative sample. Noun tokens are extracted from the sample, categorized on the basis of their case-marking, and are subsequently annotated for animacy, specificity, topicality, and a handful of other properties in order to determine which are significant predictors for Siona DCM patterns. Two types of quantitative data are considered, relating to the proportion of case-marking alternatives in the corpus sample and to the statistical analyses modeled against the extracted tokens (i.e., via the implementation of the *variable-rule* technique). This approach allows a more nuanced view of the DCM phenomenon in Siona, which had previously only been described qualitatively.

For Chapter 6, I put my *descriptivist* hat back on and complement the *preliminary description* by reporting a range of data, collected via the implementation of a

controlled elicitation-driven research program. In particular, the evidence presented in this chapter highlights the role of Siona DCM in a host of contexts traditionally associated with the expression of focus (cf. van der Wal 2015, 2016, etc.). This chapter refines the description of predictable question-answer case-matching patterns and (emphatic) contrastive uses of case-marking recognized in preceding chapters in the dissertation with data characterizing the nature of information structural DCM effects with a degree of precision lacking from earlier accounts.

Chapter 7 concludes this dissertation. On the one hand, this chapter unifies the diverse array of evidence presented in the preceding chapters and amends the *preliminary description* for Siona DCM accordingly. Additionally, this chapter enumerates a number of promising directions for how the description in this work may be furthered with future research, and for how these phenomena might be analyzed.

The remainder of this introductory chapter is organized as follows: Section 1.3 introduces the Siona language and the people who speak it. That section positions Ecuadorian Siona within the Tukanoan genealogy and within the Amazon area. Section 1.4 outlines the glossing and orthographical conventions utilized for representing Siona data in this dissertation.

1.3 The Siona people and their language

Ecuadorian Siona, or by its endonym *bãĩkohka* (the language of the people), is an endangered language spoken in the province of Sucumbíos in Ecuador. Today between two and three hundred individuals speak this variety with varying degrees of fluency (Bruil, p.c.; Mejeant 2001; Fabre 2005[4]). Bruil (2014, pp. 4-5) recognizes six small Siona communities in Ecuador, found along two converging river systems: Puerto Bolívar and Tarabëaya[5] are found in the Cuyabeno Wildlife Reservation along the Cuyabeno river to the north, and Orehuëaya, Aboquëhuira, Bi̱'a̱ña, and Sototsiaya dot the Aguarico river to the south. All of my own original fieldwork is collected in the community of Sototsiaya, supplemented by data collected online and during a four-month visit by Jamil Biaguaje to the University of Ottawa (September-December 2024). Certain other corpus materials scrutinized in this dissertation, collected by Martine Bruil, were gathered in the community of Puerto Bolívar. The map in Figure 1.1 identifies the two communities where the data for this dissertation was collected.

[4] *Diccionario etnolingüístico y guía bibliográfica de los pueblos indígenas sudamericanos* is freely consultable at http://www.ling.fi/DICCIONARIO.htm

[5] For these place-names, I utilize the native orthography conventions as used by community members. Section 1.4.3 lays out the two spelling conventions in detail.



Figure 1.1: Locations of the two Ecuadorian Siona communities where the data used here was obtained (adapted from Bruil and Stewart 2022)

There are two main motivations for which the community of Sototsiaya is well-suited as the primary site for data collection: Firstly, I am well-acquainted with several people and friends within the community, since first being introduced to them in 2018 by Martine. Secondly, this is the community where Siona enjoys its greatest degree of vitality, spoken by most community members and transmitted cross-generationally to the present day. This is in sharp contrast with other Siona communities where, in most cases, a shift to Spanish is all but complete.

1.3.1 Siona in its genealogical context

The Siona language belongs to the smaller Western branch of the Tukanoan language family. This is a medium-sized family comprising roughly two-dozen living languages, straddling the border regions between modern-day Ecuador, Colombia, Peru, and Brazil. The family is divided into two geographically-compact branches. While the members of the Western branch are strewn primarily within the Putumayo-Napo watershed, the Eastern branch is confined within the expansive Upper Río Negro (also called the Vaupés) cultural area. It is likely that the geographical distribution of each branch has remained relatively stable since before contact with European explorers and settlers (Chacon 2013, 2014). As concerns the Siona and related Western Tukanoan groups, united under the label *encabellados* (long-haired), it is well-

established that these were the dominant populations in and around these same fluvial zones in the early contact and evangelization contexts. Siona became elevated to the status of lingua franca in this area over the seventeenth- and eighteenth-centuries (cf. Wheeler 1970, 194-197; Vickers 1976, 1981; Cipolletti 1999; Carrera de la Red 2020, 2021).

Several internal classifications have been proposed for the Tukanoan family, as based upon comparative phonological and lexical datasets (e.g., Waltz and Wheeler 1972; Barnes 2006; Gomez-Imbert 2011). While the two-way division between Western and Eastern languages has been maintained, the more fine-grained classification is far from settled. There is now consensus that the putative Central Tukanoan branch, suggested in previous decades, ought to be abandoned (cf. Mason 1950; Waltz and Wheeler 1972; Malone 1987; Barnes 1999, 2006). Stenzel (2013d) provides a recent overview:

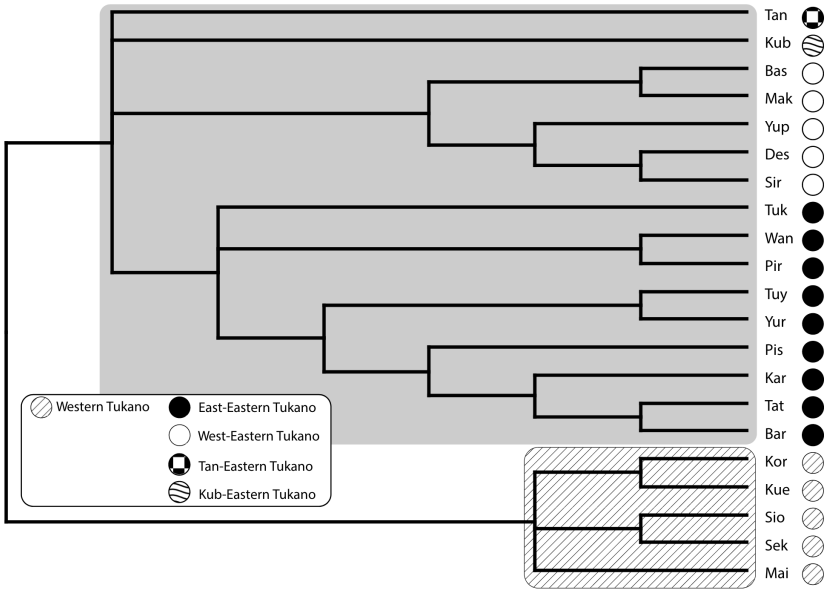


Figure 1.2: Proposed phylogenetic tree model for Tukanoan language family (Chacon and List 2015, 198, Figure 10)

Figure 1.2 reproduces the Tukanoan phylogenetic classification from Chacon and List (2015).[6] One can see that diversification is high in the Eastern branch, whereas the Western branch is more homogeneous. Setting aside the now-extinct language Kueretú (Kue), all specialists agree that Siona and Sekoya are intimately linked on the grounds of lexicostatistics and phonological innovations. A minor controversy exists as to

[6]The short-hand language names used in Figure 1.2 are as follows: Tan (Tanimuka), Kub (Kubeo), Bas (Barasano), Mak (Makuna), Yup (Yupua), Des (Desano), Sir (Siriano), Tuk (Tukano), Wan (Wanano), Pir (Piratapuyo), Tuy (Tuyuka), Yur (Yurutí), Pis (Pisamira), Kar (Karapana), Tat (Tatuyo), Bar (Bará), Kor (Koreguaje), Kue (Kueretú), Sio (Siona), Sek (Sekoya), Mai (Máíhɨ̄kì).

whether Koreguaje is more closely related to Siona-Sekoya varieties (cf. Wheeler 1992; Barnes 1999, 2006; Chacon 2014), or whether Máíhɨ̃kì is closer to these varieties and Koreguaje forms the out-group (Skilton 2013).[7] The former proposition, that Koreguaje is a closer phylogenetic relative to Siona than to Máíhɨ̃kì, is certainly borne out on the basis of the case-grammar facts laid out in this dissertation.

To elaborate slightly on this claim, this dissertation provides further support for claims regarding the development of the Western Tukanoan branch as follows. Based on attestation-type data in contemporary descriptive materials, Bruil (2018, 158-159) proposes that various components of the verbal paradigms can be reconstructed to three distinct stages. My schematization in Table 1.2 expands the original claims made by Bruil (2018):[8]

Shared innovation→	Proto-WT *Interrogative* s*-agree*	> Proto-Kor-Sio-Sek *Dependent verb paradigm*; ***Rich case system***	> Proto-Sio-Sek *Reportative -ña*
Siona (Sio)	✓	✓	✓
Sekoya (Sek)	✓	✓	✓
Koreguaje (Kor)	✓	✓	X
Máíhɨ̃kì (Mai)	✓	X	X

Table 1.2: Developmental chronology of Western Tukanoan (adapting Bruil 2018)

On the basis of this table, the attested DCM phenomena in Ecuadorian Siona analyzed in this dissertation ought to be reconstructed to a Proto-Koreguaje-Siona-Sekoya stage, after the initial split between the Western and Eastern branches. The set of cognate case-markers *-ni*, *-na*, and *-bi* could be reconstructed to the stage labeled *Rich case system* in Table 1.2. By extension, the Máíhɨ̃kì language has a strikingly simple case grammar compared to other Western Tukanoan languages − this comparison is fleshed out in Chapter 4.

On the basis of several shared innovations and a degree of mutual intelligibility (cf. Wheeler 1970; Vickers 1976), the synchronically-attested Siona and Sekoya varieties are generally conceived as a dialect continuum (Bruil 2014, 11-12). On the basis of case-marking facts, Koreguaje is remarkably similar to the Siona-Sekoya languages; although there are good grammatical and phonological reasons to exclude this language from the proposed continuum (Bruil 2014, 2018, 2019, *p.c.*).

An additional dimension is relevant to the establishment of Siona-Sekoya dialectology. A Peruvian variety of Sekoya is spoken in the vicinity of the Middle Putumayo, to the east of Ecuadorian Siona. This variety has begun to receive greater attention (e.g., Schwarz 2014; Vallejos and Schwarz 2016; Vallejos and Brown 2021), although a complete description of the case-marking facts is lacking at this time. The most crucial distinction between this variety and the others for the present purposes concerns the absence of the case-marker *-ni* (Rosa Vallejos, *p.c.*). This finding has implication for the diachronic development of the rich case-marking system found in the other Siona-Sekoya languages and in Koreguaje.

[7] Skilton (2013) bases her claims exclusively on phonological innovations.
[8] WT stands for 'Western Tukanoan'.

1.3.2 Siona in its areal context

Siona is spoken at the northwestern-most extremities of Amazonia, one of the most linguistically diverse areas in the world (Dixon and Aikhenvald 1999; Epps and Michael 2017; van Gijn and Muysken 2020). Traditionally, the top-level areal division in the study of South American languages opposes the highland (Andean) languages to those of the lowlands. In recent decades many specialists have sought criteria to identify smaller and more cohesive language areas in the lowlands (Muysken 2012; van Gijn 2014, 2017, 2020; van Gijn et al. 2023). Epps and Michael (2017) suggest the following micro-areas for lowland South America:[9]



Figure 1.3: Linguistic micro-areas across lowland South America (Epps and Michael 2017, Map 32.1)

Identifying areas is useful for delimiting domains for typological comparison. However, the Tukanoan family as a whole does not fall neatly within any one of the micro-areas proposed in the literature. On the one hand, Eastern Tukanoan languages are firmly established within the Vaupés culture-linguistic area (Aikhenvald 1999, 2001, 2003; Epps 2007; Stenzel 2008, 2013c; Zúñiga 2007; etc.). On the other hand, the status of

[9]Epps and Michael 2017 utilize the term *contact zones* for regions where lexical and grammatical traits are diffused (piggybacking on cultural diffusion). The suggested contact zones align neatly with the linguistic micro-areas presumed here.

the Western Tukanoan languages is less clear. They are not recognized in the putative Putumayo-Caquetá region in Epps and Michael (2017, §32.2.2), who enumerate the Witotoan family, Boran, Resígaro (Arawakan) and Andoke (isolate).[10] However, a plausible contact scenario between Máíhĩkì and Witotoan languages is proposed by Jolkesky (2016). Most Western Tukanoan languages are spoken further to the west of this area. Given such uncertainty regarding linguistic micro-areas, this dissertation uses a broad label, northwestern Amazonia (henceforth the NWA area), for the large geographical expanse between the Ecuadorian-Colombian-Peruvian foothills to the west and the Vaupés region to the east (following our convention in van Gijn et al. 2023).

Ecuadorian Siona in particular is spoken alongside several non-Tukanoan NWA languages within the Aguarico-Upper Napo fluvial zones − i.e., Shiwiar (Chicham), A'ingae (isolate), Wao Tededo (isolate, per Adelaar and Muysken 2004), and lowland Kichwa. Some preliminary evidence is suggestive of contact-induced lexical and grammatical patterns between Siona, Shiwiar and Kichwa (Bruil 2015a; Kohlberger 2020; Piispanen 2021; van Gijn et al. 2023), and perhaps A'ingae. Many questions remain as to patterns of contact-induced grammatical convergence, including DCM patterns, some of which are addressed in Chapter 4 of this dissertation. In any case, there is no question as to whether Siona and Sekoya communities have experienced cultural contact with Wao Tedede and A'ingae speakers for centuries (e.g., Vickers 1976; Wasserstrom 2014; de Carvalho 2022; Fischer and Hengeveld 2023).

1.4 The representation of Siona data

This section describes how Siona examples are reported in this dissertation. Section 1.4.1 considers how Siona examples are categorized on the basis of the data source and the technique employed for their collection. Then the glossing conventions followed in this dissertation are detailed in Section 1.4.2, followed by a discussion of orthographic conventions in Section 1.4.3.

1.4.1 A taxonomy of Siona example types

Siona examples used in this dissertation fall into two primary categories: NATURALISTIC data is extracted from free-flowing, semi-spontaneous Siona texts from Bruil (2012), and ELICITED data collected in the context of targeted interviews conducted with Siona speakers.

The NATURALISTIC examples are identified with an example-typing tag − i.e., NAT. This type tag precedes the metadata identifier formatted following the conventions of Bruil (2012, 2014, 2015b, 2018, 2019, *et seq.*). A subset of NAT examples reported in this dissertation are identified with an asterisk (i.e., NAT*), which indicates that the example is lifted from the narrative sample assembled for the quantitative analysis in Chapter 5, and Appendix B. (9) provides the full example identifiers for two text-based examples as an illustration:

(9) a. [NAT: 20100629slicr001.001]
b. [NAT*: 20151023orocr001.324]

[10]The putative Putumayo-Caquetá micro-area corresponds to the *People of the Centre* cultural zone (Echeverri 1997; Seifart 2005, 2007, 2011; Seifart et al. 2009; Wojtylak 2021).

The metadata tags shown in (9) are identical to those used to label the documentary raw data bundles in the ELAR database, i.e., where the original file bundles are archived.[11] Each identifier recognizes a number of details regarding the recording session as follows:[12]

- The date of data collection is recorded in YYYYMMDD format ;
- The session type identifier: (e) elicitation session, (s) staged session − including narrations, and (o) observed events ;
- The first two characters of first and last name of the (lead) consultant of the session − e.g., Ligia Criollo [licr] in (9a) ;
- Session number, for instance if there are multiple sessions with the same speaker on a given date ;
- Line number − based upon lines in transcription files provided in data bundles in Bruil 2012, as adapted upon normalization. E.g., (9b) is taken from the 324th line in the transcription of the narrative analyzed in Appendix B.

The same metadata conventions are utilized to label elicited examples. All of my elicited data is identified by the session type identifier (e); however, these data points are further compartmentalized into two sub-categories, inspired by Bochnak and Matthewson 2015. On the one hand, elicited data which is produced entirely by the speaker is labeled with the shorthand example-typing tag VOL (for 'volunteered' data). For instance, this tag is attributed to examples produced in the context of a translation task. Although the researcher provides a prompt in the communication language, i.e., Spanish, the speaker is responsible for producing the Siona sentence which forms the data point under consideration.

On the other hand, some examples considered in this dissertation involve a Siona sentence produced by the researcher, which is subsequently subjected to judgment or to manipulation by the speaker.[13] This latter type of sentential data is labeled as SUG (i.e., 'suggested' by the researcher). Such suggested-type examples are important for the presentation of negative evidence and data collected in highly-controlled contexts. Nonetheless volunteered and naturalistic examples are used wherever possible. Some instances, particularly in Chapter 6, concern volunteered sentences which are manipulated in the course of elicitation and the adapted sentence is suggested for judgment. Such instances are identified via the combined tag VOL/SUG in recognition of the split role of the consultant with respect to the production of the data in question. (10) provides example identifiers for elicited data points:

(10) a. [VOL: 20240918ejabi001.007b]
b. [SUG: 20190618emapa002.012]

[11] The full documentary dataset is found in the *Endangered Language Archive*: Bruil, Martine. 2012. "Documentation of Ecuadorian Siona". *Endangered Languages Archive*. Handle: http://hdl.handle.net/2196/00-0000-0000-000D-EA53-3.

[12] More details regarding the metadata labeling conventions is outlined in Bruil 2014, 14-17.

[13] The type of judgment, and degree of introspection involved, is a well-studied component of linguistic methodology (Chomsky 1961; Schütze 1996; Davidson 2020; and many others): e.g., grammaticality judgment, truth-conditional judgments, felicity judgment. Other types of introspective judgments provide insight into subtleties, and may fall within either the VOL or SUG category based upon the role of the consultant in gathering the concerned data point.

1.4.2 Glossing conventions: Reporting parsed Siona data

The glossing conventions followed in the present dissertation mostly replicate those found in Bruil (2014, 2018, 2019; and elsewhere). These conventions align with the best practices in the standard Leipzig Glossing Rules.[14] The tiers used in Siona glossed examples are as in (11):

(11) (A) ⇒ [**Context:** . . .]
(B) ⇒ *Transcription . . .*
(C) ⇒ Morpheme-by-morpheme parse: e.g., x-y **z**
(D) ⇒ Interlinear gloss word-affix target_word
(E) ⇒ Suggested English translation '. . . ' [(F) ⇒ Example identifier]

The Context-tier (A) is optional. In the case of certain naturalistic examples, this tier is utilized to contextualize the example in the surrounding text. In certain volunteered examples, this tier reports the suggested contexts in context-conjuring tasks (van der Wal 2016). Otherwise this tier reports speaker comments where they inform the discussion at hand.

The original transcribed sentence in (B) is presented in a legible format for the Siona public − i.e., the native orthography introduced in Section 1.4.3 below. No capitalization or punctuation conventions are used in this tier, following the format used in Bruil (2014, Appendix I). Regarding naturalistic data, this tier lifts the sentence(s) as they appear in the documentary materials. This transcription connects to its idiomatic English translation provided in tier (E), which is written following typical English punctuation conventions.

The interlinear gloss comprises morphologically-parsed Siona items, presented in *linguistic orthography* in tier (C), and the corresponding glossing labels are presented in tier (D). The linguistic orthography is also used to refer to Siona items in-text. The abbreviations used for glossing in tier (D) are presented in the list of abbreviations in the front matter of this dissertation.[15] No quotations or punctuation are found in tiers (C) or (D). (12) illustrates a fully-glossed example from the Appendix B. The material which is most relevant to the argumentation, typically nominal (sub-)constituents, is presented in **boldface** in tier (C) for ease of reference (e.g., the nominal *tsia-**re*** 'egg' in the example).

(12) [**Context:** The wife of Baina, the mythological folk hero of the Sionas, has a final meal before laying the eggs which will contain his children.]
. . . anichaojëna tsiare baoña io
ãni + chao-hi-na **tsia-re** baa-o-ña ĩ-o
eat + finish-PL.PRS.DEP-DS egg-N.SBJ have-2/3S.F.PST.N.ASS-REP 3PRO-CL:F
'. . . once (they) finished eating, she had eggs.'
[NAT*: 20150811sfryi001.112]

Turning now to the suggested English translation in tier (E), presented between single quotes '. . . ', the element corresponding to the **boldface** Siona item(s) is underlined.

[14] Guidelines accessible at https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf

[15] Following LGR conventions, < - > represents affixal concatenation, < = > represents cliticization, while head-head sequences (e.g., compounding) are indicated via < + > between individual elements.

Where relevant, quotations are presented, "...", and all English punctuation conventions are followed. Material which is either covert or tacitly understood is presented in parentheses to form an idiomatic English translation of the Siona original.

The role of DCM for encoding contrast is explored at various points in this dissertation. Where appropriate, focus is presented in the suggested English translation tier (E) as SMALL CAPS, which emulates the prosodic stress placement, encoding focus (or the emphatic contrast reading) in naturalistic English, based upon my own native speaker intuitions. The elicited example in (13) demonstrates how focal status is reported in this dissertation:[16]

(13) [**Context:** A boy comes home to find coffee and tea on the table.]
yë'ë ocore ụcusi'i cayë
jɨ'ɨ **ohko #(-re)** ũhku-si'-i kaa-jɨ
1SG water-N.SBJ drink-FUT-N3S say-N3S.PRS.ASS
'I want to drink WAter.' [VOL/SUG: 20230623ejabi001.024-025]

In addition to demonstrating the presentation of focus, (13) reports a linguistic judgment. These judgments are strictly appropriate for elicited data points, whereas naturalistic data is presumed to be grammatical and pragmatically felicitous in all cases.[17] This dissertation primarily deals with two types of linguistic violations: grammatical violations, indicated with an asterisk (*), and felicity violations, indicated with a hash symbol (#). The usage of these violation symbols aligns with the dominant conventions found in the literature.

1.4.3 Orthography matters: Presenting Siona data

There are two orthographical systems for recording Siona: (i) the native orthography, and (ii) the linguistic orthography. Siona speakers are able to read and write in the former system, but are generally unfamiliar with the latter. All preliminary documentary materials are transcribed using the native orthography. Following Bruil (2014, *et seq.*), this dissertation utilizes both orthographies, and opts for the linguistic orthography for in-text reference to Siona words or phrases.

The native orthography (originally *practical orthography*) was developed by American missionaries, Orville Johnson and his wife María Johnson, from the Summer Institute of Linguistics (SIL). This writing system was developed in the late 1950's among speakers of Siona's sister language, Sekoya. These activities were sanctioned by the regional governments across adjacent regions of Ecuador, Colombia and Peru in order to promote evangelization and assimilation among the indigenous communities of the area (cf. Vickers 1976, 1981). This writing system is found in bible translations, children's tales, work-related documentation, and other texts to the present day.

The Spanish influence of the native orthography is evident: e.g., /k/ is written as either <c> when preceding a vowel {/a/,/o/,/u/}, or their nasal counterparts, whereas it is written as <qu> before front vowels {/i/,/e/} (or before the high-mid vowel /ɨ/, written <ë>). In addition to this sixth basic vowel quality, the Johnsons

[16] Note that (13) is built around a (consultant) volunteered sentence, presented with follow-up judgment regarding the absence of *-re*; hence this example is typed as VOL/SUG accordingly.

[17] This dissertation is not concerned with false starts, self-corrections, or other dynamics of naturalistic data which may be integrated into the discussion in future research. Rather, naturalistic data are treated as a window into actual usage patterns, which are relevant to the description of certain aspects of Siona DCM in this work.

introduce the glottal stop, written as an apostrophe <'>, and the sub-litteral bar to represent the nasal quality of the vowel: e.g., <a̲> represents /ã/ (see Johnson and Levinsohn 1990, 23-24). Beyond these additions, the native orthography mirrors spelling and punctuation conventions of Spanish.

Conversely, the linguistic orthography is developed by Bruil (2014) to accomplish a more transparent phonetic representation of the language, favouring IPA conventions. This writing system is used in academic work reporting Siona data. Table 1.3 is replicated from (Bruil, 2014, 132), which maps correspondences between the native and linguistic orthography, and provides the underlying phonemic realization. I direct the reader to Bruil (2014, §3), and to Bruil and Stewart (2022), for a more elaborate discussion of various phonological and phonetic phenomena in Ecuadorian Siona:[18,19]

[18]One minor discrepancy regarding the usage of the linguistic orthography here, compared with Bruil 2014, concerns the representation of <j> and <ñ> allophony. Whereas Bruil collapses these realizations as underlying [j], this dissertation opts for a more transparent representation, without taking a theoretical stance on the issue. For instance, regarding the generic classifier, *-je* and (nasal) *-ñe* are distinct if unified in the glosses (i.e., CL:GEN).

[19]Bruil (2014 and others) suggest that the phonetic realization of the voicing contrast in the language has to do with creaky phonation. This is a crucial factor which distinguishes Ecuadorian Siona from Ecuadorian Sekoya phonology. Wiegertjes (2020) and Bruil and Stewart (2022) provides some initial phonetic analysis of this phenomenon.

Phoneme	Native orthography	Linguistic orthography
[p]	<p>	<p>
[p̰]	<b>	<b>
[t̰]	<t>	<t>
[t̰]	<d>/ <r>	<d>
[k]	<c>/ <qu>	<k>
[k̰]	<g>/ <gü>	<g>
[k^{w}]	<cu>	<kw>
[k̰^{w}]	<gu>/ <gü>	<gw>
[ʔ]	<’ >	<’ >
[s]	<s>	<s>
[s̰] / [ts]	<ts>	<z>
[ʧ]	<ch>	<ch>
[h]	<j>	<h>
*[h] (hC)	C	hC
[m]	<m>	<m>
[n]	<n>	<n>
[ɲ]	<ñ>	<ñ>
[w]	<hu>	<w>
[j]	<y>	<j>
[i]	<i>	<i>
[ĩ]	<i̱>	<ĩ>
[ɨ]	<ë>	<ɨ>
[ɨ̃]	<ë̱>	<ɨ̃>
[u]	<u>	<u>
[ũ]	<u̱>	<ũ>
[e]	<e>	<e>
[ẽ]	<e̱>	<ẽ>
[o]	<o>	<o>
[õ]	<o̱>	<õ>
[a]	<a>	<a>
[ã]	<a̱>	<ã>

Table 1.3: Overview of Siona orthographies - adpated from Bruil 2014, p. 132

*The linguistic orthography recognizes preaspiration arising before the plosive onset at certain syllable boundaries (Bruil 2014, 103-106; van ’t Veer et al. 2025), which are not represented in the native orthography.

CHAPTER 2

Differential case-marking in the typological record

The present chapter presents an overview of the literature pertaining to DCM, situating the Siona facts with respect to that literature, and establishing the conceptual backdrop for the analysis developed for Siona DCM here.

In the typological tradition, both language-specific and cross-linguistic investigations regarding DCM patterns are taken to be contributions to a single research program (Klein and de Swart 2011, 4; *contra* Carnie 2006; Haspelmath 2007):

The CHARACTERIZATION of language-specific patterns
⇳
The FORMULATION of cross-linguistic generalizations
⇳
The EXPLANATION of these cross-linguistic patterns

Section 2.2 outlines how the description of DCM patterns is conceived as comprising separate domains of inquiry that come together to achieve a 'complete' language-particular description and to facilitate typological comparisons.

Before moving to that, I begin by laying out a working definition of DCM.

2.1 DCM: A working definition

DCM refers to a range of phenomena whereby nominals of the same grammatical relation (i.e., CASE FUNCTION) are case-marked in different ways. For instance, the Sakha minimal pair in (1) juxtaposes two grammatically licit sentences. In (1a), the DIRECT OBJECT displays *ni*-marking, whereas the corresponding OBJECT-argument in (1b) is zero-marked:

(1) Sakha (*sah* — TURKIC) [adapted from Baker 2015, 1-4, exx. (1c),(5)]

a. *Erel* ***kinige-ni*** *atyylas-ta*
E.NOM book-ACC buy-PST.3SS
'Erel bought the book.'

b. *Erel* ***kinige*** *atyylas-ta*
E.NOM book buy-PST.3SS
'Erel bought a book/books.'

While both *ni-* and zero-marking are grammatically licit in the language, the choice of one or the other marking alternative has interpretative implications. Particularly, the selection of the overt case-marker, *-ni*, evokes a specific or definite reading of the OBJECT-argument, which is lacking where zero-marking is selected. Setting aside semantico-pragmatic factors underpinning case-marker selection until the following section, (2) provides a strictly formal definition for DCM, assumed here:

(2) **Differential case marking** (DCM — *working formal definition*):
The phenomenon whereby not every argument bearing a given grammatical relation is encoded with the same case-marker.
[adapted from Witzlack-Makarevich and Seržant 2018, 3, ex. (1)]

2.2 Ingredients for a holistic description of DCM

It is well-established that DCM patterns are commonplace cross-linguistically (e.g., Bossong 1985; Aissen 1999, 2003; de Swart 2007; Iemmolo 2010, 2013; Sinnemäki 2014; Haspelmath 2019, 2021). Over the decades, the DCM literature has been enriched by descriptions of languages from diverse areas of the globe. In spite of a degree of morphological and semantico-pragmatic overlap across DCM cross-linguistically, the phenomenon is far from uniform, leading to conflicting definitions.

Modern DCM theory aims to account for points of variation across DCM languages, and their implications for a unified account of this phenomenon. Rufilanchas (2014, 515) identifies three major dimensions of variation in DCM. An adequate description of DCM phenomena in a given language would need to address each of these dimensions:

(3) Definitions of [DCM] disagree in at least three aspects:

a. "The [...] morphosyntactic phenomena [involved]" ;
[i.e., the *formal* dimension — Section 2.2.1]

b. "The syntactic functions and semantic roles [subject to] the phenomenon" ; [i.e., the *functional* dimension — Section 2.2.2]

c. "The dimensions that determine [DCM] and the relevant transitivity parameters." [i.e., TRIGGERS for DCM alternations — Section 2.2.3]

This section explores each of these dimensions in turn, bringing forward data from the descriptive and typological literature, and discussing the implications for the analysis of Siona DCM. The compartmentalized discussion achieves a more refined typology than is usually seen in general descriptions.

2.2.1 The formal dimension: A typology of DCM patterns

DCM, as defined here, hinges on the presence of case-marking on at least certain nominals of a given grammatical relation. However, DCM is part of a larger family of grammatical phenomena where the coding of arguments is not strictly homogeneous for all grammatical relations. Complementary to DCM, i.e., a variable dependent-marking strategy, is its head-marking counterpart, e.g., *differential argument agreement* (DAA).[1] Swahili DAA (recently discussed in Mursell 2018) is exemplified in (4), where the presence of OBJECT-agreement morphology on the verb is correlated with a definite or topical reading of the corresponding argument:[2]

(4) Swahili (*swh* — BANTU, NORTHEAST COASTAL)
[Givón 1976, 159, cited in Coghill 2014, 338, ex. (2)]

a. *ni-li-**ki**-soma* ***kitabu***
1SG-PST-OBJ-read book
'I read the book.'

b. *ni-li-soma* ***kitabu***
1SG-PST-read book
'I read a book.'

A related phenomenon concerns covariation of so-called *clitic-doubling* and case-marking, as demonstrated in the Romanian sentences in (5). This pattern represents a combination of both dependent- and head-marking patterns and for most authors is treated as a DCM sub-phenomenon:

(5) Romanian (*ron* — INDO-EUROPEAN, ROMANCE)
[Chiriacescu and von Heusinger 2010, 301-304, exx. (3), (7a)]

a. ***l**=am* *vâzut* ***pe Ion*** / ***pe Donald Duck***
3SG.M.ACC=have seen OBJ John OBJ DD
'I have seen John / Donald Duck.'

b. *caut* ***secretarâ***
look_for secretary
'I am looking for a secretary (specific/non-specific).'

There is clear overlap between DAA and *clitic doubling* strategies and DCM, given that each phenomenon implicates an alternation between a more and a less complex marking alternative. The differences are primarily morphological in nature, and solid theoretical motivations exist for collapsing these phenomena (e.g., Chappell and Verstraete 2019, §2.3). This dissertation limits its empirical scope to pure case-marking strategies — i.e., DCM in a strict sense.

[1] Just as alternative, role-specific, labels for DCM are found in the literature — e.g., DIFFERENTIAL OBJECT MARKING (DOM), a label such as DIFFERENTIAL OBJECT AGREEMENT (DOA) is commonly found. However, like DCM, these patterns are not limited to coding the OBJECT-argument (cf. Haspelmath 2005; Coghill 2014).

[2] Similar DAA phenomena are noted across the Bantu family, and are driven by various TRIGGERS that largely align with those described for DCM in **Section 2.2.3** — e.g., Dimmendaal 2010; Zeller 2014; van der Wal 2022.

There exist several formal types of DCM patterns. Following terminology in Kittilä (2005; adopted in Chappell and Verstraete 2019),[3] an ALTERNATING-type DCM pattern is one where more than one case-marker may be selected to encode a given grammatical relation. In these alternations, no zero-marking is attested. OBJECT-oriented ALTERNATING DCM is demonstrated for Evenki in (6), and A'ingae in (7):

(6) Evenki (*evn* — TUNGUSIC)
[Nedjalkov 1997, 193, cited in Iemmolo 2013, 385, exx. (9a),(10a)]

a. ***oron-mo*** *java-kal*
reindeer-DEF.ACC take-PRS.IMP.2SG
'Catch <u>that reindeer</u>.'

b. ***oron-o*** *java-kal*
reindeer-INDEF.ACC take-PRS.IMP.2SG
'Catch yourself <u>a/any reindeer</u>.'

(7) A'ingae (*con* — *isolate*) [Fischer and Hengeveld 2023, 93, exx. (95),(97)]

a. ***matichi=ve=ta=ti=ki*** *in'jan='fa*
machete=ACC.IRR=NEW.TOP=INT=2 want=SBJ.PL
'Do you want <u>machetes</u>?'

b. *sumbu-en=jan* ***ain-fa='u=ma***
emerge-CAUS-IMP dog-CL:LAT=AUG=ACC.REAL
'Get <u>the dog</u> out.'

In contrast to these ALTERNATING-type DCM patterns are OPTIONAL-type DCM systems, where an overt case-marker alternates with zero-marking (Chappell and Verstraete 2019). The OPTIONAL-type is far more common cross-linguistically (Iemmolo 2013), and not restricted to any particular geographical area. In fact, many traditional definitions for DCM exclusively capture DCM of this formal type (e.g., Bossong 1985; Aissen 2003; Nichols and Bickel 2013; Sinnemäki 2014). A classic example of the OPTIONAL DCM pattern for Turkish OBJECT-marking is demonstrated in (8):

(8) Turkish (*tur* — TURKIC)
[von Heusinger and Kornfilt 2005, 5, cited in Sinnemäki 2014, 282, ex. (1)]

a. *Ben* ***bir kitab-ı*** *oku-du-m*
I a book-ACC read-PST-1SG
'I read <u>a certain book</u>.'

b. *Ben* ***bir kitab*** *oku-du-m*
I a book read-PST-1SG
'I read <u>a book</u>.'

There exists another possible formal DCM type, which has not been recognized previously in the literature. This dissertation proposes the label HYBRID-type DCM for the pattern which combines aspects of the two above-mentioned formal types: i.e., where zero-marking alternates with multiple overt case-markers. This pattern is relevant for

[3] An alternative terminology is utilized by de Hoop and Malchukov (2007; also Iemmolo 2013, borrowed from Dixon 1994): *symmetrical* case-marking refers to multiple overt forms in alternation, whereas *asymmetrical* case-marking refers to an alternation between an overt form and zero-marking.

the description of Siona DCM, as shown in (9) for P-oriented DCM:[4] The DIRECT OBJECT may be marked with *-ni* or with *-re*, or be zero-marked, based upon the context and the animacy class of the argument:

(9) Ecuadorian Siona (*snn* — WESTERN-TUKANOAN)
[adapted from Case & Bruil, *in progress, ms*]

a. *se̲'seni jachohuë cahuë*
sẽ'se-ni hahcho-wɨ̃ kaa-wɨ̃
peccary-N.SBJ2 shoot-N3S.PST.ASS say-N3S.PST.ASS
'I wanted to hunt a PECcary.' [NAT: 20151030oagyi001.289]

b. *. . . bacure a̲i̲ñë*
bahku-re ãĩ-ñɨ̃
pacu-N.SBJ eat-N3S.PRS.ASS
'. . . (they) are eating Pacu (fish species).'
[NAT: 20151030oagyi002.366]

c. *huë'e neni bëahuë*
wɨ'e nee-ni bɨa-wɨ̃
house make-SS stay-N3S.PST.ASS
'(I) made a hut and stayed (there).' [NAT: 20151030oagyi001.019]

I make no a priori assumptions regarding the relationship between a particular formal type and other, non-formal, properties of the DCM system (*contra*, e.g., de Hoop and Malchukov 2007; Iemmolo 2013; Chappell and Verstraete 2019; Witzlack-Makarevich and Seržant 2018). Furthermore, languages that display multiple DCM patterns may exhibit a different formal type in each pattern. This is shown to be the case for Siona in this dissertation.

2.2.2 The functional dimension: DCM across grammatical relations

DCM patterns may be attested across a range of grammatical relations in a given language. Beyond core arguments, DCM also appears in the typological literature in the encoding of spatiotemporal arguments and a handful of oblique relations.

The earliest- and best-studied DCM patterns pertain to OBJECT-marking,[5] as exemplified in (4) through (9) in the preceding section. This is the most robust DCM pattern cross-linguistically, found in every linguistic area globally (Sinnemäki 2014).

SUBJECT-oriented DCM patterns are also readily attested (de Hoop and de Swart 2008; McGregor 2010; Chappell and Verstraete 2019, 15-17), as demonstrated in the Umpithamu examples in (10),[6] and for Azhee in (11):

[4] In Case and Jeretič 2021, we label this the 'tripartite *differential object marking* system'.

[5] The present discussion uses full words to label grammatical functions (e.g., SUBJECT, GOAL); elsewhere, the short-hand labels suggested in Chapter 1 (i.e., S, P, R, L, G) are used.

[6] The Umpithamu examples in (10) exemplify the *optional ergativity* phenomenon described for languages in Australia, New Guinea, and various Tibeto-Burman languages. These are treated as a sub-type of SUBJECT-oriented DCM for the present purposes.

(10) Umpithamu (*ump* — PAMA-NYUNGAN)
[Chappell and Verstraete 2019, 2, ex. (1)]

a. ***waypala-mpal*** *maarra-n=antyangku* *motoka-nti*
whitefella-ERG take-PST=1PL.EXCL.ACC car-COM
'The whitefella took us in the car.'

b. ***waypala*** *maarra-n=antyangku*
whitefella take-PST=1PL.EXCL.ACC
'The whitefella took us.'

(11) Azhee (*yiz* — SINO-TIBETAN) [adapted from Gerner 2016, 143, exx. (7),(9a)]

a. [***lu***^{33} ***ho***^{21} **(**la***^{55})] [*go*^{33} *mo*^{33}] *tie*^{21} *bə*^{55} *wa*^{55}
hail DSM wheat hit collapse DP
'The hail destroyed the wheat.'

b. [***ŋo***^{21} (*?***la***^{55})] *ɕi*^{55} *tɕ*^{h}*e*^{33} *to*^{21} *li*^{33}
1SG DSM clothes raise rise come
'I took up the clothes.'

Shifting away from core arguments, recent work has shed light on other DCM patterns. One such pattern concerns DIFFERENTIAL RECIPIENT MARKING, referring to case-marking on the the INDIRECT OBJECT (Haspelmath 2005; Kittilä 2008).[7] (12) exemplifies animacy-driven, ALTERNATING DCM on the INDIRECT OBJECT in Finnish:

(12) Finnish (*fin* — FINNO-UGRIC) [Kittilä 2008, 256, ex. (13)]

a. *lähetti* *lähett-i* *lähettime-n* ***lähettäjä-lle***
messenger.NOM send-3SG.PST transmitter-ACC sender-ALL
'A/the messenger sent a transmitter to the sender.'

b. *lähetti* *lähett-i* *lähettime-n* ***lähetystö-ön***
messenger.NOM send-3SG.PST transmitter-ACC embassy-ILL
'A/the messenger sent a transmitter to the embassy.'

DCM patterns on other oblique arguments are typologically rare. For instance, OPTIONAL-type DCM is attested for INSTRUMENT-marking in the Western Tukanoan language, Koreguaje (Cook and Levinsohn 1985, 97), and a similar pattern is described in Tariana (Aikhenvald 2003, 154-155).

DCM phenomena are even noted in the domain of spatial case-marking (Stolz et al. 2014; Baker 2015). (13) displays an instance of OPTIONAL DCM on the static LOCATION-argument in Yidiny, and (14) shows ALTERNATING DCM in Krongo on the spatial GOAL of the directed motion predicate, distinguishing spatial nouns from human landmarks:

[7] Note that these Kittilä 2008 utilizes the term DIFFERENTIAL GOAL MARKING to refer to INDIRECT OBJECT-oriented DCM. The label GOAL is employed in this dissertation to refer to the target location of a directed motion predicate, and I eschew the usage of this term here in order to avoid confusion.

(13) Yidiny (*yii* — PAMA-NYUNGAN)
[Dixon 1977, in de Hoop and Malchukov 2008, 37]

a. *nundu djana:-n* ***naru walba:-nda***
2SG stood top stone-DAT
'It was a stone you stood on top of.'
or, 'Oh, you stood on top of the stone!'

b. *ŋayu djana:-n* ***naru walba:-***
1SG stood top stone
'I stood on top of a stone.'

(14) Krongo (*kgo* — KADUGLI-KRONGO)
[Reh 1985, 285, cited in Haspelmath 2019, 320, ex. (18)]

a. *ɲáaw* ***kú-fúunì***
he_goes LOC-street
'And he goes to the street.'

b. *n-yáaw àʔàŋ* ***àtʊ́ná ìʔìŋ***
I_go I to him
'I go to him.'

As regards temporal DCM, a handful of patterns are attested in a few disparate languages (Dabir-Moghaddam 1992; Seržant 2016; Witzlack-Makarevich and Seržant 2018, 1-4); however, a typological account of this phenomenon is currently lacking.

Recent work demonstrates that spatial DCM is in fact cross-linguistically common (Aristar 1997; Stolz et al. 2014; Haspelmath 2019). These authors find that, of the three universal spatial relations, the likelihood of zero-marking decreases from left to right as follows: GOAL > STATIC LOCATION > SOURCE. Chapter 4 in this dissertation refines the discussion regarding spatial DCM patterns within the NWA area, where Siona is spoken, which has not received significant attention to date (Roosvall 2020).

One language, multiple DCM patterns

I stated above that in languages where DCM effects are attested across multiple grammatical relations, each of these DCM patterns needs to be characterized separately. Perhaps the best-described multiple-DCM language in the literature is Hindi, where both SUBJECT- and OBJECT-oriented, OPTIONAL DCM effects are attested, as shown in (15):

(15) Hindi (*hin* — INDO-ARYAN) [Cann and Miljan 2012, 587-588, ex. (9)-(11)]

a. ***Ilaa=ne***$^{(\text{SBJ})}$ ***bacce=ko***$^{(\text{OBJ})}$ (**baccaa*) *uTaayaa*
I=ERG child-OBJ (child.NOM) lift.PERF
'Ila$^{(\text{SBJ})}$ lifted a/the child$^{(\text{OBJ})}$.'

b. ***Nadja***$^{(\text{SBJ})}$ ***kitab=ko***$^{(\text{OBJ})}$ *xarid-e-g-i*
N book=OBJ buy-3SG-FUT-F.SG
'Nadja$^{(\text{SBJ})}$ will buy a particular/the book$^{(\text{OBJ})}$.'

c. ***Nadja***$^{(\text{SBJ})}$ ***kitab***$^{(\text{OBJ})}$ *xarid-e-g-i*
N book buy-3SG-FUT-F.SG
'Nadja$^{(\text{SBJ})}$ will buy a/the book$^{(\text{OBJ})}$.'

Simplifying the facts considerably, marking of the transitive SUBJECT with the ergative marker *=ne*, is available in clauses where the verb is in the perfective aspect, e.g., (15a). In imperfective clauses the Subject is always zero-marked.[8] On the other hand, in OBJECT-oriented DCM, the dative-marker *-ko* arises on the DIRECT OBJECT under appropriate conditions, e.g., (15b), based upon referential properties of the noun phrase, including animacy and specificity status. These examples show that each core grammatical relation exhibits a unique DCM pattern, which merits description in its own right.

Several multiple-DCM languages are described in the NWA area (Aikhenvald 2015, 241), including Murui (Wojtylak 2021), Tariana (Aikhenvald 2003), and Siona. In these languages, DCM effects are noted in the encoding of core grammatical relations, including certain spatial relations.[9] Particularly regarding Siona, this dissertation analyzes principled DCM patterns across five different grammatical relations: (i) SUBJECT (S), (ii) DIRECT OBJECT (P), (iii) INDIRECT OBJECT (R), (iv) LOCATION (L), and (v) spatial GOAL (G), in addition to a few more minor DCM patterns discussed in the *preliminary description* laid out in the ensuing chapter. Although there is a degree of overlap across a subset of these patterns, each DCM pattern displays unique properties. Individual DCM patterns may vary on the basis of the formal dimension, discussed in Section 2.2.1, and/or based upon the factors governing the selection from the set of available marking alternatives in a given context of usage. These factors, which I call DCM TRIGGERS here, are surveyed in the following section.

2.2.3 The semantico-pragmatics of DCM

The case-marking alternations observed in DCM are not patterns of "free variation". Rather, various factors inform the selection of one or another alternative. This is by far the best-explored dimension of DCM in the literature, particularly in the functional-typological vein (Moravcsik 1978; Comrie 1989; Bossong 1985; Aissen 1999, 2003 — see Haspelmath 2021 for a recent overview). This dissertation employs the term TRIGGER to refer to any factor governing marking within a given DCM pattern.[10] I begin with an overview of the DCM triggers described in the typological literature. Section 2.2.3.1 introduces the notion of TRIGGER STRENGTH, and discusses its implications for the study of multidimensional DCM. Section 2.2.3.2 briefly considers the role of DCM as a disambiguation device.

At the highest level, a given DCM pattern may be governed by properties of the predicate, of the clause, or by properties of argument itself (Witzlack-Makarevich and Seržant 2018, §2; Haspelmath 2021, §8). Predicate- and clause-level TRIGGERS are not as commonly attested in the typology. The Hindi SUBJECT-oriented DCM pattern shown in (15) represents the pattern of DCM driven by clausal aspect, i.e., TAME-triggered DCM, as surveyed in Malchukov and de Hoop (2011). A related pattern

[8] In Hindi perfective clauses, ergative-marking displays an OPTIONAL-type DCM alternation where overt ergative *=ne* indicates the degree of volitionality or agentivity of the SUBJECT, as discussed in de Hoop and Malchukov 2007, and references therein.

[9] Contra Aikhenvald (2015, 241), who does not recognize spatial DCM in the languages of this region.

[10] Note that I do not follow certain authors in distinguishing between factors which TRIGGER and those which RESULT (from) the presence of a particular alternative, typically overt marking (de Swart 2007; de Hoop and Malchukov 2007; Klein and de Swart 2011), even if the only way to "encode" this meaning is via DCM. My motivation for eschewing this distinction is developed further in Chapter 6.

is demonstrated for the NWA language A'ingae in (16). The appropriate accusative case-marker is selected for the OBJECT-argument based upon the reality status of the clause:[11]

(16) A'ingae (*con — isolate*) [Fischer and Hengeveld 2023, 93, exx. (95),(97)]

a. *matichi**=ve**=ta=ti=ki* *in'jan='fa*
machete=ACC.IRR=NEW.TOP=INT=2 want=SBJ.PL
'Do you want machetes?'

b. *sumbu-en=jan* *ain-fa='u**=ma***
emerge-CAUS-IMP dog-CL:LAT=AUG=ACC.REAL
'Get the dog out.'

Another family of DCM patterns triggered by clause-level properties has garnered attention in the literature. These concern what I label thematic DCM TRIGGERS. Regarding OBJECT-oriented DCM, much literature reports that the degree of 'affectedness' factors into case-marking patterns (Næss 2004; von Heusinger and Kaiser 2011; Kizilkaya et al. 2022). A typical example is displayed for Finnish in (17). A thematic TRIGGER hinging on the degree of 'agentivity' is described for some SUBJECT-oriented DCM (de Hoop and Malchukov 2008; *contra* Fauconnier 2011; Fauconnier and Verstraete 2014), shown for Manipuri in (18):

(17) Finnish (*fin* — FINNO-UGRIC) [adapted from Næss 2004, 1203, ex. (11)]

a. *hän jo-i* ***maido-n***
s/he drink-PST.3SG milk-ACC
'S/he drank (all) the milk.'

b. *hän jo-i* ***maito-a***
s/he drink-PST.3SG milk-PART
'S/he drank (some of the) milk.'

(18) Manipuri (*mni* — SINO-TIBETAN) [de Hoop and Malchukov 2008, 571, exx. (16), (17)]

a. ***əy-nə*** *tebəl-də* thenɳi
I-ERG table-LOC touched
'I touched the table (volitionally).'

b. ***əy*** *tebəl-də* thenɳi
I table-LOC touched
'I touched the table (involuntarily).'

DCM patterns which are triggered by properties of the nominal itself, or its referent, are more common typologically (Sinnemäki 2014; Witzlack-Makarevich and Seržant 2018, §2.1; Chappell and Verstraete 2019, §2.3.1). There is a long history of analyzing these coding patterns by appealing to so-called referential prominence scales in the functionalist literature (Durant 1979; Comrie 1989; Silverstein 1986; Aissen 1999,

[11] The A'ingae pattern in (16) is reminiscent of Slavic *genitive-of-negation* constructions (e.g., Dunn and Khairov 2009, §3.3.3, regarding Russian; Bielec 2015, 92-98, for related Polish facts), which, per the definitions in this dissertation, represent predicate-triggered DCM.

2003; etc.).[12] There are three dominant argument-level triggers for DCM identified in this literature: (i) animacy status, (ii) referentiality status, and (iii) information structure (e.g., topicality status) (Haspelmath 2021). Each of these is better conceived as a trigger family (i.e., corresponding to a scale), since they collapse several implicationally related categories, which may or may not be relevant on a language-particular basis (Aissen 1999, 2003): e.g., some languages assign a given case to all animate nouns, whereas others restrict this pattern to human-denoting nominals. Following terminology in Sinnemäki (2014), these broad categories are conceived as MACRO-TRIGGERS, which facilitate the comparison of DCM patterns on a typological basis. Below each MACRO-TRIGGER is discussed in turn.

Animacy status is a common trigger for DCM patterns, particularly concerning OBJECT-marking (Bossong 1984, 1985; Aissen 2003; Sinnemäki 2014; etc.). Unlike specificity or topicality, which are generally context-sensitive properties of a given noun phrase, animacy status is an inherent property of the referent. (19) demonstrates the OBJECT-oriented DCM pattern found in Kannada, where the accusative marker, *-vannu*, is obligatory on animates, and variable on inanimate nouns. The Azhee SUBJECT-oriented DCM examples in (11) displays a similar pattern:

(19) Kannada (*kan* — SOUTH-DRAVIDIAN)
[Lidz 2006, 11, cited in de Swart 2007, 178-179]

a. *naanu* ***sekretari *(-yannu)*** *huDuk-utt-idd-eene*
1SG.NOM secretary-ACC look_for-N.PST-be-1SG
'I am looking for a secretary.'

b. *naanu* ***pustaka(-vannu)*** *huDuk-utt-idd-eene*
1SG.NOM book-ACC look_for-N.PST-be-1SG
'I am looking for a book.'

The MACRO-TRIGGER of referentiality status, following the traditional terminology in Bossong (1984), collapses the following non-inherent nominal categories: specificity, definiteness, and (pro)nominal status. The distinctive marking of pronouns and full nouns is fairly common cross-linguistically (Aissen 2003, 451-454; Haspelmath 2021, §4.1.3), as demonstrated for OBJECT-marking in the Australian language, Pitjantjatjara, in (20). (21) displays the Hebrew pattern of OBJECT-oriented DCM, namely where the distribution of the case-marker prefix, *'et-*, co-occurs with the definite article *ha-*, irrespective of noun class (Danon 2006, 2010; Hacohen et al. 2021):

(20) Pitjantjatjara (*pjt* — PAMA-NYUNGAN)
[Bowe 1990, cited in Aissen 2003, 452, exx. (23),(24a)]

a. *tjitji-ngku* ***Billy-nya*** / ***ngayu-nya*** *nya-ngu*
child-ERG B-ACC / 1SG-ACC see-PST
'The child saw Billy/me.'

b. *Billy-lu* ***tjitji*** *nya-ngu*
B-ERG child see-PST
'Billy saw the child.'

[12] Several variations on the same basic theme (enumerated in Schmidtke-Bode and Levshina 2018, 510-511): e.g., *individuation scale* (Lazard 1998), (*extended*) *animacy hierarchy* (Comrie 1981, Croft 2003), *empathy hierarchy* (DeLancey 1981), *indexability hierarchy* (Bickel 1999), *relevance hierarchy* (Sibatani 2006). This list neglects countless language-particular adaptations used in descriptive works (Haspelmath 2015).

(21) Modern Hebrew (*heb* — West-Semitic) [Aissen 2003, 453, ex. (25)]

a. *ha-seret her'a* ***'et-ha-milxama***
the-movie showed acc-the-war
'The movie showed <u>the war</u>.'

b. *ha-seret her'a* ***(*'et-)milxama***
the-movie showed acc-war
'The movie showed <u>a war</u>.'

Another well-attested context-sensitive Macro-trigger for dcm patterns pertains to the information-structural properties of the argument. Pragmatically-conditioned dcm has garnered significant attention in the more recent literature (e.g., Iemmolo 2010; Dalrymple and Nikolaeva 2011; von Heusinger and Schumacher 2019; Haspelmath 2021; Irimia and Mardale 2023),[13] and is now considered to be one of the dominant cross-linguistic dcm patterns in its own right. (22) displays topic-driven Optional dcm on the Object-argument in Persian. An instance of focus-driven Optional dcm on the Subject-argument in Tibetan is demonstrated in (23):

(22) Persian (*pes* — Indo-Aryan)
[Dalrymple and Nikolaeva 2011, 108-112, cited in Haspelmath 2021, 135, ex. (19)]

a. *man* ***sib*-i (*-râ)** *xarid-am*
1sg apple-indef-acc eat.pst-1sg
'I ate <u>an apple</u>.'

b. *ki* ***mašin-i *(-râ)*** *did*
who car-indef-acc see.pst.3sg
'Who saw <u>a car</u> (referred to previously in discourse)?'

(23) Central Tibetan (*bod* — Sino-Tibetan, Bodic)
[adapted from Tournadre 1995, 264, cited in Haspelmath 2021, 138, ex. (25)]

a. ***khōng*** *khāla' so-kiyo:re'*
he good make-ipfv.gnom
'<u>He</u> prepares the meals.'

b. ***khōng-ki'*** *khāla' so-kiyo:re'*
he-erg good make-ipfv.gnom
'<u>He</u> (i.e., not somebody else) prepares the meals.'

In summary, the bulk of dcm patterns described in the typological literature are driven by one of three argument-level Macro-triggers — i.e., the inherent animacy status of the referent, or the context-sensitive referential or pragmatic status of the argument. Besides these Macro-triggers, a handful of less common, argument-level dcm triggers are attested in the literature — e.g., number (Slavic languages (Witzlack-Makarevich and Seržant 2018, 7), Palauan (Levin 2019), P'urépecha (Maldonado 2011), Hup (Epps 2008)); and person (Abruzzese (D'Alessandro 2017);

[13] Aissen (2003, 436, footnote 2) explicitly acknowledges that topicality is often described as factoring into dcm, particularly differential object marking patterns, although she (like many others) underplays the relevance of this trigger relative to animacy- and specificity-related triggers. Although similar patterns are well-attested in the Bantu tradition — see van der Wal 2022 for recent discussion, information structure has only just begun to receive more serious attention in the dcm literature in particular.

Yindjibarndi (Wordick 1982)). The presence of these disparate TRIGGERS, along with the predicate-level triggers outlined above, highlights the importance of language-particular descriptions to further our understanding of the range of factors which drive DCM patterns cross-linguistically.

2.2.3.1 Multi-dimensionality: Ranking DCM triggers

On the one hand, it must be established *which* TRIGGERS are at play in a particular language. An equally important task is to determine *to what extent* a particular TRIGGER factors into the selection among the available alternatives in a given DCM pattern. The remainder of this section discusses two interrelated notions which are necessary to adequately describe the active factors for a particular pattern: (i) TRIGGER STRENGTH, and (ii) multi-dimensionality.

TRIGGER STRENGTH is related to the fine-grained vision of DCM triggers developed by de Hoop and Malchukov (2007, and furthered in Klein and de Swart 2011, etc.): SPLIT-type DCM effects vs. FLUID-type DCM effects.[14] The former type implicates a STRONG TRIGGER in my terminology, where one category of nominal obligatorily takes a particular marking alternative and another does not. As a case in point, consider the OPTIONAL DCM pattern displayed for Malayalam OBJECT-marking in (24). The case-marker *-e* is obligatory on the human-denoting noun in (24a), but blocked altogether on the inanimate noun in (24b):

(24) Malayalam (*mal* — SOUTH-DRAVIDIAN) [adapted from Asher and Kumari 1997, 57-59]

a. *avan* ***kuṭṭiy-e*** *aṭiccu*
3SG.M child-OBJ beat-PST
'He beat the child.'

b. *avan* ***pustakam*** ***(*-e)*** *vaayiccu*
3SG.M book-OBJ read-PST
'He read the book.'

Malayalam OBJECT-oriented DCM displays an animacy-related STRONG TRIGGER, such that failure to adhere to the marking pattern displayed in (24) results in ungrammaticality. On the other hand, WEAK TRIGGERS do not have a deterministic effect on case-marking patterns, but rather reflect marking tendencies. (25) exemplifies how specificity is a WEAK TRIGGER in Mongolian OBJECT-oriented DCM patterns. As discussed extensively in Guntsetseg 2016, §4, overt case-marking on the OBJECT favours a specific reading, as in (25a), yet both specific and non-specific readings are available where the OBJECT is zero-marked, as shown in (25b):

(25) Mongolian (*mon* — MONGOLIC) [Klein and de Swart 2011, 9, exx. (21)-(22)]

a. *Bold* ***neg ohin-ig*** *unssen*
B a girl-ACC kissed
'Bold kissed a certain girl.'

b. *Bold* ***neg ohin*** *unssen*
B a girl kissed
'Bold kissed a (certain) girl.'

[14] This terminology is adopted from earlier work regarding transitivity splits in ergative systems (cf. Dixon 1994).

The key difference regarding TRIGGER STRENGTH between (24) and (25) lies in the fact that in the latter example the overt case-marking on the nominal is not automatic given the presence of the trigger (e.g., specificity, which can be present also in the zero-marked variant). For several authors, this is a definitional property of FLUID DCM patterns (de Hoop and Malchukov 2007; Klein and de Swart 2011); I choose to talk of WEAK TRIGGERS because while the TRIGGER itself may have a strong or weak effect on the case-marking patterns, a given DCM pattern may be driven by a conspiracy of several STRONG and WEAK TRIGGERS; this is a phenomenon that is generally called 'multi-dimensional DCM' (de Hoop and Malchukov 2007; de Swart 2007; Klein and de Swart 2011), to which I now turn.

It has long been recognized that more than one TRIGGER may be active for a given DCM system (e.g., Bossong 1984, 1985; Aissen 2003), and that the way that these TRIGGERS interact with one another has implications for the distribution of marking alternatives. I briefly outline the OBJECT-oriented DCM patterns attested in Spanish and in Hindi, both of which are driven by animacy- and specificity-related MACRO-TRIGGERS, even if the status of each corresponding TRIGGER is different in each language, affecting the distribution of DCM alternatives in practice.

In broad strokes, in Spanish most OBJECT nominals are left zero-marked. However, the preposition *a* is required on a subset of human-denoting nouns, such that an OPTIONAL-type DCM pattern arises (von Heusinger and Kaiser 2003; Leonetti 2004).[15] Definite human-denoting OBJECTS are obligatorily marked with the preposition, as shown in (26a); although other non-human definite OBJECTS reject this marker. (26b) demonstrates how the preposition is unavailable with an inanimate definite OBJECT, as expected:

(26) Spanish (*spa* — INDO-EUROPEAN, ROMANCE)
[Brugè and Brugger 1996, 3]

a. *esta mañana he visto* *(***a**) ***Juan*** / *(***a**) ***la hermana de María***
this morning have.1SG seen OBJ J / OBJ the sister of M
'This morning I saw John / Maria's sister.'

b. *esta mañana he visto* (*)***a*** ***la nueva iglesia***
this morning have.1SG seen OBJ the new church
'This morning I saw the new church.'

The distribution of the preposition *a* is more complicated where non-definite, human-denoting OBJECTS are concerned. Leonetti (2004, 80) explains that DCM optionally arises with human OBJECTS with the indefinite article, as in the case of (27a), where the specificity status of the doctor is ambiguous. The preposition is rejected altogether where a bare nominal is used, excluding a specific interpretation, as shown in (27b):

(27) a. *busca* (***a***) ***un médico***
look_for.3SG OBJ a doctor
'(S)he is looking for a (particular) doctor.'

b. *necesitan* (*)***a*** ***camarero***
need.3SG OBJ water
'They need a waiter.' [adapted from Leonetti 2004, 80, exx. (2b)-(2c)]

[15] There is some dialectal variation as to the distribution of *a*-marking on animal-denoting nouns (cf. Leonetti 2004, 2008; Malchukov 2008).

Summarizing these facts, Klein and de Swart (2011, 7-9) argue that the language displays two STRONG TRIGGERS, namely, humanness and definiteness, whereas specificity acts as a WEAK TRIGGER. Elsewhere in the literature, a handful of other WEAK TRIGGERS are described as factoring into OBJECT-oriented DCM in the language: e.g., topicalization effects (von Heusinger and Schumacher 2019), affectedness effects (Leonetti 2004; von Heusinger 2008; 2011) — see Fábregas 2013 for an overview. On this basis, the questions of which TRIGGERS are active, and the relative TRIGGER STRENGTH exhibited by each, both merit separate attention.

The distribution of overt case-marking observed in OBJECT-oriented DCM in Hindi is strikingly different from the Spanish pattern outlined above, although the same primary MACRO-TRIGGERS are implicated in both languages. The basic Hindi pattern is displayed in (28). Simplifying considerably, all human-denoting nominals must bear the dative case-marker *-ko*, as shown in (28a): i.e., like Spanish, humanness is a STRONG TRIGGER for this Hindi pattern. Furthermore, this case-marker is available to animal-denoting and inanimate OBJECTS, though its usage typically evokes a definite interpretation, shown in (28b). Nonetheless, (28c) demonstrates that zero-marked inanimate OBJECTS permit both a definite and an indefinite interpretation (Monahan 1990). On this basis, definiteness is a WEAK TRIGGER for Hindi DCM:

(28) Hindi (*hin* — INDO-ARYAN)
[Monahan 1990, cited in Klein and de Swart 2011, 6, exx. (1)-(4)]

a. *Ilaa-ne* ***bacce****(***-ko***) *uthayaa*
I-ERG child-DAT lift.PF
'Ila lifted a/the child.'

b. *Ilaa-ne* ***haar-ko*** *uthayaa*
I-ERG necklace-DAT lift.PF
'Ila lifted the necklace.'

c. *Ilaa-ne* ***haar*** *uthayaa*
I-ERG necklace lift.PF
'Ila lifted a/the necklace.'

Although both animacy- and referentiality-related MACRO-TRIGGERS are active in Hindi and in Spanish, their relative TRIGGER STRENGTH yields different intersective effects in each language. With this we get a glimpse of the complexities of intersective trigger effects, which are characteristic of multi-dimensional DCM patterns.

2.2.3.2 Ambiguity avoidance in the broader DCM picture

Besides semantico-pragmatic TRIGGERS for DCM, this section recognizes a final driver for case-marking alternations: i.e., ambiguity avoidance. In many languages overt case-marking serves to maximize communicative efficiency (Kurumada and Jaeger 2015). This disambiguating function of DCM is well-established in the literature (Aissen 1999, 2003; de Swart 2007; de Hoop and Malchukov 2008; Malchukov 2008; Klein and de Swart 2011). However, many authors explicitly call for this function to be separated from the so-called 'indexing function', which comprises the semantically- and pragmatically-motivated TRIGGERS for DCM patterns outlined above. Following the logic of these authors, this dissertation assigns a special status to the disambiguation-related uses of DCM alternations.

An informative case-study is provided in (29) for Yongren Lolo, whose OBJECT-oriented DCM pattern is strictly motivated by ambiguity avoidance (Gerner 2008). The author describes how the relative ordering of core constituents is free and that a particle, $\boldsymbol{t^hie^{21}}$, may be placed after the OBJECT-argument in naturalistic speech in order to facilitate the parsing of the sentence. Contrast the unambiguous sentence in (29a) with the freely ordered, zero-marked variants in (29b), where ambiguity results:

(29) Yongren Lolo (*lol* — TIBETO-BURMAN)

[adapted from Gerner 2008, 299, ex. (1)]

a. ni^{33} $ʈs^ho^{21}$ p^ho^{21} lo^{33} go^{21} $ʑi^{33}$: $_{(i)}\boldsymbol{ŋo^{33}}$ $[_{(ii)}\boldsymbol{ɕɛ^{33}}$ $\boldsymbol{mo^{33}}$
2PL NUM:6 FG.CL:father+daughter return go 1SG snake
$\boldsymbol{t^hie^{21}}]$ $ʈs^hɔ^{33}$ $ʑi^{33}$ ne^{21} be^{33}
OBJ follow go like_that say
'(The youngest) said: "You six, (please) go back, $\underline{\text{I}}_{(i)}$ will follow $\underline{\text{the snake}}_{(ii)}$".'

b. i. ... $_{(i)}\boldsymbol{ŋo^{33}}$ $[_{(ii)}\boldsymbol{ɕɛ^{33}}$ $\boldsymbol{mo^{33}}]$ $ʈs^hɔ^{33}$ $ʑi^{33}$
1SG snake follow go
'$\underline{\text{I}}_{(i)}$ follow $\underline{\text{the snake}}_{(ii)}$.' OR '$\underline{\text{The snake}}_{(ii)}$ follows $\underline{\text{me}}_{(i)}$.'

ii. ... $[_{(ii)}\boldsymbol{ɕɛ^{33}}$ $\boldsymbol{mo^{33}}]$ $_{(i)}\boldsymbol{ŋo^{33}}$ $ʈs^hɔ^{33}$ $ʑi^{33}$
snake 1SG follow go
'$\underline{\text{I}}_{(i)}$ follow $\underline{\text{the snake}}_{(ii)}$.' OR '$\underline{\text{The snake}}_{(ii)}$ follows $\underline{\text{me}}_{(i)}$.'

In languages like Yongren Lolo, ambiguity avoidance is a dominant factor for DCM. More languages of this type are discussed by de Swart (2007, §3). Nonetheless, in the bulk of DCM languages these effects play a less significant role compared with the TRIGGER-driven alternations which concern this work. As for Siona, I take it as uncontroversial that Siona case-marking is sometimes recruited to maximize communicative efficiency and that this accounts for some of the observed variation, as is already suggested by Bruil (2014, 160) for SUBJECT-oriented DCM. However, this dissertation largely sets aside these matters for future research.

Taking stock: DCM in the typological record

This chapter has provided a cursory summary of the various dimensions of DCM, several of which are often conflated or overlooked in the broader literature. Table 2.1 schematizes the various types of DCM recognized in the typological record on the basis of formal, functional and semantico-pragmatic properties, as addressed in this chapter. This table serves as a reference for the remainder of this dissertation:

Formal dimension	
Alternating type	Two overt case-markers are grammatical to encode a particular grammatical relation (CASE$_1$ vs. CASE$_2$)
Optional type	A single case-marker may be omitted to encode a particular grammatical relation without affecting grammaticality (CASE$_1$ vs. ∅)
Hybrid type	Multiple case-markers alternate to encode a grammatical relation, but case-marking may also be absent (CASE$_1$ vs. CASE$_2$ vs. ∅)
Functional dimension	
Grammatical relations/thematic roles displaying one of the above formal DCM profiles.	
Argumental DCM	Subject, Direct object, Indirect object
Spatiotemporal DCM	Static location, Goal, Source, Path, Temporal expressions, etc.
Oblique DCM	Instrument, Companion, etc.
'Triggers' — The factors driving DCM patterns	
Multi-dimensional DCM has several active TRIGGERS, which may conspire with one another.	
Animacy status	Humanness, animate vs. inanimate (context-insensitive)
Referentiality status	Pronominality, specificity, definiteness (context-sensitive)
Discourse status	Topicality, focus, contrastivity, etc. (context-sensitive)
Predicate-level triggers	Agentivity, affectedness, etc.
Other triggers	Word order, embeddedness/clause type, ambiguity avoidance, etc.
*Trigger strength	Strong trigger: A subset of noun phrases receive a particular alternative, others do not ;
	Weak trigger: A subset of noun phrases are more likely to receive a particular alternative.

Table 2.1: Overview: DCM types for each descriptive dimension

CHAPTER 3

Siona case grammar: A preliminary description

This chapter provides the *preliminary description* of Siona case-marking facts, which is intended as a reference and as the point of departure for the ensuing chapters in this dissertation. Although the essential facts are as in previous descriptions (i.e., Bruil 2014, §4.4; Case and Jeretič 2021; Case & Bruil *forthcoming*), this preliminary description has certain features that innovate with respect to them, namely the unification of argumental and spatial DCM patterns, which is retained in the remainder of this work. I also make a handful of refinements based on data assembled from my field notes and work with texts.

Siona grammar makes extensive use of case-markers to indicate the role of a given noun phrase in its clausal context. Table 3.1 adapts the case-marker inventory as presented by Bruil (2014, 157). The usage of these markers is not always straightforward. DCM effects are attested in the coding of various grammatical relations. This table indicates the baseline *formal* and *functional* aspects of Siona DCM, and also provides a first approximation of the dominant TRIGGERS for each of the attested DCM patterns.

Case-marker	Grammatical relation	Use
-bi	Subject	focus
	Instrument	obligatory
	Source location	
-re	Direct object	specific object
	Indirect object	
	Location	
-ni	Direct object	focus
	Indirect object	
-na	Goal	specific goal
-hã'ã	Path Limit	obligatory
-hã're	Comitative	obligatory

Table 3.1: The case suffixes in Ecuadorian Siona, their functions and their use, adapted from Bruil (2014, 157, Table 4.2)

There are a few important things to note in order to frame the description. Firstly, even though in the table they are lumped, this dissertation separates argumental DCM patterns from spatial DCM patterns, in order to determine how they align and how they differ. Spatial relations are identified with grayed cells in Table 3.1.

Secondly, the table above traces the distribution of particular case-markers to one or several grammatical relations. DCM arises where the *use* of the case-marker to encode that relation is conditioned by a certain (set of) factor(s) — i.e., not obligatory. Relations that are obligatorily marked by a certain case do not exhibit DCM. Thus, two (argumental) grammatical relations in the table do not exhibit DCM. This includes the COMPANION FUNCTION (labeled Comitative in Table 3.1), encoded via the dedicated case-marker *-hã're*, and the INSTRUMENT FUNCTION, which recruits the *bi*-marker.[1] The following examples illustrate this.

(1) a. *. . . jë̲jë̲ më'ë̲bi yëto më'ja̲'re basi'i*
hɨ̃hɨ̃ mɨ'ɨ̃-bi jɨɨ-to **mɨ'ɨ̃-hã're** bah-si'-i
yes 2SG-SBJ want-COND 2SG-COM be-FUT-N3S
'. . . Yes, if you want, (I) will be (there) with you.'
[NAT*: 20150207srocr003.045]

b. *yë'ëje̲ juaquëa'ë tobë̲a̲re ñocuabi*
jɨ'ɨ-hẽ hwaa-kɨ-a-'ɨ to-bɨ-ã-re
1SG-ADD weave-CL:M-COP-N3S bag-CL:RND-PL.INAN-N.SBJ
ñohkwa-bi
chambira_palm-INST
'I (M) also often weave bags (lit. I am a bag-weaver)
with chambira (fibres).' [VOL: 20230628emabi001.004]

For other grammatical relations identified in Table 3.1, the use of the corresponding case-marker(s) is determined by either 'focus' or 'specificity'. Following the terminology established in Chapter 2, the use of certain markers is driven by either a

[1] Cook and Levinsohn 1985, 97-100, explain that the cognate *-pi* is not obligatory to encode the INSTRUMENT argument in Koreguaje, such that an OPTIONAL-type DCM pattern arises.

discourse- or referentiality-related TRIGGER. This dissertation will demonstrate that multiple TRIGGERS are active in each grammatical relation with DCM. It also claims that a discourse-related STRONG TRIGGER is active for every Siona DCM RELATION, regardless of which other TRIGGERS are active.

Thirdly, if we adopt the definition for DCM established in Chapter 2 — i.e., from (2): "The phenomenon whereby not every argument bearing a given grammatical relation is encoded with the same case-marker", two types of DCM are attested in Siona grammar.

On the one hand, there is *principled* DCM, which are primarily triggered by a conspiracy of animacy-related and information structural properties of the noun phrase. Principled DCM in Siona grammar are best viewed as pairs of marking alternatives, labeled the PLAIN and PROMINENT marking-alternatives, which display a paradigmatic distribution. Table 3.2 maps case-marking alternatives onto PLAIN (often zero-marking) or PROMINENT categories across the five principled DCM RELATIONS of Siona, as based upon animacy classes. Similar analyses are suggested for related Western Tukanoan languages — including, Colombian Siona (Wheeler 1970, 172-174, 1987) and Koreguaje (Cook and Levinsohn 1985, 92):[2]

	INANIMATE		ANIMATE ($\wedge$ $^{\uparrow}$INAN)	
	PLAIN	PROMINENT	PLAIN	PROMINENT
Argumental DCM:				
SUBJECT (S)	-∅	*-bi*	-∅	*-bi*
DIRECT OBJECT (P)	-∅	*-re*	*-re*	*-ni*
INDIRECT OBJECT (R)	*-re*/*-na*		*-re*	*-ni*
Spatial DCM:	INANIMATE		$^{\uparrow}$INANIMATE	
STATIC LOCATION (L)	-∅	*-re*	*-re*	*-ni*
GOAL (G)	-∅	*-na*	*-na*	

Table 3.2: Principled DCM RELATIONS: mapping onto PLAIN-PROMINENT forms

Principled DCM alternations in Siona tend to be driven by information structure. Consider the interpretative difference between the PLAIN zero-marked SUBJECT and its PROMINENT *-bi* marked counterpart in the minimal pair presented in (2). The overt case-marking in (2b) evokes what is called the emphatic contrast reading in this dissertation, following terminology in the optional ergative-marking literature (e.g., Tournadre 1995; LaPolla 1995). Similar effects are noted for PROMINENT case-marking across the entire paradigm traced in Table 3.2:

(2) a. *yë'ë gajeo yëco saiye*
jɨ'ɨ gahe-o jɨɨ-ko sai-je
1SG friend-CL:F want-3S.F.PRS.ASS go-CL:GEN
'My friend (F) wants to go.'

[2]There are fundamental points of micro-variation in the grammars of these Western Tukanoan languages, which makes it so that the proposed paradigms are incompatible with other languages. Some of these differences are discussed in detail in Chapter 4.

b. *yë'ë gajeobi yëco saiye*
jɨ'ɨ gahe-o-bi jɨɨ-ko sai-je
1SG friend-CL:F-SBJ want-3S.F.PRS.ASS go-CL:GEN
'My FRIEND (F) wants to go.'
[**Speaker comment:** This sounds like you are saying that my friend wants to go, but I do not.] [VOL/SUG: 20220623emipa002.009-012]

In addition to principled DCM, on the other hand, Siona also displays what I label *accidental* DCM. This term simply refers to any instance where more than one marking alternative is accepted to encode a particular grammatical relation, but which are not organized as discourse-driven PLAIN-PROMINENT pairs, as in (2). Accidental DCM refers to a heterogeneous class of patterns, which are not unified by a single TRIGGER. (3) demonstrates an instance where the use of either *-re* or *-na* is accepted on the spatial argument, but the interpretative effect is clearly not based upon information structure. To my mind, the flavour here is more akin to preposition selection in a language such as English:

(3) a. *iye mo'se iye ye'yahuë'ere ye'yeja'quëa'ë*
i-je mo'se **i-je** **je'ja + wɨ'e-re**
DEM.PROX-CL:GEN day DEM.PROX-CL:GEN teach + house-N.SBJ
je'je-hã'-kɨ-a'-ɨ
learn-PRP-CL:M-COP-N3S
'Today (I (M)) will study at the school.'
[**Speaker comment:** with *re*-marking a more specific location is signaled; i.e., the learning occurred in a particular room.]

b. *iye mo'se iye ye'yahuë'ena ye'yeja'quëa'ë*
i-je mo'se **i-je** **je'ja + wɨ'e-na**
DEM.PROX-CL:GEN day DEM.PROX-CL:GEN teach + house-GOAL
je'je-hã'-kɨ-a'-ɨ
learn-PRP-CL:M-COP-N3S
'Today (I (M)) will study at the school.'
[**Speaker comment:** with *na*-marking a more general location is signaled; i.e., not any particular room.]
[VOL/SUG: 20230619elupa002.016a-b]

This chapter describes both principled and accidental DCM patterns in order to provide a complete description of case-marking in Siona. Preliminary insights are provided regarding the set of TRIGGERS which drive each (type of) DCM alternation in the grammar. Whereas the majority of this dissertation focuses on analyzing the principled DCM type, targeted discussions regarding particular accidental patterns will resurface, especially in the context of the typological discussion in the next chapter.

The preliminary description in this chapter is organized as follows: Section 3.1 describes DCM patterns found with the SUBJECT-argument (S). This is followed by a description of non-SUBJECT DCM patterns in Section 3.2, beginning with argumental grammatical relations: i.e., DIRECT OBJECT (P) in Section 3.2.1, with the INDIRECT OBJECT (R) in Section 3.2.3, and certain extraneous argumental DCM patterns — i.e., as found with the EXPERIENCER, or the BENEFACTEE, are discussed in Section 3.2.4. The concept of 'promotion', introduced in Section 3.2.2, accounts for certain inanimate nouns which exceptionally pattern with animates. The recognition of these

seemingly exceptional nouns, represented by ↑ in Table 3.2, permits the extension of the noun class split, otherwise driven by semantic animacy classes, to account for similar marking splits across spatial DCM patterns. Section 3.3 lays out a description of spatial DCM. In Section 3.3.1, the STATIC LOCATION-argument (L) is discussed. Finally, the spatial GOAL-oriented DCM pattern is explored in Section 3.3.2. This chapter concludes with an interim summary in Section 3.4, motivating the targeted studies presented in the following chapters.

3.1 Subject-oriented DCM patterns

Siona exhibits a straightforward nominative-accusative alignment system. The grammatical SUBJECT (S) in any clause will exhibit the following properties:

(4) a. S is the only argument which governs verbal agreement.
b. S is the only argument which controls switch reference and dependent topical constructions.
c. S arises either with zero-marking (PLAIN alternative), or with the case-marker *-bi* (PROMINENT alternative).

Besides the two S-specific head-marking properties noted in (4a) and (4b); (4c) defines the FORMAL alternation characterizing Siona S-oriented DCM. Following the terminology in Chapter 2, this is an OPTIONAL-type DCM pattern, where PLAIN, zero-marking alternates with the overt PROMINENT case-marking alternative, *-bi*. The remainder of this section focuses on the TRIGGERS, which are found to drive the competition between these two marking-alternatives.

The PLAIN, zero-marked form arises in the bulk of instances where an overt SUBJECT-argument is present.[3] As pointed out by Bruil (2014, 158-159), S-marking is independent of the semantic role of the noun phrase. For instance, both naturalistic sentences in (5) represent the general case, where the S-argument is unmarked — i.e., neither the inanimate, Undergoer SUBJECT in (5a), nor the animate, Agent SUBJECT in (5b), receive PROMINENT *-bi* marking in either context:

(5) a. *jaiye joro aide'oji i̱ño*
hai-je **horo** ai + de'o-hi ĩhñ̃o
big-CL:GEN flower big + become-3S.M.PRS.ASS here
'Many flowers grow here.' [VOL: 20230623ejabi001.001]

b. *. . . i̱o tsia go'ose'e a̱coña*
ĩ-o tsia + go'o-se'e ãh-ko-ña
3PRO-CL:F egg + bunch-EXCL eat-2/3S.F.PST.N.ASS-REP
'. . . she ate the bunch of eggs, they say.'
[NAT: 20150811sfryi001.130]

PROMINENT *bi*-marking of subjects tends to occur in contexts of focus.[4] More specifically, one could evoke contrast, that is, prominent marking implicates a set of salient

[3] Per the naturalistic corpus-based analysis reported in Appendix B, and discussed in Chapter 5 of this dissertation, roughly one-in-five overt S-arguments receive nonzero PROMINENT-marking.

[4] See Cook and Levinsohn 1985 for discussion of similar effects in the related language Koreguaje.

alternative referents and then excludes those alternatives from the interpretation, as is claimed for optional ergative in Tibetan (e.g., Tournadre 1995, LaPolla 1995). This is illustrated in the sentences in (6). Note how the usage of PROMINENT *-bi* aligns with the prosodic focus in the suggested English translations.

(6) a. *më'bi yë'ni ñañe bahuë*
mɨ'ĩ-bi jɨ'ɨ-ni ñaa-ñe bãã-wɨ
2SG-SBJ 1SG-N.SBJ2 see-CL:GEN NEG.AUX-N3S.PST.ASS
'YOU didn't see ME (i.e., not vice versa).'
[NAT: 20140925salsu001.032]

b. *io nëcaco ñacona yequëbi dani hue'ecaquëña io do'rohuë*
ĩ-o$_i$ nɨhka-ko ñaa-ko-na **jehk-ɨ-bi**$_k$
3PRO-CL:F stand-S.F.PRS.DEP look-S.F.PRS.DEP-DS other-CL:M-SBJ
daa-ni we'e-kah-kɨ-ña ĩ-o do'ro-wɨ
come-SS carry-BEN-2/3S.M.PST.N.ASS-REP 3PRO-CL:F basket-CLS:CONT
'She$_i$ stood (there) and watched as the OTHER one (M)$_k$ came and brought her basket.' [NAT: 20100907slicr001.006]

Both instances presented in (6) contain a PROMINENT-marked S-nominal whose referent contrasts with other explicitly mentioned alternatives in their respective contexts. Like in other languages with similar emphatic contrast-type DCM patterns, it is also possible for a contrastive reading to arise where alternatives are not explicitly mentioned, but rather are tacitly understood in context. Such instances are relevant to the corpus-based study reported in Chapter 5 of this dissertation, but they are set aside in this preliminary description.

In addition to the emphatic contrast reading evoked by PROMINENT S-marking, shown in (6), PROMINENT-marking is also implicated in case-matching effects under particular question-answer pair configurations (henceforth Q-A). Some well-documented DCM languages — e.g., Spanish and Romanian, recently discussed in Irimia (2020, 430) — are also described as displaying such Q-A marking patterns. A similar pattern is recently described for the Kakataibo language of the southwestern Amazon (Valle 2014, 61-65). These patterns are relevant to establishing the PLAIN-PROMINENT conception of Siona DCM.

The Q-A case-matching pattern of Siona is shown for S-oriented DCM in (7). The WH-item *quei-**bi*** (who) in (7a) displays obligatory PROMINENT-marking with *-bi*; this, in turn, forces PROMINENT-marking on the focalized S-argument in the answer *yë' gajeo-**bi*** (my friend (F)). Failure to utilize PROMINENT-marking in either the question or the answer turn is judged as illicit by all speakers:

(7) a. QUESTION: *queibi saiye yëquë ?*
ke-ɨ *(-bi) sai-je jɨɨ-kɨ
WH-CL:M-SBJ go-CL:GEN want-2/3S.M.PRS.N.ASS
'Who wants to go (somewhere)?'

b. ANSWER: *yë'ë gajeobi yëco saiye*
jɨ'ɨ gahe-o #(-bi) jɨɨ-ko sai-je
1SG friend-CL:F-SBJ want-3S.F.PRS.ASS go-CL:GEN
'My FRIEND (F) wants to go.' [VOL: 20220623emipa002.008-009]

The case-matching patterns presented in (7) fall under the umbrella of Q-A congruence effects (cf. Schwarzschild 1999; Selkirk 2008; Rochemont 1998, 2013; *inter alia*). As

indicated in the suggested translation, English has a similar focus-matching requirement, expressed prosodically on the focal element in the answer in (7b). These stable Q-A congruence patterns distinguish PLAIN- from PROMINENT-marking alternatives across all principled argumental and spatial DCM patterns in the Siona language.

Although focus-driven patterns of PROMINENT *-bi* marking on the S-argument are rigid in instances like (7), this does not account for *all* instances of PROMINENT-marking. In practice, a degree of variation is observed regarding DCM patterns on non-focal S-arguments. For instance, *-bi* marking is used to flag the argument ***wãñuhmi-bi*** (anaconda) in (8), although no salient alternatives are understood in this context:[5]

(8) *. . . yë'ë a'yëmacare huạñumibi ạjạ'i*
jɨ'ɨ a'-jɨ-mahka-re **wãñuhmi-bi** ã-ha'i
1SG younger_sibling-CL:M-DIM.SG-N.SBJ anaconda-SBJ eat-3S.M.PST.ASS
'. . . The anaconda ate my brother!' [NAT: 20100907slicr002.014]

One class of non-focal PROMINENT-marking examples concerns those instances where *-bi* appears to execute a disambiguation function (Bruil 2014, 160).[6] For instance, elicited examples like (9) demonstrate how *-bi* is naturally utilized to distinguish atypicial S-arguments, which may be confused with surrounding verebal elements in the sentence, in order to maximize communicative efficiency (Kurumada and Jaeger 2015). In these particular instances, case-marking is strongly favoured, although failure to mark the noun phrase does not necessarily render the sentence ungrammatical. For this reason the absence of *-bi* marking is glossed as degraded — i.e., '?' in (9):

(9) a. *iye hui'yocobi hui'yocoa sịọcore*
i-je **wi'jo-ko** **?(-bi)** wi'jo-ko-a sĩõ-ko-re
DEM.PROX-CL:GEN open-CL:F-SBJ open-CL:F-COP.3S.F shut-CL:F-N.SBJ
'This key (∼*opener*) opens the door (lit. is the opener to the *shutter*).'
[VOL: 20230623ejabi001.004a]

b. *jëaye yo'yebi ca'raye tëtosaiji*
hɨa-je **jo'-je** **?(-bi)** ka'ra-je tɨhto + sai-hi
be_hard-CL:GEN do-CL:GEN-SBJ fear-CL:GEN strike + go-3S.M.PRS.ASS
'Bravery (lit. doing (things) with vigour) overcomes fear.'
[VOL: 20241112ejabi002.011]

To summarize, most overt S-arguments arise in the zero-marked, PLAIN form; PROMINENT *-bi* marking typically arises in one of two focus-related contexts — i.e., either of the emphatic contrast type, shown in (6), or in Q-A pair contexts, as in (7). In non-focal contexts, *-bi* often serves an amibiguity avoidance function. This appears to be a WEAK TRIGGER for S-oriented DCM whose impact on the distribution of PROMINENT *-bi* marking is not entirely clear. Whereas this dimension of the pattern is largely set aside for future research, the evidence discussed throughout this dissertation helps to

[5] Martine Bruil notes that this may be an instance of *-bi* emphasizing the whole sentence, indicating a shift of attention. This is a promising direction for research; however, it is not clear that this can be generalized, or why the PROMINENT-marking should be expected to arise on the S-argument, for that matter. Future research will establish the merit of such an analysis.

[6] Some authors would reserve the term 'disambiguation' to refer narrowly to uses of DCM where the two core arguments of the transitive clause may be confused — e.g., 'the man saw the woman' (potentially ambiguous) vs. 'the man saw the rock' (unambiguous).

disentangle these effects from other factors affecting the distribution of *-bi* marking in Siona speech.

3.2 Non-Subject-oriented DCM patterns

Siona exhibits a pattern of case marking that partly merges several non-subject functions: P-argument displays similar encoding properties to the INDIRECT OBJECT (R) and to a handful of fringe thematic roles, such as the Experiencer and the Benefactee. This division is a typical property of Tukanoan languages (cf. Barnes 1999, 2006; Gomez-Imbert 2011; Stenzel 2013d). Table 3.3 summarizes the form of P- and R-marking, and contrasts them to S-marking.

	INANIMATE		ANIMATE ($\wedge$ $^{\uparrow}$INAN)	
	PLAIN	PROMINENT	PLAIN	PROMINENT
SUBJECT (S)	*-∅*	*-bi*	*-∅*	*-bi*
DIRECT OBJECT (P)	*-∅*	*-re*	*-re*	*-ni*
INDIRECT OBJECT (R)	*-re/-na*		*-re*	*-ni*

Table 3.3: Mapping PLAIN- and PROMINENT-alternatives across (non-)SUBJECT DCM

On the one hand, besides zero-marking, no formal case-marking alternatives are shared in S-oriented and non-SUBJECT-oriented DCM RELATIONS. On the other hand, P-marking and R-marking are nearly identical, excepting the *accidental -re* vs. *-na* DCM pattern found on inanimate R-arguments — i.e., the gray shading in the corresponding cell indicates that no principled PLAIN-PROMINENT alternation is tenable for this case-marking pair. These facts are explored in detail below.

Another fundamental distinction regards how the determination of PLAIN and PROMINENT forms is conditioned by an animacy-based noun class split for *all* non-SUBJECT-oriented DCM. For instance, although the PLAIN-PROMINENT distinction is suppressed for the inanimate R category, it survives in the principled PLAIN *-re* vs. PROMINENT *-ni* alternation for the animate R category. This particular case-marking pair is found across all animate non-SUBJECT grammatical relations, inspiring the glossing conventions utilized here: *-re* (N.SBJ), *-ni* (N.SBJ2). On this basis, animacy is treated as a STRONG TRIGGER for non-SUBJECT DCM patterns, and this TRIGGER has morphological implications — i.e., establishing which alternatives are to be treated as PLAIN or PROMINENT for a given noun class. This animacy condition interacts with a focus TRIGGER, similar to that described for S-oriented DCM above. Focus effects determine the selection of the PROMINENT case-marker, which is deemed appropriate for the noun class of the noun phrase in question.

The discussion below begins with a consideration of case-marking patterns on the DIRECT OBJECT (P) in Section 3.2.1. This is followed by a description of DCM patterns in the encoding of the INDIRECT OBJECT, the R-argument, in Section 3.2.3. Other non-SUBJECT argumental case-marking patterns are outlined in Section 3.2.4.

3.2.1 Direct object-oriented DCM patterns

Three case-marking alternatives are attested to encode the P-argument: zero-marking, *-re* marking, and *-ni* marking. For this reason, elsewhere we have labeled this a 'tripartite differential object marking' system (Case and Jeretič 2021; Case & Bruil, *forthcoming*). This constitutes a HYBRID-type DCM pattern as conceived in Chapter 2. However, given any particular argument, not all three marking alternatives are possible; rather, the appropriate forms are determined by an animacy-based noun class split, reiterated in Table 3.4:

	INANIMATE		ANIMATE	
	PLAIN	PROMINENT	PLAIN	PROMINENT
DIRECT OBJECT (P)	*-∅*	*-re*	*-re*	*-ni*

Table 3.4: Mapping PLAIN and PROMINENT marking alternatives for the P-argument

Following the distribution of forms suggested in Table 3.4, this dissertation maintains that *-re* ought to be analyzed as both the PROMINENT alternative for inanimate P-arguments, and as the PLAIN alternative for animate P-arguments (henceforth, this is referred to as the SPLIT *-re* HYPOTHESIS). The sentences in (10) illustrate the PLAIN- and PROMINENT-alternatives for inanimate P and for animate P accordingly.

(10) a. *yë'ë oco ucusi' cayë*
ji'i **ohko** ũhku-si'-i kaa-ji
1SG water drink-FUT-N3S say-N3S.PRS.ASS
'I want to drink water.' [VOL: 20230623ejabi001.024]

b. *mamajëbi ga'yoyë ëcabëre*
mama-hi-bi ga'jo-ji **ihka-bi-re**
child-CL:GEN-SBJ play-N3S.PRS.ASS plastic-CL:RND-N.SBJ
'The children are playing soccer (lit. (with) a/the plastic ball).'
[NAT: 20230628emabi001.023a]

c. *bayë jaibai a'yëbaire*
baa-ji **hai-bãi** **a'j-i-bãi-re**
have-N3S.PRS.ASS many-PL.AN older_sibling-CL:M-PL.AN-N.SBJ
'(I/we/they) have
many (older) siblings.' [VOL: 20230622ejabi001.011b]

d. *më'bi yë'ni ñañe bahuë*
mi'i-bi **ji'i-ni** ñaa-ñe bãã-wi
2SG-SBJ 1SG-N.SBJ2 see-CL:GEN NEG.AUX-N3S.PST.ASS
' "YOU didn't see ME (i.e., not vice versa)." '
[NAT: 20140925salsu001.032]

The use of noun class-inappropriate alternatives is rejected outright. Consider the elicited sentence in (11), where *-re* is naturally produced on the P-argument *mihchi-**re*** (cat). When asked whether it is possible to drop this case-marker, as in (11a), this is considered to be a grammatical error.

(11) *huë'e bayë michire*
wɨ'e baa-jɨ **mihchi-re**
house have-N3S.PRS.ASS cat-N.SBJ
'(I) have a cat at home.'
a. * *huë'e bayë michi* (Intended: '(I) have a cat at home.')
[VOL/SUG: 20230629elupa001.002e-f]

Likewise, the marker *-ni* is not permitted with inanimate nouns in Ecuadorian Siona, or in related languages, such as Colombian Siona (Wheeler 1970, 1987), Ecuadorian Sekoya (Johnson and Levinsohn 1990), and Koreguaje (Cook and Levinsohn 1985; Cook and Criswell 1993). The animacy condition on *-ni* marking is demonstrated in the elicited sentences in (12), lifted from Bruil (2014, 167, ex. (64)).

(12) *nocare a̱ë'ë̱*
nohka-re ã-ɨ'ɨ
banana-N.SBJ eat-N3S.PST.ASS
'(I) ate a/the baNAna.'
a. * *noca-**ni** a̱ë'ë̱* (Intended: '(I) ate a/the banana.')
[NAT/SUG: 20120912slicr001.036-037]

The minimal pair in (12) represents the general case, where *-ni* is unavailable to inanimate nouns. Section 3.2.2 discusses a class of apparent exceptions to this generalization. Nonetheless, these exceptions are found to be systematic, applying to a class of inanimate nominalizations, labeled the *promoted* noun class in this discussion.[7]

The remainder of this section turns to focus-triggered DCM effects, setting aside animacy effects. Like in S-oriented DCM, PROMINENT-alternatives are selected in focus contexts. Instances where *-ni* and *-re* evoke the emphatic contrast reading, for animate and inanimate P respectively, are parallel with those shown for *-bi* marking on the S-argument. For instance, *-ni* is shown to trigger this prominent reading in (10d) above. As for PROMINENT *-re* marking on inanimate P, consider sentence (13), which takes (10a), where the non-focal, inanimate P has PLAIN zero-marking, and fits this argument with PROMINENT *-re*; i.e., *ohko-**re*** (water). The expected emphatic contrast reading emerges in this context:

(13) [**Context**: A boy comes into the house to find coffee and tea on the table.]
yë'ë ocore u̱cusi'i cayë
jɨ'ɨ **ohko #(-re)** ũhku-si'-i kaa-jɨ
1SG water-N.SBJ drink-FUT-N3S say-N3S.PRS.ASS
'I want to drink wAter (i.e., not some other beverage).'
a. # *yë'ë oco u̱cusi'i cayë* ('I want to drink water.' — i.e., (10a))
[VOL/SUG: 20230623ejabi001.024-025]

In addition to triggering the emphatic contrast reading, as in (13), PROMINENT-marking arises in the expected way across Q-A pairs. Consider (14), which demonstrates how *-ni* is required on the animate-presupposing WH-item *ke**ɨ**-**ni*** (who/ what

[7] In addition, as is often found cross-linguistically (cf. Corbett 1991), in Siona a handful of inanimate nouns referring to celestial bodies pattern with animate nouns: e.g., *ẽ̃sɨ* (sun), *ñañɨ* (moon), *ma'ñoko* (star), *da'ñame* (rainbow). A few other examples are noted, e.g., *jaheo* (prepared ayahuasca). Bruil 2014, 148, describes a few other such items.

(ANIM)) in (14a) and on the focal animate P argument *ja'wɨ-**ni*** (peccary) in the answer in (14b). As anticipated, the case-marker *-re* is simply unavailable to either of these arguments in such Q-A configurations:

(14) a. QUESTION: *queini huare'ne ?*
ke-ɨ-ni wa-re-'ne
WH-CL:M-N.SBJ2 kill-N2/3S.PST.N.ASS-Q
'What (animate) did you (PL) hunt?'
i. * *quei-re huare'ne ?* (Intended: 'What did you (PL) hunt?')

b. ANSWER: *ya'huëni huaë'ë*
ja'wɨ-ni wa-ɨ'ɨ
peccary-N.SBJ2 kill-N3S.PST.ASS
'(We) hunted PECCary.'
i. # *ya'huë-re huaë'ë* (Intended: 'We hunted PECCARY.' [answers (14a)])
[VOL/SUG: 20230619elupa002.001a-b]

The same pattern of obligatory PROMINENT-marking is observed in inanimate-oriented Q-A constructions, as shown in (15). For both the WH-item *kee-**re*** (what.INAN) in (15a) and the focal P-argument *jowɨ-**re*** (canoe) in the answer in (15b), zero-marking is considered ungrammatical:

(15) a. QUESTION: *quere huero'ne go'ye mo'se ?*
ke-e-re wero-o-'ne go'je + mo'se
WH-CL:GEN-N.SBJ buy-2/3S.F.PST.NASS-Q before + day
'What did you (F) buy yesterday?'
i. * *que huero'ne go'ye mo'se ?*
(Intended: 'What did you (F) buy yesterday?')

b. ANSWER: *yohuëre huerohuë*
jo-wɨ-re wero-wɨ
canoe-CL:CONT-N.SBJ buy-N3S.PST.ASS
'(I) bought a caNOE.'
i. # *yohuë huerohuë* (Intended: '(I) bought a canoe.' [answers (15a)])
[VOL: 20230617eyopa001.003a-b]

On the basis of examples (13) through (15), I maintain that information focus-effects are active for P-oriented DCM, just as they are shown to be for S-marking. The main distinction between these patterns concerns the marking split based upon animacy. However, there is another relevant distinction between S- and non-SUBJECT-marking: i.e., non-SUBJECT pronouns require overt case-markers. Compare the P-marking pattern in (16), with the corresponding S-marking pattern shown in (16b):

(16) a. *go'ye mo'se icua'ire huerohuë*
go'je mo'se **ih-ko-wa'i-re** wero-wɨ
yesterday DEM.PROX-CL:F-PL.AN-N.SBJ buy-N3S.PST.N.ASS
'I bought them (i.e., several cows) yesterday.'
i. # *go'ye mo'se icua'i-**ni** huerohuë* (Emphatic contrast reading)
ii. * *go'ye mo'se icua'i huerohuë* (Intended: I bought them yesterday.)

b. *i̱cua'i bëayë sa'nahuëna*
ih-ko-wa'i bɨa-jɨ̈ sa'nawɨ-na
DEM.PROX-CL:F-PL.AN stay-N3S.PRS.ASS inside-GOAL
'<u>They</u> (i.e., several cows) are staying inside.'

i. # *i̱cua'i-**bi** bëayë sa'nahuëna* (Emphatic contrast reading)
[VOL: 20230629elupa001.024a-e]

The case-marking patterns shown in (16) are not surprising, given that the concerned pronouns are deictic pronouns with animate referents. These patterns would equally hold for animate full nouns. However, the following naturalistic instance in (17), adapted from Bruil (2014, 349, ex. (22)), contains several instances of the deictic pronoun *hãõ* (that one) referring to individual coconuts in this context. Where this pronoun arises as the S-argument, zero-marking and *-bi* marking are attested; however, the animate PROMINENT-marker *-ni* is selected on the instance where it is the P-argument. Speakers confirm that zero-marking is impossible where this pronoun is the P-argument:

(17) [**Context:** A spirit, born out of the protagonist's hammock, orders him to go to the forest and gather coconuts — emphasizing a particular one]
. . . j̱a̱o̱ ai yequë jubë jai jubë j̱a̱o̱bi aiyo j̱a̱o̱ni huatotoj̱ë'ë̱ caoña
↑**hã-o**$_{(S)}$ ai jehk-ɨ hubɨ hai hubɨ ↑**hã-o-bi**$_{(S)}$
DEM.MED-CL:F very other-CL:M bunch big bunch DEM.MED-CL:F-SBJ
ai-o ↑**hã-o-ni**$_{(P)}$ wahtoto-hɨ̃'ɨ̃
big-3S.F.ASS DEM.MED-CL:F-N.SBJ2 take_down-IMPER
kaa-o-ña
say-2/3S.F.PST.N.ASS-REP
'. . . "<u>THAT one</u>$_{(S)}$, the other bunch, the big bunch, <u>that one</u>$_{(S)}$ is ripe — take <u>THAT one</u>$_{(P)}$ down," (she) said (so they say).'
[NAT: 20100907slicr001.022]

The instances in sentence (17) demonstrate that non-SUBJECT pronouns exhibit obligatory case-marking, appropriate for their grammatical relation. This is a general fact of Tukanoan languages, i.e., including those which do not exhibit animacy-triggered DCM (Barnes 1999, 2006; Stenzel 2008, 2013d; Gomez-Imbert 2011; etc.). For many of these Tukanoanists, this fact falls out naturally from the analysis of *-re* as a specificity marker, banking on the implicational relation between specific and pronominalized referents. In previous corpus-based work, we found that a specificity-based analysis for *-re* is untenable for Ecuadorian Siona DCM (Case & Bruil, *forthcoming*), a finding which is complemented by the corpus-based investigation reported in Chapter 5 of the present dissertation. I set aside the question of the status of putative, referentiality-related TRIGGERS for P-oriented DCM for the remainder of this chapter.

Returning to the example in (17), this dissertation suggests an alternative analysis: the pronoun *hãõ* (that one), despite being referentially inanimate, patterns with the animate noun class for case-marking purposes. Such pronouns are one of several sub-classes of noun phrases, denoting inanimate entities, which align with animate nominals in this way. I refer to this as the class of *promoted nominals*, identified here by the upward arrow shorthand (↑). Membership in this class is entirely predictable from morphological considerations. The following subsection unpacks the mechanisms

which underlie the productive noun-class shifting operation of 'promotion', accounting for a handful of case-marking splits across argumental and spatial DCM in the language.

3.2.2 Promotion and noun class mobility patterns

What I call *promotion* is a language-specific descriptive notion, which signals a group of noun phrases which have inanimate referents but which pattern with the animate noun class for the sake of case-marking. This builds upon the metaphor that inanimate nouns low on the animacy scale, as in (18), 'move up' the scale to join the animate classes under the appropriate conditions:[8]

(18) **Animacy hierarchy:** (Siewierska 2004, 149)
(i) Human < (ii) Animal < (iii) Inanimate < (iv) Abstract

Consider how the inanimate instrument nominalization in (19) arises with PROMINENT *-ni* under the conditions of the Q-A heuristic, which is generally unavailable to inanimate nouns:

(19) a. QUESTION: *quere huero'ne go'ye mo'se ?*
ke-e-re wero-o-'ne go'je + mo'se
WH-CL:GEN-N.SBJ buy-2/3S.F.PST.NASS-Q before + day
'What did you (F) buy yesterday?'

b. *huerohuë coca caconi*
wero-wɨ ↑[**kohka + kaa-ko *(-ni)**]
buy-N3S.PST.ASS word + say-CL:F-N.SBJ2
'(I) bought a CELLphone (lit. a 'speaking-device').'

i. # *huerohuë coca caco**re***
(Intended: '(I) bought a cellphone.') [answers (19a)]

ii. * *huerohuë coca caco* [VOL: 20230617eyopa001.005a-b]

Promoted nominals, such as ↑*kohka kaako* (cellphone) in (19), are easily identifiable on the basis of their morphological composition. As a condition of membership, all promoted nominals contain the nominalizer suffix *-ko*, or its reduced counterpart *-o*, which is indistinguishable from the feminine gender classifier (see Bruil 2019 for a thorough discussion of nominalization strategies). As such, promoted nouns form a morphological noun class (Corbett 1991, 33-43), which deviates from the otherwise transparent animacy-based noun class system in the language. This pattern is not unlike how diminutive nouns in German are assigned neuter gender, trumping their baseline gender class assignment: e.g., *der Bär* (the bear, masculine) → *das Bärchen* (the little bear, neuter).

This dissertation recognizes three types of promoted nominals, all unified by their usage of the morpheme *-ko* (or *-o*):

[8] Many variations on the basic hierarchy suggested by Silverstein (1986) exist in the literature, some of which conflate animacy status with other notions such as pronominality or person distinctions. For our purposes, the basic animacy hiearchy adapted from the discussion in Siewierska 2004 will suffice. Siona treats animal- and human-denoting nouns alike, such that animate classes (i) and (ii) may be collapsed (i.e., a VITALIST split system per terminology in Ortmann 1998).

1. Lexicalized deverbal nominalizations (shown in (19));
2. Post-nominal reduced relative clause nominalizations;
3. Pronouns (shown in (17)).

Before exploring each of these types in greater detail, it should be noted that promotion is subject to a *singularity condition*. Nominalizations formed with *-ko* (or *-o*) are *always* singular in number, whereas the corresponding plural (and mass-denoting) nouns are formed via the generic classifier *-je* (or its nasal allomorph *-ñe*), and simply pattern with inanimate nominals.[9] (20) illustrates the singular and plural forms for the lexicalized deverbal nominalization ↑*kohka kaako* (cellphone):

(20) a. ↑*kohka kaa-ko* (cellphone, lit. 'speaking-device')
b. *kohka kaa-je* (cellphones, lit. 'speaking-devices')

The implications of the singularity condition on promotion are demonstrated in the Q-A sequence in (21). Here another promoted nominal of the lexicalized deverbal nominalization-type is found: ↑*tojako-**ni*** (pen, lit. 'writing-device'). The correct PROMINENT form is *-ni* with the singular nominalization in (21b), but *-re* with the plural nominalization in (21c).

(21) a. QUESTION: *quere co'eco'ne ?*
ke-e-re ko'e-ko-'ne
WH-CL:GEN-N.SBJ look_for-2/3S.F.PRS.NASS-Q
'What are you (F) looking for?'

b. ANSWER: *toyaconi co'eyë*
↑**toja-ko-ni** ko'e-jɨ
write-CL:F-N.SBJ2 look_for-N3S.PRS.ASS
'(I) am looking for a PEN (lit. a 'writing-device').'
i. # *toyaco-re co'eyë*
(Intended: '(I) am looking for a PEN.' [answers (21a)])

c. ANSWER': *toyayere co'eyë*
toja-je-re ko'e-jɨ
write-CL:GEN-N.SBJ look_for-N3S.PRS.ASS
'(I) am looking for PENS (lit. 'writing-devices').'
i. * *toyaye-ni co'eyë*
(Intended: '(I) am looking for PENS.' [answers (21a)])

[VOL: 20230703ejabi001.004]

The two examples considered thus far, i.e., (19) and (21), concern the lexicalized deverbal nominalization sub-type, and fall into two semantic classes, for which I suggest the following labels: (i) INSTRUMENT NOMINALIZATIONS, and (ii) EPISODE NOMINALIZATIONS. The two examples above illustrate the former class, whereas the latter class refers (abstractly) to a particular instance of the denoted activity. Several examples of each type of nominalization are provided from my fieldnotes in (22) and (23) respectively for the sake of illustration. This is a productive procedure which ought to be considered a common lexico-genetic strategy in this variety of Siona:[10]

[9] A similar observation regarding the usage of the generic classifier with inanimate plural demonstratives, numerals, classifiers and adjectives is noted in Bruil (2014, 154-155).

[10] Note that several of many naturalistic occurrences of these nominalizations contain the

(22) Instrument nominalizations:

a. ↑*chaokɨ̃ɨ̃-ko*[11] (shovel)
[*kɨ̃ɨ̃-* (dig)] ;

b. ↑*wi'jo-ko* (key)
[*wi'jo-* (open)] ;

c. ↑*sĩõ-ko* (door)
[*sĩõ-* (shut)]

(23) Episode nominalizations:

a. ↑*sẽĩ̃-ko* (question)
[*sẽĩ̃-* (ask)] ;

b. ↑*hɨ̃hɨ̃-ko* (song)
[*hɨ̃hɨ̃-* (sing)] ;

c. ↑*je'je-ko* (lesson)
[*je'je-* (learn)]

The second type of promoted noun phrase, namely noun phrases with post-nominal relative clauses, is not altogether different from the lexicalized nominalizations outlined above. Siona exhibits two nominal modification strategies: (i) the pre-nominal modification strategy — i.e., juxtaposition modifier-head; and (ii) the post-nominal modification strategy — i.e., (head)-reduced relative clause. The second strategy implicates promotion of the complex noun phrase, where the head noun is a singular, inanimate argument. The modifier in this strategy is treated as a headless relative clause structure in apposition within the complex noun phrase, given that these 'modifiers' may freely arise without an overt head nominal.[12]

Compare the noun phrases in (24) and (25). In the case of (24), the pre-nominal modification, shown in (24a), does not trigger promotion; which is set against the post-nominal alternative in (24b), where promotion arises as expected. Neither modification strategy triggers promotion for the plural counterpart shown in (25):

(24) a. *nea toayohuë*[13]
nea toa + jo-wɨ
black fire + canoe-CL:CONT
'a/the black car'

b. ↑*toayohuë neaco*
toa + jo-wɨ nea-ko
fire + canoe-CL:CONT black-CL:F
'a/the car that is black'

(25) a. *nea toayohuë̲a̲*
nea toa + jo-wɨ-ã
black fire + canoe-CL:CONT-PL.INAN
'black cars'

b. *toayohuë̲a̲ neañe*
toa + jo-wɨ-ã nea-ñe
fire + canoe-CL:CONT-PL.INAN black-CL:GEN
'cars that are black'

resultative morpheme *-sih* — e.g., *sẽĩ̃-sih-ko* (∼that which had been asked), a possible translation for 'question'. I have not reported instances like these for the sake of simplicity; however, this morpheme, and the purposive marker (*-hã'*), and the affixal negator (*-ma'*), may freely arise within nominalizations like those presented in (22) and (23).

[11]The nominalization *chaokɨ̃ɨ̃-ko* was naturally produced with a serialization involving the aspectual verb *chao-* (finish). This is an interesting case which points to the lexical character of many of these nominalizations.

[12]A similar analysis is proposed in Epps (2008, 181,828-853) for the corresponding structure in the Hup language.

[13]The compound *toa + jo-wɨ* (lit. fire canoe) is used to refer to motorized vehicles.

The promotion status of a given noun phrase determines the PROMINENT-alternative which is appropriate under Q-A conditions. (26) demonstrates the two alternatives in (24) as focalized P-arguments in a Q-A configuration. Despite differences in case-marking, the two answers in (26b) and (26c) ought to be viewed as synonymous.[14]

(26) a. QUESTION: *quere baquë ?*
ke-e-re baa-kɨ
WH-CL:GEN-N.SBJ have-2/3S.M.PRS.N.ASS
'What does (he) (M) have?'

b. ANSWER: *jaë baji nea toayohuërе*
hã-ɨ baa-hi [**nea toa + jo-wɨ #(-re)**]
DEM.MED-CL:M have-3S.M.PRS.ASS black fire + canoe-CL:CONT
'He has a black CAR.'

c. ANSWER': *jaë baji toayohuë neaconi*
hã-ɨ baa-hi ↑[**toa + jo-wɨ**
DEM.MED-CL:M have-3S.M.PRS.ASS fire + canoe-CL:CONT
nea-ko *(-ni)/(#-re)]
black-CL:F-N.SBJ(2)
'He has a black CAR.' [VOL/SUG: 20230614erebi001.014a-b]

A complementary, naturalistic instance of promotion arising on a nominal with the appropriate post-nominal modification strategy is provided in (27), lifted from the narrative sample analyzed in Chapter 5:

(27) [**Context:** The children cover several items in the room on the bed and flee in order to fool an impostor (jaguar demon) who has taken the form of their grandmother.]
. . . sa'saji'soe tsiubë nesiconi cunijoña jaoreta'a
↑[**sã'sahĩ'soe + tsĩũ-bɨ nee-sih-ko-ni**] kũni
mate_tea + cover-CL:RND make-COMP-CL:F-N.SBJ2 bite
+ hoo-o-ña hã-o-re-tã'ã
+ cut-2/3S.F.PST.N.ASS-REP DEM.MED-CL:F-N.SBJ-CNT.EXP
'. . . (she) bit and tore at that (lit., her), at the covered mate TEA bundle they had made !' [NAT*: 20151001oolpi001.065]

Clearly, considering naturalistic instances like (27), promotion is a stable component of Siona DCM. This morphological noun class accounts for the bulk of 'exceptionally'-marked inanimate nouns in work with texts. Nonetheless, instances such as (27) are not particularly frequent in narratives even if promoted pronouns are ubiquitous.

The class of promoted pronouns with singular inanimate reference consists of two demonstratives: namely, *ihko* (this one) and *hão* (that one).[15] They are both morphologically feminine. Consider two naturalistic instances lifted from texts. (28) replicates (17) containing a P-instance of the medial demonstrative pronoun ***hãõ-ni***, which refers to an individual coconut. The sentence in (29) is lifted from the narrative sample in Chapter 5, where ***ihko-ni*** is used, deictically, to refer to a piece of charcoal:

[14]The scope possibilities for these focus effects are explored in greater detail in Chapter 6.
[15]I have also come across a few instances of the personal pronoun *ĩõ* (she) used with inanimate singular referents in my corpus work. More work is needed to check its usage and acceptability.

(28) [**Context:** A spirit, born out of the protagonist's hammock, orders him to go to the forest and collect coconuts.]
. . . ja̲o̲ ai yequë jubë jai jubë ja̲o̲bi aiyo ja̲o̲ni huatotojë'ë̲ caoña
hã-o ai jehk-ɨ hubɨ hai hubɨ hã-o-bi
DEM.MED-CL:F very other-CL:M bunch big bunch DEM.MED-CL:F-SBJ
ai-o ↑**hã-o-ni** wahtoto-hɨ̃'ɨ̃
big-3S.F.ASS DEM.MED-CL:F-N.SBJ2 take_down-IMPER
kaa-o-ña
say-2/3S.F.PST.N.ASS-REP
'. . . "THAT one, the other bunch, the big bunch, that one is ripe — take THAT one down !" (she) said (so they say).' [NAT: 20100907slicr001.022]

(29) [**Context:** A mother helps her pregnant daughter who has fallen ill.]
. . . yo'quëna i̲o̲ bëca'co̲je̲ caoña iconi gu̲ijë'ë̲. caoña. neo garabëre
jo'-ki-na ĩ-o bika'-ko-hẽ kaa-o-ña
do-S.M.PRS.DEP-DS 3PRO-CL:F parent-CL:F-ADD say-2/3S.F.PST.N.ASS-REP
↑**ih-ko-ni** gũĩ-hɨ̃'ɨ̃ kaa-o-ña
DEM.PROX-CL:F-N.SBJ2 bite-IMPER say-2/3S.F.PST.N.ASS-REP
nee-o gara-bi-re
make-3S.F.PST.ASS charcoal-CL:RND-N.SBJ
'When (he) did this, her mother said, "bite into THIS!", she had prepared charcoal.' [NAT*: 20150811sfryi001.092-093]

In summary, recognizing the morphological noun class of promoted nouns and pronouns accounts for several apparent exceptions to the otherwise transparent animacy-based noun class system relevant to non-SUBJECT-oriented DCM. Although promoted nominals have inanimate referents, the fact that they are fitted with the feminine classifier morpheme, *-ko* (or *-o*), at least where their referent is singular, has implications for case-marker selection, making them behave as animates.

3.2.3 Indirect object-oriented DCM patterns

This section describes DCM patterns as attested for the INDIRECT OBJECT. Table 3.5 summarizes the distribution of case-markers for this DCM RELATION.

	INANIMATE		ANIMATE	
	PLAIN	PROMINENT	PLAIN	PROMINENT
INDIRECT OBJECT (R)	*-re* / *-na*		*-re*	*-ni*

Table 3.5: Mapping PLAIN and PROMINENT marking alternatives for the R-argument

On the one hand, Table 3.5 demonstrates that R-oriented DCM is indistinguishable from P-oriented DCM, as concerns the alternations observed for animate arguments. On the other hand, inanimate R-marking is distinct in several respects. Firstly, all R-arguments must bear case-marking, regardless of noun class. In other words, Siona R-marking displays ALTERNATING-type DCM. Additionally, inanimate R does occur with two different case-markers, namely *-re* and *-na*; although, these markers do not

display a principled PLAIN-PROMINENT alternation, including the expected information structure effects. As such, inanimate R is conceived as displaying accidental DCM.

Like animate P, all animate R-arguments must arise with overt case-marking. The PLAIN-alternative *-re* competes with the PROMINENT-marker, *-ni*. The two naturalistic instances in (30) and (31) illustrate these alternatives.

(30) *. . . cua'coni ñatasi' mo'se neahuẽ'ñareba i̱o̱re o̱cuaoña neco*
kwa'ko-ni ñahta-si + mo'se nea-wẽ'ña-reba
cook-SS dawn-COMP + day black-LOC.DERIV-INTENS
ĩ-o-re õhkwa-o-ña nehko
DEM.PROX-CL:F-N.SBJ give_to_drink-2/3S.F.PST.N.ASS-REP *neco*_drink
'. . . after cooking (it), the next day (she) gave her *neco*[16] to drink, right at dawn.' [NAT: 20110328slicr001.023]

(31) [**Context:** A family is arguing over who gets to eat the eggs they found.]
si'a tsiago'o i̱sijẽ'ẽ yẽ'ni caoña
si'a tsia + go'o ĩhsi-hĩ'ĩ **jɨ'ɨ-ni** kaa-o-ña
all egg + bundle give-IMPER 1SG-N.SBJ2 say-2/3S.F.PST.N.ASS-REP
' "Give the whole bundle of eggs to ME !", (she) said (so they say).'
[NAT*: 20150811sfryi001.121]

Note how the usage of PROMINENT *-ni* in (31) evokes the emphatic contrast reading, as anticipated. More specifically, the character is indicating the bundle of eggs be given to her, and not to another member of the family in this context.

Applying the Q-A heuristic, the expected pattern is attested. Namely, the PROMINENT-alternative *-ni* is obligatorily realized on the WH-item and the corresponding focalized R-argument in the answer, as reported in (32):

(32) a. QUESTION: *queini i̱sija'quë'ne te'huë ocohuë ?*
ke-ɨ *(-ni) ĩhsi-hã'-kɨ-a-'ne te'-wɨ ohko-wɨ
WH-CL:M-N.SBJ2 give-PRP-CL:M-COP-Q one-CL:CONT water-CL:CONT
'Who are (you (M)) going to give a water bottle to?'

b. ANSWER: *Suteni i̱sija'quëa'ë ocohuë*
Suhte *(-ni) ĩhsi-hã'-kɨ-a-'ɨ ohko-wɨ
S-N.SBJ2 give-PRP-CL:M-COP-N3S water-CL:CONT
'(I) am going to give a bottle to SUte (nickname).'

i. *# Sute-re i̱sija'quëa'ë ocohuë*
(Intended: '(I) am going to give a bottle to Sute.' [answers (32a)])
[VOL: 20220621eerpa001.005a-b]

On the basis of these examples, for all intents and purposes animate R-marking is isomorphic to the P-oriented DCM pattern. However, R-oriented DCM of inanimate R displays a markedly different pattern. This is illustrated in the elicited sentences in (33), where inanimate R accepts both *-re* marking and marking with *-na*, which is generally reserved for the spatial GOAL-argument. The case-marker *-ni* is unavailable as anticipated. In addition, unlike P-oriented DCM (but like with animate R-oriented DCM, zero-marking is rejected on the inanimate R-argument:

[16] *Neco* refers to an anti-anemic agent brewed from locally cultivated plants.

(33) a. *yëquë isiye yëyë yijatoayohuë baidaribëre*
jɨhkɨ ĩhsi-je jɨɨ-jɨ jiha + toa + jo-wɨ
1PL.EXCL give-CL:GEN want-N3S.PRS.ASS land + fire + canoe-CL:CONT
bãĩ + dari-bɨ *(-re)
people.COL + community-CL:COL-N.SBJ
'We want to give a bus to the community.'

b. ✓ *yëquë isiye yëyë yijatoayohuë baidaribë-na*
(no noted change in meaning)

c. * *yëquë isiye yëyë yijatoayohuë baidaribë-ni*
(*-ni* unavailable on inanimate)

d. * *yëquë isiye yëyë yijatoayohuë baidaribë* (zero-marking unavailable on R)
[VOL/SUG: 20220621eerpa001.016a-e]

Admittedly, more work is needed to fully unpack the interpretative implications of *re-* and *-na* marking on the inanimate R-argument; however, as concerns (33), speakers did not report a difference. More importantly for the present purposes, the usage of *-re* and *-na* does not adhere to a PLAIN-PROMINENT alternation pattern, and is not considered a principled DCM pattern on this basis. This is one of a few instances that I have come across in my fieldwork where *-re* and *-na* may be used interchangeably on non-core inanimate arguments. For instance, (34) demonstrates this same accidental DCM alternation arising with the predicate *choi-* (invite somebody to X) — again no meaning difference is reported in this instance:

(34) a. *Pedro chojo'i i gajeibaire fiestare*
Pedro cho-ho'i ĩ-ɨ gahei-bãĩ-re **fiesta-re**
P invite-3S.M.PST.ASS 3PRO-CL:M friend-PL.ANIM-N.SBJ party-N.SBJ
'Pedro invited his friends to the party.'

b. ✓ *Pedro chojo'i i gajeibaire fiesta-na* (no change in meaning)
[VOL: 20230626eyopa001.007b]

Just as *-ni* is unavailable to (non-promoted) inanimate noun phrases, the case-marker *-na* is *never* found on animate noun phrases in Ecuadorian Siona. Whereas the animacy condition for *-ni* holds across several related Western Tukanoan varieties (i.e., at least Colombian Siona, Sekoya, and Koreguaje), the noted animacy restriction for *-na* marking in the variety described in this dissertation is not found for Colombian Siona (Wheeler 1970, 174; 1987, 126-127) or for Ecuadorian Sekoya (Johnson and Levinsohn 1990, 47-48). In those varieties, *-na* is readily attested on both animate and inanimate R-arguments. Further comparative details are provided in Chapter 4.

In summary, given that the R-argument is prototypically animate, it appears at first glance that P- and R-oriented DCM are identical. However, upon further inspection, crucial differences arise where inanimate R is concerned. Case-marking is obligatory on the R-argument, not due to animacy-related noun class conditions, but as a mandatory component of encoding this grammatical relation. It is also the case that spatial GOAL-marking with *-na* seeps into the argumental case-marking domain. Some other non-prototypical uses of *-na* are identified in Section 3.3.3.

3.2.4 Other non-Subject DCM patterns

The preceding sub-sections have demonstrated that it is possible to postulate a single, principled, non-SUBJECT DCM pattern for animate arguments, based upon a con-

sideration of P- and R-marking facts. Peripheral participants receiving at least two other thematic roles (Experiencers and the Benefactees) display the same PLAIN *-re* vs. PROMINENT *-ni* alternation. It is relevant to note that these roles also display obligatory *-re* marking across all Eastern Tukanoan languages (Barnes 1999, 2006; Gomez-Imbert 2011; Stenzel 2013d; etc.), patterning with the R-argument.[17]

Firstly, consider the case-marking on the Experiencer-argument. This role arises in impersonal constructions of the type *uu-* (be hot, burn) in the applicative desiderative construction with *-ɨa* (want to), which we discuss in van Gijn et al. 2023, and in modal constructions, discussed in Chapter 6. (35) replicates the example from Bruil (2014, 163, ex. (56)), demonstrating PLAIN *-re* marking with the impersonal predicate *uu-* 'be hot'. When a similar predicateb such as *sɨhsɨ-* (be cold), is tested against the Q-A heuristic, as reported in (36), PROMINENT *-ni* marking arises on the WH-item in the question, and on the focalized element in the answer, as anticipated:

(35) *yë're uji*
jɨ'ɨ-re uu-hi
1SG-N.SBJ be_hot-3S.M.PRS.ASS
'I'm hot (lit. it's hot to me).' [VOL: 20110302elicr001.013]

(36) a. QUESTION: *queini sësëquë ?*
ke-ɨ *(-ni) sɨhsɨ-kɨ
WH-CL:M-N.SBJ2 be_cold-2/3S.M.PRS.N.ASS
'Who is cold?'

b. ANSWER: *Palomani sësëji*
Paloma *(-ni)/(#-re) sɨhsɨ-hi
P-N.SBJ(2) be_cold-3S.M.PRS.ASS
'PaLOma (female name) is cold.' [VOL: 20230617elupa001.018a-b]

The identicial pattern is noted for the coding of the Benefactee-argument, introduced via the applicative head *-ka(i)*, presented in **boldface** in the examples below. In (37), the PLAIN *re*-marker is used; whereas PROMINENT *-ni* marking is implicated in the related Q-A construction in (38), in the expected way:

(37) *yë'ë më're goamaña necaija'quëa'ë*
jɨ'ɨ **mɨ'ɨ̃ *(-re)** goa-mahña + nee-**kai**-hã'-kɨ-a-'ɨ
1SG 2SG-N.SBJ thing-DIM.PL + do-APPL.BEN-PRP-CL:M-COP-N3S
'I am going to work for you.' [VOL/SUG: 20241112ejaba001.009a-b]

(38) a. QUESTION: *queini goamaña necaijacoa'ne më'ë ?*
ke-ɨ *(-ni)/(#-re) goa-mahña + nee-**kai**-hã'-ko-a-'ne mɨ'ɨ̃
WH-CL:M-N.SBJ(2) thing-DIM.PL + do-APPL.BEN-PRP-CL:F-COP-Q 2SG
'Who are you (F) going to work for?'

[17] Most authors describing Western Tukanoan languages which have the case-marker *-ni* do not discuss *-ni* marking on these relations. However, it is likely that this marker is available to encode, say, the EXPERIENCER-argument, since it is also available for R-marking in these languages. This remains an open question at present.

b. ANSWER: *goamaña necaijɑ̃coa'ë më' gajeini*
goa-mahña + nee-**kai**-hã'-ko-a'-**ɨ** **mɨ'ɨ̃**
thing-DIM.PL + do-APPL.BEN-PRP-CL:F-COP-N3S 2SG
+ gahe-ɨ *(-ni)/(#-re)
+ friend-CL:M-N.SBJ(2)
'(I) am going to work <u>for your FRIEND</u>.'
[VOL/SUG: 20241112ejaba001.012a-d]

The examples in (35) through (38) reinforce the idea that a single, animate non-SUBJECT DCM pattern occurs in Siona grammar, extending beyond P- and R-oriented DCM to the non-canonical Experiencer participant and the Benefactee participant.

3.3 Spatial DCM patterns

Siona grammar recruits case-marking to encode a range of spatial relations, many of which are the same markers which encode the argumental grammatical relations outlined above. These are represented in the shaded cells in Table 3.6, repeated from the opening of this chapter.

Case-marker	Grammatical relation	Use
-bi	Subject	focus
	Instrument	obligatory
	Source location	
-re	Direct object	specific object
	Indirect object	
	Location	
-ni	Direct object	focus
	Indirect object	
-na	Goal	specific goal
-hã'ã	Path	obligatory
	Limit	
-hã're	Comitative	obligatory

Table 3.6: The case suffixes in Ecuadorian Siona, their function and their use (Bruil 2014, 157, Table 4.2)

Siona exhibits a typologically common set of spatial CASE FUNCTIONS (cf. Creissels 2009; Stolz et al. 2014; Haspelmath 2019, among many others): i.e., STATIC LOCATION (L), GOAL (G), SOURCE, and PATH/LIMIT. More specific spatial relations are encoded via spatial nouns — e.g., *jɨ'oka* (below, under), *sa'niwɨ* (inside), etc., which do not concern us here.

Of the four spatial relations listed above, two display obligatory case-marking. The Source-argument is always marked with *-bi*, and speakers reject zero-marking, as demonstrated by in (39). Additionally, two instances of *-hã'ã* marking are lifted from the narrative sample analyzed in Chapter 5: (40) provides an instance where it encodes the Path-argument, and (41) the Limit-argument, respectively:[18]

[18]I have also noted a handful of argumental uses of the PATH-marking *hã'ã*. For instance,

(39) *yë'ë baⁿñë Ecuadorbi daiye*
ji'i bãã-ñi **Ecuador *(-bi)** dai-ye
1SG NEG.AUX-N3S.PRS.ASS E-SRC come-CL:GEN
'I am not from Ecuador.' [VOL/SUG: 20230526ejabi002.010a-b]

(40) *. . . daoco i airo daoquë hua'i nesi ma'aja'a oni dojaioña*
dao-ko **ĩ-i** **ai-ro** **dao-ki** **wa'i**
wander-S.F.PRS.DEP 3PRO-CL:M big-CL:LOC wander-S.M.PRS.DEP meat
nee-si **+ ma'ã-hã'ã** oni + dohai-o-ña
make-COMP + path-PATH cry + wander-2/3S.F.PST.N.ASS-REP
'. . . (she) wandered along the path where he would wander to hunt, crying as (she) wandered.' [NAT*: 20151112orapi001.013-015]

(41) *. . . sai airo saquëña huao huë'equë'roja'a*
sai-i ai-ro sah-ki-ña **wa-o**
go-S.M.DEP.PST big-CL:LOC go-2/3S.M.PST.N.ASS-REP parent_in_law-CL:F
wi'e-ki'ro-hã'ã
house-LOC.DERIV-PATH
'. . . going to the jungle (to hunt), (he) went up to his mother-in-law's house.' [NAT*: 20151001oolpi001.009]

Setting aside these obligatory case-marking patterns, spatial DCM patterns arise for both the L-argument and for the G-argument. This pattern is in keeping with the extensive typology of spatial DCM (Stolz et al. 2014; Haspelmath 2019), which establishes that relations further to the left on the hierarchy laid out in (42) are more likely to permit zero-marking than those to their right:

(42) GOAL > (STATIC) LOCATION > SOURCE (Haspelmath 2019, 324-325)

The Siona spatial DCM patterns are not strictly of the OPTIONAL-type (i.e., zero- vs. a single overt case-marker). In fact, this is one of several ways that argumental and spatial DCM align in the language. Table 3.7 maps case-marking forms to the two spatial DCM patterns, and sets these against the argumental patterns:

this case-marker is utilized to encode the VEHICLE-argument — e.g., 'go X by canoe' (. . . *tsiaja jowi-**hã'ã***). Elsewhere, the marker encodes what I label the INDIRECT INSTRUMENT function: e.g., 'call X on a phone' (. . . *kohka kaako-**hã'ã***) ∼ 'via the telephone'. It is conceivable that these are extensions of the prototypical spatial use of this case-marker.

	INANIMATE		ANIMATE (∧ ↑INAN)	
	PLAIN	PROMINENT	PLAIN	PROMINENT
Argumental DCM:				
SUBJECT (S)	-∅	*-bi*	-∅	*-bi*
DIRECT OBJECT (P)	-∅	*-re*	*-re*	*-ni*
INDIRECT OBJECT, e.g., (R)	*-re* (*-na*)		*-re*	*-ni*
Spatial DCM:	INANIMATE		↑INANIMATE	
LOCATION (L)	-∅	*-re*	*-re*	*-ni*
GOAL (G)	-∅	*-na*	*-na*	

Table 3.7: Mapping PLAIN- and PROMINENT-alternatives for all DCM RELATIONS, including spatial relations

Table 3.7 shows, firstly, that the inanimate-animate split is relevant for spatial DCM if the class of promoted inanimates is taken into account. This fact further reinforces the claim that Siona exhibits a SUBJECT vs. non-SUBJECT DCM system, since *all* animate non-SUBJECT arguments must bear overt case-marking of one form or another. In what follows, it will also be demonstrated that Q-A heuristics equally hold across spatial DCM patterns, such that the principled PLAIN-PROMINENT alternation, characteristic of Siona DCM, can be maintained across L- and G-oriented DCM patterns. The remainder of this section treats each of these patterns in turn.

3.3.1 Static location-oriented DCM patterns

According to the distribution of case-marking alternatives laid out in Table 3.7, P- and L-oriented DCM patterns are formally indistinguishable. Given that spatial arguments are typically inanimate, the general L-marking pattern is such that PLAIN zero-marking alternates with PROMINENT *-re* marking. (43) illustrates a typical instance with zero-marking in the basic locative construction, with the predicate *bai-* (to be). As shown in (44), PROMINENT *-re* marking evokes the emphatic contrast reading in an out-of-the-blue context, which is absent for PLAIN zero-marking:

(43) *te'i ë̃më huë'e ba'iji*
te'-ɨ ɨ̃mɨ **wɨ'e** ba'i-hi
one-CL:M man house be-3S.M.PRS.ASS
'The man is alone in the house.' [VOL: 20230525ejabi002.007]

(44) a. *yë' yo'jei Canadare ba'iji*
jɨ'ɨ jo'he-ɨ **Canada #(-re)** ba'i-hi
1SG younger_sibling-CL:M C-N.SBJ live-3S.M.PRS.ASS
'My younger brother lives in CAnada (i.e., not some other place).'

b. *yë' yo'jei Canada ba'iji*
jɨ'ɨ jo'he-ɨ **Canada (#-re)** ba'i-hi
1SG younger_sibling-CL:M C-N.SBJ live-3S.M.PRS.ASS
'My younger brother lives in Canada.'
[VOL/SUG: 20230619elupa002.011a-b]

In addition to displaying the expected emphatic contrast reading, as in (44a), spatial DCM displays rigid case-matching patterns as a matter of Q-A congruence effects. The Q-A heuristic is demonstrated for the L-argument in (45), where the WH-item *hero-**re*** (where) requires PROMINENT *-re* marking, and forces PROMINENT-marking on the corresponding focalized location in the answer:

(45) a. QUESTION: *jerore ba'iquëa'ne ?*
he-ro *(-re) ba'i-ki-a-'ne
WH-CL:LOC-N.SBJ live-CL:M-COP-Q
'Where do (you (M)) live?'

b. ANSWER: *yë'ë ba'iquëa'ë Canadare*
ji'i ba'i-ki-a-'i **Canada #(-re)**
1SG live-CL:M-COP-N3S C-N.SBJ
'I live in CAnada.' [VOL: 20230526ejabi002.005-006]

The examples considered thus far show that L-oriented DCM is identical to the P-oriented DCM patterns with nouns of the inanimate noun class. Like most languages (e.g., Stolz et al. 2014), conceptually animate referents are incompatible with spatial meanings. However, it is in fact possible for spatial nouns to undergo promotion and to behave like animate nouns for the sake of DCM.[19] Under the appropriate conditions — i.e., those laid out in Section 3.2.2; the promoted L-argument displays the PLAIN *-re* marking vs. PROMINENT *-ni* marking alternation. Zero-marking is thus prohibited. (46) reports elicited instances where the expected case-marking patterns arise on the singular, promoted argument, which displays the post-nominal modification strategy. (47) confirms that promotion does not arise where the argument is plural:

(46) [**Suggested context:** There are several houses in this part of town, you live in the house your grandparents built, and not, e.g., the one your uncles built.]
ba'iyë yë'ë ñicuëbai huë'e nesiconi
ba'i-ji ↑[**ji'i ñihkw-i-bãĩ wi'e**
live-N3S.PRS.ASS 1SG grandparent-CL:M-PL.AN house
nee-sih-ko-ni]
make-PERF-CL:F-N.SBJ2
'(I/we) live in the house that my GRANDparents built.'

a. # *ba'iyë yë'ë ñicuëbai huë'e nesico-re* (no contrastive reading)

b. * *ba'iyë yë'ë ñicuëbai huë'e nesico*
(zero-marking unavailable on ↑inanimate)

[19] In addition to this, Siona recruits several nominalizers to derive a spatial argument from nouns which are otherwise incompatible (Bruil 2019): i.e., *-ki'ro* and *-wẽ'ña*. The derivational device *-ki'ro* is compatible with both a Place-of-Residence readings as in (i), or with a so-called *vicinal* reading (i.e., next to X, where X is (currently) located):

(i) *ioquë'rore hueagono neni a'rimaña toani ...*
ĩ-o-ki'ro-re wea + gõno nee-ni a'ri-mahña toa-ni
3PRO-CL:F-LOC.DERIV-LOC corn + chicha make-SS small-DIM.PL grind-SS
'(They) ground up and made a little corn chicha at her place ...'
[NAT*: 20151112orapi001.137]

In the general case, *-ki'ro* is used with individuals, and *-wẽ'ña* with concepts or in deverbal nominalizations. However, further attention is warranted in order to unpack their differences.

(47) [**Suggested context:** Same as in (46), except the speaker's family lives in several houses, all built by their grandparents.]
ba'iyë yë'ë ñicuëb<u>ai</u> huë'ña nese'ere[20]
ba'i-jɨ [**jɨ'ɨ ñihk^{w}-ɨ-bãĩ wɨ'-ña**
live-N3S.PRS.ASS 1SG grandparent-CL:M-PL.AN house-PL.INAN
nee-sih-je-re]
make-PERF-CL:GEN-N.SBJ
'(We) live <u>in the HOUSes that my grandparents built</u>.'

a. # *ba'iyë yë'ë ñicuëb<u>ai</u> huë'ña nese'e* (no contrastive reading)

b. * *ba'iyë yë'ë ñicuëb<u>ai</u> huë'ña nese'e-ni*
(*-ni* marking unavailable on inanimate)

[VOL/SUG: 20241003ejabi001.003a-b]

Although instances such as those put forth above are not particularly common in semi-spontaneous speech, again, promoted pronominals are commonplace in Siona narration. There are two classes of spatial pronouns, which are used almost interchangeably in the language: (i) inherent spatial pronouns, and (ii) promoted spatial pronouns. In fact, the latter class are the exact same pronominals *borrowed* from the feminine-singular paradigm in the argumental domain:

(48) a. **Inherent spatial pronouns:**
ĩhño/ĩñe (here), *hã-ro* (there), *heh-to* (over there, or its reduced form *to*)

b. **Promoted spatial pronouns:**
$^{\uparrow}$*ih-ko* (this one (F) ≃ here), $^{\uparrow}$*hã-o* (that one (F) ≃ there)

In addition to these spatial pronouns, the category of spatial pronouns also contains demonstrative pronominals derived with the spatial classifier *-ro/-to*. Any element derived with this morpheme behaves as inanimate, as with the interrogative pronoun *hero* (where). A handful of naturalistic instances of spatial L-pronouns are illustrated in (49), demonstrating the available PLAIN- and PROMINENT-alternatives accordingly. Firstly, inherent spatial pronouns are shown in (49), and then promoted pronouns are shown in (50):

(49) a. *ñadojai ñaëña jero cani ñatoj<u>e</u> — tocato j<u>a</u>re s<u>o</u>quëñë<u>a</u> beoëña*
ñaa + dohai ñaa-ɨ-ña he-ro kaa-ni
see + wander see-2/3S.M.PST.N.ASS-REP INTER-CL:LOC say-SS
ñaa-to-hẽ **to=kato** hãre sõhkɨ-ñɨ-ã
see-COND-ADD ANA.LOC=TOP like_that wood-CL:TREE-PL.INAN
beo-ɨ-ña
NEG.AUX-2/3S.M.PST.N.ASS-REP
'. . . (he) took a look around, seeing where (he) was speaking of — when (he) looked, there were no trees <u>there</u>.'

[NAT*: 20151023orocr001.108]

[20]This dissertation advocates for the commonly found resultative nominalization *-se'e* (Bruil 2019) to be treated as underlyingly formed of the perfective *-si(h)* and the generic classifier *-je*. Recall that *-je* is used with both plural- and mass-denoting concepts, hence its wider distribution.

b. [**Context:** The speaker is asked about a blanket beside him.]
iñorebare uiji
ihño-reba-re ũi-hi
here-INTENS-N.SBJ be_lying-3S.M.PRS.ASS
'(It) is lying right HERE.' [NAT: 20140805serpi003.093]

(50) a. *... ba'ina jaore saniti'ani abita mëni ocuereña*
ba'i-ɨ-na ↑**hã-o-re** sani + tĩ'ã-ni abita
be-S.M.PST.DEP-DS DEM.MED-CL:F-N.SBJ go + find-SS abita_fruit
mɨni + õhkwe-re-ña
go_up + slurp-N2/3S.PST.N.ASS-REP
'...(it, the fruit) was up there, so (they) climbed up (there) to find the *abita* fruit and to eat (it) up.' [NAT*: 20150811sfryi001.450-452]

b. *... iconi ja'që baquë caëña jamaca*
↑**ih-ko-ni** ha'-kɨ bah-kɨ
DEM.PROX-CL:F-N.SBJ2 parent-CL:M be-2/3S.M.PST.N.ASS
kaa-ɨ-ña hãmahka
say-2/3S.M.PST.N.ASS-REP then
'...then (he) said, "Dad was HERE." ' [NAT*: 20150811sfryi002.666]

A final naturalistic data point is provided in (51) to round off the present discussion regarding L-oriented DCM. The pair of sentences in (51) is lifted from the narrative sample in Chapter 5, where a naturalistic Q-A pair arises in dialogue. In his answer turn in (51b), the protagonist uses the promoted spatial pronoun ↑*ihko-**ni*** (this one (F) ≃ here), as expected:

(51) a. *jerore baquë më' huë'e ?*
he-ro-re bah-kɨ mɨ'ɨ̃ wɨ'e
WH-CL:LOC-N.SBJ be-2/3S.M.PST.N.ASS 2SG house
'Where was your house?'

b. *iconi baja'i huë'e, jemacare huë'e baisihko beo*
↑**ih-ko-ni** ba-ha'i wɨ'e hẽ-mahka-re
DEM.PROX-CL:F-N.SBJ2 be-3S.M.PST.ASS house DEM.DIST-DIM.SG-N.SBJ
↑[wɨ'e ba'i-sih-ko] beo-ɨ
house be-COMP-CL:F NEG.EXIS-S.M.PRS.DEP
'The house was HERE, (and now) the house that was there is gone.'
[NAT*: 20151023orocr001.127-128]

The data considered in this section has demonstrated that, once promotion is taken into account, P- and L-oriented DCM patterns display a full formal syncretism. It is also the case that a focus-related STRONG TRIGGER may be maintained for L-oriented DCM, as it is shown to arise in argumental DCM. The latter pattern also holds as concerns G-oriented DCM, discussed in what follows.

3.3.2 Goal-oriented DCM patterns

Like all other principled DCM patterns concerning inanimate arguments, G-oriented DCM displays an OPTIONAL-type DCM pattern where PLAIN zero-marking alternates with overt PROMINENT-marking, namely the case-marker *-na*. The G-argument refers

to the target location in directed motion predicates, e.g., *sai-* (go somewhere),[21] also found with change-of-posture predicates, e.g., *hã'ru-* (sit down somewhere). Additionally, certain three-argument predicates implicate a G-argument — e.g., *aja-* (fill X into Y); *saa-* (bring X to Y). Two prototypical instances of G-arguments are lifted from the sample in Chapter 5 in (52), illustrating both marking alternatives:

(52) a. *ja̲ë̲hua'iquë'ro saco'ë*
hã-ɨ-wa'i-kɨ'ro sah-ko'ɨ
DEM.MED-CL:M-PL.AN-LOC.DERIV go-3.S.F.PST.ASS
'(She) went to where they were.' [NAT*: 20151112orapi001.164]

b. *yëiñëquë'rona sanëca bã'quëña*
jɨi-ñi-kɨ'ro-na saa + nɨhka-a
cotton-CL:TREE-LOC.DERIV-GOAL bring + stand-NEG
bah-kɨ
be-2/3S.M.PST.N.ASS
'... (he) brought (them) to the kapok tree and stood (there).'
[NAT*: 20151023orocr001.237]

Although it is not clear from the sentences in (52), the emphatic contrast reading is associated with PROMINENT-marking, as expected, on the basis of elicited examples as in (53). In this context of a CONTEXT-CONJURING task (e.g., van der Wal 2016), the speaker provides a context in which PROMINENT-marking is deemed appropriate where the particular location contrasts with another alternative location:

(53) [**Suggested context:** A child is in trouble and is going to lie in the hammock. The mother does not want them to lie there, she instructs them to sit next to her.]
ja'rujë'ë̲ iñona !
ha'ru-hɨ̃'ɨ̃ **ihño-na**
sit_down-IMPER here-GOAL
'Sit HERE (i.e., not anywhere else) !'

a. *# ja'rujë'ë̲ iño !*
('Sit here !' ; no contrastive reading) [VOL: 20230530ejabi001.002a-b]

As a definitional component of any principled DCM pattern in the language, obligatory PROMINENT-marking is found on both the WH-item, and the focalized G-argument in the answer of a Q-A pair configuration:

(54) a. QUESTION: *jerona maibi saiye ?*
he-ro *(-na) mai-bi sai-je
WH-CL:LOC-GOAL 1PL.INCL-SBJ go-N2/3S.PRS.N.ASS
'Where are we (INCL) going?'

b. ANSWER: *saiyë ye'yahuë'ena*
sai-jɨ **je'ja + wɨ'e #(-na)**
go-N3S.PRS.ASS learn + house-GOAL
'(We) are going to the SCHOOL (lit. teaching house).'
[VOL: 20230619elupa002.018a-b]

[21] In Siona, manner-of-motion predicates have the same argument structure as plain directed motion predicates: e.g., *wɨ'wɨ-* (run), *tõme-* (fall).

The data considered thus far demonstrates that G-oriented DCM displays the principled PLAIN-PROMINENT alternation, characteristic of principled DCM patterns in the language. However, when considering the class of promoted nominals, DCM alternations are suppressed. Rather, promoted G-arguments display obligatory *-na* marking. This fact fits neatly into the analysis developed in this chapter, since animate, non-SUBJECT nouns require overt case-marking; however, crucially, the case-marker *-na* simply lacks a PROMINENT counterpart (i.e., unlike *-re*, which alternates with *-ni*).

The obligatory *-na* marking hypothesis for promoted G-arguments is borne out empirically. Firstly, consider the naturalistic instance of (obligatorily) *na*-marked promoted spatial pronouns in (55) as a preliminary illustration. No unmarked promoted spatial nominals are attested in my textual research. Next, consider the elicited instance reported in (56) — related to (46) in the previous section, which confirms the ungrammaticality of zero-marking on the promoted modified nominal via elicitation:

(55) [**Context:** The demon *cuẹhuati* smells the protagonist hiding in the tree.]
. . . jaọna dahuë'huësicobi guịnaobi se go'ico ñacona jạ icona mëco'ë naso caoña
↑hã-o-na daa + wɨ'wɨ-sih-ko-bi
DEM.MED-CL:F-GOAL bring + run-COMP-CL:F-SBJ
gwĩna-o-bi se go'i-ko
the_one_who_does_it_again-CL:F-SBJ SE return-S.F.PRS.DEP
ñaa-ko-na hã **↑ih-ko-na** mɨh-ko'ɨ
see-S.F.PRS.DEP-DS DISC.PART DEM.PROX-CL:F-GOAL go_up-3S.F.PST.ASS
nahso kaa-o-ña
wooly_monkey say-2/3S.F.PST.N.ASS-REP
'. . . (she) came back running to there — (she) returned and looked around, "aha! the monkey climbed up HERE", (she) said.'
[NAT*: 20151112oespa001.101]

(56) *saiyë yë' ñicuëbaị huë'e nesicona*
sai-jɨ **↑[jɨ'ɨ ñihk^{w}-ɨ-bãĩ wɨ'e nee-sih-ko *(-na)**]
go-N3S.PRS.ASS 1SG grandparent-CL:M-PL.AN house make-PERF-CL:F-GOAL
'(I/we) am going to the house (or HOUSE) that my grandparents built.'

a. # *saiyë yë'ë ñicuëbaị huë'e nesico-**re*** (L-marking is inappropriate here)

b. * *saiyë yë'ë ñicuëbaị huë'e nesico*
(zero-marking unavailable on ↑inanimate)

[VOL/SUG: 20241003ejabi001.006a-b]

Given that, unlike *-re* and *-ni*, the marker *-na* does not alternate with a formally distinct PROMINENT-alternative, *-na* obligatorily marks the promoted G argument in (56). On the paradigmatic view espoused in this dissertation, inanimate and animate (i.e., promoted) G arguments display case-marker syncretism. This is demonstrated in Table 3.8:

	INANIMATE		↑INANIMATE	
	PLAIN	PROMINENT	PLAIN	PROMINENT
LOCATION (L)	-∅	*-re*	*-re*	*-ni*
GOAL (G)	-∅	*-na*	*-na*	

Table 3.8: Mapping PLAIN and PROMINENT marking alternatives for the R-argument

On the basis of Table 3.8, once promotion is adequately considered, the generalization that non-SUBJECT arguments of the animate class require overt case-marking is maintained. Spatial DCM patterns display the same conspiracy of animacy status and information structure STRONG TRIGGERS outlined for (non-SUBJECT) argumental DCM above.

3.3.3 Extended uses of the Goal-marker, -na

In the spirit of completing the *preliminary description* of Siona case-marking in this chapter, a brief discussion is in order regarding certain non-canonical uses of *-na* marking. These observations complement the recognition of certain restricted uses of *-na* on inanimate R-arguments, a component of the accidental DCM pattern outlined in Section 3.2.3. This subsection discusses two patterns in turn, atypical spatial *-na* marking, and what I label pseudo-spatial *-na* marking.

As concerns atypicial *-na* marking in the spatial domain, it is observed that some spatial nouns, which thematically align with the STATIC LOCATION-argument (i.e., L), accept *-na*. The distribution of *-na* on L-arguments is constrained by the predicate at hand. For instance, certain predicates, such as *je'je-* (learn, study), shown in (57), permit *-re* and *-na* on the place of study. Speakers report subtle meaning differences, as shown below; however, more work is required to fully unpack the meaning difference and the range of predicates which display this pattern. For the present purposes these are conceived as an accidental DCM alternation:[22]

(57) a. *iye mo'se iye ye'yahuë'ere ye'yeja'quëa'ë* [23]
i-je mo'se **i-je** **je'ja + wɨ'e-re**
DEM.PROX-CL:GEN day DEM.PROX-CL:GEN teach + house-N.SBJ
je'je-hã'-kɨ-a'-ɨ
learn-PRP-CL:M-COP-N3S
'Today (I (M)) will study at the school.'
[**Speaker comment:** with *re*-marking a more specific location is signaled; i.e., the learning occurred in a particular room.]

[22] Martine Bruil suggests that *-na* may be acceptable here due to the purposive morpheme, *-hã'-*, which contributes the future-oriented reading to the examples in (57). However, *-na* is available in other temporal conjugation patterns.

[23] Note that this is not the same type of *specificity* suggested for the marker *-re* elsewhere. Rather specificity as a trigger for DCM alternations is suggested to be relevant to both *-re* on L and *-na* on G equally, see Table 3.6.

b. *iye mo'se iye ye'yahuë'ena ye'yeja'quëa'ë*
i-je mo'se **i-je** **je'ja + wɨ'e-na**
DEM.PROX-CL:GEN day DEM.PROX-CL:GEN teach + house-GOAL
je'je-hã'-kɨ-a'-ɨ
learn-PRP-CL:M-COP-N3S
'Today (I (M)) will study at the school.'
[**Speaker comment:** with *na*-marking a more general location is signaled; i.e., not any particular room.]

c. ✓ *iye mo'se iye ye'yahuë'e ye'yeja'quëa'ë*
[VOL/SUG: 20230619elupa002.016a-c]

Whereas some predicates allow either *-re* or *-na*, with a subtle semantic effect, as in (57), a second set of predicates select *-na* more rigidly on their L-like arguments. This holds for a heterogeneous class of predicates — e.g., *gãjo-* (play), *guja-* (bathe), *ai de'o-* (grow), *bɨa-* (stay), *goa nee-* (work), *kã-* (sleep),[24] which systematically disfavour *re*-marking on the L-argument. Illustrative corpus examples with the predicate *bɨa-* (stay) and *ai de'o-* (grow) are provided in (58) and (59) accordingly. The latter instance involves a promoted nominal, such that *-na* marking ought to be considered obligatory, for the reasons signaled above:

(58) *ĩ jamu ti ja'o dasiquë jare sa'nihuëna bëaëña*
ĩ-ɨ hãmu ti ha'o daa-sih-kɨ hãre **sa'niwɨ-na**
3PRO-CL:M armadillo ANA mud bring-COMP-CL:M like_that inside-GOAL
bɨa-ɨ-ña
stay-2/3S.M.PST.N.ASS-REP
'The armadillo, who brought the mud, stayed inside (the river) (so they say).'
[NAT*: 20151023orocr001.098]

(59) *. . . sa'nihuë jai cua'coro sɨ̃usicona aide'ouña*
↑[**sa'niwɨ hai kwa'ko-ro** **sĩũ-sih-ko-na**] ai
inside big cook-CL:RECEP cover-COMP-CL:F-GOAL big
+ de'o-ɨ-ña
+ become-2/3S.M.PST.N.ASS-REP
'. . . (they (i.e., referring to eggs)) grew inside the big covered pot.'
[NAT*: 20150811sfryi001.147]

In order to confirm that the usage of *-na* is forced by the predicate in such instances, consider the Q-A pair presented in (60). In this instance the predicate *bɨa-* (stay) is used, and, as such, PROMINENT *-na* marking is required on both the WH-item and the focalized L-argument in the answer:

(60) a. QUESTION: *jerona mësarubi bëaye iye ñami ?*
he-ro *(-na) mɨhsaru-bi bɨa-je i-je
WH-CL:LOC-GOAL 2PL-SBJ stay-N2/3S.PRS.N.ASS DEM.PROX-CL:GEN
ñahmi
night
'Where are you (PL) staying tonight?'

[24] A similar list is mentioned in an end-note regarding Koreguaje [*coe*], by Cook and Levinsohn (1985, end-note 10): i.e., they mention 'sleep', 'dawn', 'rest', and 'cook', and indicate that there are others.

i. * *jero-**re** mësarubi bëaye iye ñami ?*
(Intended: 'Where are you (PL) staying tonight?')

b. ANSWER: *mai bëayë më' gajei huë'ena*
mai bɨa-jɨ [**mɨ'ɨ̃ gahe-ɨ wɨ'e *(-na)**]
1PL.INCL stay-N3S.PRS.ASS 2SG friend-CL:M house-GOAL
'We (INCL) are staying <u>at your friend's house</u>.'

i. * *mai bëayë më' gajei huë'e-**re***
(Intended: 'We are staying at your friend's house.' [answers (54a)])
[VOL: 20230619elupa002.031a-c]

To summarize, it can be stated that certain L-arguments are encoded as G-arguments, as determined by the selectional properties of the predicate at hand. More work is needed to determine the exact set of predicates which display this atypical spatial case-marking; however, as will be demonstrated in Chapter 4, other languages in the region display similar patterns, based upon a similar set of predicates. Ultimately, this quirk is treated as an accidental sub-pattern of L-oriented DCM.

The final case-marking pattern which concerns this *preliminary description*, regards so-called pseudo-spatial uses of *-na* marking. This refers to certain thematic roles which are obligatorily case-marked via the case-marker *-na*: i.e., (i) the Place-of-Putting argument, and (ii) the Point-of-Contact argument. A natural instance of the former is presented in (61), which displays the predicate *hẽña-* (stick X <u>to Y</u>) in (61). The latter is presented in (62), with the predicate *to'te-* (sting X <u>on the Y</u>):[25]

(61) *<u>i</u>o be<u>j</u>etubëna ja'o je<u>ñ</u>aquë ba<u>'</u>quëña*
ĩ-o **bẽhẽ-tubɨ-na** ha'o hẽña-kɨ
DEM.PROX-CL:F trunk-CL:BRANCH-GOAL clay stick-CL:M
ba-∼'-kɨ-ña
be-RMT.PST-2/3S.M.PST.N.ASS-REP
'He stuck clay <u>onto the tree trunk</u>.' [NAT*: 20151023orocr001.061]

(62) [**Context:** an evil child spirit turns into a wasp in order to kill a tapir.]
cacani <u>i</u>re sa'nahuë ahuëna to'tehueaña ba<u>'</u>ë <u>i</u>re yi'yebi
kahka-ni ĩ-ɨ-re + sa'nawɨ **a-wɨ-na**
enter-SS DEM.PROX-CL:M-N.SBJ + inside heart-CL:CONT-GOAL
to'te + wẽã-a ba-ɨ'ɨ ĩ-ɨ-re ji'je-bi
sting + kill-NEG be-N3S.PST.ASS DEM.PROX-CL:M-N.SBJ wasp.species-SBJ
'(He) entered inside him, and the wasp stung him <u>on the heart</u>.'
[NAT*: 20150811sfryi001.541]

These patterns are noted for the sake of descriptive completeness in the present chapter. However, the pseudo-spatial instances of *-na* marking cannot be treated as an extension of the G-oriented DCM patterns — i.e., unlike the atypical L-marking patterns above, since the case-marker is obligatory on the relevant arguments in instances such as (61) and (62). Ultimately, certain uses of the case-marker *-na* are attested beyond strict spatial GOAL-marking. Similar extended uses of this marker and its cognates are attested across Western Tukanoan languages, as will be shown in the next chapter.

[25] We provide a more elaborate of some of these complex predicates in van Gijn et al. 2023, in a comparative perspective.

3.4 Interim summary: Siona DCM as a paradigm

The *preliminary description* laid out in this chapter primes the reader for the more targeted studies reported in the ensuing chapters of this dissertation. Table 3.9 schematizes the primary DCM facts outlined in this chapter, including the mapping of PLAIN- and PROMINENT-alternatives across all principled DCM patterns, and the SUBJECT vs. non-SUBJECT split, which underpins the animacy-based marking split:

	INANIMATE		ANIMATE (∧ ↑INAN)	
	PLAIN	PROMINENT	PLAIN	PROMINENT
Argumental DCM:				
SUBJECT (S)	-∅	*-bi*	-∅	*-bi*
DIRECT OBJECT (P)	-∅	*-re*	*-re*	*-ni*
INDIRECT OBJECT (R)	*-re/-na*		*-re*	*-ni*
Obligatory case-marking:	*-bi* (Instrument); *-hã're* (Companion)			
Spatial DCM:	INANIMATE		↑INANIMATE	
LOCATION (L)	-∅	*-re*	*-re*	*-ni*
[*some L] GOAL (G)	-∅	*-na*	*-na*	
Obligatory case-marking:	*-bi* (Source); *-hã'ã* (Path, Limit)			

Table 3.9: Siona case-inventory, incl. (principled) DCM patterns

The remainder of this dissertation addresses several components of the complex case-marking system described here, such that each chapter applies different methodologies to further the description. Though the primary objective is descriptive, theoretical insights are postulated where appropriate.

Chapter 4 takes the *preliminary description* established in the present chapter, and situates several of these case-marking properties among the case-marking patterns observed in other Tukanoan languages and in other languages across the putative NWA area (defined in Section 1.3.2). This chapter discusses the results of a comparative survey comprising a modest number of languages in the region, as developed via the implementation of a typological questionnaire, presented in full in Appendix A. This chapter provides various insights into the diachronic development of Siona DCM.

Chapter 5 (along with Appendix B) reports a corpus-based investigation regarding the distribution of case-marking alternatives in a sample of Siona narratives. In addition to establishing the marking-rate for S-, P-, and L-oriented DCM patterns in a naturalistic sample, a series of *variable-rule* analyses are performed on a pool of nominal tokens. These analyses contribute to the determination (and ranking) of TRIGGERS for the concerned DCM patterns. This is an important contribution to the characterization of various plausible TRIGGER candidates, necessary for a complete description of multidimensional Siona DCM.

Finally, Chapter 6 reports the findings of an elicitation-driven research program, exploring the focus effects identified in this chapter in greater detail. Given that focus-encoding varies considerably from language-to-language (Büring 2009; van der

Wal 2011, 2015, 2016, 2022; Kratzer and Selkirk 2020; etc.), this chapter contributes considerably to the description of focus-triggered DCM patterns in Siona. Of course, in light of patterns like Q-A heuristics, a proper understanding of focus effects has implications for the development of an adequate formal account of focal DCM patterns. Together these types of novel evidence achieve a holistic description of Siona DCM.

CHAPTER 4

A comparativist study: Siona case-marking in typology

For the present chapter, I put on my *comparativist* hat in order to situate the case-marking facts described for (Ecuadorian) Siona in Chapter 3 among those described for languages spoken throughout a proposed northwestern Amazonian (NWA) language area. To date, regarding Tukanoan languages in this region, the literature has privileged the better-described Eastern branch (cf. Duarte and de Oliveira Lopes 2020; Gomez-Imbert 2011; Ramirez 2019[1997], 164-170; Stenzel 2008, 2013b, 2013c; Zúñiga 2007; among others).

With the express aim of placing the Siona case-marking facts in their regional context, I have developed the NWA case-marking dataset, which underpins the discussion in the present chapter. The dataset is presented in full in Appendix A.

This chapter is organized as follows: Section 4.1 outlines the methodology used to develop the NWA dataset, and how the dataset is used for comparison. Section 4.2 describes the most interesting typological insights revealed by the NWA dataset. This discussion is compartmentalized according to the FORMAL, FUNCTIONAL, and REALIZATIONAL domains of inquiry for case-marking and DCM patterns, as conceived in Chapter 2 of this dissertation. Section 4.3 closes the chapter by integrating the findings into the broader typological context and by considering their possible diachronic implications.

Siona in its genetico-areal context: The NWA dataset

The NWA dataset consists of information regarding dependent marking in 21 languages of the region, plotted in the map in Figure 4.1.[1] In the map below, and in the

[1]The languages are identified by their three-letter ISO codes: Siona [*snn*], Sekoya [*sey*], Koreguaje [*coe*], Máíhɨ̄kì [*ore*], Kubeo [*cub*], Tukano [*tuc*], Wanano [*gvc*], Desano [*des*], Barasano [*bsn*], Tanimuka [*tnc*], A'ingae [*con*], Shiwiar [*acu*], Tena Kichwa [*quw*], Imbabura Kichwa [*qvi*], Awa Pit [*kvi*], Murui [*huu*], Bora [*boa*], Tariana [*tae*], Kakua [*cbv*], Hup [*jup*], Yuhup

ensuing discussion, Western Tukanoan languages are identified with the symbol ♦, Eastern Tukanoan languages by ◊, and the Ecuadorian Siona language, targeted in this dissertation, with an asterisk (*). The ISO code for Ecuadorian Siona is presented as ([E]*snn*) to distinguish this language from the Colombian Siona language ([C]*snn*) described by Wheeler (1962, 1970, *et seq.*).[2] The Vaupés cultural-linguistic contact zone, discussed in Section 1.3.2, is traced roughly with a dotted gray line. The relevance of this sub-area for the typological discussion in this chapter is made explicit in what follows.

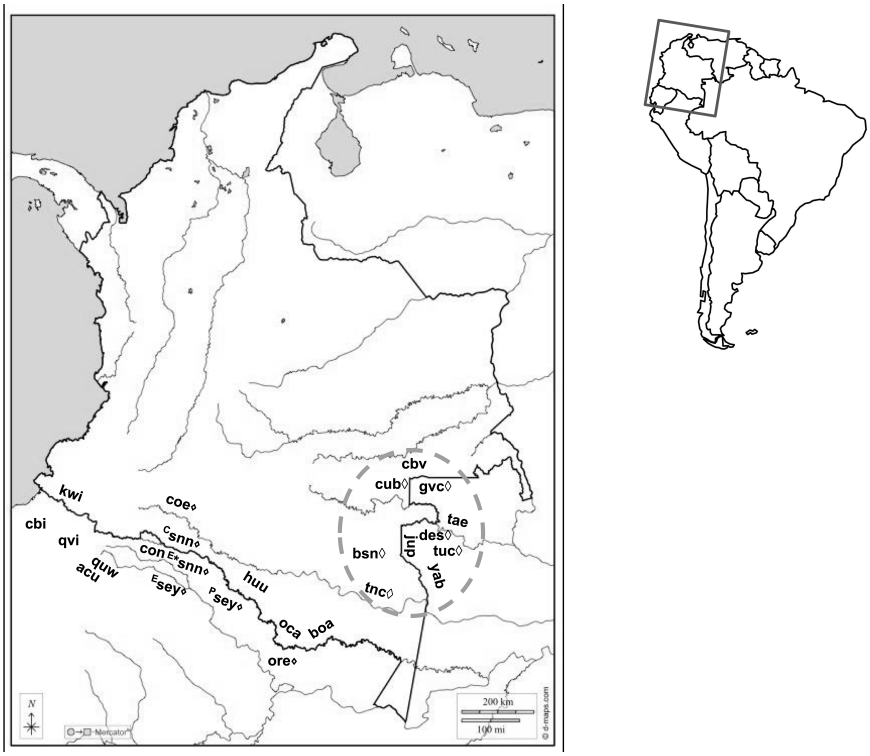


Figure 4.1: Geographical distribution of languages . Map template downloaded from *https://d-maps.com*.

For the sake of this chapter, the following languages may be viewed as forming a continuum: Ecuadorian Siona [[E]*snn*], Colombian Siona [[C]*snn*], Ecuadorian Sekoya [[E]*sey*], Peruvian Sekoya [[P]*sey*], and Koreguaje [*coe*]. Máíhɨ̃kì [*ore*] is the sole Western Tukanoan not considered to belong to this continuum on the basis of several notably different case-marking properties. The continuum is referred to collectively as the Siona+ languages in this chapter, such that one of its objectives is to outline certain ways in which the case-marking systems across Siona+ languages converge and diverge.

[*yab*]. For more information on the languages, see Table A.1 in Appendix A.

[2] I utilize the same convention to distinguish Ecuadorian Sekoya ([E]*sey*) from Peruvian Sekoya ([P]*sey*) in Figure 4.1.

The language sample was chosen to facilitate the comparison with the Siona+ facts in several respects. Firstly, the reported dataset comprises several Western Tukanoan languages, some of which are shown to display case-marking properties that are remarkably similar to those of Siona+. Secondly, the sample includes several non-Tukanoan languages, both from the Vaupés region, where they are spoken alongside the Eastern Tukanoan languages, and from the vicinity of the Western Tukanoan languages. On the one hand, there is ample evidence for linguistic and cultural diffusion within the Vaupés region (cf. Aikhenvald 1999, 2001; 2003; Barnes 1999, 2006; Epps and Michael 2017; van Gijn and Muysken 2020; Stenzel 2008, 2013c; Zúñiga 2007). This includes convergence on the basis of various case-marking properties and DCM patterns. With regards to the Siona+ context, there is some preliminary evidence for contact-induced grammatical convergence with languages in and around the Napo-Putumayo area (Bruil 2015a; van Gijn 2014, 2017; van Gijn et al. 2023), and for some lexical diffusion (Jolkesky 2016; Piispanen 2021).[3] The NWA dataset permits the comparison of the Ecuadorian Siona case-marking facts against a genealogical and areal backdrop simultaneously, contributing to the broader enterprise of tracing areal patterns in Amazonia.

4.1 Methodology: Case-marking questionnaire

The primary goal of linguistic typology is to identify structural similarities and differences across languages. In order to facilitate cross-linguistic comparisons, typologists must navigate descriptions of specific languages to extract comparable features. As such, an important component of typological research is the development of COMPARATIVE CONCEPTS (Bickel 2007, 2009; Croft 2001, 2003, 2009; Haspelmath 2005, 2007, 2008, 2010; Newmeyer 2002). With that objective in mind, the discussion found in this chapter abstracts away from certain phenomena which are very specific to Ecuadorian Siona DCM patterns, but which are not appropriate for comparing this language to others in an informative way. For instance, it is not clear whether the phenomenon of 'promotion', described in Section 3.2.2, is even relevant to case-marking in Siona+ languages, and is therefore set aside here.

The NWA dataset consists of the answers to a questionnaire for each of the languages in Figure 4.1. The structural questionnaire is a tool implemented by many typologists (e.g., Haspelmath 2005, 2008; Krasnoukhova 2012, Birchall 2014; Janic and Haspelmath 2023). This approach permits the researcher to unify generalizations from disparate languages by developing meaningful, abstract questions and constraining the possible answers. The full case-marking questionnaire and the instructions for extracting data from descriptive sources to establish the NWA dataset are spelled out in Section A.1.2 of Appendix A.

According to Birchall (2014, §2.2.2), although structural questionnaires vary greatly with respect to length and details, the resulting comparative datasets ought to be conceived as being organized into the three layers in (1):[4]

[3] Along the Middle Putumayo, several authors provide evidence for a cultural-linguistic contact zone, which is typically called the 'People of the Centre' (cf. Echeverri 1997; Seifart 2005, 2007, 2007, 2009, 2011). This refers to a narrower region within the NWA area than the non-Vaupés sample established in the present NWA dataset. Nonetheless, some of these languages are studied in this sample: i.e., Murui [*huu*] and Bora [*boa*].

[4] Birchall in fact suggests the terms DOMAIN, FEATURE, and VALUE; however, I avoid the

(1) Domain > Question > Value

Domains refer to bundles of related Questions, which are in turn attributed a set of possible Values (i.e., answers to the Questions comprising the questionnaire). The case-marking questionnaire developed to formulate the nwa dataset is divided into three Domains, which align with the three primary descriptive dimensions outlined in Chapter 2 of this dissertation:

- The Formal domain pertains to members of the case-marker inventory and to properties relating to the expression of case-marking morphology ;
- The Functional domain pertains to the distribution of formal case-markers across case-marking grammatical relations (i.e., Case functions) ;
- The dcm domain investigates the Triggers influencing dcm alternations and the formal types of alternations attested in the sample.

By partitioning the questionnaire into these domains, a more nuanced typology of case-marking and dcm properties emerges. For instance, it is conceivable that the Ecuadorian Siona case-marking facts may align with sampled Tukanoan languages regarding certain formal properties, but not regarding their distribution across functions, or other dcm-related properties. These comparisons fall out from the nwa dataset by design.

4.2 Typologizing Siona case-marking

In this section I place Ecuadorian Siona case-marking within the typology that emerges from the nwa dataset. The discussion laid out here does not go into every interesting pattern revealed by the dataset, but rather emphasizes the patterns which shed light on Ecuadorian Siona in its genealogical or areal context.

Each of the three case-marking Domains is discussed in a separate subsection.

4.2.1 Typologizing Siona case: The Formal domain

As shown in Chapter 3, Ecuadorian Siona has six case-marking morphemes in complementary distribution: i.e., *-bi*, *-re*, *-ni*, *-na*, *-hã'ã*, and *-hã're*. Three traits of the inventory of the languages in the sample are compared in this section: (i) the size of the inventory, (ii) the members of the inventory, and (iii) the ban against multiple case-marking.

4.2.1.1 The case-marker inventory (size)

The case-marker inventory size is measured directly in the nwa dataset (i.e., Question *1.1*), replicating data collection guidelines in Iggesen 2013, (wals-49). The inventory size of sampled languages ranges from the very small, two-member inventory attested in Máíhĩkì [*ore*], to the very large, nine-member inventories attested in A'ingae [*con*] and Awa Pit [*kwi*]. The six-member case-marker inventory attested in Ecuadorian Siona roughly aligns with the mean of 5.43.

term Feature for the present discussion given its several uses in linguistic theory. In this chapter, and in Appendix A, I utilize the terms Question, and presume that the implications are connected to typological "traits" or "properties" for the discussion in this chapter.

The distribution of inventory sizes is plotted onto the map in Figure 4.2. Inventory size groups and their labels are developed following the types suggested in Velupillai (2012, 176, footnote 112):[5,6]



Figure 4.2: Areal distribution of inventory size types (*Q1.1*) across NWA sample { SMALL (2-3 members); MEDIUM-SIZED (4-5) ; LARGE (6-7) ; VERY LARGE (8-9) }

[5]Note that the case-marking questionnaire developed in this dissertation does not count zero-marking as a bona fide alternative, which is included to calculate inventory size. Velupillai 2012, based upon the Iggesen 2013 original, includes zero-marking. Therefore, to compare directly with these works, the VALUES presented here ought to be considered as $N + 1$.

[6]The language specimens within the NWA dataset are grouped as follows, with the number of members in each inventory indicated between parentheses, (N):
SMALL INVENTORY: Máíhɨ̄kì♦ [*ore* (2)] ; MEDIUM-SIZED INVENTORY: Koreguaje♦ [*coe* (4)], Tukano◊ [*tuc* (4)], Wanano◊ [*gvc* (4)], Desano◊ [*dsn* (4)], Barasano◊ [*bsn* (4)], Kakua^{V} [*cbv* (4)], Hup^{V} [*jup* (4)], Kubeo◊ [*cub* (5)], Shiwiar◊ [*acu* (5)], Bora◊ [*boa* (5)], Tariana^{V} [*tae* (5)]; LARGE INVENTORY: *Ecuadorian Siona♦ [^{E}*snn* (6)], Colombian Siona♦ [^{C}*snn* (6)], Murui [*huu* (6)], Tanimuka◊ [*tnc* (7)], Imbabura Kichwa [*qvi* (7)], Yuhup^{V} [*yab* (7)]; VERY LARGE INVENTORY: Tena Kichwa [*quw* (8)], A'ingae [*con* (9)], Awa Pit [*kwi* (9)].

The Siona+ languages display a LARGE-sized inventory,[7] with more members than are attested in other Tukanoan languages.[8] In fact, the six-member inventory in Ecuadorian Siona is larger than that described for Koreguaje [*coe*], for Máíhĩkì [*ore*], and for all Eastern Tukanoan languages, with the exception of Tanimuka [*tnc*], which has a seven-member inventory.

On the other hand, most non-Tukanoan languages spoken in the vicinity of Ecuadorian Siona display inventories with six or more members: i.e., belonging to the LARGE-sized or the VERY LARGE-sized type. These include the Murui language [*huu*], A'ingae [*con*], Awa Pit [*kwi*], and the Quechuan languages, Imbabura Kichwa [*qvi*] and Tena Kichwa [*quw*]. The bulk of these larger inventories are attested along the western reaches of the proposed NWA area. Conversely, the dominant inventory size found among the non-Tukanoan languages spoken within the Vaupés region is the MEDIUM-sized inventory.[9]

Based upon the above generalizations, the Ecuadorian Siona language aligns more closely with the languages in its surroundings than with the general Tukanoan pattern as regards case-marker inventory size. In general Tukanoan languages exhibit modest inventories of multi-functional case-markers with complex distributions (cf. Barnes 1999, 2006; Stenzel 2008, 2013d; Zúñiga 2007). Other languages within the NWA area display richer case-marker inventories, which generally include a subset of members which execute a dedicated FUNCTION. Ecuadorian Siona occupies an intermediate position with regards to the case inventory size typology.

4.2.1.2 The case-marker inventory (members)

Although the NWA case-marking questionnaire does not ask for the individual members of case-marker inventories, this section briefly considers the distribution of cognates for the Ecuadorian Siona case-markers in the Tukanoan family.

Table 4.1 directly compares the members of the Ecuadorian Siona case-marker inventory with their cognates in Western Tukanoan languages included in the dataset. The Ecuadorian Sekoya inventory is excluded from this table since its members are identical to those described for Ecuadorian Siona for all intents and purposes (Johnson and Levinsohn 1990, 45-51). Certain items listed in the table are identified a gray X, e.g., X, which indicates that they are excluded from the inventory size metric. These items do not meet the criteria as bona fide case-marker affixes in this study (see also Section A.1.2.1 in Appendix A), although they are relevant to the discussion below:

[7] This includes Ecuadorian Sekoya, according to Johnson and Levinsohn 1990, 45-51, which is described as having a six-member inventory like Ecuadorian Siona.

[8] The mean inventory size in the NWA sample, as restricted to Tukanoan languages, has **4.65 members**. This represents te MEDIUM-SIZED INVENTORY type as conceived here.

[9] Yuhup [*yab*] is an exceptional language with regards to inventory size — i.e., it exhibits a seven-member inventory. This is due in large part to the rich system of spatial case-marking where different markers are made available on the basis of deictic categories (Ospina Bozzi 2002, 159-161; 2013). A similar system is described for Awa Pit (Curnow 1997, 135-140).

*[E]Siona♦	[C]Siona♦	Koreguaje♦	Máíhɨ̃kì♦
-re (N.SBJ)	***-re*** (OBJ.FOC)	***-re*** (N.SBJ.SPEC)	***-re*** (NON.SUBJ)
-ni (N.SBJ2)	*-ni* (OBJ.EMPH.FOC)	*-ni* (PAT)	N/A
-bi (SBJ,SRC,INST)	*-bi* (SBJ.FOC,ABL,INST)	*-pi* (SBJ,SRC,INST)	N/A
-na (GOAL)	*-na* (GOAL.FOC)	*-na* (GOAL)	N/A
-hã'ã (PATH,LIMIT)	*-ja'an* (PATH)	[X]*jatëca* (LIM)	N/A
-hã're (COM)	[X]*naconi* (COM)	[X]*ja'me* (COM)	*-jànù* (COM,INST)
	[X]*tëca* (LIMIT)		
[X]*=kato* (TOP)	*-ga* (SBJ.EMPH.FOC)		[X]*-ga* (TOP)
($N = 6$)	($N = 6$)	($N = 4$)	($N = 2$)

Table 4.1: Cognate candidate sets across the Western Tukanoan languages

The first relevant observation on the basis of Table 4.1 concerns the four case-markers identified in the first four rows. These comprise what is labeled the *rich case-marking* system in Section 1.3.1 of the introductory chapter to this dissertation. This set of markers is found across the Siona+ continuum and they display remarkably similar distributional properties, including DCM properties, which are touched upon at a later point in this chapter. In this chapter, I refer to this set as the Siona+ primary case-markers.

Besides the primary case-marker set, Table 4.1 identifies certain cognates to the case-markers *-hã'ã* and *-hã're* across the Western Tukanoan languages. Wheeler (1987, 132-133) describes a cognate marker *-ja'an* in Colombian Siona, but this marker only displays the PATH function, as shown in (2a). A separate marker is recruited to encode the LIMIT function in this language, as in (2a), which finds a cognate in Koreguaje [*coe*], *jatëca*, illustrated in (3). No corresponding marker is described for Máíhɨ̃kì [*ore*]:

(2) Colombian Siona♦ ([C]*snn* — WESTERN-TUKANOAN)
[W87:112;W70:76]

a. ***sabi-ya ma'a-ja'an*** *gani-ni ti'an-huë*
S-CL:RIV path-ALONG go-SEQ.SS arrive-PST
'(They) went along the Sabiya gorge path and arrived (there).'

b. ***Pedro wɨ'e-se'e-ga tɨ́ká*** *saí-yɨ*
P house-EXCL-GA LIM go-N3S.PRS.ASS
'(I) am going only toward Pedro's house.'

(3) Koreguaje♦ (*coe* — WESTERN-TUKANOAN)
[Cook 1990, 98, *Viaje a Bogotá*.033]

ɯmɯ-ca caje-cheja jatɯ'ka *mɯ'ɯ-re te'e-ɯ-re sesenta*
high-stay land-place up_to 2SG-N.SBJ.SPEC one-M.SG-N.SBJ.SPEC sixty
pesos ro'i-mɯ chii-si'-kɯ-a'-mɯ
pesos pay-M.SG say-PERF-NMLZ.M.SG-VBLZ-M.SG
' "It will cost you $60 pesos to go only up to the airport", (he) told (me).'

The case-like LIMIT-markers shown in (2b) and (3) are not fully integrated into the corresponding nominal words. This criterion suffices to exclude these markers as members of the respective case-marker inventories in the NWA dataset. Importantly, similar

conclusions are drawn in the descriptive sources — i.e., Wheeler (1970, 76; 1987, 132-133) identifies these markers as *function words* in Colombian Siona, whereas Cook and Criswell (1993, 51-52) postulate an independent class of *post-positions*. All authors classify such markers separately from bona fide case-markers on various grounds. The present chapter labels these morphemes as 'pseudo-case-marker'.

According to Table 4.1, Colombian Siona and Koreguaje also recruit pseudo-case-markers to encode the COMPANION function — i.e., separately from the INSTRUMENT function, which is encoded via cognates of *-bi* across the Siona+ group. Colombian Siona [^{c}*snn*] has a marker *naconi*, shown in (4a), whereas Koreguaje [*coe*] exhibits a clear cognate with Ecuadorian Siona *-hã're* (i.e., *jɑ̲'me*), shown in (4b). Notably, another cognate candidate is attested in Máíhɨ̃kì [*ore*], namely the case-marker *-jànù*, illustrated in (4c). However, this final marker encodes both COMPANION and INSTRUMENT functions (e.g., Neely 2012):

(4) a. Colombian Siona♦ (^{c}*snn* — WESTERN-TUKANOAN) [W70:76]
wa'í maú-si'-kɨ́ ***nakóni*** *koká* *ká-wɨ́*
fish catch-COMP-CL:M COM word say-N3S.PST.ASS
'(I) conversed with the one who had caught the fish.'

b. Koreguaje♦ (*coe* — WESTERN-TUKANOAN) [CC93:51]
i-na-re ***jã'me*** *pa-sa* *ɟ̌ii-mɨ́* *ɟ̌ɨ'i*
this-PL.ANIM-N.SBJ COM live-DES.SG want-M.SG 1SG
'(I (M)) want to live with them.'

c. Máíhɨ̃kì♦ (*ore* — WESTERN-TUKANOAN) [Neely 2012, ex. (7) in BERK12]
sáí-yí *Tótóyà* ***Mámàsò-jànù***
go-1SG.PRS T M-COM
'I am going to Totoya with Mamaso.'

The examples shown in (2) through (4) demonstrate that an element, /*hã*/, is the basis for several cognates within the Western Tukanoan languages in the sample. This fact is independent of whether or not a given marker has become fully grammaticalized as a bona fide case-marker in a particular language. For instance, *jã'me* co-occurs with the case-marker *-re* in (4b), which is other disallowed in Siona+ grammars. The fact that a cognate marker is also attested as the instrumental-comitative marker in Máíhɨ̃kì, shown in (4c), points to an early shared innovation in the Western Tukanoan branch; which is unrelated to the development of the Siona+ primary case-markers.

This hypothesis finds support in the fact that no plausible cognates are noted among the Eastern Tukanoan languages, though a very similar marker is attested in Shiwiar [*acu*] — =*hʲãĩ̃* (COM,INST), as described by Kohlberger (2020, 263-264).

Whereas the case-marker *-re* is common to all Tukanoan languages, the origins of the remaining members of the set of primary case-markers remains unclear (i.e., *-ni*, *-na*, and *-bi*).[10] Table 4.2 recognizes a handful of cognate candidates from non-Siona+ languages in the NWA dataset, particularly from Tanimuka [*tnc*], Kubeo [*cub*], and the Witotoan language, Murui [*huu*].

[10]It is relevant to note that the case-marker *-ni* is not attested in Peruvian varieties of Sekoya (Rosa Vallejos, p.c.), spoken further down-river in the Middle Putumayo region. However, cognates to *-bi* and *-na* are attested in this language, and they exhibit similarly complex distributions (e.g., Vallejos and Brown 2021).

	*[E]Siona♦	Tanimuka◊	Kubeo◊	Murui
	-*re* (N.SBJ)	**-*re*** (DEIC1)	**-*re*** (N.NOM)	N/A
	-*ni* (N.SBJ2)	N/A	N/A	-*ri* (PATH)
	-*bi* (SBJ)	-*ɸí* (INST, ABL)	N/A	N/A
	-*na* (GOAL)	-~*rá* (ALL)	-*rã* (FOC.LOC)	-*na* (N.S/A.TOP)
			-*ta* (PATH)	
	...	...	...	...
INVENTORY SIZE (*Q1.1*):	($N = 6$)	($N = 7$)	($N = 5$)	($N = 6$)

Table 4.2: Cognate candidates for Siona+ primary case-marker set

The case-marker -*ni* has a distribution that is tied closely to that of -*re* across the Siona+ languages. The Witotoan language Murui [*huu*] has a spatial marker -*ri* (PATH), also used to encode the Benefactor role (Wojtylak 2021, §8.1.2.4). The nearby Bora language [*boa*] also has a marker -*rì* (glossed OBL.IN, or 'inanimate oblique') which has spatial functions and encodes the INSTRUMENT-argument (Thiesen and Weber 2012, 290-297). Nearer to the Siona+ languages is the isolate Ainga'e [*con*], which has a spatial marker =*ni* (glossed LOC). It is unclear whether these markers are related; however, it is possible that a spatial marker developed into the case-marker -*ni* in Siona+ languages, acquiring its constrained distribution at a later time. Despite some uses of spatial -*ni* on promoted argumentas in Ecuadorian Siona, the dominant function of this marker is to encode animate, non-SUBJECT arguments, with pragmatic overtones.[11] Given that the marker -*ni* is not attested in Peruvian varieties of Sekoya (Vallejos, *p.c.*), spoken further to the east, in the Middle Putumayo region, it is possible that this marker represents a more recent innovation.

A plausible cognate candidate to Siona+ -*bi* is found in Tanimuka [*tnc*] — i.e., -*ɸí*. No other (Eastern) Tukanoan language displays a cognate to this term. In fact, the Tanimuka marker displays both INSTRUMENT- and spatial SOURCE-encoding functions (Eraso 2015, 280), as is noted with the Siona+ cognate. Focusing on its spatial function, the set of spatial case-markers identified for Tanimuka is similar to the spatial functions noted for the primary case-markers across Siona+ languages (*id.*, §6.5.4):[12]

(5) Tanimuka◊ (*tnc* — EASTERN-TUKANOAN)
[E15:210-211, exx. (151a)-(151c)]

a. ***wi'í-a-re'ká*** *ɟi-ĩbé*
maloca_house-N-LOC 1SG-COP
'I am <u>in the maloca house</u>.' [Siona+ L-marker: -*re*]

b. ***wi'í-a-~rá*** *ɟi-káka-ɟú*
maloca_house-N-GOAL 1SG-enter-PRS
'I am entering <u>the maloca house</u>.' [Siona+ GOAL-marker: -*na*]

c. ***wi'í-a-ɸi*** *ɟi-ɸó-ɟú*
maloca_house-N-SRC 1SG-leave-PRS
'I am leaving <u>from the maloca house</u>.' [Siona+ SOURCE-marker: -*bi*]

[11] It is also noteworthy that the Vaupés languages, Kakua [*cbv*] and Yuhup [*yab*] exhibit the markers =*di'* (OBJ) and -*dìh* (*accusatif*) respectively.

[12] In Tanimuka [*tnc*] that -*re* appears with axial nouns and pronouns, whereas the complex marker -*re'ka* arises on plain nouns representing static LOCATION-arguments

The final marker which concerns this discussion is Siona+ *-na*, which finds a likely cognate in Tanimuka, as shown in (5b). A nearly identical case-marker is described in the Kubeo [*cub*] spatial case-marking system — i.e., *-rã̃*, labeled the FOCAL LOCATIVE by Morse and Maxwell (1999, 120) and Chacon and Genetti (2019, 413). Another form described in Kubeo is the PATH-marker *-ta* (Morse and Maxwell 1999, 121-122), which comes closer to the meaning of Siona+ *-na*:

(6) Kubeo$^{\diamond}$ (*cub* — EASTERN-TUKANOAN) [MM99:122, ex. (392)]
ke *te-Rĩ̃* *ape-kɨ̃* *ɨ̃bĩ̃* *'bA-kɨ̃-RE*
like_that do-GER other-CL:M man be-NMLZ.N.FUT.M.SG-N.SBJ
xaro-RExa-Ibã̃-jA ***boa-I-Rõ̃-I-ta***
send-RMT.PST-3PL-REP kill-I-NMLZ.SG.INAN-LOC-PATH
'Doing it so, so they say, a long time ago they sent another man, who has died since then, up to that killing spot.'

Setting aside possible cognacy with other Tukanoan languages, it is relevant to acknowledge several markers which formally resemble the item *-na*: e.g., *-na* (OBJ.PRO) in Tariana [*tae*]; *-na* (N.S/A.TOP) in Murui [*huu*]; $=n^{j}a$ (OBJ) in Shiwiar [*acu*]; and perhaps even the accusative marker *-ta* in Quechuan languages and in Awa Pit [*kwi*]. Given that these markers all encode the DIRECT OBJECT (P), one might ask whether a common source may be claimed for all of these markers. Despite the relative ubiquity of case-markers resembling *-na*, a plausible diachronic source for the Siona+ marker, with likely cognates in Tanimuka [*tnc*] and Kubeo [*cub*], is its development from the motion predicate *da(i)-* (come), attested in each of these languages (cf. Wheeler 1992).

From the hypotheses put forward in this section regarding the history of Siona+ case markers, I consider particularly compelling that at least *-bi* and *-na* developed at a relatively early stage, accounting for the cognates recognized in Table 4.2 in Tanimuka [*tnc*] and in Kubeo [*cub*]. Most specialists agree that these languages belong to a unified Eastern branch of the Tukanoan family; however, these are precisely those languages which were once proposed to belong to a putative Central branch (e.g., Mason 1950; Waltz and Wheeler 1972; Malone 1987; Barnes 1999) — see discussion in **Section 1.3.1** of the introduction to this dissertation. This is not to suggest that the Central Tukanoan hypothesis be revived per se, since there are convincing motivations for its abandonment (cf. Stenzel 2013c; Chacon 2013, 2014, 2016; Chacon and List 2015). However, case-marking is not taken into account in these sources. Alternatively, it is possible that Siona+ groups and putative Central Tukanoan groups experienced contact with similar groups in the Putumayo region (e.g., Gomez-Imbert 2011; Chacon 2012; Rose et al. 2017; Arias et al. 2022).

4.2.1.3 Single versus multiple case-marking patterns

The presence or absence of the multiple case-marking phenomenon is represented explicitly in the NWA dataset (i.e., QUESTION *1.2*).[13] The multiple case-marking property is attributed to languages which allow multiple case marks to arise on a single nominal word. This pattern is attested in certain languages of the Vaupés region, including certain Eastern Tukanoan languages (cf. Stenzel 2008, 2013a), as shown for Desano [*des*] in (7) and for Kakua [*cbv*] in (8):

[13] A full presentation of the relevant facts is provided in Table A.7 in Appendix A.

(7) Desano$^{\diamond}$ (*des* — EASTERN-TUKANOAN) [M99:32, ex. (119)]
ero-ge-re *gɨa* *árĩ-bɨ*
there-LOC-N.SBJ 1PL.EXCL be-N3.PST
'We were there, at that place.'

(8) Kakua$^{\vee}$ (*cbv* — KAKUA-NUKAK) [B16:206, exx. (71)-(72)]

a. *webit* ***'ɨ'=hĩ'=di'*** *bû'jup=di'* *mâw=na=ka*
child blow_gun=INST=N.SBJ hummingbird=N.SBJ kill=DECL=ASS
'ĩ=tʃãhãp
3PL=PROG
'The children are killing hummingbirds with the blow gun.'

b. *hiw* *'ã=ñap=na=be* ***nin=bŭ=di'=hé'***
jaguar 3S.M=jump=DECL=REC.PST this=LOC=N.SBJ=INTENS
'The jaguar jumped right here.'

In the case of Eastern Tukanoan languages, multiple case-marking patterns are generally restricted to spatial case-marking contexts (Ramirez 2019[1997], 169-170; Miller 1999, 30-34). Similar patterns are described for Tariana (Aikhenvald 2003, 157-159). However, the Kakua pattern is slightly more permissive, as shown in (8a), where the INSTRUMENT bears the typical marker *=hĩ*, as well as the non-subject marker *=di'*. These patterns are limited to the Vaupés region in the NWA sample consulted here, and are likely the result of contact.[14]

This finding is reported in the current discussion in order to demonstrate that Ecuadorian Siona, and the other Siona+ languages, do not pattern with Eastern Tukanoan languages in this regard. In other words, Ecuadorian Siona exhibits the dominant NWA pattern where case-markers are in complementary distribution.

4.2.2 Typologizing Siona case: The Functional domain

In the *preliminary description* I noted how several Ecuadorian Siona case markers have complex distributions. More concretely, there is not always a one-to-one mapping from case-marker to CASE FUNCTION, such that certain syncretisms emerge. This section compares the Ecuadorian Siona FORM-FUNCTION mapping with the NWA sample. This section closes with a consideration of certain functional gaps — i.e., typical case-marking FUNCTIONS, attested in the area, which do not recruit overt case-markers in Ecuadorian Siona grammar.

4.2.2.1 Situating Siona case-marker correspondences

Table 4.3 traces the FORM-FUNCTION mapping for Ecuadorian Siona case-markers based upon the description laid out in Chapter 3. The upper portion of the table presents argumental CASE FUNCTIONS, corresponding to a mixed set of grammatical

[14] A related phenomenon pertains to the presence of *complex case-markers* — such as the Ecuadorian Siona comitative case-marker *-hã're*, which is presumably a combination of *-hã'*(*ã*) and *-re*. A similar marker is attested in Kubeo, for instance, where the PATH-marker requires the presence of a locative marker, *-i*, to host the dedicated marker *-ta*; an example is provided in (17) in Appendix A. Although the presence such forms may indicate multiple case-marking diachronically, this study finds that Vaupés-type multiple case-marking is synchronically rare in the NWA area.

functions and oblique relations, and the lower portion enumerates the attested spatial FUNCTIONS. The case-markers, *-hã'ã* and *-hã're* are presented in gray text since neither of these markers displays a complex distribution. The table confirms that all and only the Siona+ primary case-markers discussed in the previous section (i.e., *-re*, *-ni*, *-bi*, and *-na*) exhibit synctretic distributions. The FUNCTIONAL DOMAIN within the NWA case-marking questionnaire is developed, in part, to facilitate the comparison of similar complex case-marker distributions in the sample. In particular, QUESTIONS *2.1* through *2.5* implement the notion of case-marker correspondence, adapted from Stolz et al. (2013, WALS-52), in order to compare the distribution of case-markers across sets of CASE FUNCTIONS between sampled languages.

CASE FUNCTIONS	DISTRIBUTION OF FORMS
SUBJECT (S)	*-bi*
DIRECT OBJECT (P)	*-re*, *-ni*
INDIRECT OBJECT (R)	*-re*, *-ni*, *-na*
INSTRUMENT (INST)	*-bi*
COMPANION (COMP)	*-hã're*
LOCATION (L)	*-re*, *-ni*, *-na*
GOAL (G)	*-na*
SOURCE (SRC)	*-bi*
PATH	*-hã'ã*

Table 4.3: Ecuadorian Siona case-markers mapped onto FUNCTIONS

Stolz et al. (2013) identify three types of correspondences, which may arise across a given set of FUNCTIONS. Examples of each type, as well as the corresponding coding procedure, are outlined in Section A.1.2.2 of Appendix A:

- The IDENTITY-type uses the same markers across these FUNCTIONS ;
- The DIFFERENTIATION-type uses different markers across these FUNCTIONS ;
- The MIXED-type displays partial overlap in the markers recruited across both FUNCTIONS, but also additional markers are available to a certain FUNCTION.

The tested correspondences in the FUNCTIONAL domain of the questionnaire are selected based upon the patterns recognized in Table 4.3, described particularly for Ecuadorian Siona. The selected sets of FUNCTIONS for comparison are outlined in Table 4.4, which also identifies the corresponding QUESTION number and the respective Ecuadorian Siona correspondence types, to be compared in the ensuing discussion.

QUESTION	TARGETED FUNCTIONS	SIONA TYPE
Q2.1	SUBJECT-INSTRUMENT	IDENTITY type
Q2.2	DIRECT-INDIRECT OBJECT	MIXED type
Q2.3	COMPANION-INSTRUMENT	DIFFERENTIATION type
Q2.4	OBJECT-LOCATION	MIXED type
Q2.5	LOCATION-GOAL-SOURCE	MIXED type

Table 4.4: Tested case-marker correspondences in the NWA dataset

Turning firstly to SUBJECT-INSTRUMENT case-marking correspondences (*Q2.1*), Ecuadorian Siona displays the IDENTITY type. This type is rarely attested in the sample, according to the full dataset for this QUESTION, as presented in Table A.8. Most languages in the NWA sample do not display any SUBJECT-marking whatsoever, such that the dominant correspondence type is the DIFFERENTIATION type (N=17). Nevertheless, in all Siona+ languages, cognates of *-bi* encode both concerned FUNCTIONS (cf. Cook and Criswell 1993, 47-48, for Koreguaje; Wheeler 1970, 74, for Colombian Siona; Johnson and Levinsohn 1990, 45-46, for Ecuadorian Sekoya). This pattern is demonstrated by the pair of Koreguaje [*coe*] sentences in (9):[15]

(9) Koreguaje♦ (*coe* — WESTERN-TUKANOAN) [CC93:47-48]

a. ***jɨ̀ʔɨ́-pi*** *ǰo'o-ra*
1SG-SBJ do-DES.SG
'I (M) (i.e., nobody else) will do (it).'

b. ***wa'tʰɨ́-pi*** *ǰo'o-mɨ̀* *jɨ̀ʔɨ́*
knife-INST do-M.SG 1SG-SBJ
'I (M) will do (it) with a knife.'

The only non-Siona+ language in the NWA sample which displays the IDENTITY type for this correspondence is Tariana [*tae*] (Aikhenvald 2003, 143-145), spoken in the Vaupés region. None of its Eastern Tukanoan neighbours display this marking pattern. Consider the instances in (10), where the case-marker *-ne* encodes the SUBJECT in (10a), and the INSTRUMENT in (10b) accordingly:

(10) Tariana^{V} (*tae* — ARAWAKAN) [A03:141,152, exx. (7.4),(7.50)]

a. *kiya-ku* *hiniri* *di-wha-ka* *di-ka* *di-niwa*
strong-PERS ukuki_fruit 3S.N.FUT-fall-SUB 3S.N.FUT-see 3S.N.FUT-collect
di-wake-ta *di-yã̀-nhi-pidana* [***diha***
3S.N.FUT-join+CAUS1-CAUS2 3S.N.FUT-stay-ANT-RMT.PST.REP ART
nawiki-ne]
man-SBJ.FOC
'The man saw that the ukukí fruit was falling down a lot (and) he was collecting (them) ...'

b. *di-wapa* ***mawipi-ne***$_{(\text{INST1})}$ [***diha*** ***pupawa***
3S.F-wait hook-COM ART poison
di-sue-ta-nipe-ne]$_{(\text{INST2})}$
3S.F-lay+CAUS1-CAUS2-NMLZ-COM
'He (the Makú man) was waiting (for the demon) with the stick$_{(\text{INST1})}$ (and) with the poison which he had prepared$_{(\text{INST2})}$.'

Although such disparate instances of SUBJECT-INSTRUMENT IDENTITY are uncommon in the the proposed NWA area, similar patterns are attested elsewhere. In particular, patterns of instrumental-ergative syncretism are well-described within several linguistic areas (cf. McGregor 2009; Chappell and Verstraete 2019, §4.1.1): e.g.,

[15] Technically, given that Wheeler (1970, *et seq.*) analyzes *-ga* as the emphatic SUBJECT-marker, which is unavailable to the INSTRUMENT, Colombian Siona displays the MIXED type. However, the relevant observation that *-bi* is distributed to both SUBJECT and INSTRUMENT functions still holds.

Tibeto-Burman languages (LaPolla 1995), Australian languages (Dixon 2002; Palancar et al. 2012), and even in some South American languages, such as Shipibo-Konibo and Sanumá (Borgman 1990).[16] Ultimately, the ergative-instrumental case-marking syncretism, displayed in the Darma language [*drd*] in (11), is strikingly similar to that observed in the Ecuadorian Siona sentence in (12):

(11) Darma (*drd* — Sino-Tibetan)
[Willis 2011, cited in Chappell and Verstraete 2019, 23]
niŋ su$_{(S)}$ ***pharsa su***$_{(Inst)}$ *nadu* *pyel-n-su*
1pl erg ax inst dem.neut chop-1pl-pst
'We$_{(S)}$ chopped it with an ax$_{(Inst)}$.'

(12) Ecuadorian Siona$^{\blacklozenge}$ (^{E}*snn* — Western-Tukanoan)
i̯o̯bi i̯re tsë'bobi tëtojoña i̯re
ĩ-o-bi$_{(S)}$ ĩ-i̵-re **tsi̵'bo-bi**$_{(Inst)}$
3pro-cl:f-sbj 3pro-cl:m-n.sbj ax-inst
ti̵hto-hoo-o-ña ĩ-i̵-re
hit-split-2/3s.f.pst.n.ass-rep 3pro-cl:m-n.sbj
'She$_{(S)}$ killed him with an ax$_{(Inst)}$, they say.' [nat: 20101123slicr001.065]

The next correspondence laid out in Table 4.4 concerns the Direct-Indirect object case-marking correspondence (*Q2.2*). All Siona+ languages display the Mixed type, whereas the remaining Tukanoan languages, and all other languages spoken in the Vaupés region, display the Identity type. The Awa Pit [*kwi*] and Shiwiar [*acu*] languages also display this type — i.e., the dominant pattern in the region (N=13). The fact that the case-marker *-re* encodes both Object functions across the Tukanoan family is well-described (e.g., Ramirez 2019[1997], 164-166; Barnes 1999, 2006; Stenzel 2013d), and Birchall (2014, 139-140) demonstrates that this pattern is also characteristic of the Nadahupan and Arawakan languages spoken in the vicinity of the Eastern Tukanoan languages.[17] The Identity type is demonstrated for Wanano [*gvc*] in (13), for Kakua [*cbv*] in (14), and for Hup [*jup*] in (15):

(13) Wanano$^{\Diamond}$ (*gvc* — Eastern-Tukanoan) [S13:336, ex. (54)]
$^{\sim}$*bichá-ré* *tí-kó-ró* ***hó-ré***$_{(P)}$ $^{\sim}$***bak-ú-***$^{\sim}$***dá-ká-ré***$_{(R)}$
today-n.sbj ana-cl:f-sg banana-n.sbj child-cl:m-dim-pl-n.sbj
wá-ré
give-vis.perf.2/3
'Today she gave the little boys$_{(R)}$ banana(s)$_{(P)}$.'

(14) Kakua$^{\vee}$ (*cbv* — Kakua-Nukak) [B16:189, ex. (15)]
Hãmu=di'$_{(R)}$ ***ma=ni̵m=di'***$_{(P)}$ *ma=wĭ-í́*
H=obj 2sg-daughter=obj 2sg=give-imper
'Give Hamu$_{(R)}$ your daughter$_{(P)}$!'

[16]The similar phenomenon of *causative-instrumental* applicative syncretism, described across various Bantu languages, displays possible functional overlap. See Jerro (2017) for a recent discussion.

[17]Birchall (2014, 139-140) remarks that Tukanoan, Nadahupan, and Arawakan languages display *neutral* alignment, such that (di)transitive p and r are coded indiscriminately. Albeit, Desano likely displays the *indirective* type, as illustrated in **Section A.2.2** in Appendix A.

(15) Hup[V] (*jup* — NADAHUPAN) [E08:170, ex. (21)]
núp ***hõp-ăn***$_{(P)}$ *'ăn* *no'-'ũh-'áy* *hám*
DEM.PROX fish-OBJ 1SG.OBJ give-APPL-VENT.IMPER go.IMPER
pawdína-ăn$_{(R)}$
P-OBJ
'Go give this fish$_{(P)}$ to Paulina$_{(R)}$ for me!'

There is a degree of overlap between the Siona+ MIXED-type pattern and the Eastern Tukanoan pattern, illustrated by the Wanano [*gvc*] example in (13). The case-marker *-re* is implicated in encoding both concerned functions. The main difference lies in the fact that additional Siona+ case-markers display this overlap; i.e., the case-markers *-ni* and *-na*. However, crucially, the latter marker may encode the R-argument, whereas this marker is strictly unavailable to the P-argument in all described varieties. (16) illustrates various marking possibilities in ditransitive sentences in Sekoya [[E]*sey*]:

(16) Ecadorian Sekoya[♦] (*e-sey* — WESTERN-TUKANOAN) [JJ90:48-49]

a. ***yë'ë*** ***yo'je-i̱-re***$_{(R)}$ ***coca***$_{(P)}$ *quëa-huë*
1SG younger_sibling-CL:M-N.SBJ word tell-N3S.PST.ASS
'I told my brother$_{(R)}$ the information$_{(P)}$.'

b. ***yë'ë-ni***$_{(R)}$ ***coca-re***$_{(P)}$ *ca-jë* *da-ë'ë*
1SG-N.SBJ2 word-N.SBJ say-PL.PRS.DEP come-N3S.PST.ASS
'When they came, they told ME$_{(R)}$ the message$_{(P)}$.'

c. ***ja'-quë-na***$_{(R)}$ ***më'ë*** ***do'ro-huë-re***$_{(P)}$ *i̱si-jë̱'ë̱*
parent-CL:M-GOAL 2SG basket-CL:CONT-N.SBJ give-IMPER
'Give your basket$_{(P)}$ to dad$_{(R)}$.'

The pattern shown in (16) is nearly identical to that described for Ecuadorian Siona, except regarding the distribution of *-na*, as described in Section 3.2.3 of the preceding chapter. More specifically, the marker *-na* is only attested on the R-argument in a handful of elicited instances with inanimate RECIPIENTS; which constitutes an accidental DCM pattern as conceived in this dissertation. As such, sentence (16c), would need to be amended so that the R-argument bears *-re* or *-ni* accordingly in Ecuadorian Siona.

This slight discrepancy marks a point of micro-variation across the Siona+ grouping, regarding animacy conditions on certain case-markers. On the one hand, the case-marker *-ni* is strictly available to animate non-SUBJECT arguments across all described varieties.[18] On the other hand, Ecuadorian Siona restricts the distribution of *-na* to inanimate nouns. This restriction does not hold for Ecuadorian Sekoya, as shown in (16), nor does it hold in Colombian Siona [[C]*snn*] (Wheeler 1970, 174; 1987, 126-127), as demonstrated in (17) below:

[18]The notion of promotion, laid out in detail in **3.2.2** in this dissertation, accounts for certain apparent exception to this animacy restriction; albeit, promotion is not described for other Siona+ languages. Some similar observations regarding the derivational functions of the feminine classifier, *-ko*, are laid out in Vallejos 2021, regarding Peruvian Sekoya. However, she does not discuss the implications for case-marking in the language. It is also relevant to note that Peruvian Sekoya does not have a case-marker *-ni*.

(17) Colombian Siona♦ (*c-snn* — WESTERN-TUKANOAN) [W87:127]
ja'yë-na$_{(R)}$ ***yë'e-re***$_{(P)}$ *insi-jë'ën*
older_brother-GOAL basket-OBJ give-IMPER
'Give (your) older brother$_{(R)}$ the basket$_{(P)}$!'

On this basis, the marker *-na* enjoys a broader distribution in Siona+ varieties besides Ecuadorian Siona.[19] Setting aside animacy-based restrictions, the general pattern is one where the spatial GOAL-marker, *-na*, can encode the R-argument in ditransitive constructions. A highly similar MIXED correspondence type is also described for various Quechuan languages (cf. Grzech 2016, 103-104, van Gijn et al. 2023), including the sampled variety spoken nearby to Ecuadorian Siona, Tena Kichwa [*quw*], as shown in (18):

(18) Tena Kichwa (*quw* — QUECHUAN) [G16:103, ex. (2.107)]
Pablo ***sisa-guna-ta***$_{(P)}$ *kuya-n* ***Maria-ma/Maria-ta***$_{(R)}$
P flower-PL-ACC give-3 M-DAT/ACC
'Pedro gave flowers$_{(P)}$ to Maria$_{(R)}$.'

Besides Quechuan languages, a few other languages in the NWA sample recruit spatial case-markers to encode the R-argument, including Murui [*huu*] (Wojtylak 2021, 270-273) and Bora [*boa*] (Thiesen and Weber 2012, 271-279). It remains an open question whether the Siona+ pattern is the product of contact-induced grammar change, since this differs from the IDENTITY type found in all Eastern Tukanoan languages. It is relevant to note that these languages generally exhibit rich sets of spatial case-markers from which additional R-markers could be drawn (cf. Stenzel 2008, 2013a). What the NWA dataset does demonstrate, however, is that the primary spatial marker, *-na*, found across the Siona+ languages, displays both argumental and spatial FUNCTIONS.

The third tested case-marking correspondence concerns the COMPANION-INSTRUMENT correspondence (*Q2.3*), per Table 4.4.[20] The Siona+ languages align in demonstrating the DIFFERENTIATION type, where the COMPANION-argument is encoded with an oblique case-marker (i.e., *-hã're* in Ecuadorian Siona) or a 'pseudo-case-marker', as discussed in Section 4.2.1.2. Despite differences in the forms used to flag the COMPANION, all Siona+ (and Tanimuka) languages have a cognate of the marker *-bi* to encode the INSTRUMENT function. The Koreguaje [*coe*] sentences in (19) represent the general Siona+ pattern, described for Ecuadorian Siona in Chapter 3:

(19) Koreguaje♦ (*coe* — WESTERN-TUKANOAN) [CC93:47-48]

a. ***wa't^{h}i-pi*** *ǰo'o-mɨ* *ǰɨ'ɨ*
knife-INST do-M.SG 1SG-SBJ
'(I (M)) will do (it) with a knife.'

b. ***i-na-re*** ***jã'me*** *pa-sa* *ǰii-mɨ* *ǰɨ'i*
this-PL.ANIM-N.SBJ COM live-DES.SG want-M.SG 1SG
'(I (M)) want to live with them.'

[19]The role of *-na* marking in Koreguaje [*coe*] is not explicitly discussed in the consulted sources. On this basis of this omission, it is determined that *-na* marking is not recruited for the R-argument.

[20]In the original chapter in *WALS*(52), Stolz et al. (2013) discuss *Comitatives and Instrumentals*, referring to conventional labels for the case-markers, rather than the CASE FUNCTIONS, as conceived in the assembly of the NWA dataset.

The DIFFERENTIATION type demonstrated in (19) is the most common correspondence type in the extensive global survey in Stolz et al. (2013, *WALS(52)*; i.e., 66.15% of 322-language sample). Nonetheless, in the NWA survey reported in Appendix A, this is not the dominant type. Besides the Siona+ grouping and Tanimuka, the DIFFERENTIATION type only applies to Bora [*boa*] (Thiesen and Weber 2012, 290-295) and Yuhup [*yab*] (Ospina Bozzi 2002, 158-161). No languages of the MIXED type are attested in the NWA sample. As such, all remaining Eastern Tukanoan languages, Máíhɨ̃kì [*ore*], and other sampled languages, display COMPANION-INSTRUMENT IDENTITY (N=15), as illustrated for the Barasano [*bsn*] language below:

(20) Barasano$^{\Diamond}$ (*bsn* — WESTERN-TUKANOAN)
[JJ91:68-69, exx. (189)-(190)]

a. ***bĩsi-bedo-rãka*** *gate* *bʉto* *asi* *bahi-ro-dẽ* *hʉ̃a-ko-aka-hʉ* *ti*
vine-ring-INST toast very.much hot be-NMLZ-SPCR crisp-FF-MOT-PST-3 3PRO.PL
'Toasting (leaves) with a vine-ring, when they are very hot they become crisp.'

b. ***yʉ-rãka*** *wa-rʉa-be-a-ti* *bʉ*
1SG-COM move-DES-NEG-PRS-Q 2SG
'Don't you (SG) want to come with me?'

For each QUESTION in the FUNCTIONAL domain discussed to this stage, Siona+ languages diverge from the Eastern Tukanoan languages. However, turning to the next tested correspondence outlined in Table 4.4, a pan-Tukanoan pattern is detected, where the case-marking on argumental P (i.e., DIRECT OBJECT) and on spatial L (i.e., STATIC LOCATION) are taken into account (*Q2.4*). Every Tukanoan language in the NWA dataset (except for Máíhɨ̃kì [*ore*])[21] displays the MIXED type, where the P-marker, *-re*, is available to L-arguments, alongside additional spatial case-markers. Table 4.5 spells out the distribution of case-markers across the two concerned FUNCTIONS in all tested Tukanoan languages:

[21]Note that spatial arguments are simply always zero-marked in the Máíhɨ̃kì (cf. Velie 2008[1975], 15-17; Farmer 2015, §3.3; Neveu 2012a, 2012b). This is the only Tukanoan language displaying this pattern to my knowledge, including those excluded from the NWA dataset.

Language	Direct object (P)	Static location (L)
*Ecuadorian Siona♦ [E*snn*]	**-*re***, **-*ni***	**-*re***, **-*ni***, -*na*
Colombian Siona♦ [C*snn*]	**-*re***, -*ni*	**-*re***, -*na*
Koreguaje♦ [*coe*]	**-*re***, -*ni*	**-*re***, -*na*
Máíhĩkì♦ [*ore*]	**-*re***	N/A
Kubeo◇ [*cub*]	**-*re***	**-*re***, -*i*, -*rã*
Tanimuka◇ [*tnc*]	**-*re***	**-*re***, -*re'ká*
Tukano◇ [*tuc*]	**-*re***	**-*re***, -*i*, -*pʉ*
Desano◇ [*des*]	**-*re***	**-*re***, -*ge*
Barasano◇ [*bsn*]	**-*re***	**-*re***, -*hʉ*
Wanano◇ [*gvc*]	**-*re***	**-*re***, -*i*, -*pʉ*

Table 4.5: Case-markers on P and L Functions, across Tukanoan sub-sample

Table 4.5 demonstrates how essentially all Tukanoan languages are unified with regards to the distribution of the case-marker -*re* across argumental and spatial domains (cf. Barnes 1999, 2006; Gomez-Imbert 2011; Stenzel 2008, 2013a, 2013d). In several Eastern Tukanoan languages, -*re* may co-occur with other dedicated spatial markers, as considered in Section 4.2.1.3, whereas this pattern never arises in the Siona+ languages. Besides Tukanoan languages, similar MIXED-type correspondences are only described for Hup [*jup*] (Epps 2008, 166-184)[22] and for Awa Pit [*kwi*] (Curnow 1997, 135-140), where L-marking is more complex than P-marking. All other sampled languages display the DIFFERENTIATION type, where there is no overlap between the case-marking found across the concerned Functions.[23]

Returning particularly to the Siona+ MIXED-type pattern, the dominant case-marker used to encode the L-argument, and the P-argument, is -*re*. I set aside patterns of spatial -*ni* attested in Ecuadorian Siona, described in Section 3.3.1 of the preceding chapter, since these patterns are not described for the other Siona+ languages. As regards instances of -*na* marking attested on the L-argument, and *never* on the P-argument, these patterns are illustrated for all Siona+ languages in their respective sources, although they are only recognized in prose comments. For instance, consider instances where -*re* and -*na* are utilized to encode the L-arguments in Koreguaje [*coe*] and Colombian Siona [C*snn*] in (21) and (22) respectively:

(21) Koreguaje♦ (*coe* — Western-Tukanoan) [CC93:48-49; CL85:105]

a. ***wɨ'e-re*** *pa'i-mo repa-o*
house-N.SBJ be-F.SG that-F.SG
'She is <u>in the house</u>.'

b. ***wɨ'e-na*** *k^{h}ãi-me j̃ĩi*
house-GOAL sleep-PL children
'The children are sleeping <u>in the house</u>.'

[22] In fact, Epps (2008, §4.3.2) considers the directional oblique marker, -*an*, to be historically related, but distinct from the stressed, Object case-marker, -*ăn*. Even if these are only historically related, this is a deparature from other non-Tukanoan languages sampled in the NWA dataset, where entirely different markers are selected for the concerned Functions.

[23] Consult **Table A.11** in Appendix A for examples of both attested types.

(22) Colombian Siona♦ ([C]*snn* — WESTERN-TUKANOAN) [W70:62]

a. ***wɨ'é-de*** *ba'ɨ́-hi*
house-N.SBJ be-3S.M.PRS.ASS
'(He) is <u>in the house</u>.'

b. ***yógu-té*** *yũ'ɨ́-hi*
canoe-N.SBJ be_seated-3S.M.PRS.ASS
'(He) is seated <u>in the canoe</u>.'

c. ***yógu-na*** *kãɨ́-hɨ́* *daí-ma-té*
canoe-GOAL sleep-PL.PRS.DEP come-RMT.PST-N2/3S.PST.N.ASS
'(We) came as (we) slept <u>in the canoe</u>.'

The sentences (21b) and (22c) demonstrate how the STATIC LOCATION (L) is encoded with *-na* with the predicate 'to sleep' in both languages. As discussed in Section 3.3.3 in the preceding chapter, this is the same pattern noted for Ecuadorian Siona with this predicate. This fact is shown in (23), where the speaker outright rejects the case-marker, *-re*, on the quesiton word, *hero* (where), where the predicate *kã(ĩ)-* (sleep) is concerned:

(23) Ecuadorian Siona♦ ([E]*snn* — WESTERN-TUKANOAN)

a. *jerona caquë'ne go'ye ñami ?*
he-ro-na *kãh-kɨ́-'ne* go'je ñami
WH-CL:LOC-GOAL sleep-2/3S.M.PST.N.ASS-Q before night
'<u>Where</u> did (you) (M) sleep last night?'

b. ****he-ro-re*** *kãh-kɨ́-'ne* go'je ñami
WH-CL:LOC-N.SBJ sleep-2/3S.M.PST.N.ASS-Q before night
Intended: '<u>Where</u> did (you) (M) sleep last night?'
[VOL/SUG: 20240928ejabi002.020a-b]

Ultimately, this dissertation treats the above-noted pattern as an instance of a predicate-specif DCM pattern, similar to the *-re* vs. *-na* alternation attested on inanimate P-arguments. Following the terminological conventions laid out in Chapter 3, this ought to be labeled an accidental DCM, since these forms do not represent a PLAIN-PROMINENT pair in the relevant sense. Nonetheless, more work is needed to see which predicates take *-na* on their (static) spatial arguments in other Siona+ languages.

A strikingly similar pattern is attested in the nearby lowland Quechuan language, Tena Kichwa [*quw*], as shown in (24b) for the predicate, *traba-* (work). The same pattern is observed with the corresponding predicate in Ecuadorian Siona, *goame nee-* (work, lit. 'do tasks'):

(24) Tena Kichwa (*quw* — QUECHUAN) [G16:68,62; exx. (2.47),(2.29)]

a. *Maria kaspi-wa wajta-n Juan-ta* ***wasi-pi***
M stick-INST hit-3 J-ACC house-LOC
'Maria hits John with a stick <u>in the house</u>.'

b. *Maria* ***chagra-ma*** *traba-nga ra-w-n*
M field-DAT work-FUT make-PROG-3
'Maria will work <u>in the field</u>.'

No patterns of the above type are described for Eastern Tukanoan languages. The predicate has little bearing on the selection of spatial case-marking in these languages. On this basis, the pattern shown above may have arisen due to contact-induced convergence between Siona+ languages and Quechuan languages.

The final correspondence tested in the NWA case-marking questionnaire concerns the spatial FUNCTIONS: LOCATION-GOAL-SOURCE (*Q2.5*). Section 4.2.1.2 already establishes that Siona+ languages recruit distinct markers for each of these FUNCTIONS. Table 4.6 demonstrates that the Siona+ pattern of spatial marking diverges sharply from the dominant Eastern Tukanoan pattern (Stenzel 2008, 2013a) — i.e., corresponding to the IDENTITY type:

LANGUAGE	LOCATION	GOAL	SOURCE	TYPE
*Ecuadorian Siona♦ [E*snn*]	*-re*, *-ni*, ***-na***	*-na*	*-bi*	MIXED
Colombian Siona♦ [C*snn*]	*-re*, ***-na***	*-na*	*-bi*	MIXED
Koreguaje♦ [*coe*]	*-re*, ***-na***	*-na*	*-pi*	MIXED
Tanimuka◇ [*tnc*]	*-re*, *-re'ká*	*-rã*	*-ϕi*	DIFFERENTIATION
Kubeo◇ [*cub*]	*-re*, *-i*, *-rã*	*-re*, *-i*, *-rã*	*-re*, *-i*, *-rã*	IDENTITY
Tukano◇ [*tuc*]	*-re*, *-i*, *-pʉ*	*-re*, *-i*, *-pʉ*	*-re*, *-i*, *-pʉ*	IDENTITY
Desano◇ [*des*]	*-re*, *-ge*	*-re*, *-ge*	*-re*, *-ge*	IDENTITY
Barasano◇ [*bsn*]	*-re*, *-hʉ*	*-re*, *-hʉ*	*-re*, *-hʉ*	IDENTITY
Wanano◇ [*gvc*]	*-re*, *-i*, *-pʉ*	*-re*, *-i*, *-pʉ*	*-re*, *-i*, *-pʉ*	IDENTITY

Table 4.6: Case-markers mapped onto spatial FUNCTIONS, across Tukanoan subsample

The Eastern Tukanoan pattern shown in Table 4.6 is labeled the 'generalized spatial case-marker pattern' for this discussion: i.e., a uniform set of spatial case-markers (including *-re*), encodes all primary spatial FUNCTIONS indiscriminately. Contrast the Ecuadorian Sekoya [E*sey*] examples in (25) with the generalized spatial case-marking pattern illustrated for Desano [*des*] in (26), which also displays an alternation between locative *-ge* and *-re*:

(25) Ecuadorian Sekoya♦ (E*sey* — WESTERN-TUKANOAN) [JJ90:45-49]

a. ***dayá-huë-té*** *huahuá-quë* *pa-jï'i* *huẽqu'ë*
pool-CL:NOM-REF float-CL:M be-3S.M.PST.ASS tapir
'The tapir was floating in the pool.'

b. *jã* *co'a-ye-re* ***tsia-ya-na*** *jẽo-jẽ'ë*
DEM.MED be_bad-CL:GEN-REF river-CL:NOM-GOAL throw-IMPER
'Throw that bad thing into the river!'

c. ***tẽ'të-pa-pi*** *caje-ni* *cuya-huë*
riverbank-CL:NOM-SRC descend-SS bathe-PST
'(I) came down from the riverbank and bathed.'

(26) Desano$^{\diamond}$ (*des* — EASTERN-TUKANOAN) [M99:59-61, ex. (241),(245),(254)]

a. *gɨa* *yoa-ri-bohe* *árĩ-bɨ* ***ero-ge-re*** ***iri***
1PL.EXCL be_long-DBV-time be-N3.PST there-LOC-OBJ DEM.PROX
bãkã-ge-re
town-LOC-OBJ
'We were a long time <u>there</u> <u>in this town</u>.'

b. *gɨa* ***õ-re*** *era* *pɨ'rɨ* *bu'a-bɨ* *pare* ***finka-ge***
1PL.EXCL here-N.SBJ arrive after go_down-N3.PST finally farm-LOC
'We arrived <u>here</u>, after which we went down <u>to a farm</u>.'

c. ***gahi-rã-ge*** *ari-biri-bã*
other-AN.P-LOC come-NEG-3PL
'(People) did not come <u>from other places</u>.'

The fact that the generalized spatial marking is attested across the Eastern Tukanoan languages, but differentiated spatial case-marking is the general patterns in Siona+ languages, highlights a fundamental divide between these groupings. Although both types implicate multiple spatial case-markers, their distribution across FUNCTIONS is radically different.

The generalized pattern is found across the Vaupés region,[24] whereas non-Vaupés languages in the NWA sample display either the MIXED or the DIFFERENTIATION type. The common denominator in the non-Vaupés group is that SOURCE is always encoded distinctly from LOCATION and GOAL.[25] This pattern matches those described for other Amazonian areas (cf. Wälchli and Zúñiga 2006; Roosvall 2020). Again, the Siona+ pattern matches the Quechuan pattern, displaying some limited LOCATION-GOAL overlap, and a dominant system of dedicated spatial case-markers. The Tena Kichwa [*quw*], examples in (27) are a case in point:

(27) Tena Kichwa (*quw* — QUECHUAN)
[G16:62-68, exx. (2.47),(2.28),(2.29),(2.30)]

a. *Maria* *kaspi-wa* *wajta-n* *Juan-ta* ***wasi-pi***
M stick-INST hit-3 J-ACC house-LOC
'Maria hits Juan with a stick <u>in the house</u>.'

b. *Maria* ***chagra-ma*** *traba-nga* *ra-w-n*
M field-DAT work-FUT make-PROG-3
'Maria will work <u>in the field</u>.'

c. *Pablo* ***ista-ma*** *shamu-n*
P party-DAT come-3
'Pablo came <u>to the party</u>.'

d. ... *kay* *awa* ***luma-manda***
DEM.PROX high hill-SRC
'...<u>from here above, from the hill</u>.'

[24] Technically Hup [*jup*] exhibits the MIXED correspondence type, since more markers are available to encode L than the other spatial FUNCTIONS (Epps 2008, 367-369). However, the primary spatial case-marker, *-an*, does display a *generalized* distribution in typical Vaupés fashion.

[25] Table A.13 in Appendix A provides illustrative instances of both the Vaupés and the non-Vaupés patterns in the NWA dataset.

In summary, Siona+ grouping patterns differently from Eastern Tukanoan as far as the case-marking correspondences analyzed here go. These typological findings are bolstered by the recognition of a handful of CASE FUNCTIONS, which are *not* encoded via case-marking in Ecuadorian Siona, but which do recruit case-marking in other NWA languages. For the sake of the present discussion, I label these FUNCTIONAL GAPS and discuss them briefly in the following section.

4.2.2.2 Functional gaps in Siona case-marking

The remaining questions in the FUNCTIONAL DOMAIN of the questionnaire address two particular FUNCTIONAL GAPS: (i) the POSSESSOR FUNCTION (*Q2.6*), and (ii) the TIME FUNCTION (*Q2.7*).

The POSSESSOR FUNCTION refers to the dependent POSSESSOR element within the possessive noun phrase, as opposed to the POSSESSEE, which is the head of this noun phrase. As such, the NWA questionnaire is designed to categorize languages based on whether they display overt POSSESSOR-marking (i.e., genitive case-marking), as shown to be the case in Kubeo [*cub*] in (28), or not:[26]

(28) Kubeo$^{\diamond}$ (*cub* — EASTERN-TUKANOAN) [MM99:118-119, exx. (381b)-(381c)]

a. ***Vícto-I*** *kɨ̃rãbĩ*
V-POSS house
'Victor's house.'

b. ***kai-wɨ*** ***i-Rã*** ***bue-I-xiejo-wA-I*** *papera*
all-PL.AN DEM.PROX-PL.AN student-I-child-PL.AN-POSS paper
'all of the students' papers.'

In stark contrast to the Kubeo [*cub*] overt POSSESSOR-marking strategy, the Ecuadorian Siona strategy consists in juxtaposition of the POSSESSOR and the POSSESSEE, in that order. This is a common strategy for encoding phrasal possession within the NWA sample (N=9). Consider the Ecuadorian Siona possessive noun phrase shown in (29) alongside a similar instance from Murui [*huu*] in (30):

(29) Ecuadorian Siona$^{\blacklozenge}$ (*e-snn* — WESTERN-TUKANOAN)
yë' gajei yohuë
jɨ'ɨ **gaje-ɨ** jo-wɨ
1SG friend-CL:M canoe-CL:CONT
'my friend's canoe.' [VOL/SUG: 20240928ejabi002.014]

(30) Murui (*huu* — WITOTOAN) [W21:8, ex. (1.7)]
Lusio *yoe-fai*
L metal-CL:SHORT.THICK
'Lucio's machete.'

The juxtaposition strategy illustrated in (30) and (29) is common to all Western Tukanoan languages.[27] Regarding non-Tukanoan languages, the neighbour to Murui

[26] This determination is made based upon the definition established by Iggesen (2013), as laid out in the methodological discussion in Section A.1.2.1 of Appendix A.

[27] The juxtaposition strategy for encoding possession is discussed in the following sources: See Wheeler (1987, §2.3.3.5) for the decription for Colombian Siona [^{C}*snn*]; Cook and Criswell (1993, 51) regarding Koreguaje [*coe*]; Johnson and Levinsohn (1990, 41) for Ecuadorian Sekoya [^{E}*sey*]; and Velie (2008[1975], 19-20) regarding Máíhɨ̃kì [*ore*].

[*huu*], Bora [*boa*] (Thiesen and Weber 2012, 251), also displays the juxtaposition strategy, as does A'ingae [*con*] (Fischer and Hengeveld 2023, §4).[28]

Besides the languages demonstrating the juxtaposition strategy outlined above, the dominant pattern in the area is the overt POSSESSOR-marking strategy shown in (28). This latter pattern is attested in all Eastern Tukanoan languages and surrounding non-Tukanoan Vaupés languages (Stenzel 2013b), and in various other languages in the sample, including the Quechuan languages (i.e., see Grzech (2016, 64-65) regarding Tena Kichwa [*quw*]; and Jake (1983, 23) regarding Imbabura Kichwa [*qvi*]), Shiwiar [*acu*] (Kohlberger 2020, 240-255), and Awa Pit [*kwi*] (Curnow 1997, 122-126).[29] Ultimately, the languages for which POSSESSOR-marking is not attested happen to form a cluster at the geographical centre of the proposed NWA area, subsuming the Siona+ languages and their closest neighbours.

A similar clustering of languages emerges as concerns the coding of TIME nominals (*Q2.7*), as defined in Section A.1.2.1 of Appendix A. As exemplified for Tukano [*tuc*] in (31), Eastern Tukanoan languages display overt case-marking, whereas the corresponding TIME-argument is not case-marked in the Western Tukanoan, as shown in (32) for Ecuadorian Sekoya [^{E}*sey*]:

(31) Tukano$^{\diamond}$ (*tuc* — EASTERN-TUKANOAN) [R97:237, ex. (580)]
dĩ'ká-de *bu'ê-dã*
today-N.SBJ study-IMPER
'Let's study TODAY!'

(32) Ecuadorian Sekoya$^{\blacklozenge}$ (*sey* — WESTERN-TUKANOAN) [JJ90:43]
ñami jopo-repa *sëta-huë*
night middle-PREC awaken-N3S.PST.ASS
'I woke up at midnight exactly.'

The zero-marking pattern demonstrated for Ecuadorian Sekoya in (32), matches the findings in the corpus-based investigation for Ecuadorian Siona, reported in Chapter 5 in this dissertation.[30,31] However, although there is no general TIME-marking in Koreguaje [*coe*] (Cook and Criswell 1993, §2.3.7), in another work, Cook and Levinsohn

[28]Fischer and Hengeveld (2023, 80) recognize a secondary possession encoding strategy where the POSSESSOR is postponed to the head nominal, i.e., POSSESSEE, in which case the POSSESSOR must be marked with the benefactive marker, -=*mbe*. The authors express that the unmarked, pre-nominal POSSESSOR configuration is the dominant pattern, matching the juxtaposition strategy discussed here.

[29]Two Vaupés languages are technically classified as POSSESSOR-marking NOT-ATTESTED based on the definitions for this marking outlined in Iggesen 2013: i.e., Wanano$^{\diamond}$ [*gvc*] and Tariana [*tae*], for similar reasons. Although these languages have over case-markers involved in encoding possession, and these markers are cognate with -*ya* found in other Eastern Tukanoan languages, these display different morphosyntactic behaviour from the other sampled languages. The relevant facts are laid out in Stenzel 2013b, 374-375, regarding Wanano, and a few other Eastern Tukanoan languages, and Aikhenvald 2003, §6.5.1 regarding Tariana.

[30]In the Siona narrative dataset in Chapter 5, of all coded nominal TIME tokens (N=142), the bulk of tokens were zero-marked (N=130, 91.55%), and, crucially, no tokens are marked with -*re*, as shown with the Tukano cognate -*de* in (31). Instead, a handful of Ecuadorian Siona temporal tokens in the corpus are marked with -*bi* (N=7), with a 'temporal starting point' reading, or with -*na* (N=3), receiving a 'temporal end point' reading. Similar marking patterns are noted regarding -*bi* in Colombian Siona (Wheeler 1987, 125-126), translated as 'since' (*desde* in Spanish), with both spatial and temporal readings.

[31]Regarding Colombian Siona, Wheeler 1987, 125-126 recognizes that, surprisingly, -*bi* may also be used to encode a temporal reading, 'during the whole X', but only with two

(1985, §4) do note a few instances of *-na* to express focus on a particular temporal interval, as shown in (33). Such patterns are not described for other Siona+ languages:

(33) Koreguaje♦ (*coe* — WESTERN-TUKANOAN) [CL85:113, ex. (49-i)]
6:30 pa'i-to oracha sa'aro meta jʉosina'me ***ñami-na*** ...
6:30 when_be Maticurú port come_down_river docked.PL night-GOAL
'(We) came downriver and docket at Maticurú when it was 6:30 <u>at NIGHT</u>...'

To summarize, despite a few disparate instances of temporal *-bi* and *-na* (but never *-re*) across the Siona+ languages, the dominant pattern is one where TIME nominals are left unmarked. Similar patterns are described for Tanimuka [*tnc*] (Eraso 2015, 293-294, §7.6.3), for A'ingae [*con*] (Fischer and Hengeveld 2023, 82, 93-96), and for Bora [*boa*] (Thiesen and Weber 2012, §10.4.6, §17.3). This is a similar grouping of languages to those which do not exhibit overt POSSESSOR-marking.

On the other hand, sampled Eastern Tukanoan languages are described as taking spatial case-markers for the sake of encoding (certain types of) TIME arguments. This is demonstrated for Tukano [*tuc*] in (31); and below for Kubeo [*cub*] in (34) and Wanano [*gvc*] in (35):

(34) Kubeo◊ (*cub* — EASTERN-TUKANOAN) [MM99:114,118; exx. (361b),(378)]
a. *oko-Rɨ̃bɨ̃-A-RE*
water-CL:TEMP-PL.INAN-N.SBJ
'in the rainy season...'
b. *ke te-Rɨ̃* ***ape-xãrãwɨ-I*** *dɨ̃-karã* ...
like_that do-GER other-CL:DAY-LOC go-1PL.EXCL
'So we went <u>another day</u>...'

(35) Wanano◊ (*gvc* — EASTERN-TUKANOAN) [S08:6 ex. (17),(15)]
a. *tu'sʉ* [***ti*** ***~yabi-re*** ***~hi'da***] *~kha'a-ro-pʉ-re ti-ro-re*
finish ANA night-OBJ EMPH dream-SG-LOC-OBJ ANA-SG-OBJ
ya'u-a
warn-ASSERT.PERF
'When (the man) was done, <u>that very night</u>, (someone) warned him in a dream.'
b. *~phado-pʉ-re* ...
long_ago-LOC-OBJ
'In the olden days...'

It is relevant to note that it is not only the case-marker *-re*, which exhibits a TIME-marking FUNCTION in these languages, but rather the full arsenal of spatial case-markers may be employed in this way. For instance, temporal uses of the so-called plain locative-marker are attested in Kubeo [*cub*], *-I*, as shown in (34b); and the Wanano item in (35b) displays the typical spatial double-marking pattern discussed

lexical items: *umuguse* (day) and *ñami* (night). The exact same observations are put forth for Ecuadorian Sekoya (Johnson and Levinsohn 1990, 47), and some similar examples are provided for Koreguaje (Cook and Levinsohn 1985, 113). This is similar to a few noted restricted instances of temporal *-na* in Ecuadorian Siona, pointed out in Chapter 5, which require further attention.

in Section 4.2.1.3. Similar facts are attested in the descriptions of Desano [*des*] (Silva 2012, 97-99, 161) and Barasano [*bsn*] (Jones and Jones 1991, §5.5); and described for Quechuan languages (see Grzech (2016, 68) regarding Tena Kichwa [*quw*]; and Cole (1982, 125) regarding Imbabura Kichwa [*qvi*]), Shiwiar [*acu*] (Kohlberger 2020, 264) and Awa Pit [*kwi*] (Curnow 1997, 340-346). To date, temporal case-marking is relatively under-explored and poorly-described compared with other FUNCTIONS, and more targeted data collection is warranted.

The discussion in the present section has demonstrated that the Siona+ languages, and a handful of other languages spoken around the geographical centre of the proposed NWA region, converge on two FUNCTIONAL GAPS: (i) POSSESSOR-marking, and (ii) TIME-marking. A contact-induced explanation for these patterns is conceivable, involving neighbouring languages like A'ingae [*con*] and/or the languages of the *People of the Centre* (i.e., including Murui [*huu*] and Bora [*boa*]).

4.2.3 Typologizing Siona DCM properties

Table 4.7 identifies the relevant properties for each of the five Ecuadorian Siona DCM patterns based upon the *preliminary description* in Chapter 3. Each of these patterns is attributed a QUESTION in the NWA dataset, as indicated in the table. The ensuing discussion typologizes these DCM patterns on a FUNCTION-by-FUNCTION basis:[32]

QUESTION	DCM FUNCTION	AVAILABLE FORMS	ACTIVE MACRO-TRIGGERS			
			ANIM	REF	DISC	OTHER
Q3.1	SUBJECT (S)	{-∅, *-bi*}			✓	
Q3.2	DIRECT OBJECT (P)	{-∅, *-re*, *-ni*}	✓	✓	✓	
Q3.3	INDIRECT OBJECT (R)	{*-re*, *-ni*, *-na*}	✓		✓	
Q3.4	LOCATION (L)	{-∅, *-re*, *-ni*, *-na*}		(✓)	✓	✓ PRED
Q3.5	GOAL (G)	{-∅, *-na*}		(✓)	✓	

Table 4.7: Overview of Ecuadorian Siona DCM patterns — comparative benchmark

4.2.3.1 Situating Siona Subject-oriented DCM

I established earlier that S-oriented DCM in Ecuadorian Siona displays the OPTIONAL formal type, such that the marker *-bi* alternates with zero-marking. This alternation is driven almost entirely by discourse-related properties of the S-argument. As demonstrated in Table 4.8, S-oriented DCM patterns of any type are rare in the NWA sample (*Q3.1*), which includes highly similar patterns in all three included Siona+ languages.

[32]The short-hand labels utilized for MACRO-TRIGGERS are as follows: ANIM refers to animacy-related TRIGGERS, REF refers to referentiality-related TRIGGERS (e.g., specificity, pronominality), and DISC refers to discourse-related TRIGGERS. The OTHER category is reserved for any other TRIGGERS which do not easily fit into one of these MACRO-TRIGGERS.

Language	Forms	Macro-triggers			
		anim	ref	disc	other
*Ecuadorian Siona♦ [Esnn]	{-∅, -*bi*}			✓foc	
Colombian Siona♦ [Csnn]	{-∅, -*bi*, -*ga*}			✓foc	
Koreguaje♦ [*coe*]	{-∅, -*pi*}			✓contr	
Tanimuka◊ [*tnc*]	{-∅, -*re*}	✓hum			
Murui [*huu*]	{-∅, =*dɨ̀*}			✓top	
Vtariana [*tae*]	{-∅, -*n*(*h*)*e*}			✓top	
no dcm on Subject:	*ore*♦ *cub*◊ *tuc*◊ *des*◊ *bsn*◊ *gvc*◊ *con boa quw qvi acu kwi* V*cbv* V*jup* V*yab*				

Table 4.8: Subject-oriented dcm patterns in the nwa dataset

S-oriented dcm is not attested in any Eastern Tukanoan language, excepting Tanimuka [*tnc*], which displays an exceptional pattern where s- and p-marking overlap (Eraso 2015, 229-231). Some examples for this pattern are provided in Section A.2.3 of Appendix A. Among non-Tukanoan languages in the nwa sample, only two languages display s-oriented dcm patterns: Murui [*huu*] (Wojtylak 2021, 245) and Tariana [*tae*] (Aikhenvald 2003, §7.2.1, 140-143).

Both the Murui and the Tariana patterns align neatly with Siona+ descriptions for s-oriented dcm with regards to the discourse-related triggers that drive the realization of the appropriate case-markers. For instance, compare the the naturalistic contrastive focal usage of s-marking in Tariana in (36) with a similar Ecuadorian Siona corpus example in (37), lifted from the sample in Chapter 5:

(36) Tariana[V] (*tae* — Arawakan) [A03:141, ex. (7.4)]
kiya-ku *hiniri* *di-wha-ka* *di-ka* *di-niwa*
strong-persist ukukí_fruit 3s.f-fall-sub 3s.f-see 3s.f-collect
di-wake-ta *di-yã-nhi-pidana* ***diha nawiki-ne***
3s.f-join+caus1-caus2 3s.f-stay-ant-rmt.pst art man-foc.a/s
'The man saw that the ukukí fruit was falling down a lot (and) he was collecting (them) (while the evil spirit was trying to steal the fruit).'

(37) Ecuadorian Siona♦ (*e-snn* — Western-Tukanoan)
[**Context:** The black vulture and another personified vulture (condor) are searching for the traps that were set for trapping animals.]
. . . jãre tacarobi ña bã'co — ñani nëcadojaiyo ñani. . .
hãre **tãhkaro-bi** ñaa-a bah-ko — ñaa-ni
like_that black_vulture-sbj see-neg be-2/3s.f.pst.n.ass — look-ss
nɨhka + dohai-yo ñaa-ni
be_standing + wander-cl:f look-ss
'. . . likewise the black vulture was looking out. She was wandering on foot (i.e., on the ground) and looking around (while the condor searched from the sky). . . ' [nat*: 20151112orapi001.175-176]

The focus- and/or contrastivity-related uses of overt s-marking noted in (36) and (37),

are described for all Siona+ languages.[33] Although Aikhenvald (2003) and Wojtylak (2021) employ the term 'topicality' to describe the discourse-related TRIGGERS they observe, contrastive focus and disambiguation are overtly recognized as sub-patterns of topical S-marking in the relevant portions of these descriptions. These patterns are highly similar, as Aikhenvald (2015) recognizes in a later work.

One point of apparent micro-variation within the Siona+ grouping concerns the HYBRID formal type attested for Colombian Siona in Table 4.8 — i.e., *-bi* and zero-marking alternate with another overt marker, *-ga* (Wheeler 1970, 173-176, 1987, 131). The selection of the appropriate case-marker is determined on the basis animacy conditions[34] and the type of focus — i.e., *-bi* expresses 'plain focus', whereas *-ga* expresses 'emphatic focus'. The three marking alternatives are illustrated in (38).

(38) Colombian Siona♦ (*snn* — WESTERN-TUKANOAN) [W70:47]

a. ***yɨ'ɨ-ga*** *saí-yɨ*
1SG-SBJ.EMPH go-N3S.PRS.ASS
'And I, I am going.'

b. ***ča'kádi-wa'-na-bi*** *bá-gɨ-na* *gahe-ni*
fire_ant-ANIM-PL-SBJ PRO.HUM-CL:M-GOAL go_down-SS
sɨ'á-wɨ
attach-N3S.PST.ASS
'The fire ants came down to him and latched on.'

c. ***'okó*** *ka'ní-hi*
rain gather-3S.M.PRS.ASS
'The rain is gathering.'

Although Ecuadorian Siona does not have a S-marker *-ga*, a cognate exists — i.e., *=kato*, which is analyzed as a contrastive topic marker in Bruil (2014, *et seq.* — see Table 4.1). Interestingly, a formally similar topic marker is described for several surrounding non-Tukanoan languages as well: i.e., *-ka* in Murui [*huu*] (Wojtylak 2021, 245-246); *=ga* in Imbabura Kichwa (cf. Cole 1982, 26-29) and in Tena Kichwa (Grzech 2016, 157-159); $=k^{j}a$ (TOP/Q) in Shiwiar [*acu*] (Kohlberger 2020, 198, 269-273); and *=ka* (EMPH) in Awa Pit [*kwi*] (Curnow 1997, 378-379). To my view, this is a strong candidate for contact-induced convergence, although a full discussion is inappropriate in the present analysis of case-marking. Wheeler (1970, 1987, etc.) analyzes the corresponding marker as belonging to the case-marker inventory. The Ecuadorian Siona discourse marker *=kato* will resurface in the next two chapters, but not as a case marker.

Setting aside the slight formal discrepancy between Colombian and Ecuadorian varieties of Siona, discourse-driven S-oriented DCM is common to the Siona+ languages. Strikingly similar patterns are described for Murui [*huu*] and Tariana [*tae*]. Otherwise, S-marking is typologically rare in the NWA region. An entirely different typology is found regarding the case-marking patterns described for encoding the

[33] Although similar contrastivity-related effects are noted for S-marking elsewhere in the Siona+ sources, a vareity of labels are used: see Wheeler (1970, 1987) regarding Colombian Siona [^{C}*snn*], Johnson and Levinsohn 1990 regarding Ecuadorian Sekoya [^{E}*sey*], and Cook and Levinsohn (1985) and Cook and Criswell (1993) regarding Koreguaje.

[34] According to Wheeler 1970, 174, the SUBJECT-marker *-ga* is strictly available to animates as an 'emphatic focus marker'. This fact is surprising given that the cognate, *=kato*, does not exhibit animacy-based restrictions in Ecuadorian Siona.

DIRECT OBJECT, i.e., the P-argument, which is described for essentially all sampled NWA languages.

4.2.3.2 Situating Siona Direct object-oriented DCM

Patterns of P-oriented DCM are ubiquitous in the NWA region, as they are in nearly all regions of the world (cf. Bossong 1984, 1985; Sinnemäki 2014). The only language in the NWA dataset which is described as displaying obligatory P-marking (i.e., with no DCM pattern for this FUNCTION) is Imbabura Kichwa [*qvi*] (Cole 1982, 103-104).

To guide the ensuing discussion of DCM and the expression of DIRECT OBJECT-marking, Table 4.9 compares the available case-marking forms and the active MACRO-TRIGGERS for Ecuadorian Siona, and for all other sampled Tukanoan languages (*Q3.2*). A full chart, including non-Tukanoan languages, is provided in Table A.20 in Appendix A. This table distinguishes active STRONG TRIGGERS from WEAK TRIGGERS, by representing the latter type in parentheses, (✓), following the distinction laid out in Chapter 2 (inspired by Klein and de Swart 2011). On the one hand, STRONG TRIGGERS rigidly divide the set of P-arguments as receiving a particular case-marking alternative or not, whereas WEAK TRIGGERS represent marking tendencies. This determination adds nuance to the discussion of micro-variation across related languages that display (often subtly) different DCM patterns, including Siona+ languages:

LANGUAGE	FORMS	MACRO-TRIGGERS			
		ANIM	REF	DISC	OTHER
*Ecuadorian Siona♦ [E*snn*]	{-∅, **-re**, -*ni*}	✓ANIM	✓PRO,(SPEC)	✓FOC,(TOP)	✓COORD
Colombian Siona♦ [C*snn*]	{-∅, **-re**, -*ni*}	✓(ANIM)	✓PRO,(SPEC)	✓FOC,(TOP)	
Koreguaje♦ [*coe*]	{-∅, **-re**, -*ni*}	✓(ANIM)	✓PRO,(SPEC)	✓FOC,(TOP)	
Máíhɨ̃kì♦ [*ore*]	{-∅, **-re**}	✓ANIM	✓(SPEC)		
Kubeo◇ [*cub*]	{-∅, **-re**}	✓(ANIM)	✓PRO,(REF)		✓COORD
Tanimuka◇ [*tnc*]	{-∅, **-re**}	✓HUM		✓CONTR	
Tukano◇ [*tuc*]	{-∅, **-re**}	✓(ANIM)	✓PRO,DEF		
Desano◇ [*des*]	{-∅, **-re**}		✓PRO,SPEC		
Barasano◇ [*bsn*]	{-∅, **-re**}		✓PRO,DEF		
Wanano◇ [*gvc*]	{-∅, **-re**}	✓(ANIM)	✓PRO,REF		✓POSITION

Table 4.9: Tukanoan DIRECT OBJECT-oriented DCM patterns in the NWA dataset

Table 4.9 demonstrates that the Siona+ languages display a remarkably stable set of P-oriented DCM patterns, both on formal ground and regarding the TRIGGERS, which drive the observed alternations. On the one hand, *all* Tukanoan languages display at least a zero-marking vs. -*re* marking alternation for this function (Barnes 1999, 2006; Gomez-Imbert 2011; Stenzel 2008, 2013c, 2013d; Zúñiga 2007) — i.e., corresponding to the OPTIONAL formal type, the most common DCM type cross-linguistically (cf. Aissen 2003; de Hoop and Malchukov 2007; Chappell and Verstraete 2019).

On the other hand, each of the Siona+ languages possesses an additional overt marker -*ni*, discussed in the preceding sections of this chapter. See Wheeler (1987, 126) regarding the usage of -*ni* in Colombian Siona [C*snn*]; Cook and Criswell (1993, 48-49) and Cook and Levinsohn (1985, 100-104) regarding Koreguaje [*coe*]; and Johnson and Levinsohn (1990, 47) regarding Ecuadorian Sekoya [E*sey*]. These patterns align neatly

with the distribution of Ecuadorian Siona *-ni* on the DIRECT OBJECT (P-argument), as described in Section 3.2 of the preceding chapter. As such, on formal grounds, the Siona+ languages exhibit the typologically rare HYBRID type, where multiple markers alternate with zero-marking,[35] which is only attested otherwise once in the NWA sample for P-oriented DCM: i.e., in Tariana [*tae*] (Aikhenvald 2003, 145-148), as demonstrated in (80) in Appendix A. The Colombian Siona sentences in (39) illustrate the three marking alternatives available to the P-argument in Siona+ languages:

(39) Colombian Siona♦ (*snn* — WESTERN-TUKANOAN) [W70:43-44]

a. ***jo'ó*** *yo'ó-wɨ*
work do_work-N3S.PST.ASS
'(I) did the work.'

b. ***'ãõ-de*** *k^{w}a'kú-yɨ*
food-N.SBJ cook.TRS-N3S.PRS.ASS
'(I) am cooking food.'

c. ***ya'ó-ni*** *hu-ɨ́'ɨ* 'aídu-na
peccary-N.SBJ2 kill-N3S.PST.ASS forest-GOAL
'(I) shot a peccary in the woods.'

Besides the discrepancy noted regarding formal types of DCM, Siona+ languages are described as displaying complex, multi-dimensional DCM (Klein and de Swart 2011). More concretely, animacy-related, referentiality-related, and discourse-related TRIGGERS must each be considered in order to accurately capture the distribution of the case-marking alternatives shown in (39). In what follows, the relevance of each of these MACRO-TRIGGERS is discussed and contextualized within the NWA sample accordingly.

Firstly, the selection of the appropriate case-marker in Ecuadorian Siona is inextricably connected to the animacy class of the P-argument, as described in Section 3.2 in Chapter 3. More concretely, two related animacy-related patterns are noted: *-ni* may only arise on animate (or promoted) nouns, as in (39c); and these same nouns reject zero-marking. The former pattern is common to all Siona+ languages; however, Table 4.9 highlights a slight discrepancy as to the obligatory case-marking on animate nouns. Unlike the Ecuadorian Siona pattern described throughout this dissertation, zero-marked animate P-arguments are permitted in Koreguaje [*coe*] sentences such as (40) — see Wheeler 1970, 145 regarding similar instances in Colombian Siona [^{C}*snn*]. Ecuadorian Siona speakers outright reject these unmarked instances, as demonstrated in the corresponding elicited sentences in (41):

(40) Koreguaje♦ (*coe* — WESTERN-TUKANOAN) [CL85:105]

a. ***pãi*** *k^{h}u'e-aso-mʉ* *kã'ʉ*
people look_for-PST-M.SG he
'He is looking for people.'

b. ***ocho*** ***pãi-re*** *soni-aso-mʉ* *repa-ʉ*
Carijona people called-PST-M.SG that-M.SG
'He called Carijona people.'

[35] In Peruvian varieties of Sekoya, it is the case that the case-marker *-ni* is unattested (Rosa Vallejos, p.c.). Therefore, this variety of Sekoya also displays an OPTIONAL formal type of P-oriented DCM.

(41) Ecuadorian Siona♦ (*snn* — WESTERN-TUKANOAN)

a. *i̲ ba̲i̲re coeji*
ĩ-i̷ **bãĩ-re** *kwee-hi*
3PRO-CL:M people-N.SBJ look_for-3S.M.PRS.ASS
'He is looking for people.'

b. * *i̲ ba̲i̲ coeji*
ĩ-i̷ **bãĩ** *kwee-hi*
3PRO-CL:M people look_for-3S.M.PRS.ASS
Intended: 'He is looking for people.'

[VOL/SUG: 20230721eyopa001.009a-b]

On the basis of examples such as (40) and (41), Koreguaje [*coe*] is treated as displaying animacy as a WEAK TRIGGER, whereas animacy is a STRONG TRIGGER in Ecuadorian Siona. No matter the determination of TRIGGER STRENGTH, the animate-inanimate distinction is central to capturing P-oriented DCM patterns across the Western Tukanoan languages, and is the primary factor determining the distribution of *-re* on the DIRECT OBJECT in Máíhɨ̃kì [*ore*] (Farmer 2015, 93-97).

As regards the Eastern Tukanoan sample, the role of animacy in determining case-marking is mixed. On the one hand, Tanimuka [*tnc*] is the only language where this is the primary TRIGGER for P-marking (Eraso 2015, 296). A handful of other authors recognize animacy as a secondary, WEAK TRIGGER among the sampled languages — i.e., Kubeo [*cub*] (Chacon and Genetti 2019, 412), Tukano [*tuc*] (Ramirez 2019[1997], 164-166), and in Wanano [*gvc*] (Stenzel 2008). Among the non-Tukanoan languages in the NWA dataset however, animacy is a STRONG TRIGGER in Awa Pit [*kwi*] (Curnow 1997, 72-74), in Bora [*boa*] (Thiesen and Weber 2012, 273-274); and, interestingly, animacy is the primary driver of P-oriented DCM in various languages spoken among the Eastern Tukanoan languages in the Vaupés region: i.e., Kakua [*cbv*] (Bolaños 2016, 203-205), Hup [*jup*] (Epps 2008, 174-177), and Yuhup [*yab*] (Ospina Bozzi 2002, 141-146). It is relevant to note that the non-Tukanoan, Vaupés pattern draws a human-vs.-non-human distinction, whereas the animacy-based DCM patterns outlined for Western Tukanoan distinguish all animates from their inanimate counterparts. This discrepancy may point to a divergent development of the attested patterns.

Ultimately, as Sinnemäki (2014) finds in his extensive survey, animacy-based DCM patterns are commonly attested in all linguistic areas, including the proposed NWA region. However, the NWA sample underpinning the present discussion traces a marked distinction between Western Tukanoan languages, where animacy is one of the most relevant, STRONG TRIGGERS, for these patterns, and the Eastern Tukanoan languages where this tends to take a secondary position to other TRIGGERS. The most relevant class of TRIGGERS, as concerns the Eastern Tukanoan patterns, pertains to referentiality-related TRIGGERS, including pronominality, specificity, definiteness, etc. On the one hand, it is well-established that pronominal P-arguments require case-marking in all instances, as a generalizable pan-Tukanoan trait (cf. Barnes 1999, 2006; Gomez-Imbert 2011; Ramirez 2019[1997]; Stenzel 2013c); a constraint which is also attested in Awa Pit [*kwi*] (Curnow 1997), and several languages of the Vaupés region (Zúñiga 2007): i.e., Hup [*jup*] (Epps 2008, 172); Kakua [*cbv*] (Bolaños 2016, 193-194); and Tariana [*tae*] (Aikhenvald 2003, 146-147). The dominant pattern across all of these (non-)Tukanoan may be stated as a generalization: non-SUBJECT personal pronouns require overt case-marking.

Despite the pronoun-marking generalization, on the other hand, the status of other referential properties as TRIGGERS for DCM in the NWA area is a less straightforward question. A sharp divide is noted in the literature across the branches of the Tukanoan family. Firstly, referentiality-related TRIGGERS, e.g., specificity and individuation, are described as STRONG TRIGGERS for Eastern Tukanoan languages (Barnes 1999, 2006; Waltz and Waltz 1997), excepting Tanimuka [*tnc*].[36] For instance, the (near) minimal pairs in (42) and (43) are used by Duarte and de Oliveira Lopes (2020) to argue for a definiteness-based analysis for such patterns in Tukano [*tuc*]:[37]

(42) Tukano$^{\diamond}$ (*tuc* — EASTERN-TUKANOAN)
[DL20:30-31, exx. (33)-(34),(38)-(40)]

a. ***ni'ká wi'i*** *da'rê-gɨ'* wee-mí
one house make-SS.M.SG do-PRS.VIS.3S.M
'(He) is making a house.'

b. * ***ni'ká wi'i-re*** *da'rê-gɨ'* wee-mí
one house-N.SBJ make-SS.M.SG do-PRS.VIS.3S.M
Intended: '(He) is making a house.'

(43) a. ***ohô*** *ba'â-ya*
banana eat-IMPER
'Eat (some) banana!'

b. ***ohô-poro-re*** *ba'â-ya*
banana-CL:OBL-N.SBJ eat-IMPER
'Eat the banana!'

c. * ***ohô-re*** *tɨá-yɨ*
banana-N.SBJ eat-IMPER
Intended: 'Eat some/the banana!'

A STRONG TRIGGER analysis for referentiality-based DCM patterns, based upon examples such as (42) and (43), is untenable for Siona+ languages. The corresponding Ecuadorian Siona sentences below demonstrate the starkly different case-marking facts, concerning the availability of *-re* marking on non-specific (inanimate) P arguments. Firstly, the sentence in (44) demonstrates an instance where *-re* is naturally produced with the numeral one, *te'e*, where it receives a non-specific interpretation:[38]

[36] Authors describe referentiality-related STRONG TRIGGERS in their respective sourcese — i.e., Tukano [*tuc*] (Ramirez 2019[1997], 164-167; Duarte and de Oliveira Lopes 2020); Desano [*des*] (Miller 1999, 57; Silva 2012, 171-173); Barasano [*bsn*] (Jones and Jones 1991, 65-66); and Wanano [*gvc*] (Stenzel 2008). Authors tend to bridge the referentiality-based patterns for the distribution of *-re* with the obligatory pronoun-marking generalization, and invoke the *Referentiality hierarchy* (or other incarnations thereof, e.g., the *Individuation scale*) to account for Eastern Tukanoan P-oriented DCM effects.

[37] In order to unpack (43), it is relevant to note that Tukano, like all Tukanoan languages, exhibits a rich system of nominal classification. For Duarte and de Oliveira Lopes 2020, the overt classifier, *-poro*, for oblong objects, serves an individuating function, thereby licensing the referentiality-sensitive *re*-marker and achieving the desired definite reading.

[38] Note that the numeral 'one' is used sparingly in such cases in naturalistic speech. Generally bare nominals are employed with either a specific or a non-specific reading — see the corpus analysis in Chapter 5 and Appendix B of the present dissertation for further details.

(44) Ecuadorian Siona♦ (*e-snn* — WESTERN-TUKANOAN)
yë' jaquëbaibi te'e huë'ere bayë
jɨ'ɨ ha-kɨ-bãĩ-bi **te'-e** **wɨ'e-re** baa-jɨ
1SG parent-CL:M-PL.AN-SBJ one-CL:GEN house-N.SBJ have-N3S.PRS.ASS
'My parents have a house (lit. one house).' [VOL: 20230525ejabi001.023]

The next set of sentences, shown in (45), demonstrates that the presence of the classifier does not have a direct impact on DCM in Ecuadorian Siona. On the basis of Q-A congruence effects (see Chapter 3), obligatory *-re* marking is observed on the focalized DIRECT OBJECT in the answer in (45), independently of the presence of a classifier. Additionally, to further demonstrate the irrelevance of classification on case-marking, the non-classified P-argument, *nohka* (banana), receives a definite reading in (46), but is naturally produced without case-marking:

(45) Ecuadorian Siona♦ (*e-snn* — WESTERN-TUKANOAN)

a. *quere baco mẽ' jëtëre ?*
ke-e-re baa-ko mɨ'ɨ̃ hɨ̃hti-re
WH-CL:GEN-N.SBJ have-2/3S.F.PRS.N.ASS 2SG hand-N.SBJ
'What do you (F) have in your hand?'

b. *nocare bayë*
nohka-re baa-jɨ
banana-N.SBJ have-N3S.PRS.ASS
'I have a baNAna.'

c. *nocamore bayë*
nohka-mo-re baa-jɨ
banana-CL:CYL-N.SBJ have-N3S.PRS.ASS
'I have a baNAna.' [VOL: 20230704eyopa001.015a-c]

(46) Ecuadorian Siona♦ (*e-snn* — WESTERN-TUKANOAN)
[**Context:** A monkey comes and steals the speaker's banana off of the table.]
nasore noca aiquëni co'eyë
[nahso-re **nohka** ãĩ-kɨ-ni] ko'e-jɨ
monkey-N.SBJ banana eat-CL:M-N.SBJ2 look.for-N3S.PRS.ASS
'I am looking for the monkey that ate the banana.'
[VOL: 20231906elupa001.002a]

It is clear from instances such as (44) through (46) that referentiality does not play the same role in Ecuadorian Siona as it does for Eastern Tukanoan languages, such as Tukano [*tuc*]. We reach a similar conclusion in recent corpus-based work (Case & Bruil, *forthcoming* — also see Chapter 5 in this dissertation). It remains to be seen whether referentiality-related TRIGGERS have no bearing on the attested patterns, or whether they are better classified as WEAK TRIGGERS, and similar conclusions are drawn regarding other Siona+ languages: i.e., Colombian Siona [^{C}*snn*] (Wheeler 1987, 127-128), Ecuadorian Sekoya [^{E}*sey*] (Johnson and Levinsohn 1990, 48), and Koreguaje [*coe*] (Cook and Levinsohn 1985). What is clear for all Siona+ languages is that referentiality-related TRIGGERS for P-oriented DCM in these languages is secondary to animacy- and discourse-related TRIGGERS.

Similar analyses regarding referentiality as a WEAK TRIGGER are put forth regarding Máíhɨ̃kì [*ore*] (Farmer 2015, 95-97), and for Kubeo [*cub*] (Chacon and Genetti

2019, 412). Several non-Tukanoan languages in the NWA sample are also claimed to display a WEAK TRIGGER for referentiality: i.e., Murui [*huu*] (Wojtylak 2021, §8.1.15), Awa Pit [*kwi*] (Curnow 1997, 72-74), Kakua [*cbv*] (Bolaños 2016, 200-202), and Hup [*jup*] (Zúñiga 2007; Epps 2008, 174-177). Ultimately, the typology is such that Eastern Tukanoan languages, excluding Kubeo [*cub*] and Tanimuka [*tnc*], exhibit P-oriented DCM patterns where referentiality-related notions, i.e., definiteness and individuation, are the primary drivers for case-marking alternations.

In stark contrast to the dominant Eastern Tukanoan pattern, setting aside the role of animacy, Siona+ languages are better described as displaying discourse-driven DCM patterns. This is already established for Ecuadorian Siona in preceding chapter, based upon examples such as (45) above, demonstrating rigid QUESTION-ANSWER congruence effects, and other patterns of focus marking (see elaborate discussion in Chapter 6). Very similar analyses are advanced for other Siona+ languages: i.e., Colombian Siona [^{c}*snn*] (Wheeler 1970, 43-44) and Koreguaje [*coe*] (Cook and Levinsohn 1985). Compatible descriptions are also made regarding Tariana [*tae*] (Aikhenvald 2003, 145, 2015) and Murui [*huu*] (Wojtylak 2021, 245-258). The focus-driven usage of the appropriate P-marker is shown for Murui and Koreguaje in (47) and (48) respectively:

(47) Murui (*huu* — WITOTOAN) [W21:256,267]
aros$_{i}$ *ati-ñe-iti-o* ***asukar-na***$_{ii}$ *ati-iti-o*
rice bring-NEG-FUT.LK-2SG sugar-N.SBJ.TOP bring-FUT.LK-2SG
'You won't bring <u>rice</u>$_{i}$. You will bring <u>the SUGAR</u>$_{ii}$.'

(48) Koreguaje$^{\blacklozenge}$ (*coe* — WESTERN-TUKANOAN) [CC93:49, ex. (201)]
k^{h}ura-wa'-i-ni *k^{h}u'e-mo* ***ji i***
chickens-creature-M.SG-N.SBJ2 look.for-F.SG 1SG
'I (F) am looking for <u>a chicken</u> (a particular one, not just any).'

Although these authors do recognize focus-driven DCM patterns in their descriptions, they also recognize DCM effects relating to topicality. Each of these authors, regarding Siona+ and non-Tukanoan languages, indicate that case-marking is related to highlighting arguments whose referents will be important to the ongoing narration.[39] For these languages, specialists converge on describing such discourse-related notions as STRONG TRIGGERS for P-oriented DCM. Otherwise, a few claims for corresponding WEAK TRIGGERS are described for Kakua [*cbv*] (Bolaños 2016, 203-205) and for Kubeo [*cub*] (Chacon and Genetti 2019, 412).

To round off the discussion typologizing DCM as found on the DIRECT OBJECT in Ecuadorian Siona, a final sub-pattern is observed. This concerns the restriction of case-marking within coordinated arguments — identified as an OTHER-type TRIGGER in Table 4.9. The quirkiness of case-marking under coordination is relevant to certain DCM-related discussions in the literature (e.g., Kalin and Weisser 2018; Irimia 2024); however, there are little insights into this dimension of DCM within the NWA sources consulted here. The Ecuadorian Siona pattern is such that no case-marking is found on coordinated P-arguments, including animate nouns which otherwise require marking. The example in (50), lifted from a narrative in Chapter 4, illustrates the general pattern. A related pattern is described for Kubeo [*cub*] (Morse and Maxwell 1999, 111-113), where only the first conjunct is *-re* marked, as shown in (49):

[39] cf. Cook and Levinsohn 1985, §2 regarding such uses of *-ni* in Koreguaje, and *ibid*, §3, regarding focal *-re*; also see Wojtylak 2021, §8.1.1.3 regarding the topical role of the case-marker *-na* in referent management and narrative advancement.

(49) Kubeo$^{\diamond}$ (*cub* — Eastern-Tukanoan) [MM99:112-113, ex. (353)]
boa-'bI ***jupari-Rɨ-re*** *ape-ko* ***dẽĩdoki-ko*** ***warĩ-ko***
kill-3S.M rabo.colorado-CL:3D-OBJ other-CL:F ñacundá-CL:F jacha-CL:F
pidubã-ko *boa-'bI* *ĩ̃* *bɨkɨ-kɨ*
tucunaré-CL:F kill-3S.M 3PRO.S.M old-CL:M
'The elderly man recently caught a *rabo colorado*, (and also) a *ñacundá*, a *jacha*, and a *tucunaré* (guan).'

(50) Ecuadorian Siona$^{\blacklozenge}$ (*snn* — Western-Tukanoan)
. . . yo'oquë sai uye naso huasaiya baquë . . .
jo'o-kɨ sai-ɨ **uje** **nahso** wa + sai-a
do-S.M.PRS.DEP go-S.M.PRS.DEP turkey wolly.monkey kill + go-NEG
bah-kɨ
be-RMT.PST-N2/3S.PST.N.ASS
'. . . doing (this), (he) went out to hunt turkey, monkey(, etc.).'
[NAT*: 20151023orocr001.324]

This discussion has situated the Ecuadorian Siona, and the highly similar Siona+, Direct object-oriented DCM patterns within the NWA area. Besides the presence of the marker *-re*, and the non-Subject pronoun-marking generalization, the Siona+ DCM patterns are distinct from Eastern Tukanoan languages in almost all relevant respects: including, the number of marking alternatives, and the role of the primary Macro-triggers (i.e., animacy, referentiality, and discourse). Although DCM likely developed prior to the Eastern-Western split in the family, the two branches have since diverged sharply. The Siona+ pattern has far more in common with the rich, discourse-based DCM patterns described for Tariana [*tae*] and Murui [*huu*] (cf. Aikhenvald 2015). Future work will delve into these details to determine the plausibility of a contact-induced grammar shift as a diachronic explanation for the complex Siona DCM system.

4.2.3.3 Situating Siona Indirect object-oriented DCM

Per Table 4.7 above, Ecuadorian Siona exhibits DCM patterns on the Indirect object (i.e., the R-argument), as described in Section 3.2.3 of Chapter 3. The R-argument is obligatorily marked, and the dominant pattern concerns animate Recipients, as in (51), where both *-re* and *-ni* are equally available:

(51) Ecuadorian Siona$^{\blacklozenge}$ (*snn* — Western-Tukanoan)

a. *yë' ja'co curani isisi'i yë' gajeire*
jɨ'ɨ ha'-ko kura-ni ĩsi-si'-i **jɨ'ɨ** **gahe-ɨ-re**
1SG parent-CL:F chicken-N.SBJ2 give-FUT-N3S 1SG friend-CL:M-N.SBJ
'(I) will give my mother's chicken to my friend (M).'
[VOL: 20230625ejabi001.014]

b. *yëquëni iye cocare quëajë'ë tijubë*
jɨhkɨ-ni i-je kohka-re kɨa-hɨ̃'ɨ̃ tihũbɨ
1PL.EXCL-N.SBJ2 DEM.PROX-CL:GEN story-N.SBJ tell-IMPER again
'Tell us (EXCL) the story again!' [VOL: 20230525ejabi001.024]

Besides the dominant R-oriented DCM pattern shown in (51), an accidental DCM pattern is also attested where *-na* is produced on an inanimate RECIPIENT, in alternation with *-re*, as in (52), replicated from Chapter 3. Together these examples illustrate that all R-arguments exhibit ALTERNATING-type DCM, i.e., excluding zero-marking, such that the available case-markers are determined by animacy-based noun classes:

(52) Ecuadorian Siona♦ (*snn* — WESTERN-TUKANOAN)
yëquë isiye yëyë yijatoayohuë baidaribëna
jɨhkɨ ĩhsi-je jɨɨ-jɨ jiha + toa + jo-wɨ
1PL.EXCL give-CL:GEN want-N3S.PRS.ASS land + fire + canoe-CL:CONT
bãĩ + dari-bɨ *(-na/-re)
people.COL + community-CL:COL-GOAL/-N.SBJ
'We want to give a bus to the community.'
[VOL/SUG: 20220621eerpa001.016a-c]

In the preceding discussion in this chapter concerning the FUNCTIONAL DOMAIN, in Section 4.2.2.1, it was determined that the same case-markers are active across Siona+ languages, despite more permissive patterns of *-na* marking in Colombian Siona [C*snn*] (Wheeler 1987, 126-127) and Ecuadorian Sekoya [E*sey*] (Johnson and Levinsohn 1990, 48-49). In particular, see examples (16) and (17). Such R-oriented DCM patterns are rare in the NWA sample (i.e., *Q3.3*), since most languages simply recruit a single overt case-marker to encode the INDIRECT OBJECT-argument. The fact that Eastern Tukanoan languages, and Máíhɨ̃kì [*ore*], display obligatory *-re* marking is a well-established generalization in the literature (cf. Barnes 1999, 2006; Gomez-Imbert 2011; Ramirez 2019[1997], 164-166; Stenzel 2013c, §5.2.2). The full typology based upon the NWA dataset is reported in Table 4.10:

LANGUAGE	FORMS	MACRO-TRIGGERS			
		ANIM	REF	DISC	OTHER
*Ecuadorian Siona♦ [E*snn*]	{***-re***, *-ni*, (*-na*)}	✓ANIM		✓FOC	
Colombian Siona♦ [C*snn*]	{***-re***, *-ni*, *-na*}	✓(ANIM)		✓FOC	
Koreguaje♦ [*coe*]	{***-re***, *-ni*}	✓ANIM		✓CONTR	
Murui [*huu*]	{*-∅*, *-mo*}		✓PRO	✓CONTR	
Tena Kichwa [*quw*]	{*-ta*, *-ma*}				✓TBD
VTariana [*tae*]	{*-na*, *-nuku*}		✓PRO		
OBLIGATORY R-MARKING:	*ore*♦ (*-re*) *cub*◊ (*-re*) *tnc*◊ (*-re*) *tuc*◊ (*-re*) *des*◊ (*-re*) *bsn*◊ (*-re*) *gvc*◊ (*-re*) *con* (*=nga*) *qvi* (*-man*) *boa* (*-vu*) *acu* (*=n^{y}a*) *kwi* (*=ta*) V*cbv* (*=di'*) V*jup* (*=ăn*) V*yab* (*-dih*)				

Table 4.10: INDIRECT OBJECT-oriented DCM patterns

Table 4.10 demonstrates that the bulk of the sampled languages (N=15) require overt case-marking on the R-argument, and there is simply one available case-marker. In fact, Murui [*huu*] is the only language which permits zero-marking on pronominal and personal name INDIRECT OBJECTS (Wojtylak 2021, 269), such as the instance shown in (53). This results in an exceptional OPTIONAL DCM pattern, whereas the other

five NWA languages, including the Siona+ languages, display ALTERNATING DCM, as exemplified for Tena Kichwa [*quw*] in (54):[40]

(53) Murui (*huu* — WITOTOAN) [W21:269, exx. (8.95),(8.96)]

a. *dɨo-kaɨ*$_{(P)}$ ***kue***$_{(R)}$ *ine*
tobacco-CL:STEM 1SG give.IMPER
'Give <u>me</u> a cigarette! (normal reading)'

b. *dɨo-kaɨ*$_{(P)}$ ***kue-mo***$_{(R)}$ *ine*
tobacco-CL:STEM 1SG-LOC give.IMPER
'Give <u>ME</u> a cigarette! (abrupt reading, brusque)'

(54) Tena Kichwa (*quw* — QUECHUAN) [G16:103, ex. (2.107)]
Pablo sisa-guna-ta$_{(P)}$ *kuya-n* ***Maria-ma/-ta***$_{(R)}$
P flower-PL-ACC give-3 M-DAT/ACC
'Pedro gave flowers$_{(P)}$ <u>to Maria</u>$_{(R)}$.'

Setting aside the formal types, the Murui pattern demonstrated in (53) aligns with Ecuadorian Siona in that the selection of case-marking alternatives is driven by discourse. More concretely, the overt marking of the R-argument in (53b) evokes a contrastive focal reading (Wojtylak 2021, 269). Chapter 3 provides ample evidence suggesting that such discourse-related notions also drive Ecuadorian Siona DCM and, as stated elsewhere, similar claims are made regarding Siona+ languages (Wheeler 1970, 144-148; Cook and Levinsohn 1985), independently of animacy-based contraints. This analysis is made explicit in the analysis of Colombian Siona [^{c}*snn*] DCM, reflected in Table 4.11, adapted from Wheeler (1970, 174) (cell shading and boldface text not in original. Wheeler does not explicitly define the different focus categories):

	EMPHATIC FOCUS	NORMAL FOCUS	NON-FOCUS
Animate			
S	*-ga*	*-bi*	*-∅*
P	*-ni*	*-de*	*-∅*
R	***-na***	*-ni*	*-de*
Inanimate			
S	X	*-bi*	*-∅*
P	X	*-de*	*-∅*
R	X	***-na***	*-de*

Table 4.11: Focus-dependent case marking alternatives in Colombian Siona (adapted from Wheeler 1970, 174, *Chart XI*)

Side-stepping complications regarding the terminology chosen by Wheeler (1970, *et seq.*), Table 4.11 encapsulates the two primary factors at play in R-oriented DCM across the Siona+ group. On the one hand, animacy-based restrictions on the distribution

[40]Several Quechuan languages display the ALTERNATING DCM pattern shown for Tena Kichwa in (54). However, according to Cole (1982, 104-105), "[t]he use of dative [*-man*] for [R] is obligatory in [Imbabura Kichwa [*qvi*], unlike other Quechuan languages]".

(or concomitant interpretation) of certain markers is reflected by the EMPHATIC FOCUS category — i.e., a category which does not extend to inanimate nouns. On the other hand, the degree of focus is represented in scalar fashion (EMPHATIC FOCUS > NORMAL FOCUS > NON-FOCUS), reflecting the status of this information structural dimension as a STRONG TRIGGER in the language. The former dimension, regarding noun class restrictions, in present in Murui [*huu*], as shown in (53), and is the ony relevant factor in Tariana [*tae*]: i.e., *-na* arises with pronouns, whereas all other arguments take *-nuku* (Aikhenvald 2003, 143-144). The emphatic reading achieved via overt marking in Murui in (53b) aligns fairly neatly with Siona+ focal case-marking patterns.

In summary, R-oriented DCM patterns are rare in the NWA region. The Siona+ languages pattern with other Tukanoan languages in banning zero-marking on the INDIRECT OBJECT; however, variation is observed as to which overt case-marker is selected. The selection of case-markers in this domain is similar to that noted for other argumental DCM patterns in Siona+ languages, such that animacy factors into determining which markers are available to a given nominal, and then discourse-related factors settle the selection of a particular alternative. In this way, the Tariana and Murui languages align with the Siona+ languages, and separately from other surrounding languages, as regards discourse-related factors as STRONG TRIGGERS across the attested argumental DCM patterns.

4.2.3.4 Situating Siona spatial DCM patterns

Up to this stage, this section has typologized the Ecuadorian Siona argumental DCM patterns, and the remainder of the discussion focuses on the spatial DCM patterns attested in the language. Per Table 4.7, there are two spatial CASE FUNCTIONS, which display DCM patterns in the language: i.e., the STATIC LOCATION (L) (*Q3.4* in the NWA questionnaire — described in Section 3.3.1 in the *preliminary description*), and the GOAL-argument (G) (*Q3.5* in the NWA questionnaire — as described in Section 3.3.2). The present discussion collapses these spatial DCM patterns for the purposes of situating the Ecuadorian Siona facts within the NWA dataset.

To guide the discussion below, Table 4.12 spells out the relevant spatial DCM patterns within the Tukanoan languages in the dataset — consult Table A.22 in Appendix A for the full dataset. Since certain languages display the 'generalized spatial-marking' pattern discussed in Section 4.2.2.1 of the present chapter, the CASE FUNCTION(S) are indicated in the table below. Note that the animacy MACRO-TRIGGER is not considered here since this dimension is typically incompatible with spatial roles on semantic grounds:

Language	Function	Form(s)	Macro-triggers		
			REF	DISC	OTHER
*Ecuadorian Siona♦ [ᴱ*snn*]	L	{-∅, ***-re***, (-*ni*, -*na*)}	✓(PRO, SPEC)	✓FOC	✓(PRED)
	G	{-∅, -*na*}	✓(PRO, SPEC)	✓FOC	
Colombian Siona♦ [ᶜ*snn*]	L	{-∅, ***-re***, (-*na*)}	✓PRO, (SPEC)	✓TOP	✓(PRED)
	G	{-∅, -*na*}	✓PRO, (SPEC)		
Koreguaje♦ [*coe*]	L	{-∅, ***-re***, (-*na*)}	✓SPEC	✓TOP	✓(PRED)
	G	{-∅, -*na*}	✓SPEC	✓TOP	
Kubeo◊ [*cub*]	L=G	{***-re***, -*i*, -*rã*}	✓SPEC	✓(FOC)	
Tukano◊ [*tuc*]	L=G	{-∅, ***-re***, -*pʉ*}	✓SPEC	✓(TOP)	
Desano◊ [*des*]	L=G	{-∅, ***-re***, -*ge*}	✓SPEC	✓(FOC)	✓(PROTO)
Barasano◊ [*bsn*]	L=G	{-∅, -*hʉ*, (***-re***)}	✓(SPEC)	✓(FOC)	
Wanano◊ [*gvc*]	L	{***-re***, -*i*, -*pʉ*}	✓REF		✓(DIST)
	G	{***-re***, -*pʉ*}	✓(SPEC)		
No spatial case-markers:		Máíhɨ̃kì♦ [*ore*] (-∅(L/G))			
Obligatory case-marking:		Tanimuka◊ [*tnc*] (-*re'ká*(L); -*rã*(G))			

Table 4.12: Tukanoan spatial DCM patterns attested

It is immediately clear from the data presented in Table 4.12 that *all* Tukanoan languages exhibit rich spatial case-marking patterns. Besides Máíhɨ̃kì [*ore*], overt case-marking is required in Tanimuka [*tnc*], in Kubeo [*cub*] (Morse and Maxwell 1999, 120), and in Wanano [*gvc*] (Stenzel 2013c, 161-162,338, 2013a, 89). Alternating DCM patterns are noted in the latter two languages where multiple case-marking alternatives are available. Formally similar DCM patterns are noted in some surrounding non-Tukanoan languages in the Vaupés region (i.e., Yuhup [*yab*] (Ospina Bozzi 2002, 159-161) and, particularly for L-marking in Hup [*jup*] (Epps 2008, 367-368)), and Alternating spatial DCM is also described for Awa Pit [*kwi*] (Curnow 1997, 135-140) and in Quechuan languages (i.e., L-marking in Tena Kichwa [*quw*] (Grzech 2016, 68), and G-marking in Imbabura Kichwa [*qvi*] (Jake 1983, 22-23)).

Regarding the Siona+ languages, as detailed in Section 3.3 of the preceding chapter for Ecuadorian Siona, zero-marking alternates with multiple case-marking in the context of L-marking, based upon the conspiracy of a principled DCM pattern and an accidental pattern, concerning the marker -*na*. Conversely, zero-marking alternates with the case-marker -*na* in the context of G-marking. As defined in this dissertation, formally, Ecuadorian Siona exhibits Hybrid L-oriented DCM and Optional G-oriented DCM accordingly. The latter Hybrid pattern is accounted for on the basis of certain instances of -*na* marking on the L-argument, as determined by the predicate (i.e., the Other trigger, labeled PRED in Table 4.12), already discussed in the context of the Functional domain, in Section 4.2.2.1 of the current chapter. Nearly identical patterns are described across the Siona+ group. As shown in Table 4.12, zero-marking is also attested for the generalized spatial-marking in certain Eastern Tukanoan languages, representing the Hybrid type: i.e., as described for Tukano [*tuc*] (Ramirez 2019[1997], 161-169), for Desano [*des*] (Silva 2012, 159-161, 174-175), and for Barasano [*bsn*] (Jones and Jones 1991, 69-72), as well as a few other Eastern Tukanoan languages beyond the NWA dataset discussed here (cf. Stenzel 2013a).

Regarding the active Triggers in spatial DCM across the Eastern Tukanoan and the Siona+ sample, a similar pattern is observed to the discussion of argumental DCM. More concretely, these marking patterns tend to be triggered by referentiality-

based factors in most Eastern Tukanoan languages, particularly regarding the usage of *-re*, which overlaps with its distribution on non-SUBJECT arguments (cf. Stenzel 2013a, 2013c, 161-163, 338; Zúñiga 2007). For several authors, this semantic overlap motivates the usage of the label 'referential'- or 'specificity'-marker for the case-marker *-re* (Barnes 1999, 2006; Waltz and Waltz 1997; Ramirez 2019[1997]; Stenzel 2008, 23), and is a primary, STRONG TRIGGER for the various argumental and spatial CASE FUNCTIONS for this case-marker. Nonetheless, a handful of Eastern Tukanoan spatial case-markers are described as being selected to encode discourse-related notions, e.g., the focal locative case-markers, *-rã̃* in Kubeo [*cub*] and *-pʉ̵* in Tukano [*tuc*].[41] A similar claim is made regarding the contrastive spatial marker *-có* in the nearby Nadahupan language, Hup [*jup*] (Epps 2008, 367-368). The sentences in (55) demonstrate how the degree of specificity of the concerned locations factors into the selection of spatial case-marking in Tukano [*tuc*]:[42]

(55) Tukano$^{\diamond}$ (*tuc* — EASTERN-TUKANOAN)
[R97:168-170,237; exx. (339),(580),(340)]

a. ***a'tó*** *kãrigɨ́* *dĩĩ-á-pɨ*
here sleep-NOM.M.SG be-REC.PST-1SG
'I usually sleep here.'

b. ***a'tó-de*** *dõ'ó-pɨ* *kãdí-gɨ́* *dĩĩ-a-ti*
here-N.SBJ where-FOC sleep-NOM.M.SG be-REC.PST-Q
'Here, where did you (M) sleep? (here in the city, where did you sleep?)'

c. ***a'tó-pɨ*** *a'tî-a*
HERE-FOC come-IMPER
'Come HERE!'

d. ***toó-pɨ-de*** *peêdu* *sĩ'dí-wɨ*
ANA.LOC-FOC-N.SBJ caxiri_drink drink-REP.PST.VIS
'I will drink *caxiri* THERE (an established location).'

Whereas Eastern Tukanoan languages tend to display referentiality-related TRIGGERS for spatial DCM, aligning with their argumental DCM patterns, Siona+ languages also display similar sets of TRIGGERS for spatial and argumental DCM — i.e., discourse-related STRONG TRIGGERS are the principle drivers of the observed alternations. This is demonstrated for Ecuadorian Siona in Section 3.3 of the *preliminary description.* In a similar vein, Cook and Levinsohn (1985) demonstrate that the distribution of spatial case-marking is determined primarily by topicality in Koreguaje [*coe*]. The authors provide the stretch of discourse in (56), which implicates several G-arguments, and the usage of overt *-na* in (56b) stresses the importance of this location for narrative progression:

(56) Koreguaje$^{\blacklozenge}$ (*coe* — WESTERN-TUKANOAN) [CL85:111, ex. (45i)-(45iii)]

a. ***kee*** ***ja'o-cha*** *coʉ̵* *sai-na'-me* *chʉ̵kʉ̵na*
DEM.MED leaf-CL:RIV turtle go-PRP-N3S 1PL.EXCL
'We (EXCL) go to the Caquetá River for turtles . . . '

[41] It is relevant to note that Stenzel (2013c, §10.5) claims that the cognate locative marker, *-bʉ̵* indicates a plain, nonreferential location; whereas *-re*, or the combination *-bʉ̵-re*, indicates an established, well-defined, referential location.

[42] Similar examples of specificity- and discourse-driven spatial DCM patterns are provided for Kubeo [*cub*] and Desano [*des*] in (103) and (104) respectively in Appendix A.

b. ***meja-huʉ-na*** *coʉ* *mai-na'-me*
sand-CL:CONT-GOAL turtle come-PRP-N3S
'... the turtles come up <u>on the beach</u> ...'

c. *jainʉko* ***meja-huʉ*** *coʉ* *mai-me*
many.PL.ANIM sand-CL:CONT turtle come-N3S
'... many are the turtles which come up <u>on the beach</u>.'

Certain authors also hint at referentiality-related factors for spatial case-marking in Siona+ langauges: i.e., Ecuadorian Sekoya (Johnson and Levinsohn 1990, 47-48), and Koreguaje (Cook and Criswell 1993, 34-37); and similar generalizations are claimed regarding *-na* marking on G in Ecuadorian Siona (Bruil 2014, 170-171). Ultimately, based on the available descriptions, like argumental DCM, discourse-related TRIGGERS for spatial DCM appear to take precedence over referentiality-related WEAK TRIGGERS, as conceived in this dissertation.

At this stage, it is relevant to note that the overlap observed across the Tukanoan language family concerning the set of active TRIGGERS for spatial and argumental DCM is not attested in the non-Tukanoan languages in the NWA sample. Just like the P-L correspondence pattern discussed in the FUNCTIONAL DOMAIN (*Q2.4*), this appears to be a pan-Tukanoan trait, rather than a common property in the NWA area. Outside of the Tukanoan family, most spatial DCM patterns are rather distinct from their argumental counterparts. Two common spatial DCM patterns are noted across the non-Tukanoan sub-sample.

Firstly, as shown in the full dataset for *Q3.4-Q3.5* displayed Table A.22 in Appendix A, some languages restrict zero-marking to inherently spatial concepts (i.e., the OTHER-type TRIGGER, labeled PROTO). This patterns is described for Murui [*huu*] (Wojtylak 2021, 273-275), where DCM is restricted to spatial pronouns or common spatial concepts like, *iyɨ* (jungle), and highly similar patterns are described for Shiwiar [*acu*] (Kohlberger 2020, 256-262), shown in (107) and (108) in Appendix A respectively. Another language with this pattern is Tariana [*tae*] (Aikhenvald 2003, 148-150). Conversely, in most Tukanoan languages, including all Siona+ languages, *all* spatial noun phrases are treated equally as regards DCM.[43]

Per Table A.22, the second type of spatial DCM, which is distinct from the Tukanoan-type patterns outlined above, concerns languages where spatial case-marking is selected on the basis of deictic distance distinctions (i.e., the TRIGGER is labeled DIST). This type of system is attested in Nadahupan languages spoken in the Vaupés region, i.e., Hup [*jup*] (Epps 2008, 367-368) and Yuhup [*yab*] (Ospina Bozzi 2002, 156-162),[44] and is also described for Awa Pit [*kwi*] (Curnow 1997, 135-140 — shown in (110) in the appendix). Similar distance-related uses of spatial markers are described as a WEAK TRIGGER for the Tukanoan language, Wanano [*gvc*] (Stenzel 2013c, 338). Otherwise, such deictic-based spatial DCM patterns are unattested in the Tukanoan languages within the NWA dataset.

Ultimately, as demonstrated in the FUNCTIONAL DOMAIN, particularly *Q2.3*, the dominant Tukanoan spatial DCM system demonstrates significant overlap with argu-

[43] A possible exception to the generalization that all spatial noun phrases exhibit the same DCM properties in Ecuadorian Siona concerns the class of spatial promoted pronouns described in Section 3.3.1 in Chapter 3. As described there, these pronouns, like any promoted L-nominal, must be marked with either *-re* or *-ni*, or obligatorily with *-na* where this refers to a G-argument. Ultimately, this is not a property of spatial marker per se, but rather a property of *promotion*, which is active across all non-SUBJECT arguments.

[44] An illustrative example of this spatial DCM system is provided in (109) in Appendix A.

mental DCM. This generalization holds on formal grounds, since the marker *-re*, and *-na*, as concerns Siona+ languages, are distributed in both types of DCM. This also holds with respect to the active TRIGGERS which drive case-marker selection in subtypes of DCM. The NWA dataset demonstrates that this generalization holds to the extent that Eastern Tukanoan languages favour referentiality-driven DCM, whereas Siona+ languages, including Ecuadorian Siona, favour discourse-driven DCM across all concerned DCM FUNCTIONS. These dominant STRONG TRIGGERS differ across the groups of Tukanoan languages analyzed here (excluding Máíhɨ̃kì [*ore*]); however, there are other WEAK TRIGGERS, which add nuance to the typologization of spatial DCM that merit further investigation.

4.3 Conclusion: Siona case-marking in typology

For this chapter, I put on my *comparativist* hat with the aim of typologizing the Ecuadorian Siona case-marking and DCM properties outlined in Chapter 3. The NWA dataset, reported in Appendix A, is designed to contextualize various case-marking properties, corresponding to the three primary pillars of the characterization of DCM outlined in Chapter 2: comprising the FORMAL DOMAIN, the FUNCTIONAL DOMAIN, and the REALIZATIONAL DOMAIN. This chapter finds that the Ecuadorian Siona facts are consistent with those described across all Siona+ languages, excepting a few points of micro-variation, and strikingly different from the corresponding Eastern Tukanoan facts. On the other hand, there are reasons to believe that certain traits may have arisen due to contact-induced convergence. However, among the candidates for points of grammatical convergence and cognacy, some of the languages are spoken to west of the Siona+ cluster, in the foothills and into the Andes proper (i.e., Quechuan languages, Shiwiar [*acu*] — cf. Bruil 2015a; van Gijn et al. 2023); whereas others implicate potential contact to the east, involving the languages of the *People of the centre* (i.e., Murui [*huu*] and Bora [*boa*] — cf. Echeverri 1997; Epps and Michael 2017; etc.).[45] Ultimately, Ecuadorian Siona displays typical NWA case-marking traits, including an admixture of Tukanoan and non-Tukanoan properties. This section summarizes the targeted discussion in this chapter.

Firstly, as concerns the FORMAL case-marking properties laid out in Section 4.2.1, Ecuadorian Siona exhibits a LARGE case-marker inventory (N=6), which aligns with the mean inventory size in the NWA sample (i.e., 5.43 members). This inventory comprises the set of 'primary' Siona+ case-markers (i.e., *-re*, *-ni*, *-bi*, and *-na*), and two additional, historically related, markers, *-hã'ã* (LIMIT/PATH) and *-hã're* (COM). Whereas these latter two markers find cognates across the Siona+ grouping, displaying varying degrees of grammaticalization, the primary case-marker set is more stable across the group. However, *-ni* likely developed more recently in the Upper Putumayo-Napo Siona+ varieties, given its absence in Peruvian Sekoya, spoken in the Middle Putumayo region (Rosa Vallejos, p.c.).

[45] It is possible that contact-induced convergence with the *People of the centre* groups may have occured indirectly, via contact with Tanimuka [*tnc*] and/or Kubeo [*cub*], if indeed an earlier connection between these particular Eastern Tukanoan and the Siona+ languages can be established. Of course this also introduces the possibility for the extensively-documented contact between Arawakan groups and these Eastern Tukanoan groups as a source for certain cognate forms, or other patterns. However, these languages are not considered in the NWA dataset developed here, so further work is merited.

Based upon several candidates for cognacy discussed in Section 4.2.1.2, it is conceivable that *-bi* and *-na* are retained from a period prior to the east-west split. In particular, the presence of cognate markers, executing similar CASE FUNCTIONS, particularly in Tanimuka [*tnc*] and Kubeo [*cub*], spoken at the western fringes of the Vaupés region, is suggestive of a shared innovation. Admittedly, the relatively impoverished case-marking system in the Western Tukanoan language, Máíhɨ̃kì [*ore*], complicates this picture. Nonetheless, even considering a few non-Tukanoan candidates which may factor into the equation in some way, it is likely that all six Ecuadorian Siona case-markers are retained from earlier stages of development. Further research is awaited to refine the chronology of the establishment of the inventory attested today.

Besides the six members of the Ecuadorian Siona inventory, this chapter also touched upon one marker, i.e., the contrastive topic marker, *=kato* (following Bruil 2014), whose cognate *-ga* is analyzed as a bona fide case-marker in Colombian Siona [^{c}*snn*] (Wheeler 1970, 173-176, 1987, 131). According to the author, *-ga* is the SUBJECT-marking equivalent to the non-SUBJECT marker *-ni*. There are several reasons for which this analysis is incompatible with the Ecuadorian Siona facts, which are discussed in the ensuing chapters of this dissertation (see especially Section 6.4.2 of Chapter 6). However, *=kato* is the most frequently occurring nominal discourse marker in the language,[46] and, more importantly, it displays certain unique co-occurence restrictions with particular case-markers (see Section 6.4.2). Ultimately, as is found with the development of PATH- (or LIMIT-) and COMPANION-markers in Siona+ languages, it is certainly plausible that *-ga* became grammaticalized as a case-marker in Colombian Siona, but developed separately as a discourse marker in Ecuadorian Siona. Yet again, setting aside questions of analysis, these markers are cognate, representing a shared innovation across the Western Tukanoan branch.

Turning to the FUNCTIONAL DOMAIN, discussed in Section 4.2.2, the NWA questionnaire is designed to bias certain case-marking correspondences, and FUNCTIONAL GAPS, which are attested in Ecuadorian Siona grammar. The first tested correspondence (*Q2.1*), found that the familiar SUBJECT-INSTRUMENT IDENTITY-type noted across the Siona+ group, is only otherwise attested in Tariana [*tae*], which is itself unique among the languages in the Vaupés region (cf. Aikhenvald 2015). This observation is compatible with the fact that S-marking is typologically rare in the NWA area; however, Siona+ languages appear to have developed an S-marking system by adapting the oblique marker *-bi*, whose cognates also encode both INSTRUMENT and SOURCE, also encoded by *-ϕí* in Tanimuka [*tnc*] (Eraso 2015, 280).[47]

A related correspondence addressed in the NWA dataset concerns the INSTRUMENT- COMPANION correspondence (*Q2.3*). In none of the languages with a cognate for *-bi*, including all Siona+ languages and Tanimuka [*tnc*], is this marker used to

[46] **Table B.10** in Appendix B spells out the distributional facts pertaining to discourse markers arising in the corpus sample for Chapter 5.

[47] Although no Eastern Tukanoan languages, besides Tanimuka [*tnc*], exhibit S-marking at all; several authors describe discourse markers, which favour S- and spatial argument-marking: i.e., Tukano [*tuc*] (*-pʉ*, glossed FOC in Ramirez 2019[1997]); *-pʉ* in Desano [*des*] (Silva 2012, 162); and *-bɨ* in Máíhɨ̃kì [*ore*] (cf. Neely 2012). In fact, the latter marker is also attested in INSTRUMENT-oriented interrogatives (Neely 2012); suggestive of a supple SUBJECT-INSTRUMENT correspondence. Like *-ga/=kato*, the role of discourse markers may be relevant to the broader enterprise of comparative inquiry into the diachrony of Western Tukanoan nominal morphology. It is unclear whether these facts are directly relevant to the discussion of case-marking at hand.

encode both of these FUNCTIONS, such that the DIFFERENTIATION-type is observed. As demonstrated in Table A.10 in Appendix A, all remaining Tukanoan languages display the dominant IDENTITY-type in the NWA region. This latter type is also displayed by Máíhɨ̃kì [*ore*], where the marker *-jànù* executes both FUNCTIONS, despite the fact that it is historically related to the dedicated COMPANION-marker *-hã̃'re* in Ecuadorian Siona and its cognates in other Siona+ languages.

The remaining tested case-marking correspondences (*Q2.2, Q2.4-2.5*) point to two pan-Tukanoan properties; namely where the Ecuadorian Siona facts align with all Tukanoan patterns revealed in the NWA dataset:

- The DIRECT OBJECT (P) *may* arise with zero-marking, whereas the INDIRECT OBJECT (R) requires overt case-marking ;
- There is a certain degree of overlap regarding the forms available to argumental and spatial case-marking (i.e., *-re*).

Inasmuch as these case-marking properties are common to all Tukanoan languages, Siona+ languages deviate from Eastern Tukanoan languages with respect to a handful of details.

The most relevant discrepancy for the present discussion concerns the role of GOAL-marking, via *-na*, which is available to both the R-argument, and to certain STATIC LOCATION (L-)argument, to varying degrees among the Siona+ languages (discussed especially in Section 4.2.3.3). In none of these languages is the GOAL-marker is permitted on the P-argument. Regarding L-marking, these languages have a relatively stable set of predicates which display (optional) *-na* marking (e.g., 'sleep', 'stay', 'work'), which overlaps with certain GOAL-marker-recruiting predicates attested in neighbouring Quechuan varieties. Although Eastern Tukanoan languages also display rich spatial case-marking systems, none of them display the same non-prototypical uses common to the Siona+ languages — of course, most of these languages do not have dedicated GOAL-marking in any case (see discussion in Sections 4.2.1.3, 4.2.3.4, etc.).

In addition to the discrepancies outlined above, Section 4.2.2.2 discusses a few FUNCTIONAL GAPS (*Q2.6* and *Q2.7*), which also trace a clear divide between the Siona+ group and the Eastern Tukanoan languages. The Siona+ languages do not recruit overt case-marking to encode the POSSESSOR or the TIME FUNCTIONS, as defined here, whereas these patterns are commonplace among the sampled Eastern Tukanoan languages. In this way, Siona+ again patterns more closely with other non-Tukanoan languages, like Murui [*huu*], Bora [*boa*], and A'ingae [*con*].

Finally, the discussion regarding DCM DOMAIN further established the divide between the Siona+ languages and the Eastern Tukanoan family. Above it is noted that SUBJECT-oriented DCM is rare in the NWA area — however, the Siona+ languages pattern fairly closely with Murui [*huu*] and Tariana [*tae*] in displaying discourse-driven DCM for the SUBJECT-argument. In abstract terms, regarding DIRECT OBJECT-oriented DCM, all Western Tukanoan languages, including Máíhɨ̃kì [*ore*] (Farmer 2015, 97), display animacy-triggered patterns; whereas referentiality is the primary TRIGGER across the Eastern Tukanoan languages considered here. It is also the case that discourse-related TRIGGERS are active for all argumental and spatial DCM patterns in the Siona+ languages; whereas only a few disparate discourse-driven DCM patterns are described across the Eastern Tukanoan languages, generally attributed a secondary status to referentiality-related TRIGGERS. Ultimately, although the overlap between argumental and spatial DCM is a general characteristic of both groups of Tukanoan

languages, each sub-type maintains a distinctly Siona+ or Eastern Tukanoan flavour. Various traits particular to, at least, Máíhɨ̃kì [*ore*], and to Tanimuka [*tnc*], complicate the typological and diachronic picture.

In summary, the NWA case-marking dataset presented in Appendix A, and discussed in the present chapter, corroborates the claim in Bruil (2014, 10-12), that Ecuadorian Siona belongs to a continuum of highly related language varieties. Based upon the case-marking properties analyzed here, the Koreguaje [*coe*] also displays compatible patterns. On the one hand, some Tukanoan languages align more closely with Siona+ case-marking than others — particularly Tanimuka [*tnc*] and Kubeo [*cub*]. On the other hand, the NWA data also highlights some aspects of the Ecuadorian Siona case-marking, which align more with neighbouring, non-Tukanoan languages than with other Tukanoan languages. These points of convergence arise from languages spoken both to the west and to the east from the regions where Siona+ languages are spoken today. On this basis, it is plausible that some case-marking traits are the result of contact-induced convergence. The discussion in this chapter is a contribution to the study of the typology of the NWA area and to the diachrony of Tukanoan case-marking, which opens several doors for future scrutiny. The following chapters return to the analysis of DCM patterns particular to Ecuadorian Siona.

CHAPTER 5

A variationist study: DCM in a narrative sample

For the present chapter, I put on my *variationist* hat in order to ascertain the actual distribution of case-marking alternatives across a corpus of naturalistic Siona data — i.e., referred to as the Siona narrative dataset. Therefore, this chapter contributes quantitative data, which complement the qualitative description established elsewhere in this dissertation. More concretely, the actual PROMINENT-marking rates for various Siona DCM RELATIONS is ascertained, while the application of statistical techniques probes the activity of plausible TRIGGERS for Siona DCM effects. The statistical models themselves are reported in full, alongside lifted illustrative examples, in Appendix B, which accompanies the discussion developed in this chapter. Ultimately, this quantitative chapter is a welcome contribution to analyzing the REALIZATIONAL DOMAIN, as a component of the holistic characterization of DCM patterns in the language, as pertains to the set of active TRIGGERS, which drive case-marking patterns in actual usage.

This chapter is organized as follows: Section 5.1 outlines the Siona narrative sample developed for this chapter. Section 5.2 outlines the methodological approach adopted to assemble and code the narrative sample. This section also introduces the variationist techniques utilized to establish the statical models developed to analyze the sample. These models are spelled out in full in Appendix B, which accompanies the present chapter. Section 5.3 discusses the quantitative results on a grammatical relation-by-grammatical relation basis, considering DCM patterns for argumental and for spatial DCM patterns in turn across the narrative sample. Section 5.4 closes the present chapter by integrating the quantitative data into the broader characterization of Siona DCM purported in this dissertation.

5.1 The Siona narrative dataset

The Siona narrative dataset analyzed in this chapter, and Appendix B, comprises five narratives from the ELDP documentary corpus (Bruil 2012),[1] as enumerated in Table 5.1. On the one hand, the length of each narrative is measured in "Time", corresponding to the duration of the raw video recording in the archival materials; whereas "Word count" is determined based upon the lifted transcriptions, which form the basis for the token-based analysis reported here. Each narrative is attributed an English title for convenience, although the analysis purported here is performed upon the pooled dataset:

STORY TITLE [Recording code in Bruil 2012]	TIME (HH:MM:SS)	WORD COUNT	NOMINAL TOKENS *N* (% of total words)
Siona Genesis story [20151023orocr001]	21:33	1714	376 (21.94%)
The demon kwẽẽwahti [20151112oespa001]	15:48	1291	327 (25.33%)
The mother of jaguars [20151001oolpi001]	14:12	981	251 (25.59%)
The mother of the moon [20151112orapi001]	16:16	1350	331 (24.52%)
The children of Baina [20150811sfryi001/2]	33:25	2742	629 (22.94%)
POOLED TOTAL	**1:41:04**	**8078**	**1914 (23.69%)**

Table 5.1: Composition of the Siona narrative dataset

The corpus outlined in Table B.1 is designed in order to provide a balanced and representative sample of the traditional oral folktale genre (cf. Biber et al. 1998, Jones and Waller 2015). In addition to the semi-spontaneous nature of these narratives, the included folktales are appropriate for token-based, statistical inquiry because they implicate a set of characters, which are tracked across longer stretches of discourse. As such, this genre is an ideal testing ground for certain plausible DCM TRIGGER candidates, which have been suggested elsewhere — e.g., specificity and topicality.

To ensure the development of a balanced narrative sample, several factors were taken into consideration. Firstly, each story is told by a different native-speaker narrator: including two middle-aged males, one elderly male, one middle-aged female, and one elderly female.[2] Secondly, the narrative sample is fairly balanced as con-

[1] Project ID IPF0211SG0067 [https://www.elararchive.org/dk0184]. All materials were collected by Martine Bruil between 2010-2015 under affiliation with Leiden University and University of California, Berkeley. Stories are archived as multimedia bundle folders, which may be accessed via the ELAN (EUDICO linguistic annotator) tool: ELAN (Version 6.9) [Computer software]. (2024). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from https://archive.mpi.nl/tla/elan.

[2] Given the degree of intermarriage with neighbouring communities, it is not always the case that all speakers are equally confident in their Siona language competency. This fact was taken seriously as Prof. Bruil and I determined which narratives to include in this corpus

cerns story length, as shown in Table B.1, with a mean word count of 1615 word forms. This ensures that no narrator is represented disproportionately in the pooled dataset. Additionally, the sample size utilized in this analysis is in keeping with other manually-coded quantitative (DCM) studies in the literature (e.g., Shain and Tonhauser 2010; Tonhauser and Colijn 2010; Bruil 2014; Stenzel 2015; Grzech 2016). Crucially, this sample size is also suitable for the application of statistical techniques (i.e., the *variable*-rule model), whereas smaller datasets do not yield informative results. More methodological details are provided in the following section.

5.2 Methodology: The variationist program

The present chapter adopts a *variationist* methodology for collecting and testing quantitative data, based upon the Siona narrative dataset outlined in Table B.1 above. The family of variationist techniques is unified by the assumption that natural language contains *linguistic variables*, which are components of language which exhibit structured variation in their distribution (cf. Tagliamonte 2006, 2012). The conception of DCM adopted in this dissertation is compatible with the conventionalized notion of *linguistic variable* — i.e., "the phenomenon whereby not every argument bearing a given grammatical relation is encoded with the same case-marker" (lifted from (2) in Chapter 2). For instance, consider the two variant forms of the spatial GOAL-argument, *airo* (jungle), shown in (1) — i.e., the PLAIN zero-marked variant arises in (1a), lifted from *The mother of jaguars* narrative; whereas the PROMINENT *-na* marked variant arises in (1b) from *The children of Baina* narrative:

(1) a. *... ba'iquëbi dos dia jesǫ mo'seña baja'i ba'iquëbi airo saiyë caëña*
ba'i-ki-bi dos dia hesõ mo'se-ña ba-ha'i
be-CL:M-SBJ two day how_many day-PL.INAN be-3S.M.PST.ASS
ba'i-ki-bi **ai-ro** sai-jɨ kaa-i-ña
be-CL:M-SBJ big-CL:LOC go-N3S.PRS.ASS say-2/3S.M.PST.N.ASS-REP
'... he was (there) two days, or however many days it was, he wanted to go to the forest (to go hunting).' [NAT*: 20151112oespa001.076-77]

b. *... airona sani tǫmejë huahuaroquë'rona sanitǫmereña*
ai-ro-na sani + tõme-hɨ wawaro-kɨ'ro-na
big-CL:LOC-GOAL go + fall-PL.PRS.DEP cockroach-DERIV.LOC-GOAL
sani + tõme-re-ña
go + fall-N2/3S.PST.N.ASS-REP
'... (they) landed in the forest, they landed where the COCKroaches live.' [NAT*: 20150811sfryi001.352-356]

The proportion of PROMINENT-marking may be established in a systematic way. Firstly, and more straightforwardly, it determines the rate at which PROMINENT-marking arises in the narrative sample — i.e., simply referred to as proportional evidence. Secondly, it is not immediately clear why PROMINENT *-na* marking arises in (1b), but not in (1a), for instance. This chapter develops statistical models in order to ascertain which of a series of predictor variables drive the selection of PROMINENT-markers for a given grammatical relation, each of which is conceived as a plausible DCM TRIGGER candidate — i.e., contributing statistical evidence.

sample.

This analysis is performed on each grammatical relation displaying principled DCM alternations, as described in the *preliminary description*, and reiterated in Table 5.2:

	INANIMATE		ANIMATE (∧ ↑INAN)	
	PLAIN	PROMINENT	PLAIN	PROMINENT
Argumental DCM:				
SUBJECT (S)	-∅	*-bi*	-∅	*-bi*
DIRECT OBJECT (P)	-∅	*-re*	*-re*	*-ni*
INDIRECT OBJECT (R)	*-re* (*-na*)		*-re*	*-ni*
Obligatory case-marking:	*-bi* (Instrument); *-hã're* (Companion)			
Spatial DCM:	INANIMATE		↑INANIMATE	
LOCATION (L)	-∅	*-re*	*-re*	*-ni*
[*some L] GOAL (G)	-∅	*-na*	*-na*	
Obligatory case-marking:	*-bi* (Source); *-hã'ã* (Path, Limit)			

Table 5.2: Siona DCM paradigm, PLAIN-PROMINENT mappings for corpus annotation

For the remainder of this chapter, each DCM RELATION is identified by the shorthand identifier in Table 5.2 — e.g., S-tokens refer to those categorized as SUBJECT nominal tokens. The analysis addresses two research questions for each tested grammatical relation: (i) how often PROMINENT alternatives are selected in the sample, and (ii) what TRIGGERS factor into the selection of PROMINENT, over PLAIN, forms in a given context. Both proportional and statistical evidence are welcome quantitative contributions to the characterization of Siona DCM.

The remainder of this section outlines the how variationist techniques were applied to ascertain this quantitative evidence against the Siona narrative dataset. Section 5.2.1 outlines how the pool of nominal tokens is assembled, and introduces the DCM TRIGGER candidates, which were converted into factors (i.e., predictor variables) for the sake of statistical analysis. Section 5.2.2 introduces the *variable rule* model, which is the statistical technique applied in the study reported in this chapter.

5.2.1 From tales to tokens: Coding the narrative sample

The variationist methodology adopted in this chapter is applied against the narrative sample, which comprises 1914 nominal tokens in its raw form. These tokens within the pooled set are subsequently sorted into their respective grammatical relation categories, which are enumerated in Table 5.2. Once sorted on the basis of grammatical relation, the case-marking alternative (incl. zero-marking) found on each token is recorded. In post hoc fashion, for each individual token, the 'PROMINENT-marking status' category is determined, which is the dependent variable for this study.

Compare the instances of P-tokens reported in (2). Each of the highlighted tokens in these instances bears the case-marker *-re*. However, per the PLAIN-PROMINENT mapping summarized in Table 5.2, the inanimate P in (2a) belongs to the PROMINENT-

marked category; whereas its animate counterpart in (2b) is in the PLAIN-marked category accordingly:

(2) a. [**Context:** The children of Baina have killed the mother of the household and put her in the stew, but the father is not yet aware.]
. . . ñajëna yë'ë biada'care cusi'i caëña
ñaa-hi-na ji'i **bia** **+ da'ka-re** ũhku-si'-i
see-PL.PRS.DEP-DS 1S aji_pepper + liquid-N.SBJ drink-FUT-N3S
kaa-i-ña
say-2/3S.M.PST.N.ASS-REP
'. . . (they) watched (him) as (he) said "I want (some) soup." '
[NAT*: 20150811sfryi001.264]

b. *jaje yo'co aco'ë tsidohuëre yureta'a quëaja'cua'ia'ë ja'quëre careña*
hãhẽ jo'-ko ãh-ko'i **tsĩ-dowi-re**
like_that do-S.F.PRS.DEP eat-3S.F.PST.ASS child.COL-PL.AN-N.SBJ
jure-tã'ã kia-hã'-ko-wa'i-a'-i ha'-ki-re
now-CNT.EXP tell-PRP-CL:F-PL.AN-COP-N3S parent-CL:M-N.SBJ
kaa-re-ña
say-N2/3S.PST.N.ASS-REP
' "(She) did this and ate the children, now we will tell father", (they) said.'
[NAT*: 20151001oolpi001.067]

The examples in (1) and (2) suffice to demonstrate how tokens are identified, extracted, and sorted. Previous descriptive work, assembled in the *preliminary description* in Chapter 3, permits the identification of grammatical relations for a given token and the determination of its PROMINENT-marking status. Once these initial stages are undertaken, the proportional evidence may be calculated in a straightforward way.

The next stages lay the groundwork for the implementation of statistical techniques — i.e., the *variable-rule* model, introduced in Section 5.2.2. The initial step concerns the systematic coding of nominal tokens on the basis of factors, or, conceptually, trigger candidates. In the variationist literature, this is typically referred to as *circumscribing* the independent variables (Sankoff 1988; Tagliamonte 2006, 2012; etc.). An elaborate overview of the coding protocols is provided in Section B.1 of Appendix B; whereas, this section is concerned with motivating the independent variables selected for the statistical analysis reported in this chapter.

The independent variables tested in this chapter correspond to well-established, cross-linguistic TRIGGERS for DCM patterns, discussed in Chapter 2. Conceptually, each of these variables is a DCM TRIGGER CANDIDATE, viewed as a plausible motivation for selecting PROMINENT-marking in a given context. Table 5.3 enumerates and provides a working definition for each TRIGGER CANDIDATE tested in this chapter, whereas details regarding their operationalization for coding purposes are given in the appropriate places in Appendix B:[3]

[3]These variables are assembled into groups by convention, these add conceptual coherence to the discussion found in this chapter.

Trigger candidate (predictor variable)	Working definition
	Sentence-level Triggers
Word order	The nominal token appears before or after the main verbal word in its clause.
Thematic status	The referent of the nominal token is attributed a particular thematic role based upon the predicate: (i) *agentive* or *non-agentive* (s-argument); (ii) *affected* or *unaffected* (p-argument); or (iii) *subcategorized* or *non-subcategorized* (l-argument)
	Referent-level Triggers
Animacy status	The nominal token denotes an animate or inanimate entity.
Specificity status	Pronominal tokens are separated from full nominal tokens. Full nominal tokens refer to specific or non-specific entities.
	Discourse-level Triggers
Contrastivity status	The referent of the nominal token contrasts with another entity (explicit or implicitly) in the narrative.
Topicality status	(1) The referent of the nominal token has a recent, overt mention in the discourse (Look-back topic) (2) The referent of the nominal token persists as a topic in immediately ensuing discourse (Look-ahead topic)

Table 5.3: Working definitions for annotated factors in Siona narrative sample

The first group of trigger candidates tested in this chapter concern sentence-level factors — precise annotation guidelines are provided in Section B.1.1. In certain languages, case-marking is obligatory where the noun is in one sentential position, but variable in another. For instance, certain Eastern Tukanoan languages display such patterns for Direct object-marking, where dcm strictly arises in the immediate pre-verbal position in the sentence (Stenzel 2013c, 333-5). Although these particular constraints on dcm are not observed in Siona or other Western Tukanoan languages, this study tests for any statistical correspondence between the surface position of the nominal token (i.e., either as pre- or post-verbal) and case-marker selection. Put another way, this study measures the impact of deviation from the canonical sov word order on dcm, such that both canonical s- and p-tokens arise pre-verbally.

A second family of trigger candidates concern the impact of thematic properties of an argument on its case-marking. As discussed in Section 2.2.3, in various languages, the 'agentivity' status of the Subject impacts case-marking (Malchukov 2008; Fauconnier 2011); and similar patterns are observed regarding the 'affectedness' status of the Direct object (Næss 2004; von Heusinger and Kaiser 2011; Kizilkaya et al. 2022). Both of these factors are tested systematically in this chapter. This study also pilots the testing of whether the sub-categorization status of spatial arguments affects dcm — i.e., distinguishing implied Locations in a predicate like 'we live in the house', from adjunctive Locations, e.g., 'we are playing in the park'. This chapter adapts the technique used to test for affectedness effects in Spanish *differ-*

ential object marking found in von Heusinger and Kaiser (2011) — the technique is spelled out in detail in Section B.1.1.

The third group of tested trigger candidates are referent-level factors, which form the best-studied set of DCM TRIGGERS in the literature. The *preliminary description* provided in Chapter 3 established unequivocally that Siona (non-SUBJECT) DCM is shaped by animacy-based noun classes, as reflected in Table 5.2 above. Whereas this fact clearly factors into the mapping of *formal* alternatives, the statistical analysis developed in this chapter determines whether notional animacy class is also a relevant factor regarding the proportion of PROMINENT-marking. This aspect of Siona DCM has not been addressed in previous work, although such WEAK TRIGGERS are described for various (non-)Tukanoan languages in the NWA area, as shown in Chapter 4.

An important contribution of the quantitative analysis developed in this chapter is to test whether specificity factors into Siona DCM, which is at best an unsettled issue up to this stage. nominal tokens are sorted into pronominal, specific, and non-specific categories (consult Section B.1.2 for annotation details), and it is determined whether these classes display different PROMINENT-marking rates. This is a dominant TRIGGER according to many descriptions for Tukanoan languages; and, although this dissertation has demonstrated, at least qualitatively, that specificity is not a STRONG TRIGGER for Siona DCM, it remains to be seen whether it has any bearing whatsoever — i.e., as a WEAK TRIGGER for (P-oriented) DCM in the language.

The final group of trigger candidates tested in this chapter pertains to discourse-related TRIGGERS. Firstly, CHAPTER 3 establishes that PROMINENT-marking is related to certain focus effects in Siona DCM, such as QUESTION-ANSWER constructions and what is labeled the *emphatic contrast* reading, which are expounded considerably in Chapter 6. Since there were not a significant number of questions or answers in the narrative sample, these are not directly coded for in the sample. Instead this chapter contributes to this broader discussion by testing whether the contrastivity status of the referent for a nominal token is relevant to the selection of PROMINENT-marking, as inspired by similar text-based techniques (Chafe 1987; Myhill and Xing 1996; Vallejos Yopán 2009; Vallduví 2016, §23.5). More details regarding the precise coding protocols are outlined in Section B.1.3.

This chapter tests for a topicality-related TRIGGER for Siona DCM via the implementation of two corpus-driven approaches adapted from elsewhere in the literature. The first topicality metric used in this study measures 'look back' topicality, as conceived in the seminal works of Givón (1976, 1983, *et seq.*): i.e., a nominal token is more topical if it is mentioned more recently in discourse. A related vein of literature concerns referent activation (e.g., Chafe 1976, 1994; Ariel 2014 — recently summarized in Arnold 2010); on this view this chapter analyzes the recency of (re)activation as a factor for DCM. This chapter opts for the term LOOK-BACK TOPICALITY. Following recent analyses for topicality effects in *differential object marking* (i.e., Chiriacescu and von Heusinger 2010, Shain and Tonhauser 2010), the mean distance between mentions is established for each grammatical relation in the sample, corresponding to the number of intervening main clauses between mentions of a particular referent. A nominal token is topical if it is found below this distance threshold, whereas a nominal token is non-topical if the number of intervening clauses exceeds this threshold. This technique is applied to detect topicality-triggered *differential object marking* in Paraguayan Guaraní (Shain and Tonhauser 2010).

The second topicality metric concerns LOOK-AHEAD TOPICALITY, following the terminological conventions in Givón's work. Put simply, the prediction is that PROMI-

NENT-marking is favoured in instances where the referent of the nominal token persists as an at-issue aboutness topic in immediately ensuing discourse (i.e., the following main clauses in the narrative). This notion includes many instances of topic shift in the discourse; however, some tokens are mentions of referents which are already activated at that moment in discourse, rendering the term *shift* less appropriate. This is compatible with the notion of (re)activation of a given referent as a discourse topic in the activation literature mentioned above. To measure LOOK-AHEAD TOPICALITY, chapter implements a similar technique to that argued to demonstrate significant topicality effects for Romanian *differential object marking* by Chiriacescu and von Heusinger (2010).

In summary, this chapter makes an important contribution to the holistic analysis of Siona DCM undertaken in this dissertation. Conceived as a language which exhibits several multi-dimensional DCM systems, it is expected that several TRIGGERS are active and account for the complex distribution of PLAIN and PROMINENT marking alternatives to varying degrees. The variationist techniques employed in this chapter are appropriate for testing which of these plausible trigger candidates are relevant to Siona DCM, in particular, by refining the characterization of the REALIZATIONAL DOMAIN based upon a naturalistic dataset.

5.2.2 The variable-rule technique: An overview

This chapter employs the *variable rule analysis* technique (henceforth *varbrul*) in order to develop a statistical model to test the trigger candidates identified in Table 5.3, across the relevant grammatical relations in Siona grammar. The *varbrul*-analysis is generally accepted as the gold standard for variationist linguistics, following its inception in the pioneering works of Labov (1966, 1970, *et seq.*; see also Sankoff and Labov 1979; Sankoff 1988). On statistical grounds, this technique produces a linear regression model, based upon a maximum likelihood algorithm. Today several digital tools are available to create *varbrul* models, and the models developed for the Siona narrative sample were generated via the open-source *Rbrul* tool.[4]

This section briefly unpacks the inner-working of the *varbrul* model, how to interpret its results, and how this technique is implemented to ascertain statistical evidence for the present chapter. In fact, the present chapter develops three independent *varbrul* models, analyzing the proportion of PROMINENT-marking across S-oriented, P-oriented, and L-oriented DCM patterns respectively. Whereas this dissertation recognizes principled DCM patterns for the R- and the G-argument, the modest Siona narrative sample consulted here yields insufficient tokens for the development of an informative *varbrul* model for these grammatical relations.[5] Only initial proportional evidence is reported for these relations accordingly. The three full *varbul* models are reported in Appendix B, whereas condensed reports are provided in this chapter in order to focus on the informative results.

[4]Accessible via online manual at http://www.danielezrajohnson.com/rbrul.html. This tool operates within the statistical engine *R* (R Core Team 2021). Alteranative open-source tools exist: e.g., *GoldVarb*, [up-to-date version via http://individual.utoronto.ca/tagliamonte/goldvarb.html]; and *R-varb* (see Palillo 2002).

[5]In statistical parlance the *goodness of fit* metric determines how informative the significance and factor weight measures are for a given linear regression model (i.e., *varbrul*-analysis). Both the number of included tokens and the degree of variation determine the goodness of fit for resulting statistical model. Further details regarding these statistical matters are unpacked in Section B.1 of Appendix B.

On conceptual grounds, the *varbrul* technique is applied to a coded token-based dataset, where the concerned linguistic variables display structured variation — i.e., where the PLAIN marking alternative varies with the PROMINENT-marker. This variation is structured, since these forms are not in free variation. Each token is attributed a series of property values, based upon a series of pre-selected predictor variables, called factors, which correspond to the trigger candidates in Table 5.3. So long as factors are not implicationally related to one another, several factor groups may be tested simultaneously on a given dataset — e.g., if definite tokens are *always* specific, it is inappropriate to test for two independent factors. Each factor represents a hypothesis of the type "S-oriented DCM is driven by *animacy status*", where the *varbrul* technique executes hypothesis testing on a factor-by-factor basis, based upon the dataset at hand.

On statistical grounds, the output yielded by the *varbrul* model reveals three pieces of statistical evidence for each tested factor (Tagliamonte 2006, 2012). Firstly, it addresses the hypothesis directly as to whether a given factor is statistically significant in the dataset, generating a p-value accordingly. A significant result is yielded where the observed proportion of alternatives is determined not to have arisen by chance. By convention, a p-value of < 0.001 is considered to be highly significant, and a p-value between $0.001 - 0.05$ is significant. If the p-value is > 0.05, then it is insignificant. For the sake of presentation; the condensed *varbrul* reports in this chapter collapse all significant factors, and insignificant effects are acknowledged and set aside.

Secondly, where a factor is significant, a FACTOR WEIGHT value is calculated for each factor-internal category. For instance, concerning the word-order factor, a weight value between 0 and 1 is determined for the pre-verbal category and for the post-verbal category on the basis of the proportion of PROMINENT-marking, category frequency and variability metrics. FACTOR WEIGHT is calculated by establishing the difference between the highest and lowest values within the factor-internal categories, which is often referred to as the *Range* value (Tagliamonte 2012). The *Range* value reflects the effect size for a given factor, where consequential differences in proportion are recorded across the internal categories. For the sake of the discussion in this chapter, numerical *Range* values are interpreted as effect sizes as shown in Table 5.4 — full *Range* values are reported in the complete models in Appendix B:

EFFECT SIZE	FACTOR WEIGHT RANGE	*e.g., calculation from varbrul*
LARGE EFFECT SIZE	*Range*: > 50	CAT_1 (0.71) − CAT_2 (0.15) = 56
MODERATE EFFECT SIZE	*Range*: $20 - 50$	CAT_1 (0.56) − CAT_3 (0.23) = 33
SMALL EFFECT SIZE	*Range*: < 20	CAT_1 (0.48) − CAT_3 (0.39) = 9

Table 5.4: Conversion chart: From *varbrul Range* values to effect sizes

The third contribution of the *varbrul* model is related to the second. On the view that FACTOR WEIGHT RANGE, as shown in Table 5.4, corresponds to the notion of TRIGGER STRENGTH in multi-dimensional DCM (Klein and de Swart 2011), a ranking of significant factors may be established for each tested grammatical relation. In addition to converting *Range* values to EFFECT SIZE categories for the discussion in this chapter, the *varbrul* results also permit an internal ranking of, for example, which factors with an MODERATE EFFECT SIZE have a larger impact than the others. Significant factors are represented in their order of EFFECT SIZE in this chapter to

reflect their relative import on the distribution of PROMINENT-marking in the Siona narrative dataset.

In summary, this chapter provides both proportional evidence and different types of statistical evidence. The latter family of evidence ultimately demonstrates which tested trigger candidates are significant, and how big of an impact the significant CANDIDATES have on the observed DCM patterns. For the remainder of this chapter, the descriptive term, TRIGGER CANDIDATE is favoured over the technical term factor.

5.3 Results: DCM in the narrative dataset

This section reports the quantitative evidence yielded by the variationist analysis of the Siona narrative dataset. Firstly, the discussion opens with a consideration of the global distribution of case-marking across all nominal tokens in the corpus in Section 5.3.1 − i.e., considering all 1914 nominal tokens identified in the pooled sample. This initial analysis concerns both grammatical relations with DCM patterns and those without. The ensuing sections report and interpret the statistical evidence based upon the *varbrul* results for each tested DCM RELATION in the dataset in particular: In the first place, Section 5.3.2 considers the results for S-oriented DCM patterns. Next, Section 5.3.3 considers the P-oriented DCM patterns. Finally, the results for L-oriented DCM are reported in Section 5.3.4.

5.3.1 Distribution of marking alternatives in the sample

Table 5.5 categorizes all 1914 nouns tokens in the Siona narrative dataset based upon the intersection of grammatical relation and surface case-marking. The term 'marking alternatives' is used to include all overt case-marking *and* zero-marking for the purposes of the present discussion. This proportion of marking alternatives arising in the full annotated sample is referred to as the global case-marking distribution in this chapter. It is important to note that the raw case-marking figures reported here do not reflect proportional facts per se, given that case-markers are not attributed PLAIN or PROMINENT status in the global marking dataset. Nonetheless, the mapping of included tokens on the basis of PROMINENCE STATUS is crucial to the statistical evidence considered in the ensuing sections.

The patterns of case-marking alternatives laid out in Table 5.5 do not present any real challenges to the generalizations in the *preliminary description* in Chapter 3. Nearly all instances of obligatory case-marking, and principled and accidental DCM patterns, arise with an expected marking alternative according to their grammatical relation (and animacy class, where appropriate). Therefore, the primary contribution of this chapter pertains to the *variation* between the available marking alternatives in the narrative sample. A few initial observations are in order before moving to a discussion of the proportional and statistical results.

	-∅	*-bi*	*-re*	*-ni*	*-na*	*-hã'ã*	*-hã're*	**Total (%marked)**
Subject (s)	618 (75.27%)	203 (24.73%)	0	0	0	0	0	**821 (24.73%)**
Direct object (p)	209 (43.63%)	0	224 (46.76%)	46 (9.6%)	0	0	0	**479 (56.37%)**
Indirect object (r)	0	0	53 (75.71%)	17 (24.29%)	0	0	0	**70 (100%)**
Point-of-Contact	2 (7.14%)	0	3 (10.71%)	1 (3.57%)	21 (75%)	1 (3.57%)	0	**28 (92.86%)**
Place-of-Putting	1 (10%)	0	2 (20%)	0	6 (60%)	1 (10%)	0	**10 (90%)**
Instrument	0	**29**	0	0	0	0	0	**29 (100%)**
Companion	0	0	0	0	0	0	**2**	**2 (100%)**
Static location (l)	65 (40.12%)	0	53 (32.72%)	22 (13.58%)	22 (13.58%)	1 (0.62%)	0	**162 (59.88%)**
Goal (g)	44 (34.38%)	0	2 (1.56%)	0	74 (57.81%)	8 (6.25%)	0	**128 (65.63%)**
Source	0	**25**	0	0	0	0	0	**25 (100%)**
Path/Limit	0	0	0	0	0	**21**	0	**21 (100%)**
Time	130 (91.55%)	7 (4.93%)	0	0	3 (2.11%)	1 (0.7%)	0	**142 (7.75%)**
(*N*)	*-∅* (1070)	*-bi* (264)	*-re* (333)	*-ni* (86)	*-na* (126)	*-hã'ã* (33)	*-hã're* (2)	**Total (1914 tokens)**

Table 5.5: Raw distribution of nominal tokens in Siona narrative dataset, by case-marking and grammatical relation

Firstly, concerning argumental case-marking, R-tokens are never zero-marked in the sample, but rather receive *-re* or *-ni* marking, as anticipated based upon the description in Section 3.2.3. There are no instances of inanimate R-tokens in this dataset, such that *-re* and *-ni* correspond to PLAIN- and PROMINENT-alternatives accordingly. On this basis, a PROMINENT-marking rate of 24.29% is observed for R-oriented DCM, which aligns with other argumental DCM patterns in the sample.[6] As regards Point-of-Contact and Place-of-Putting arguments, described in Section 3.2.4 — setting aside a few outliers in the sample, these tokens are marked with *-na*.[7] All attested INSTRUMENT and COMPANION tokens bear their respective obligatory case-markers.

Turning to the spatiotemporal case-marking facts outlined in the global raw data in Table 5.5, a few observations are in order. Firstly, setting aside the full statistical analysis of L-marking for Section 5.3.4, an interesting pattern may already emerges based on the distributional figures alone. As pertains to G-oriented DCM, ignoring a few outlier tokens marked with the PATH-marker, *-hã'ã*, PROMINENT *-na* is selected at a rate of 62.71%.[8] This is markedly higher than the PROMINENT-marking rate observed for R-marking above. In fact, as will be established in this chapter, this discrepancy represents a broader, novel generalization: Argumental DCM displays PROMINENT-marking at a far lower rate than spatial DCM.

Another set of observations emerge from Table 5.5 as concern case-marking and temporal tokens. This nominal category does not arise in the *preliminary description* in Chapter 3, since Siona generally does not exhibit temporal case-marking. This generalization figured in the typological discussion in Chapter 4. The distributional facts considered in this chapter confirm this generalization. (3) displays two typical instances of zero-marked TIME tokens from the narrative sample:

(3) a. *i̲o neato ñami ti̲'a̲ni a̲isicobi a̲nichaoni i̲recato a̲coña*
ĩ-o **nea-to** **ñami** tĩ'ã-ni ãĩ-sih-ko-bi ãni
PRO3-CL:F be_dark-CL:LOC night arrive-SS eat-COMP-CL:F-SBJ eat
+ chao-ni ĩ-i̶-re=kato ãh-ko-ña
+ finish-SS 3PRO-CL:M-N.SBJ=TOP eat-2/3S.F.PST.N.ASS-REP
'She arrived later in the night, and she (lit. the one who had eaten), finished up (the other one) and then ate HIM.'
[NAT*: 20151001oolpi001.092-093]

b. *bacoña i̲o de̲baocato. ja̲macarebacato ë̲jë̲re baoña bacona . . .*
bah-ko-ña ĩ-o dẽbao=kato
be-2/3S.F.PST.N.ASS-REP 3PRO-CL:F vulture(spec)=TOP
hãmahka-reba=kato ĩhĩ̶-re baa-o-ña
then-INTENS=TOP husband-N.SBJ have-2/3S.F.PST.N.ASS-REP

[6] As indicated above, the sub-sample of R-tokens in the Siona narrative dataset does not meet the threshold for quantity (of internal categories) to merit the application of the *varbrul* technique, following the same coding protocols as the tested DCM RELATIONS analyzed in this chapter. Several examples of coded R-tokens are provided in Section B.2.4 of Appendix B as a matter of reference.

[7] Some typical and outlier Point-of-Contact tokens are illustrated in Section B.2.4.

[8] Note that these baseline figures do not take into account the fact that a small handful of promoted G-tokens (i.e., $N = 8$) are invariable, and ought to be excluded from any *varbrul* model. On this basis, the sub-sample of G-tokens is even smaller than anticipated. Some illustrative G-tokens are lifted in Section B.2.4 in Appendix B.

bah-ko-na
live-S.F.PST.DEP-DS
'There was the Dẽbao. At that time (she) lived on and had a husband...'
[NAT*: 20151112orapi001.005-007]

A handful of exceptionally case-marked TIME tokens are identified in the sample, which reflect certain sub-types of temporal nouns which have not been discussed elsewhere. These concern the few instances of temporal *-hã'ã* marking and *-bi* marking noted in Table 5.5.[9] These appear to reflect two fringe categories of nominals, which likely correspond to temporal extensions of spatial concepts — i.e., what might be labeled, the temporal Start-Point and End-Point relations, aligning with the spatial SOURCE and PATH relations accordingly. As such, these would correspond to instances of obligatory, semantic case-marking, although further work is required to properly analyze the encoding and interpretation of these temporal nominals. An illustrative instance of each is lifted from *The mother of jaguars* narrative in (4):

(4) a. *...yureta'a jamacabi te'e cocabi cañu'u caquë yequë ëjacoca, oracoca, yecua'i goasi'ahua'i*
jure-tã'ã **hãmahka-bi** te'e kohka-bi kaa-ñu'ũ
now-CNT.EXP then-SRC one language-INST speak-HORT
kaa-kɨ jehk-ɨ ɨha + kohka ora + kohka
say-S.M.PRS.DEP other-CL:M foreign + language kichwa + language
jehk-wa'i goasi'a-wa'i
other-PL.AN any- PL.AN
'...and now, from then on, he wanted them to speak but one language; but one (was speaking) Spanish (lit. foreigner language), and others yet the Kichwa language.' [NAT*: 20151001oolpi001.353-354]

b. *...yureta'a goade'ohuesëyë. jamacaja'a baja'i Baina caquë ba'isico*
jure-tã'ã goade'o + wehsɨ-jɨ **hãmahka-hã'ã**
now-CNT.EXP worsen + not_know-N3S.PRS.ASS then-LIM
ba-ha'i Baina kaa-kɨ ba'i-sih-ko
be-3S.M.PST.ASS Baina say-S.M.PRS.DEP be-COMP-CL:F
'...and now (we) don't know (our own language). Up to here, this was what Baina is said to have done.' [NAT*: 20151001oolpi001.396-398]

Besides these few observations, case-marking in the narrative sample aligns with the generalizations laid out in Chapter 3 in this dissertation. The remainder of this discussion undertakes a targeted discussion of the DCM patterns analyzed for S-, P-, and L-arguments in the sample, focusing on both proportional and statistical evidence yielded in this analysis. Each of the ensuing sub-section outlines the grammatical relation-particular sub-sample utilized to develop the statistical models, and then navigates the results of each model.

[9]Regarding the few instances of TIME tokens marked with *-na*, there is a handful of temporal expression which occur with this marker, although its semantic contribution is more opaque. For example, items like *jure-**na*** (now) or *neato-**na*** (tomorrow at dawn) have a more emphatic reading, although this is less likely in the common Siona greeting, *ñami-**na** ñañu'u* (let's see (each other) tomorrow). An instance from the sample is shown in (52) in Appendix B.

5.3.2 S-marking in the Siona narrative dataset

This section is concerned with analyzing patterns of S-oriented DCM which arise in the Siona narrative dataset. As demonstrated in Table 5.5, this is the best-represented grammatical relation in the dataset — i.e., accounting for 821 of the total 1914 nominal tokens identified in the pooled sample. Roughly 35% of main clauses in the dataset arise with an overt S-argument, which is comparable to the rate noted in work with texts in the Eastern Tukanoan language, Wanano [*gvc*] (Stenzel 2015, 222).[10] Nonetheless, a first order of business is to revise the sub-sample particular to S-arguments for the sake of statistical inquiry, such that only those tokens which exhibit the PLAIN zero-marking vs. PROMINENT *-bi* marking alternation are included. This procedure of *cleaning* the data, in order to eliminate *invariable* tokens, is a precursor to performing variationist analyses on any token-based sample (Tagliamonte 2006, 2012). For this chapter, the sub-samples particular to each tested grammatical relation are cleaned prior to implementing the *varbrul* technique to produce statistical evidence.

This section is organized as follows: Section 5.3.2.1 outlines two classes of S-tokens which are excluded from the statistical analysis, and which serves to refine the description of S-oriented DCM accordingly. Section 5.3.2.2, reports and discusses the proportional and statistical evidence results of the *varbrul* analysis.

5.3.2.1 Excluding invariable S-tokens

This chapter recognizes two classes of excluded S-tokens, which do not display the principled PLAIN-PROMINENT alternation, defining S-oriented DCM per **Section 3.1** of the *preliminary description* in this dissertation. The first class pertains to S-arguments participating in a particular dependent-clause construction, which we label SLDC (i.e., *subject-like dependent construction*) in other work (Bruil and Case 2023).[11] The second class of excluded S-token concerns those which are marked with the contrastive topic marker, *=kato*, whose impact on case-marking is discussed in greater detail in the next chapter. Each of these classes is discussed in turn in what follows.

Turning firstly to the SLDC construction, which we identify in a recent talk (Bruil and Case 2023); the crucial fact for the present purposes pertains to the fact that the implicated S-argument displays obligatory *-bi* marking. As such, this construction suppresses the principled DCM alternation characteristic of the S-argument. The SLDC construction refers to a clause-chaining strategy, which is commonplace in Siona narratives, where a de-verbal nominalization serves as the S-argument of a main clause. The referent is interpreted as the S-argument of the de-verbalized event and of the main clause event — i.e., a sort of same-subject interpretation arises. A canonical example is shown in (5), as lifted from *The mother of jaguars* narrative, where the S-argument in (5b) is excluded for the current purposes since zero-marking simply never arises in this construction. This chapter eliminates 48 such tokens from the S-marking sub-sample for statistical analysis:

[10] Consult Table B.9, which provides a more detailed breakdown regarding the distribution of overt and covert S-arguments found in the sample.

[11] In Bruil and Case 2023, we note an instance of PROMINENT *ni*-marking on an SLDC instance with an EXPERIENCER SUBJECT. So far only one instance is attested and requires further investigation.

(5) a. *jo̲maca i̲jo̲ bainajo̲ nëcaquë ñaëña . . .*[12]
hãmahka [ĩ-i̵-hẽ Baina-hẽ$_i$] nɨhka-kɨ̵
then 3PRO-CL:M-ADD B-ADD stand-S.M.DEP.PRS
ñaa-i̵-ña
look-2/3S.M.PST.N.ASS-REP
'From there he, Baina$_i$, was also standing there, looking out . . . '

b. *. . . nëcaquëbi jaiñë jujujë̲'ë̲ ame yëiñë caëña*
nɨhka-kɨ̵-bi$_i$ hai-ñɨ huhu-hĩ'ĩ̵ a-me jɨi-ñɨ
stand-CL:M-SBJ big-CL:TREE grow-IMPER COP-? cotton-CL:TREE
kaa-i̵-ña
say-2/3S.M.PST.N.ASS-REP
'. . . (he)$_i$ was standing there (lit. the standing one (M)) and said "grow, grow" and the giant kapok tree grew (they say).'
[NAT*: 20151001oolpi001.092-093]

The second category of excluded S-tokens concerns those arising with the contrastive topic clitic *=kato* (glossed TOP). There turns out to be a general co-occurrence restriction between this nominal discourse marker and the PROMINENT case-marker *-bi*, which is supported by elicitation-based evidence in Section 6.4.2 of the following chapter.[13] (6) illustrates a contiguous stretch of discourse with two S-tokens from *The demon kwẽ̃ẽwahti* narrative. The narrator marks the S-argument in (6b) with the contrastive topic marker *=kato*, which never co-occurs with *-bi* marking on S-nominals. This chapter eliminates 46 such S-tokens from the statistical analysis:

(6) a. *nocabo jë̲o̲saia baquë i̲ i̲o̲ ë̲jë̲maca . . .*
nohka-bo hɨ̃õ + sai-a bah-kɨ̵ [**ĩ-ɨ̵**
banana-CL:ENCLOS clean + go-NEG be-2/3S.M.PST.N.ASS 3PRO-CL:M
ĩ-o **ɨ̃hĩ-mahka**$_i$]
3PRO-CL:F husband-DIM.SG
'He, her husband$_i$ went to tend to the banana plantation . . . '

b. *. . . i̲o̲macacato a̲o̲re sëoco ñucoña*
ĩ-o-mahka=kato$_k$ ãõ-re sɨo-ko
3PRO-CL:F-DIM.SG=TOP cassava-N.SBJ fry-S.F.PRS.DEP
ñuh-ko-ña
sit-2/3S.F.PST.N.ASS-REP
'. . . as for her, she$_k$ was sitting, frying casSAva.'
[NAT*: 20151112oespa001.164-165]

On the basis of the two excluded categories of S-token established above, this chapter develops proportional and statistical evidence for S-oriented DCM against a sub-sample which consists of 727 included, *variable* S-tokens.

[12]Baina is the primary benevolent cultural hero in Siona (and Sekoya) mythology and he is ubiquitous in the folktale genre, as discussed in various places in the Vickers (1976) ethnography.

[13]See Table B.10 in the Appendix for a full breakdown regarding the distribution of *=kato* and other discourse particles in the narrative sample.

5.3.2.2 Exploring the *varbrul* results for S-oriented DCM

Table 5.6 provides the condensed *varbrul* report based upon the S-marking sub-sample ($N = 727$), which corresponds to the full *varbrul* report in Table B.6, found in Section B.2.1 of Appendix B. In this sub-sample PROMINENT *-bi* arises on 21.18% of these tokens — i.e., which will be referred to as the 'global PROMINENT-marking rate' in this chapter. This global rate acts as proportional evidence for the present discussion, which is established prior to categorizing the dataset of the basis of TRIGGER CANDIDATE categories. Statistical evidence is gleaned on the basis of TRIGGER CANDIDATE-internal categories, and their respective PROMINENT-marking rates, as displayed in Table 5.6, for significant factors:

TRIGGER CANDIDATE	*N* [% PROMINENT]
GLOBAL S-MARKING RATE	727 [21.18%]
CONTRASTIVITY STATUS	(LARGE EFFECT SIZE)
Contrastive S	83 [95.18%]
Non-contrastive S	644 [11.6%]
SPECIFICITY STATUS	(MODERATE EFFECT SIZE)
Pronominal S	346 [17.6%]
Specific S	273 [26.4%]
Non-specific S	108 [19.4%]
WORD ORDER	(MODERATE EFFECT SIZE)
Pre-verbal S	554 [24%]
Post-verbal S	173 [12.1%]

Table 5.6: Significant factors for S-oriented DCM in *varbrul* sub-sample

As a point of departure, Table 5.6 demonstrates that three tested trigger candidates are highly significant in the S-marking sub-sample (i.e., $p =< 0.001$). By extension, this chapter provides statistical support that the non-significant trigger candidates do *not* factor into Siona S-oriented DCM patterns (i.e., $p => 0.05$). Therefore, these quantitative findings generally complement observation made in earlier descriptions (Chapter 3 in this dissertation; Bruil 2014, §4.4): Neither AGENTIVITY STATUS nor ANIMACY STATUS are active S-marking TRIGGERS in the language.[14] This chapter also finds that neither tested topicality metric is relevant to these patterns — i.e., deviating from topic-driven S-marking patterns with the cognate *-pi* described for the related language, Koreguaje (Cook and Levinsohn 1985, §1).

Turning to the active TRIGGERS in the S-marking sub-sample, the only tested TRIGGER CANDIDATE with a large statistical effect concerns CONTRASTIVITY STATUS. Table 5.6 indicates the nearly all contrastive S-tokens are marked with *-bi*; whereas, of

[14] A handful of examples of both PLAIN zero-marking and PROMINENT *-bi* marking across these non-significant categories are provided in Section B.2.1 in Appendix B.

the 644 non-contrastive S-tokens, only 11.6% of tokens receive PROMINENT *-bi* marking. A clear instance of contrastive S-marking is demonstrated in (7), from *The mother of the moon* narrative, where two established personified bird characters go out to check their traps for food. In this instance, the narrator describes the activities undertaken by the *pɨ'pɨri* (vulture) character in (7a), which happens to be marked with *=kato*, and contrasts these activities directly with those undertaken by the *tãkaro-**bi*** (black vulture, species) character in (7b):

(7) a. *pë'përi ba'iji iñore mëabecato ba'i ñabi. jo'ni ja̲maca i̲o̲maca ba'iona ...*
pɨ'pɨri$_i$ ba'i-hi [ihño-re mɨabe=kato] ba'i-ɨ
vulture be-3S.M.PRS.ASS here-N.SBJ above=TOP be-S.M.PRS.DEP
ñaa-bi yo'-ni hãmahka **ĩ-o-mahka**$_i$ ba'i-o-na
see-3S.M.PST.ASS do-SS then 3PRO-CL:F-DIM.SG be-S.F.PST.DEP-DS
'The VULture$_i$ was up HERE, looking out. Doing this, as she$_i$ was (there)...'

b. *...ja̲re ta̲carobi ña baco. ñani nëcadojaiyo ñani*
hãre **tãhkaro-bi**$_k$ ñaa-a bah-ko ñaa-ni
like_that black_vulture-SBJ see-NEG be-2/3S.F.PST.N.ASS look-SS
nɨhka + dohai-o ñaa-ni
be_standing + wander-S.F.PST.DEP look-SS
'...likewise the BLACK vulture$_k$ was looking out. (She$_k$) wandered on foot (i.e., on the ground) to go on the look out ...'

[NAT*: 20151112orapi001.174-176]

This chapter reports statistical evidence for a very strong correlation between CONTRASTIVITY STATUS and the selection of the PROMINENT-marker, *-bi*; however, this does not mean that PROMINENT-marking is blocked on non-contrastive tokens. In fact, a non-trivial set of naturalistic, non-contrastive S-tokens arise with PROMINENT *-bi* marking. It is also relevant to note that the *varbrul* model, reported in Table 5.6, yields no statistically significant intersective effects; which is taken to indicate that the other noted significant effects shed light on the factors which drive the selection of non-contrastive *-bi* marking. More concretely, the analysis demonstrates that two tested trigger candidates display moderate effect sizes — i.e., SPECIFICITY STATUS and WORD ORDER. Each is discussed in turn in what follows.

Beginning with SPECIFICITY STATUS, this analysis finds a meaningful discrepancy regarding the proportion of PROMINENT *-bi* across the three specificity categories, as conceived here: pronominal S, specific full noun S, and non-specific full noun S. In particular, specific full noun S-tokens bear PROMINENT-marking at an above-average rate, compared with the other categories.[15] This subtle effect is unanticipated — given what we know based upon the *preliminary description* in Chapter 3, specificity is not expected to factor into S-oriented DCM patterns.

For instance, unlike non-SUBJECT DCM, pronominal S is not obligatorily case-marked. (8) juxtaposes two post-verbal personal pronoun S-tokens from the sample.

[15]By design, this statistical study conflates specific indefinites with definites — for the reasons laid out in Section B.1.2 of Appendix B. Impressionistically, the noted discrepancy across these sub-categories is subtle. It is not clear whether the corpus-based variationist approach adopted here is appropriate to study definiteness and its, possibly subtle, effects on case-marking. A more targeted semantically-driven study is likely in order.

In (8a) from *Siona Genesis story* the pronoun is zero-marked; whereas the corresponding S-pronoun from *The demon kwẽẽwahti* narrative in (8b) bears PROMINENT *-bi* marking. In its context, the latter S-token is coded as non-contrastive:

(8) a. *. . . cajë saijëna be'teiña i̱*
kaa-hɨ̱ sai-hɨ̱-na be'te-ɨ̱-ña
say-PL.PRS.DEP go-PL.PRS.DEP-DS follow-2/3S.M.PST.N.ASS-REP
ĩ-ɨ̱
3PRO-CL:M
'. . . (they) spoke and left, and he followed.'
[NAT*: 20151023orocr001.108]

b. *i̱o ëjë cañeje̱ caoña i̱obi*
[ĩ-o ɨ̃hɨ̃ kaa-je-hẽ] kaa-o-ña
3PRO-CL:F husband say-CL:GEN-SIM say-2/3S.F.PST.N.ASS-REP
ĩ-o-bi
3PRO-CL:F-SBJ
'She (i.e., the *kwẽẽwahti* demon) spoke like her husband speaks.'
[NAT*: 20151112oespa001.170]

Besides pronouns, I am unaware of any claims regarding *-bi* as a specificity marker for S-nominals, either in earlier descriptions of Ecuadorian Siona (Bruil 2014, §4.4.1), or for related languages with cognate S-markers, as discussed in Chapter 4.[16] Rather, as discussed at length in the preceding chapter, specificity is claimed to be an active TRIGGER in P-oriented DCM in many Western and Eastern Tukanoan languages. As will be shown, the quantitative analysis in this chapter arrives at the surprising finding that PROMINENT-marking is favoured on specific full noun tokens for both S- and P-oriented DCM patterns, with a compatible MODERATE-SIZE EFFECT across both tested argumental grammatical relations.

The weakest significant effect for S-oriented DCM reported in Table 5.6 concerns WORD ORDER. Although Siona displays a dominant SOV word order, as laid out in Chapter 3, the order of arguments is relatively flexible. This chapter finds statistical evidence that the proportion of PROMINENT-marking on the S-argument is affected by its surface position, relative to the inflected verbal word in the clause. Most S-tokens in the sample arise in a pre-verbal position — i.e., 554 tokens, following the dominant pattern and arising with *-bi* marking at roughly the global marking rate.[17] However, PROMINENT *-bi* marking only arises on 12.1% of the 173 remaining S-tokens in post-verbal position.

To my knowledge, the correlation between word order and S-oriented DCM has not been established in previous research, for Ecuadorian Siona, or for other Western Tukanoan languages. It is not immediately clear how to interpret the correlation between word order and DCM. Nonetheless it is certainly possible that certain factors which are relevant to determining PROMINENT-marking also factor into variable sentential word order patterns.[18] Some hypotheses for this correlation are laid out in

[16] A handful of examples of S-tokens falling within each specificity category is provided in Section B.2.1 of Appendix B.

[17] Table B.9 in Appendix B demonstrates that 76.2% of overt S-tokens in the sample, in main clauses, arise in pre-verbal position.

[18] It may also be relevant to note that most arguments are simply dropped in naturalistic speech. An adequate variationist account of Siona word order effects likely accounts for argument-dropping in some way.

the final section of this chapter, which merit refinement via other techniques. Interestingly, however, this chapter detects a similar WORD ORDER effect, as a TRIGGER CANDIDATE for P-oriented DCM, which points to another commonality across argumental DCM patterns which has previously gone unnoticed. The following section explores the proportional and statistical evidence for P-oriented DCM in the dataset.

5.3.3 P-marking in the Siona narrative dataset

This section is concerned with analyzing patterns of P-oriented DCM which arise in the Siona narrative dataset. Table 5.5 identifies 479 P-tokens in the sample. However, as is established in detail in Section 3.2 of the *preliminary description*, the determination of the proportion of PROMINENT-marking for this sub-sample may not be ascertained by a straightforward mapping of particular surface case-marker onto PLAIN and PROMINENT categories. Following the terminological conventions laid out in Chapter 2, Siona P-oriented DCM exhibits the formal HYBRID-type, where three formal alternatives are mapped onto PLAIN and PROMINENT categories, as determined by an animacy-based noun class split — i.e., zero-marking, *-re*, and *-ni*. On this basis, Figure 5.1 schematizes the distribution of surface case-marking alternatives across animate and inanimate P-tokens in the full Siona narrative dataset:

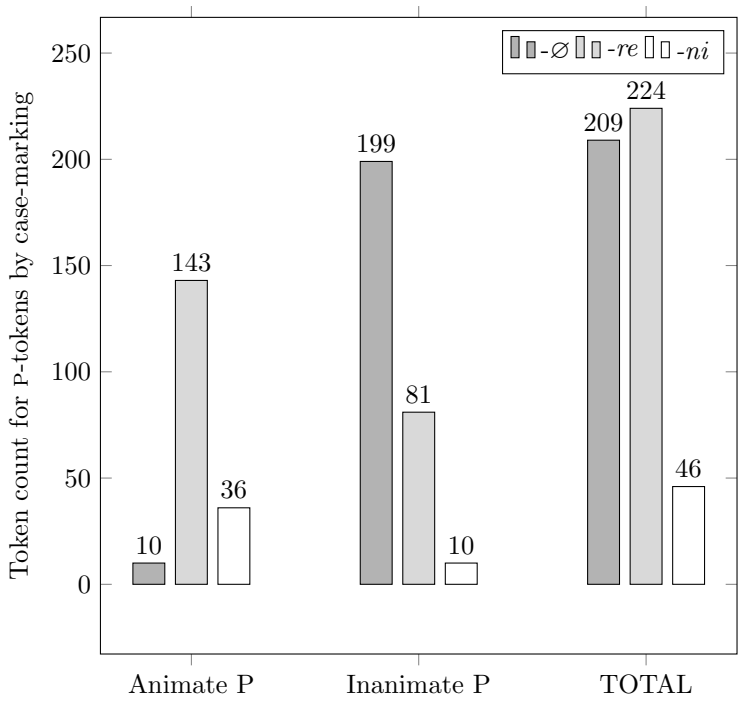


Figure 5.1: P case-marking alternatives across notional animacy categories

Firstly, the small subset of 10 inanimate P-tokens with *-ni* marking are anticipated, given the phenomenon of promotion detailed in Section 3.2.2. Predictably, promoted

inanimate nominals pattern with the animate noun class for the purposes of case-marking. (9) lays out the post hoc conversion process, from surface case-marking to PROMINENT-marking status, which acts as the dependent variable for the analysis developed here — as such, promoted P-tokens are not excluded from the sub-sample:

(9) a. Animate P ⇒ PLAIN *-re* vs. PROMINENT *-ni*
b. ↑Inanimate P ⇒ PLAIN *-re* vs. PROMINENT *-ni*
c. Inanimate P ⇒ PLAIN *-∅* vs. PROMINENT *-re*

Promoted P-tokens display the same principled, *variable* DCM pattern analyzed for canonical P-tokens. (10) provides two instances of promoted pronominal P-tokens: An instance of PROMINENT *-ni* marking from *The children of Baina* narrative is shown in (10a); whereas (10b) from the *Siona Genesis story* demonstrates PLAIN *-re* marking:

(10) a. *... yo'quëna io bë'cacoje caoña iconi guijë'ë caoña. neo garabëre*
jo'-ki-na ĩ-o bɨ'ka-ko-hẽ
do-S.M.PRS.DEP-DS 3PRO-CL:F parent-CL:F-ADD
kaa-o-ña ↑**ih-ko-ni** gũĩ-hɨ̃'ɨ̃
say-2/3S.F.PST.N.ASS-REP DEM.PROX-CL:F-N.SBJ2 bite-IMPER
kaa-o-ña nee-o gara-bɨ-re
say-2/3S.F.PST.N.ASS-REP make-3S.F.PST.ASS charcoal-CL:RND-N.SBJ
'When (he) did this, her mother said, "bite into THIS!", she had made charcoal.' [NAT*: 20150811sfryi001.092-093]

b. [**Context:** All of the animals in the forest gather to fell a great tree.]
icore cueni jeũñu'u cueni jeona jaobi taitota'a tsiaya ba'ija'coa
↑**ih-ko-re** kwẽẽ-ni hẽõ-ñu'ũ kwẽẽ-ni
DEM.PROX-CL:F-N.SBJ take_down-SS leave_behind-HORT take_down-SS
hẽõ-ɨ-na ↑hã-o-bi tãi-to-tã'ã
leave_behind-S.F.PST.DEP-DS DEM.MED-CL:F-SBJ fall-COND-CNT.EXP
tsia-ja ba'i-hã'-ko-a
river-CL:RIV be-PRP-CL:F-COP.3S.F
' "Let's take it (i.e., the tree) down, when (it) comes down, it (lit. she) will be the river." ' [NAT*: 20151023orocr001.251-253]

To confirm, as shown in Figure 5.1, promoted tokens are coded as belonging to the notional inanimate class for the sake of testing ANIMACY STATUS as a TRIGGER CANDIDATE. Both concerned tokens in (10) have inanimate referents, independently from the coding of the PROMINENT-marking status for a particular token.

Whereas this chapter includes all instances of promoted P-tokens (*N=16*); there are two small classes of zero-marked animate P-tokens, identified in Figure 5.1, which are eliminated for the sake of statistical analysis in the present chapter. Section 5.3.3.1 briefly discusses these sets of tokens. Section 5.3.3.2 explores the proportional and statistical evidence produced against the P-marking sub-sample.

5.3.3.1 Excluding invariable and outlier P-tokens

Figure 5.1 identifies ten zero-marked animate P, whose case-marking deviate from the dominant, principled DCM patterns, and which are eliminated from this analysis.

This set of tokens is divided evenly into two sub-categories, each of which informs the broader characterization of P-oriented DCM.

Firstly, as alluded to in Chapter 4, the class of coordinated P-tokens is *invariable* since they simply resist case-marking altogether and are excluded on this basis. A few such instances arise in the sample (N=3), where nominals are coordinated via juxtaposition.[19] Consider (11) lifted from the *Siona Genesis* narration where the corresponding non-conjoined animate P-tokens would require case-marking if they had surfaced independently:

(11) *yo'oquë sai uye naso huasaiya baquë . . .*
jo'o-kɨ̵ sai-ɨ̵ **uje** **nahso** wa'i + sai-a
do-S.M.PRS.DEP go-S.M.PRS.DEP turkey wolly_monkey kill + go-NEG
bah-kɨ̵
be-N2/3S.PST.N.ASS
'. . . doing (this), (he) went out to hunt turkey, monkey(, etc.).'
[NAT*: 20151023orocr001.324]

The second set of excluded P-tokens (N=7) concern the small class of lexical generics (e.g., *hoja* (dogs, domesticated animals), *tsĩ̃ĩ* (children)), whose case-marking facts require further investigation. These nouns do not *require* case-marking, like regular animate nouns in non-SUBJECT positions; albeit they readily accept both *-re* and *-ni* elsewhere. Therefore, following best practices in variationist research, these few unclear instances are eliminated as outliers for the present purposes.[20,21] (12) juxtaposes P-tokens of the zero-marked lexical generic *tsĩ̃ĩ* with the canonical animate *mamakɨ-re* (child (M)) in a single episode of *The demon kwẽ̃ẽwahti* narrative:

(12) a. *cani jɑ̃maca nasobi mëco'ë tsɨ̃ gachaco ɨ̃oje. tsɨ̃ hueo yo'ye ba'io. . .*
kaa-ni hãmahka nahso-bi mɨh-ko'ɨ̵ **tsĩ̃ĩ̃**
say-SS then wolly_monkey-SBJ climb-3S.F.PST.ASS child
gahcha-ko ĩ̃-o-hẽ̃ **tsĩ̃ĩ̃** wee-o
put_down-3S.F.PRS.ASS 3PRO-CL:F-ADD child carry-S.F.PRS.DEP
jo'-je ba'i-o
do-CL:GEN be-2/3S.F.PRS.N.ASS
'After (she) said "then the monkey went up (there)", she put down the child$_i$ – she was also carrying a child$_i$.'

[19]The same blocking effect arises where speakers realize coordination with the overt conjunction *kwĩ'ne* (also), which is commonplace in elicited data. Simply put, the presence of coordination bleeds the obligatory case-marking requirement found on animate non-S arguments. This generalization is corroborated by elicited data in my field notes.

[20]In other corpus-driven work, we discuss lexical generics in slightly greater detail (Bruil & Case, *forthcoming*). On the basis of preliminary fieldwork, this pattern is only observed for a small class of nouns, like those mentioned above; whereas generic animates otherwise pattern with animate nouns in general for case-marking purposes (e.g., *wa'i* (fish); *bĩ'ã* (birds)).

[21]One hypothesis: the textual instances which do arise with these generics appear to be in immediately pre-verbal position, and may belong to a separate construction of pseudo-noun incorporation of the DIRECT OBJECT — i.e., not DCM, as discussed by Driemel 2023 for a variety of languages.

b. *. . . jetere hueco <u>io</u> mamaquëre hueco*
hehte-re wee-ko [**ĩ-o** **mama-ki-re**]
back-N.SBJ carry-3S.F.PRS.DEP PRO3-CL:F child-CL:M-N.SBJ
wee-ko
carry-3S.F.PRS.ASS
'(She) is carrying <u>her child</u>$_i$ on (her) back.'
[NAT*: 20151112oespa001.102-103]

The handful of eliminated tokens discussed here hint at certain questions which remain to be addressed in future research, regarding classes of P-arguments which deviate from the principled DCM outlined in Section 3.2.1. The resulting sub-sample for P-marking consists of 469 P-tokens, classified as bearing their appropriate PLAIN- or PROMINENT-alternative accordingly. The following section considers the proportional and statistical evidence which is yielded against this P-marking sub-sample.

5.3.3.2 Exploring the *varbrul* results for P-oriented DCM

In addition to the relative complexity of P-oriented DCM on formal grounds (HYBRID-type DCM), Chapter 3 recognizes an array of plausible TRIGGERS for the observed DCM patterns, which are not necessarily relevant to S-oriented DCM. For instance, previous descriptions indicate that both specificity- and discourse-related factors conspire to influence case-marking of the P-argument (i.e., Bruil 2014, §4.4.2). This chapter provides statistical evidence which confirms that P-oriented DCM is more complex in this regard. The condensed *varbrul* model in Table 5.7, corresponding to the full report in Table B.7 in Appendix B, demonstrates that more trigger candidates are active in P-oriented DCM, as compared to the S-marking patterns analyzed in the previous section.

Table 5.7 strictly considers significant factors, corresponding to active trigger candidates, in the sample; however, it is also relevant to note that this *varbrul* model provides evidence *against* two tested trigger candidates — i.e., AFFECTEDNESS STATUS and LOOK-BACK TOPICALITY, as coded in this chapter yield insignificant effects (i.e., $p => 0.05$).[22] Therefore, on the one hand, it is not the case that the degree of affectedness of the P-argument affects DCM. On the other hand, regarding topicality, this analysis finds that LOOK-BACK TOPICALITY STATUS — i.e., how recently the referent of a given token is mentioned in the narrative ('look back' topicality per Givón 1976, *et seq.*) is *not* relevant to P-oriented DCM. Conversely, the LOOK-AHEAD TOPICALITY metric is siginificant in the sample, as discussed further below:

[22]The reader is referred to Section B.2.2 in Appendix B for a handful of examples of pertaining to these inacive factors.

Trigger candidate	*N* [% Prominent]
Global p-marking rate	469 [24.27%]
Contrastivity status	(Large effect size)
Contrastive p	33 [87.9%]
Non-contrastive p	436 [19.6%]
Look-ahead topicality status	(Moderate effect size)
Local topical p	45 [51.1%]
Local non-topical p	424 [21.5%]
Specificity status	(Moderate effect size)
Pronominal p	106 [20.2%]
Specific p	162 [32.7%]
Non-specific p	201 [19.6%]
Word order	(Moderate effect size)
Pre-verbal p	383 [26%]
Post-verbal p	86 [19.7%]
Animacy status	(Small effect size)
Animate p	179 [19.7%]
Inanimate p	290 [27.2%]

Table 5.7: Significant factors for p-oriented dcm in *varbrul* sub-sample

Firstly, as pertains to the proportional facts, Table 5.7 recognizes a global prominent-marking of 24.27% on the p-marking sub-sample in the Siona narrative dataset. This global marking rate is comparable to the 21.18% marking rate observed in the s-marking sub-sample. This analysis finds that five of the seven tested factors have significant effects on the distribution of prominent-marking, which includes all three which are active in the s-marking sub-sample, in addition to two other effects: including Look-ahead topicality status and Animacy status. The relative impact of each of these Triggers in the sample is outlined in what follows.

Firstly, as was found with s-oriented dcm patterns in the dataset, Contrastivity status stands alone as the only independent variable with a large effect size (i.e., a highly significant result, $p =< 0.001$). The small class of contrastive p-tokens are prominent-marked at an especially high rate, as compared with non-contrastive p-tokens.[23] The sentences in (13) demonstrate clear prominent-marking on contrastive

[23]Interestingly, non-contrastive p is marked at roughly the global marking rate, unlike non-contrastive s. There may be several reasons for this fact; however, there are clearly more factors, which impact p-marking generally. The Factor weight (i.e., *varbrul Range*) is higher in the s-marking sub-sample, compared to p-marking, albeit both clearly have large effect sizes in their respective analyses. It is not clear at this stage how much weight, if any, to place on this statistical discrepancy.

P-tokens: (13a) illustrates an instance of *-ni* marking from *The mother of the moon* narrative, where the human referent is contrasted with the other members of the husband's family. (13b) provides an example, from *The children of Baina* narrative, where *-re* marking contrasts one type of meat with other types of food:[24]

(13) a. [**Context:** A recently deceased husband transforms into a wasp to bring his family to the Upper World to join him in the afterlife.]
. . . ja ñacona. ja ioni jëasi'i cani huajëconi jëasi'i cani yo'quëna . . .
ñaa-ko-na hã **ĩ-o-ni** hɨa-si'-i
see-S.F.PRS.DEP-DS DISC.PART 3PRO-CL:F-N.SBJ2 bring_across-FUT-N3S
kaa-ni **wahɨ-ko-ni** hɨa-si'-i kaa-ni yo'-kɨ-na
say-SS alive-CL:F-N.SBJ2 bring_across-FUT-N3S say-SS do-S.M.PRS.DEP-DS
'. . . when (she) saw (this), oh, (he) wanted to bring HER, the living woman, (to the other side). (He) did (this) and then. . . '
[NAT*: 20151112orapi001.081-082]

b. [**Context:** In order to ascend and rejoin their father in the sky, the children of Baina must drink yage.]
. . . hueo yë'cato soquëtene hua'ire aicoa'ë caco jaëhua'ini. iocato goeye baco
wee-o jɨ'ɨ=kato **sõhkɨtẽne + wa'i-re**
lie_in_hammock-S.F.PST.DEP 1S=TOP catfish(spec) + meat-N.SBJ
ãĩ-ko-a'-ɨ kaa-ko hã-ɨ-wa'i-ni
eat-PRP-CL:F-COP.3S.F say-S.F.PRS.DEP DEM.MED-CL:M-PL.AN-N.SBJ2
ĩ-o=kato gwee-je bãã-ko
3PRO-CL:F=TOP not_want-CL:GEN NEG.AUX-S.F.PRS.DEP
'. . . (she) lay (there, in the hammock) and said to them, "As for ME, I eat (only) CATfish meat". She didn't want (i.e., yage).'
[NAT*: 20151001oolpi001.065]

The second statistical effect noted in the P-marking sub-sample concerns another discourse-level TRIGGER CANDIDATE, LOOK-AHEAD TOPICALITY STATUS (i.e., $p =< 0.001$). Unlike the inactive LOOK-BACK TOPICALITY STATUS, which distinguishes tokens with recently mentioned referents, from those with a more distantly mentioned referent; there is statistical evidence for an active LOOK-AHEAD TOPICALITY STATUS effect, demonstrating that P-tokens whose referents persist in immediately ensuing lines of the narrative favour PROMINENT-marking. This medium-sized effect is markedly distinct from that noted for CONTRASTIVITY STATUS above, although, taken together, these demonstrate that the selection of PROMINENT-marking is largely driven by discourse-related TRIGGERS.

Below are two canonical instances of PROMINENT-marked P-tokens, which are non-contrastive local topics, per the definitions assumed in this chapter. On the one hand, (14) demonstrates an instance from *The children of Baina* narrative, where PROMINENT *-ni* is found on the P-argument, which persists as at-issue argument in the following sentence. A similar instance is lifted from *The mother of the moon* narrative in (15): PROMINENT *-re* is used where the *pɨ'pɨri wa'tijo* (the vulture's blade) is retained as an argument in the following sentence, important to the narrative development:

[24]Various instances of P-tokens sorted on the basis of all tested discourse-level trigger candidates are provided in Section B.2.2 of Appendix B.

(14) a. *ja bë'ë tsiayacua'ini tsi'choa baquë . . .*
hã [**bɨ'ɨ tsia-ja-ah-ko-wa'i-ni**] si'cho-a
DISC.PART caiman river-CL:RIV-COP-CL:F-PL.AN-N.SBJ2 gather.TRS-NEG
bah-kɨ
be-2/3S.M.PST.N.ASS
'He gathered the caimans that lived in the (the) river. . . .'

b. *. . . sanijeñaquë caquë saquëña jaona teore cuni tayobi. . .*
sani + hẽẽ + ñaa-kɨ kaa-kɨ
go + cross + see-S.M.PRS.DEP say-S.M.PRS.DEP
sah-kɨ-ña hã-o-na te'-o-re kũni
go-2/3S.M.PST.N.ASS-REP DEM.MED-CL:F-GOAL one-CL:F-N.SBJ bite
+ tajo-bi
+ tear-3S.M.PST.ASS
'. . . (he) said (he) would go and look, so (he) went there and killed bit one (of them) and tore (them) up.' [NAT*: 20150811sfryi001.427-438]

(15) a. *iconi isisi'i pë'përi hua'tiyore caëña. . .*
↑**ih-ko-ni** ĩhsi-si'-i [**pɨ'pɨri wa'ti-jo-re**]
DEM.PROX-CL:F-N.SBJ2 give-FUT-N3S condor blade-CL:LONG-N.SBJ
kaa-ɨ-ña
say-2/3S.M.PST.N.ASS-REP
'He wanted to give IT (to her), the vulture's sword. . . '

b. *. . . isiña jare isina jamaca caquëna jamaca io go'ihuëocona. . .*
ĩhsi-ɨ-ña hãre ĩhsi-ɨ-na hãmahka
give-2/3S.M.PST.N.ASS-REP like_that give-S.M.PST.DEP-DS then
kaa-kɨ-na hãmahka ĩ-ɨ go'i + wɨo-ko-na
say-S.M.PST.DEP-DS then 3PRO-CL:M return + begin-S.F.PRS.DEP-DS
'. . . (he) gave (it) like that and then, after giving (it), after saying (that), he started to return (to the terrestial world).'
[NAT*: 20151112orapi001.290-294]

Besides these two discourse-level trigger candidates, the next largest effect sizes are found for SPECIFICITY STATUS (i.e., a highly significant factor, $p =< 0.001$) and WORD ORDER (i.e., a significant factor, $p = 0.001$-0.05); demonstrating highly compatible findings to those discussed for the S-marking sub-sample in Section 5.3.2. On the one hand, as regards SPECIFICITY STATUS, the same pattern emerges as discussed above, such that specific full noun P-tokens are marked at a slightly higher rate than their non-specific and pronominal counterparts. This finding indicates that, although several classes of pronouns undergo promotion, which has implications for surface case-marking (see Section 3.2); a stable distribution of PLAIN and PROMINENT alternatives is observed across all types of pronominal tokens in the sample — example (10), shown earlier in this section, demonstrates instances of PLAIN *-re* and PROMINENT *-ni* on promoted pronoun *ihko* (this one/it — lit. this one (F)).

On the other hand, concerning WORD ORDER effects, the same pattern is noted as found with S-marking, although the effect on P-marking is slightly less robust. Given the dominant SOV word order, pre-verbal P-tokens arise more frequently than post-verbal P, as anticipated. More importantly, PROMINENT-marking is found at a higher

rate for the former category, than for the latter.[25] Further remarks about possible explanations for this pattern, particular to argumental DCM patterns in this dataset, are set aside for the discussion section of this chapter in Section 5.4.

There is one more statistically significant TRIGGER CANDIDATE in the P-marking sub-sample, as demonstrated in Table 5.7: i.e., ANIMACY STATUS. The *varbrul* model finds that inanimate P is slightly more likely to be PROMINENT-marked than animate P. The proportional discrepancy across animacy categories is nearly identical to that found for word order categories, despite differences in effect size.[26] To my view, there are two main ways this statistical effect can be interpreted.

Firstly, a *direct* interpretation takes the notional animacy status of a token's referent to influence the proportion of PROMINENT-marking, along the lines of the effect that specificity status is shown to have. Of course, this type of WEAK TRIGGER is described for several related Tukanoan languages, discussed in Chapter 4. Nevertheless these other Tukanoan languages display the more robust cross-linguistic pattern in which the animate noun class favours overt (i.e., PROMINENT-)marking (Moravcsik 1978; Bossong 1984, 1985; Aissen 1999, 2003; Sinnemäki 2014; etc.); whereas the opposite pattern is detected in the P-marking sub-sample for Siona analyzed in this chapter.

A second interpretation of this finding is possible, which I consider to be *indirect* since the mapping of case-marking alternatives to abstract PLAIN and PROMINENT categories, motivated in Chapter 3 is, in a sense, suspended. The idea is that this slight proportional discrepancy across inanimate and animate categories, not found in S-marking, has to do with the fact that canonical inanimate P displays OPTIONAL-type DCM (i.e., PLAIN zero-marking vs. PROMINENT *-re* marking), whereas animate P displays ALTERNATING DCM (i.e., PLAIN *-re* vs. PROMINENT *-ni*).[27] In light of this formal split, it is conceivable that inanimate P-tokens are PROMINENT-marked in some instances in order to avoid zero-marking, for the sake of disambiguation (see Section 2.2.3.2). I return to this hypothesis in the discussion section of this chapter.

Ultimately, even if we are to accept the *indirect* interpretation of this final finding, it is relevant to recall that the effect of ANIMACY STATUS on the proportion of PROMINENT-marking is a subtle one. This fact is not taken for granted a priori. Although there is evidence for uniting PROMINENT *-re* and *-ni* as a single abstract category shown elsewhere in this dissertation — e.g., QUESTION-ANSWER pairs in Chapter 3; this does not indicate that these two markers necessarily display compatible distributions in a naturalistic sample. As such, this chapter is viewed as supporting the analysis of Siona principled DCM patterns as pairs of PLAIN and PROMINENT forms, which goes beyond a simple formal mapping of case-marking forms. Taken together with the previous section on S-marking patterns in the sample, the quantitative analyses considered thus far have demonstrated that a conspiracy of discourse-related TRIG-

[25]Consult Table B.9 in Appendix B for a more detailed breakdown regarding argument-dropping and word order patterns in the narrative sample.

[26]The ANIMACY STATUS effect detected in the P-marking sub-sample is a significant effect (p=0.001−0.05) — i.e., not a *highly* significant statistical effect as found with the other tested trigger candidates in this analysis (p=< 0.001). This fact is not relevant to the discussion at hand, although it further supports the notion that ANIMACY STATUS has only a slight effect on the distribution of PROMINENT-marking in the sample. Crucially, no intersective effects between ANIMACY STATUS and other tested factors are found in the dataset. This finding merits further discussion.

[27]For all intents and purposes, promotion is irrelevant, since it accounts for only a small subset of P-tokens in the sub-sample (i.e., *N=16*).

GERS, SPECIFICITY STATUS and WORD ORDER factor into argumental DCM in Siona narrations. The following section reviews the results of the L-marking sub-sample.

5.3.4 L-marking in the Siona narrative dataset

The final *varbrul* model developed for the present chapter concerns the STATIC LOCATION, the L-argument. According to the raw distributional facts laid out in Table 5.5 at the opening of this results section, there are a total of 162 L-tokens identified in the Siona narrative dataset. Although the set of L-tokens displays a range of surface case-marking patterns, these all conform to the description of spatial DCM spelled out in the *preliminary description* in Chapter 3 — especially Section 3.3. For this reason, as is the case with the P-marking sub-sample in discussed in the previous section, the mapping from surface case-marking to PROMINENT-marking status has a few steps.

Firstly, the dominant pattern for L-oriented DCM is syncretic with the P-oriented pattern. Canonical, non-promoted L-tokens arise with either PLAIN zero-marking, or with PROMINENT *-re* marking. (16) displays instances of canonical L-tokens: (16a) provides a zero-marked, complex L-argument from *The demon kwẽẽwahti* narration; whereas the noun, *ma'ã-**r**e* (on the path) bears PROMINENT *-re* marking in the instance in (16b), lifted from *The children of Baina* narrative:

(16) a. *mëabebi jaro mëabe tuina naso yai cha'caquë cha'caquë daquëña*
mɨabe-bi **hã-ro** **mɨabe** tui-ɨ-na
above-SRC DEM.MED-CL:LOC above sit_on_top-S.M.PST.DEP-DS
nahso + jai cha'ka-kɨ cha'ka-kɨ
wolly_monkey + jaguar jump-S.M.PRS.DEP jump-S.M.PRS.DEP
dah-kɨ-ña
come-2/3S.M.PST.N.ASS-REP
'As (he) sat up there, the monkey jaguar (species) came, jumping and jumping, from above.' [NAT*: 20151112oespa001.115-116]

b. [**Context:** Baina helps his pregnant wife retrieve arari fruit from the trees]
. . . ñacona arari darëse'e baquëña ma'are
ñaa-ko-na arari da'rɨ-si-je
see-S.F.PRS.DEP-DS arari_fruit fall_to_ground-COMP-CL:GEN
bah-kɨ-ña **ma'ã-re**
be-2/3S.M.PST.N.ASS-REP path-N.SBJ
'. . . (she) saw the arari fruits that had fallen, (they) were on the path.' [NAT*: 20150811sfryi001.073]

Another relevant class of L-tokens concerns promoted tokens. Like animate P, promoted L-tokens arise either with PLAIN *-re* marking, as in (17a) from *The children of Baina* narrative; or with PROMINENT *-ni* as in the promoted spatial pronoun in (17b), lifted from a different episode in the same narrative:

(17) a. *. . . ba'ina jaore saniti'ani abita mëni ocuereña*
ba'i-ɨ-na ↑**hã-o-re** sani + tĩ'ã-ni abita
be-S.M.PST.DEP-DS DEM.MED-CL:F-N.SBJ go + find-SS abita_fruit
mɨni + õhkwe-re-ña
go_up + slurp-N2/3S.PST.N.ASS-REP
'. . . (it, the fruit) was up there, so (they) climbed up (there) to find the *abita* fruit and to eat (it) up.' [NAT*: 20150811sfryi001.450-452]

b. *. . . iconi ja'që baquë ? caëña jamaca*
↑**ih-ko-ni** ha'-kɨ bah-kɨ
DEM.PROX-CL:F-N.SBJ2 parent-CL:M be-2/3S.M.PST.N.ASS
kaa-ɨ-ña hãmahka
say-2/3S.M.PST.N.ASS-REP then
'. . . then (he) said, "was Dad HERE?" ' [NAT*: 20150811sfryi002.666]

In a similar way to the P-marking sub-sample, when the distinction between canonical L-tokens, as in (16), and promoted L-tokens, as in (17), is taken into account, the PROMINENT-marking status of these tokens may be ascertained. Nonetheless, one final class of L-arguments exists, which do not arise in the context of P-marking, as detailed in Section 3.3.3. These concern L-arguments which display *-na* marking, instead of the typical L-marking patterns outlined above. This quirky case-marking pattern is determined by the predicate — as observed in other languages in the region, per the typological discussion in Chapter 4. (18) demonstrates an instance with PROMINENT *-na*, as found with the predicate *bɨa-* (stay), arising in the *Siona Genesis story* narrative:

(18) *i jamu ti ja'o daasiquë jare sa'nihuëna bëaëña*
[ĩ-ɨ hãmu ti ha'o daa-sih-kɨ] hãre **sa'niwɨ-na**
3PRO-CL:M armadillo ANA mud bring-COMP-CL:M like_that inside-GOAL
bɨa-ɨ-ña
stay-2/3S.M.PST.N.ASS-REP
'The armadillo, who brought the mud, stayed inside (the river).'
[NAT*: 20151023orocr001.098]

PROMINENT *-na* marking as in (18) alternates with PLAIN zero-marking, such that a principled DCM pattern is maintained (consult Section 3.3.3 for further details). By extension, where promoted L-nominals are concerned, zero-marking is blocked, as anticipated, and *-na* marking is forced where one of the appropriate predicates arises. Consider the promoted spatial pronoun in (19) from the same narrative as above, which exhibits obligatory *-na* marking with the predicate *guja-* (bathe):[28]

(19) *. . . go'ini caquë ba'quëña baja'i tsiaya jaona. jaëhua'i guyayë*
go'i-ni kaa-kɨ ba-∼'-kɨ-ña
return-SS say-S.M.PRS.DEP be-RMT.PST-2/3S.M.PST.N.ASS-REP
ba-ha'i tsia-ja ↑**hã-o-na** hã-ɨ-wa'i
be-3S.M.PST.ASS river-CL:RIV DEM.MED-CL:F-GOAL DEM.MED-CL:M-PL.AN

[28]This is one of the predicates which has been tested in elicitations. Speakers do not accept *-re* on the L-argument with the predicate *guja-* (or by extension *-ni* where the L-argument is promoted).

guja-ji̵
bathe-N3S.PRS.ASS
'. . . (he) came back and said, "There is a river. They are bathing THERE." '
[NAT*: 20151023orocr001.240-242]

Given that promoted L-arguments, with an appropriate predicate, bear obligatory *-na* marking, these tokens are *invariable* and ought to be excluded from the statistical analysis. Only 9 such tokens are eliminated from this study, including the pronominal instance shown in (19).[29] All other classes of L-tokens are included. Table 5.8 lays out the post hoc conversion from surface case-marking to PROMINENT-marking status utilized for the categorization of tokens in the L-marking sub-sample:

	PLAIN alternative	PROMINENT alternative
LOCATION $_{(-na)}$	*-∅*	*-re/-na*
↑LOCATION	*-re*	*-ni*
↑LOCATION$_{-na}$	*-na* (excluded token)	

Table 5.8: Mapping from noun class to PLAIN-PROMINENT alternatives for L-tokens

The conversion outlined in Table 5.8 is applied against the 152 included L-tokens, which comprise the L-marking sub-sample for this chapter. The following section explores the proportional and statistical evidence yielded by the *varbrul* analysis.

5.3.4.1 Exploring the *varbrul* results for L-oriented DCM

As a point of departure, Table 5.9 recognizes the significant factors from the *varbrul* model developed for the L-marking sub-sample, corresponding to the full model reported in Table B.8 in Appendix B. However, before assessing these results, a disclaimer is in order: Due to the relatively small number of tokens in this sub-sample, the statistical model developed here is found to represent a moderate fit for the data, at best.[30] For the present purposes, this means that the *statistical* evidence reported here is preliminary and ought to be taken with a grain of salt. A future research program, with a larger, targeted dataset, is in order so as to confirm and refine the results outlined in this section. In any case, these preliminary quantitative findings contribute to the broader description of Siona DCM aspired to in the present chapter.

Firstly, and less controversially, Table 5.9 demonstrates that the appropriate PROMINENT-marking arises in the L-marking sample at a global marking rate of 52.63%. This marking rate is markedly higher than those observed for argumental DCM patterns in their respective sub-samples: i.e., S-tokens displays a PROMINENT-marking rate of 21.18%, and P-tokens, 24.27%. As hinted at in the introduction of

[29]Table 5.5 identifies a single outlier L-token which bears *-hã'ã* (PATH/LIMIT) marking. This token is also excluded from the study per variationist conventions.

[30]On statistical grounds, this is determined by the *input rate* metric, reported in the full model at Table B.8 in Appendix B. Given that this rate is relatively close to 1 (as opposed to 0), the *varbrul* model is deemed to provide *weak* statistical evidence for any of the findings which it yields. It is also likely that the higher global marking rate implies a lower degree of *variability* in the dataset, which necessitates the assessment of further data to detect effect size.

this chapter, this fact points to a previously unnoticed pattern where spatial DCM exhibits PROMINENT-marking at a far higher rate than argumental DCM in the language. I return to this generalization in the discussion in Section 5.4.

Setting aside the proportional discrepancy between argumental and spatial DCM, Table B.8 reports preliminary statistical evidence for remarkably similar effects for the tested trigger candidates in the L-marking sample and discussed for argumental patterns in the previous sections of this chapter. The TRIGGER CANDIDATE, LOOK-AHEAD TOPICALITY, is retained for the sake of this discussion. However, every L-token coded as topical (based upon the *topic persistence* metric) bears PROMINENT-marking in the sub-sample, I take this finding to be relevant for the present discussion. For this reason, this TRIGGER CANDIDATE is indicated separately from the other significant factors in the table below — i.e., in gray text with the symbol '!' :[31]

TRIGGER CANDIDATE	*N* [% PROMINENT]
GLOBAL L-MARKING RATE	152 [52.63%]
CONTRASTIVITY STATUS	(LARGE EFFECT SIZE)
Contrastive L	20 [90%]
Non-contrastive L	132 [47%]
SPECIFICITY STATUS	(MODERATE EFFECT SIZE)
Pronominal L	65 [46.2%]
Specific L	67 [62.7%]
Non-specific L	20 [19.4%]
!LOCAL TOPICALITY STATUS	(INVARIABLE)
Local topical L	8 [100%]
Post-verbal L	144 [50%]

Table 5.9: Significant factors for L-oriented DCM in *varbrul* sub-sample

Firstly, Table B.8 demonstrates that this *varbrul* model did *not* detect an effect for LOOK-BACK TOPICALITY (i.e., $p => 0.05$); which distinguishes tokens with locations mentioned more recently, from those with distant mentions in a given discourse. Likewise, the thematic effect tested for L-oriented DCM, SUBCATEGORIZATION STATUS, is not statistically significant (i.e., $p => 0.05$). This result points to the fact that L-marking is not significantly different for L-arguments which are implied by the predicate, and for those which are not. Finally, this analysis finds that WORD ORDER does not impact L-oriented DCM.[32]

[31] Many variationists would elect to eliminate this factor from the report, since these are technically *invariable* tokens. Given that this fact is determined *a posteriori*, and is not expected based upon the grammar, per Chapter 3, this type of invariability is informative for the present purposes.

[32] Table B.9 in Appendix B outlines the distribution of core and non-core arguments relative to the main, conjugated, verbal word in main clauses in the Siona narrative dataset. Spatial arguments are collapses together with temporal and adjunctive arguments; however, there is

Turning to trigger candidates with active statistical effects in the L-marking sub-sample, CONTRASTIVITY STATUS exhibits a large effect size (i.e., a highly significant result, $p=< 0.001$). (20) illustrates a pair of prototypical contrastive L-tokens from *The demon kwẽ̃wahti* narrative. In this episode, while the demon attacks one man in his home, the shaman protagonist lies in a hammock in the corner of the room. The narrator uses PROMINENT *-ni* on the two promoted L-tokens to contrast the locations of each of the victims:

(20) *a̲coña i̲re ba̲i̲re te'i yequëje̲ baquëña turihuë hueosiconi. ja̲o̲ni u̲quëña yequë*
ãh-ko-ña ĩ-i-re bãĩ-i-re te'-i
eat-2/3S.F.PST.N.ASS-REP 3PRO-CL:M-N.SBJ person-CL:M-N.SBJ one-CL:M
jehk-i-hẽ bah-ki-ña ↑[**turi-wi**
other-CL:M-ADD be-2/3S.M.PST.N.ASS-REP corner-CL:CONT
weo-sih-ko-ni] ↑**hã-o-ni**
lay_down.TRS-COMP-CL:F-N.SBJ2 DEM.MED-CL:F-N.SBJ2
ũh-ki-ña jehk-i
be_lying-2/3S.M.PST.N.ASS-REP other-CL:M
'(She) ate him, the man. Also another man was in the CORner where he laid down. The other one (M, i.e., the protagonist) was lying THERE.'
[NAT*: 20151112oespa001.024-028]

Example (20) resembles the prototypical instances of contrastive S- and P-tokens presented in the preceding sections. In a similar vein, Table B.8 a similar, highly significant effect is detected for SPECIFICITY STATUS (i.e., $p=< 0.001$), resembling such effects in the argumental DCM sub-samples. Specific full noun L-tokens bear PROMINENT-marking at a markedly higher rate than pronominal and non-specific tokens. This TRIGGER CANDIDATE displays a similar medium-sized effect within the L-marking sub-sample. As such, the propensity for specific noun tokens to favour PROMINENT-markers is common to all statistically tested grammatical relations in this chapter.

Returning to LOOK-AHEAD TOPICALITY STATUS, given that all 8 local topical tokens coded in this modest sample bear PROMINENT-marking, many variationists would eliminate this factor from the *varbrul* report — i.e., technically, no significance value may be ascertained. Given that the *varbrul* model developed for the L-marking sub-sample is preliminary anyhow, this is treated as a relevant finding to be refined in future work. This is the same brand of topicality effect which is active in the P-marking sub-sample — i.e., such that L-tokens which refer to locations that are at-issue in immediately ensuing lines of narrative bear PROMINENT-marking. This pattern is illustrated in the following excerpt of *The children of Baina* narrative, lifted in (21):

(21) a. *iñore ba'ico cani sani saëña . . .*
ihño-re ba'i-ko kaa-ni sani + saa-i-ña
there-N.SBJ live-3S.F.PRS.ASS say-SS go + bring-2/3S.M.PST.N.ASS-REP
' "(She) lives THERE", (he) said, and (he) brought (her there). . . '

still a propensity for nominal tokens to arise before the verb in the sentence, as was found with core arguments. The difference is that this does not appear to have implications for PROMINENT-marking on spatial nouns.

b. *. . . cani saëna tɨ̲'a̲o̲ña – tɨ̲'a̲cona jai tsio baquëña joborare huë'e baquëña*
kaa-ni saa-ɨ-na tĩ'ã-o-ña
say-SS bring-S.M.PST.DEP-DS arrive-2/3S.F.PST.N.ASS-REP
tĩ'ã-ko-na hai tsio bah-kɨ-ña
arrive-S.F.PRS.DEP-DS big garden be-2/3S.M.PST.N.ASS-REP
hobo-reba-re wɨ'e bah-kɨ-ña
halfway-INTENS-N.SBJ house be-2/3S.M.PST.N.ASS-REP
'. . . after (he) said that and brought (her) (there), (she) arrived. (She) arrived and there was a big garden, and there was a house in the middle.'
[NAT*: 20150811sfryi001.450-452]

In spite of the need for further data to strengthen the statistical findings presented in this section, the analysis of the L-marking sub-sample of the Siona narrative dataset contributes preliminary quantitative findings to the broader discussion in this chapter. On the one hand, L-oriented DCM in the sample deviates from argumental DCM patterns as regards their global PROMINENT-marking rates. It is the case that spatial arguments tend to bear overt, PROMINENT-marking, in more non-contrastive instances than is noted for argumental DCM. On the other hand, all grammatical relations which have undergone statistical analysis here demonstrate a strong effect of CONTRASTIVITY STATUS, and a medium-sized effect based upon SPECIFICITY STATUS. These findings are integrated into the broader analysis of Siona DCM in the discussion section below.

5.4 Discussion: Siona DCM in usage

This section summarizes the proportional and statistical evidence assembled in the variationist study in this chapter and bridges these findings into the broader discussion in this dissertation. As a point of departure, Table 5.10 provides a summary for the global PROMINENT-marking rates (i.e., proportional evidence) and the effects detected for each tested TRIGGER CANDIDATE in the present analysis (i.e., statistical evidence), as found in the S-, the P-, and the L-marking sub-samples of the Siona narrative dataset. trigger candidates found to be insignificant are identified as displaying $^{!}$NO EFFECT (i.e., $p => 0.05$); whereas significant trigger candidates are identified by their respective effect sizes. A subscript numeral ranks these effects relative to one another for each given grammatical relation (e.g., 1, 2. . . n):[33]

[33] As is the case throughout this chapter, this discussion collapses significant and highly significant factors, as yielded by the *varbrul* models (p-value $= < 0.05$ are treated as significant). For the present purposes, this distinction is not relevant — however, the reader is directed to the full reports in Appendix B for a more fine-grained demonstration of significance values.

Trigger candidate [Global marking rate]	s-marking [21.18%]	p-marking [24.27%]	l-marking [52.63%]
		Sentence-level factors	
Word order	Moderate effect$_3$	Small effect$_4$	!No effect
Thematic status	!No effect	!No effect	!No effect
		Referent-level factors	
Animacy status	!No effect	Small effect$_5$	n/a
Specificity status	Moderate effect$_2$	Moderate effect$_3$	Moderate effect$_2$
		Discourse-level factors	
Contrastivity status	Large effect$_1$	Large effect$_1$	Large effect$_1$
Look-ahead topicality status	!No effect	Moderate effect$_2$	!Invariable
Look-back topicality status	!No effect	!No effect	!No effect

Table 5.10: Summary of proportional and statistical evidence from corpus study

Firstly, Table 5.10 summarizes the proportional evidence for this analysis, based upon global prominent-marking rates. As suggested in the preceding sections, a non-trivial discrepancy is detected regarding the proportion of prominent-marking found with argumental dcm patterns, compared with spatial dcm in the sample. Concretely, the l-marking sub-sample explored in Section 5.3.4 exhibits a global marking rate which is roughly twice as high as those found for the s-marking sub-sample analyzed in Section 5.3.2, or for p-marking in Section 5.3.3.

As suggested at the opening of Section 5.3, based upon the raw surface case-marking distributions in Table 5.5, further support for the proportional discrepancy between argumental and spatial dcm is found where Indirect object (r) and the spatial Goal (g) case-marking are taken into account. Each of these grammatical relations displays a principled, plain-prominent dcm pattern in its own right — as laid out in Section 3.2.3 and Section 3.3.2, respectively, in the *preliminary description*. Although the corresponding sub-samples in this analysis were not extensive enough to permit the informative application of the *varbrul* technique, on the basis of the raw distributional facts alone, the division between argumental and spatial dcm regarding prominent-marking rates is borne out. The preliminary proportional figures for these grammatical relations are reiterated in Table 5.11, indicated with the symbol '!' (based upon datasets which are not thoroughly 'cleaned'). These are placed alongside the proportional facts for the tested grammatical relations, and animate and inanimate p-marking rates are presented independently, for the sake of the ensuing discussion:[34]

[34]As regarding the r-tokens in this sample, only animate r is recorded — according to Section 3.2.3, any inanimate r-tokens would reject zero-marking and be excluded in any case. The g-marking figures demonstrated here do *not* account for excluded promoted tokens, which would require *-na* marking in all instances. These rough figures are strictly intended to highlight the discrepancy as a generalization.

Grammatical relation	N [% Prominent]
s-marking	727 [21.18%]
p-marking	478 [24.27%]
Animate p	188 [19.68%]
Inanimate p	290 [27.24%]
! r-marking	70 [24.29%]
l-marking	152 [52.63%]
! g-marking	118 [57.8%]

Table 5.11: Proportion of prominent-marking in sample, argumental vs. spatial
(! indicates that no corresponding *varbrul* analysis is developed in the current analysis)

The motivation for the above-noted proportional discrepancy is not immediately clear. In elicitation-based research, like that considered in Chapter 3 and Chapter 6, speakers report highly compatible judgments for argumental and spatial dcm: For instance, the use of overt, prominent-marking in out-of-the-blue contexts tends to trigger the *emphatic contrast* reading. Nonetheless, in the naturalistic narrative dataset analyzed in this chapter, spatial nouns differ markedly from argumental nouns on the basis of proportional evidence. Since l- and g-tokens have spatial referents, on conceptual grounds, it is unlikely that their propensity for overt, prominent-marking is driven by pressures to disambiguate them from surrounding (argumental) nominals. At present, the proportional discrepancy summarized in Table 5.11 is taken to be a descriptive generalization, to be more fully unpacked in future research.

The remainder of this discussion considers the statistical evidence put forth in this chapter, and recapitulated in Table 5.10. In light of patterns of overlap regarding the activity of particular trigger candidates, as conceived in this chapter; a discussion by Factor groups (rather than by grammatical relation) is in order. Section 5.4.1 delves into the findings regarding discourse-related Triggers. Then, Section 5.4.2 considers referent-level Triggers. This discussion closes with a consideration of sentence-level Triggers in Section 5.4.3.

5.4.1 Discourse-related Triggers in the dataset

The largest statistical effect found for dcm patterns on s-, p-, and l-oriented nominal tokens concerns Contrastivity status. This analysis adds statistical evidence for contrastivity as a Strong trigger for Siona dcm across-the-board, which is furthered with elicitation-based evidence in the ensuing chapter of this dissertation. To be certain, the inherent connection between prominent-marking and discourse-related Triggers aligns with earlier text-based analyses for dcm in related Siona+ languages, conceived as 'focal', vs. 'plain', case-markers accordingly (Cook and Levinsohn 1985 for Koreguaje [*coe*]; Wheeler 1970, etc., for Colombian Siona, who includes a so-called *emphatic* focal marker category); and the present analysis lends quantitative support to this conception of dcm, which has been adopted from the outset of this dissertation for the analysis of the target variety of Ecuadorian Siona.[35]

[35]In this chapter I am not committed to making a distinction between contrast and focus as the Strong trigger for Siona dcm, since Focus status is not tested here per se. The patterns noted in this quantitative study are largely compatible with those arrived at via

Besides highlighting the uniquely strong effect of CONTRASTIVITY STATUS on the distribution of PROMINENT-marking, this chapter assembles statistical evidence to test for effects related to TOPICALITY STATUS, as recognized in the broader literature. The summary provided in Table 5.10 points to a correlation between topicality and PROMINENT-marking in both the P- and the L-marking sub-samples. As such, the role of topicality in Siona DCM transcends the argumental-spatial divide noted on the basis of proportional evidence, and reinvigorates the SUBJECT vs. non-SUBJECT divide, established in Chapter 3.

Section 5.2.1 (and Section B.1 in Appendix B) lays out how two text-based metrics for testing for topicality effects are implemented in this study: On the one hand, LOOK-BACK TOPICALITY STATUS, based upon the distance between overt mentions, is not significant in the sample. Conversely, what I label LOOK-AHEAD TOPICALITY STATUS, measures (re)activation and persistence of a referent across clauses (related to the effect detected in Romanian *differential object marking*, Chiriacescu and von Heusinger 2010). The latter topicality metric accounts for the second strongest statistical effect in the P-marking sub-sample, and *all* local topical L-tokens are PROMINENT-marked.

A relevant question emerges on the basis of the topicality effects detected in the Siona narrative dataset: Why is LOOK-AHEAD TOPICALITY not active in the S-marking sub-sample? To my view, this discrepancy has to do with the fact that the topicality of the S-argument is encoded via non-DCM mechanics of Siona grammar — i.e., switch-reference markers on dependent verbs, and the related, S-particular, SLDC construction introduced in Section 5.3.2 (see Bruil and Case 2023). As such, strictly non-SUBJECT grammatical relations recruit DCM to signal that their referent is important for the immediately ensuing discourse[36] — i.e., a mechanism to facilitate referential access in the sense of Gernsbacher (1989). It is relevant to note that Cook and Levinsohn (1985) report a similar effect of 'importance to narrative development' as a TRIGGER for 'focal' case-marking in Koreguaje [*coe*]. Though they claim that this effect equally obtains for S-marking via the cognate *-pi* marker, based upon illustrative text-based examples. Nonetheless, as regards non-SUBJECT DCM, highly compatible findings are reported for Siona in this chapter. Discourse-related TRIGGERS outweigh others in Siona DCM.

5.4.2 Referentiality-related Triggers in the dataset

Besides the robust discourse-related DCM effects found here, SPECIFICITY STATUS demonstrates a remarkably stable effect on PROMINENT-marking across each tested sub-sample, independently of other factors. After discourse-related TRIGGERS, the next largest effect in all sub-samples is one where specific full noun tokens are fitted with PROMINENT-marking at a slightly higher rate than do pronominal or non-specific tokens.[37] Whereas this observation is made impressionistically for *-re* marking on inanimate P in Bruil (2014, §4.4.2),[38] the generality of this effect in the Siona narra-

qualitative techniques, where the notion of contrastive focus is at the fore.

[36]It is not clear whether this analysis finds non-SUBJECT PROMINENT-marking to be a bona fide topic shift device. This analysis is perhaps untenable given that only a medium-sized effect is detected in the relevant sub-samples.

[37]As laid out in Section B.1 of Appendix B, this analysis finds no meaningful distinction in the coding of specific indefinite nouns and definite noun phrases, as defined here. This indirectly supports the notion that definiteness is *not* a bona fide category in Siona grammar.

[38]Bruil 2014, §4.4.3, hypothesizes that specificity may also impact the distribution of *-na* marking on the G-argument.

tive dataset is unanticipated. For instance, specificity-triggered SUBJECT-marking is generally taken to be cross-linguistically rare (de Hoop and Lamers 2006; Chappell and Verstraete 2019; cf. Haspelmath 2021 regarding pronoun-full noun distinction). It may be the case that specific nouns are more likely to be important, or for their referents to be contrasted with others; although, crucially, no intersective effects are detected with these factors in the respective *varbrul* models developed for the present analysis. Rather, like the strong contrastivity effect, this subtle specificity effect seems to characterize the PROMINENT-marking category itself, as a component of principled DCM patterns in Siona grammar. This conception is at odds with specificity-related effects noted in the *differential object marking* literature (Moravcsik 1978; Bossong 1985; Aissen 1999, 2003; Sinnemäki 2014; etc.). SPECIFICITY STATUS corresponds to a general WEAK TRIGGER for Siona DCM, which merits further investigation via different techniques.

As regards ANIMACY STATUS, Table 5.10 confirms that no effect is observed in the S-marking sample. This finding aligns with the description for S-oriented DCM laid out in the *preliminary description* (Section 3.1). However, as regards P-oriented DCM, a notably small effect is noted such that inanimate P favours PROMINENT-marking over animate P (reiterated in Table 5.11). In Section 5.3.3.2, it is claimed that this finding likely points to the fact that inanimate P sometimes recruits (PROMINENT) case-marking for the sake of disambiguation, given that the dominant pattern is one where this overt PROMINENT-marker competes with PLAIN zero-marking (disregarding promoted P-tokens: N=16) — i.e., unlike animate P. Of course, the design of this chapter only arrives at this hypothesis indirectly. A targeted study of disambiguation effects in Siona DCM is certainly warranted, since several questions arise which cannot be addressed here: e.g., Are these effects also expected elsewhere where an OPTIONAL-type DCM pattern is attested (i.e., S-marking)? Does disambiguation depend upon the presence of *overt* competitors in the sentence? The role of disambiguation is avoided by design in this dissertation, although it merits further investigation as a component of a holistic description of Siona DCM.

5.4.3 Sentence-level Triggers in the Siona narrative dataset

This chapter tests for two sentence-level trigger candidates: i.e., WORD ORDER and THEMATIC STATUS. Firstly, regarding the latter, the quantitative analysis carried out here finds that there is no correlation between the proportion of PROMINENT-marking and the grammatical relation-appropriate thematic categories. Therefore, as concerns the S-marking sub-sample, the agentive S and the non-agentive S categories demonstrate compatible DCM patterns. The same is found in the P-marking sub-sample, regarding affected and non-affected P categories; and in the L-marking sub-sample as concerns those spatial arguments which are subcategorized for by their predicate, and those which have an adjunctive function. These findings corroborate impressionistic generalizations described for S-oriented DCM and P-oriented DCM in Chapter 3. This finding for L-oriented DCM is a novel one, but fits the following generalization: The thematic properties of the predicate do not have any bearing on DCM patterns found on its arguments in Siona DCM. This generalization matches similar calls for skepticism for thematically-driven DCM patterns in the literature (Fauconnier 2011; Fauconnier and Verstraete 2014).

Turning to the effect of WORD ORDER on Siona DCM, this TRIGGER CANDIDATE is found to be active for both argumental grammatical relations tested in this sample —

i.e., no such effect is detected in the L-marking sub-sample. As regards the S-marking sub-sample, this is a medium-sized effect, but the weakest of the active effects for the tested factors. A small effect size is noted in the P-marking sub-sample, such that only the above-noted effect of ANIMACY STATUS is found to have a smaller effect. These effects are highly similar, such that nouns arising before the main verbal word in the sentence display a higher proportion of PROMINENT-marking. In the general case, this corresponds to PROMINENT-marking being *disfavoured* where nominal tokens arise after the verb. Given the dominant SOV word order in the language, post-verbal argumental noun tokens are always outside of their dominant sentential positions.

To my view, the above-noted subtle word order effect on argumental DCM signals two things about Siona DCM. Firstly, it is conceivable that noun phrases arising outside of their canonical positions might display PROMINENT-marking at an above-average rate given that their grammatical relation is potentially ambiguous. In a sense, the speaker is eliminating one cue for identifying the role of a noun token. The hypothesis is not borne out since the pattern arises in the opposite direction in the present narrative sample.

Another way to interpret the small, yet significant, word order effects is that pre-verbal nominals tend to be those which are more important in one way or another. (Given that there are no intersective effects in the present study, this effect must be rather indirect.) On these grounds, however, it might be expected that similar effects arise for spatial nouns. In any case, a more detailed investigation of the motivations for variation in word order is merited, so that the link between pre-verbal sentential position and PROMINENT case-marking in narration may be better established. Ultimately, word order is taken to be a WEAK TRIGGER for argumental DCM in naturalistic settings, which indicates a fickle connection between DCM and other variable components of the grammar.[39]

In summary, this chapter has shown that a unified account of Siona DCM is tenable, on the basis of a PLAIN-PROMINENT analysis. Certain quirks separate argumental and spatial DCM, particularly on proportional grounds; whereas, in other respects, the SUBJECT vs. non-SUBJECT divide in Siona DCM also has subtle implications for the relevance of particular TRIGGERS. As such, this latter division goes beyond the activity of a formal split based upon animacy classes, as established in Chapter 3, and set aside for the present discussion. Inasmuch as the animacy-conditioned split is a STRONG TRIGGER for Siona DCM, following the conception of multi-dimensional DCM codified by Klein and de Swart (2011); it is different in kind from the STRONG TRIGGER of contrastivity. Whereas the former determines morphological case-marking splits, the latter is the strongest predictor for selecting the appropriate PROMINENT-marking in a given context. The quantitative analysis in this chapter corroborates similar claims made on the basis of elicited data presented in the following chapter.

[39]The conspiracy of these grammatical devices, and probably patterns of argument-dropping, are often viewed as components of discourse-configurationality in grammar. Perhaps a broader investigation along these lines is in order.

CHAPTER 6

Information structure and Siona DCM

The previous chapters in this dissertation establish that information structure plays an indispensable role in Siona DCM patterns. Firstly, the *preliminary description* in Chapter 3 employs a QUESTION-ANSWER (Q-A) congruence heuristic to establish the paradigmatic distribution of PLAIN vs. PROMINENT case-markers in the grammar. This aligns with the notion of the interrogative type of information focus in the literature (e.g., Dik et al. 1981, 1997; Kiss 1998). The Q-A pair presented in (1) exemplifies the general pattern of obligatory case-matching in information questions:

(1) a. QUESTION: *quebi iño aide'oquë ?*
ke-e *(-bi) ihño ai + de'o-kɨ
WH-CL:GEN-SBJ here big + become-2/3S.M.PRS.N.ASS
'What grows here?'

b. ANSWER: *jaiye jorobi aide'oji iño*
hai-je **horo #(-bi)** ai + de'o-hi ihño
big-CL:GEN flower-SBJ big + become-3S.M.PRS.ASS here
'Many FLOWers grow here.' [VOL/SUG: 20230622ejabi002.003a-b]

Secondly, Chapter 3 also demonstrates that the emphatic contrast reading is typically evoked by the presence of PROMINENT case-marking in non-interrogative contexts. This function of DCM finds counterparts in other languages spoken in the putative NWA region, as discussed in Chapter 4 — including Colombian Siona [^{c}*snn*] (Wheeler 1970, 1987), Koreguaje [*coe*] (Cook and Levinsohn 1985), and the non-Tukanoan languages Tariana [*tae*] (Aikhenvald 2003) and Murui [*huu*] (Wojtylak 2021). The minimal pair presented in (2) illustrates the interpretative effects associated with PROMINENT-marking on the Experiencer argument. Whereas the use of PLAIN *-re* marking in (2a) produces an unmarked reading; PROMINENT *-ni* marking in (2b) evokes an interpretation where the denoted referent is contrasted with salient alternatives:

(2) a. *iye cocabi yë're jëaji*
i-je kohka-bi **ji'i-re** hia-hi
DEM.PROX-CL:GEN language-SBJ 1SG-N.SBJ be_hard-3S.M.PRS.ASS
'This languages is difficult <u>for me</u>.'

b. *iye cocabi yë'ni jëaji*
i-je kohka-bi **ji'i-ni** hia-hi
DEM.PROX-CL:GEN language-SBJ 1SG-N.SBJ2 be_hard-3S.M.PRS.ASS
'This languages is difficult <u>for ME</u> (i.e., not for somebody else).'
[**Speaker comment:** *-ni* means it is only hard for me.]
[VOL: 20230525ejabi002.005a-b]

(2a) and (2b) assert the same content; however, the latter evokes a set of alternative propositions, which are subsequently excluded for the sake of interpretation. This effect meets the generally accepted definition for contrastive focus (e.g., Rooth 1985, 1992; Selkirk 2008).

Baseline definitions aside, there is general consensus that there exists no monolithic notion of focus which is cross-linguistically valid (van der Wal 2015, 2016, 2022; Kratzer and Selkirk 2020). An important question arises as to whether the Q-A pattern shown in (1) may be analyzed together with the contrast-related pattern shown in (2). This dissertation has already established that both are important components of any adequate analysis of information structure-driven DCM in Siona. Another immediate question which arises is whether PROMINENT-marking is also implicated in other typical focal contexts — e.g., corrective focus, selective focus, etc. (Dik et al. 1981, 1997 — see Krifka 2011 for a recent overview). The distribution of PROMINENT-markers under these different contexts provides insights into their precise semantico-pragmatic function. After all, Matić and Wedgwood (2013) argue that the distribution of a certain marker may be related to focus, without encoding focus per se.

The present chapter investigates whether the above-noted information structure-driven patterns of Siona DCM may be unified under a single analysis — i.e., PROMINENT-marking is obligatorily used in overt Q-A constructions and to induce the emphatic contrast effect. For the sake of the present discussion, these are referred to as the information focus effects and the contrastive focus effects of Siona PROMINENT-marking. On this analysis, Siona DCM might be argued to be a focus-encoding strategy in its own right (Büring 2003, 2009). For instance, such an analysis would be compatible with Siona DCM information structure effects as driven by underlying (implicit) question-answer sequences across-the-board (e.g., the *question-under-discussion* approach developed in Roberts 2012, etc.). It will be demonstrated that this analysis is likely too strong since it (wrongly) predicts PROMINENT-marking to occur in a number of contexts where it is absent.

The alternative view is that the information focus-related function of PROMINENT-marking in Q-A contexts ought to be separated from its contrastivity-encoding function. Basically, in non-interrogative contexts, PROMINENT-marking is found primarily in contrastive contexts. In the terminology of this dissertation, the latter analysis maintains that Siona DCM has two strong TRIGGERS — i.e., one for information focus and another for contrastive focus, or perhaps contrastivity more broadly. It is relevant to recognize that PROMINENT-marking is not utilized to encode contrastive topicality; which recruits a dedicated marker *=kato* (glossed TOP). Section 6.4.2 identifies the distribution and function(s) of this and other information structure-related nominal morphology. Ultimately the aim of this chapter is to describe the distribution

of PROMINENT case-marking as conditioned by information structure. This analysis implements an elicitation-driven methodology (inspired by van der Wal 2015, 2016).

The remainder of this chapter is organized as follows: Section 6.1 outlines the elicitation techniques employed to derive the data for this chapter. Section 6.2 situates the PROMINENT case-marking strategy in the typology of focus-encoding strategies, corresponding to the notion of FOCUS REALIZATION (cf. Büring 2009; Zimmermann and Onea 2011; Aannestad 2021). Section 6.3 addresses questions relating to FOCUS SIZE, which analyzes the scopal properties of the PROMINENT-marking strategy. Section 6.4 assesses the distribution of PROMINENT-marking strategy across typical focus contexts. This section also provides a description of the interaction of PROMINENT-marking with discourse markers in the language. Section 6.5 concludes this chapter.

6.1 Methodology: Controlled elicitation

The data presented in this chapter are collected via elicitation-based techniques designed to target focus-encoding strategies, as recently outlined in van der Wal (2015, 2016). The incorporation of elicitation-derived data into the description of Siona DCM is appealing on several grounds: i.e., these techniques permit for fine-grained researcher-control over variables and context, and for the production of negative examples (Matthewson 2004, 2022). Three classes of elicitation-based heuristics inform the present analysis: (i) Q-A heuristics; (ii) CONTEXT-CONJURING heuristics; and (iii) CO-TEXT heuristics.

Q-A heuristics are a classical diagnostic used to assess the size of the focalized constituent at interpretation (Dik 1997; Rooth 1985, 1992 Krifka 2011; Roberts 2012; and others — see van der Wal 2016, §3). In the domain of WH-questions, one particular type of focus is targeted, labeled COMPLETIVE FOCUS in Dik (1981, 1997),[1] where the listener is presumed to be aware of all information besides the focalized element. The formalization in (3) demonstrates how the Q-A heuristic can be operationalized:

(3) **Q-A test:**
If a question asks for some X (X being a syntactic category), in a direct answer to this question, the constituent which corresponds to X is [focalized].
[as formalized in Kasimir 2005, 12; cited in van der Wal 2016, 265]

Van der Wal goes on to provide the illustrative instances from English adapted in (4):

(4) a. [What kind of juice did Little Tiger drink?]
He drank APPle juice.
b. [What did Little Tiger drink?]
He drank APPle juice.
c. [What did Little Tiger do?]
He drank APPle juice.

The sentences in (4) demonstrate that the same form is used to answer various questions. Each question targets a different (sub-)constituent of the sentence, such that the

[1] As discussed in Krifka and Musan 2012, many earlier accounts presume that Q-A diagnostics involve *information focus*, rather than *contrastive focus*. Following Kratzer and Selkirk 2020, and others, this chapter does not place weight on this distinction for the analysis of Siona DCM patterns.

corresponding focalized elements in the answers are of various sizes. Q-A diagnostics are exploited to assess the focus sizes available for the interpretation of PROMINENT-marking in Siona in Section 6.3.

The Q-A heuristic is particularly useful for the study Siona DCM, in light of the Q-A congruence effects noted in Chapter 3. Compare the nearly identical sentences in (5) and (6b) below. In the former, out-of-the-blue instance, the speaker deems PROMINENT-marking to be infelicitious on the SUBJECT-argument. However, in the latter case, PROMINENT-marking is obligatory where this is the focalized constituent in the ANSWER turn, just as PROMINENT-marking arises obligatorily on the corresponding WH-item in the QUESTION turn in (6a):[2]

(5) *jaiye joro aide'oji iño*
hai-je **horo (# -bi)** ai + de'o-hi ihño
big-CL:GEN flower-SBJ big + become-3S.M.PRS.ASS here
'Many flowers grow here.' [VOL: 20230622ejabi001.012]

(6) a. QUESTION: *quebi iño aide'oquë ?*
ke-e-bi ihño ai + de'o-ki̶
WH-CL:GEN-SBJ here big + become-2/3S.M.PRS.N.ASS
'What grows here?'

b. ANSWER: *jaiye jorobi aide'oji iño*
hai-je **horo #(-bi)** ai + de'o-hi ihño
big-CL:GEN flower-SBJ big + become-3S.M.PRS.ASS here
'Many FLOWers grow here.' [VOL/SUG: 20230622ejabi002.003a-b]

On the basis of the rigid Q-A congruence effects shown in (6), and outlined at length in Chapter 3, the implication of Q-A heuristics ought to be treated with care. PROMINENT case-marking is obligatory in instances such as (6b), with an overt question (i.e., (6a)), which is better treated as a construction effect. This analysis takes this fact to show that information focus is not a TRIGGER for DCM per se, excluding instances with overt Q-A pairs like these. Conversely, this analysis amasses evidence to demonstrate that contrastive focus *is* a STRONG TRIGGER for Siona DCM via additional focus heuristics.

The second class of heuristic employed to analyze information structural effects on DCM concerns CONTEXT-CONJURING (van der Wal 2016, §2). This technique consists in the researcher presenting the speaker with the target construction, and asking for their intuition about where such a sentence might be most naturally employed.[3] (7) illustrates an accepted instance with PROMINENT *-re* marking arises on the predicative complement. Beyond simply acknowledging the grammaticality of the sentence, the speaker conjures a context which elucidates the semantic implications of the PROMINENT-marking strategy:

[2]This chapter retains the typographical conventions established in **Chapter 1** for the sake of consistency. Underlining serves to link an element of the suggested English translation to the corresponding **boldface** item in the parsed Siona original. Focus is indicated in the English translation tier via SMALL CAPS, which emulated English prosodic stress placement: cf. 'canoe' and 'caNOE'.

[3]Note that these are classified as volunteered instance (VOL), since the speaker is responsible for producing the contexts in a semi-spontaneous manner — i.e., unlike grammaticality judgments produced entirely by the researcher.

(7) *huëhuëyere yëyë*
wɨwɨ-je #(-re) jɨɨ-jɨ
run-CL:GEN-N.SBJ want-N3S.PRS.ASS
'(I) want to RUN.' ⇒ {I want to hunt, I want to fish,...}
[**Context suggested by speaker**: You prefer to run over participating in the other activities planned for a particular event.]
[VOL: 20241104ejabi001.015]

The final heuristic employed in this chapter concerns what van der Wal (2016, §6) labels CO-TEXT diagnostics. These diagnostics are implemented as a class of translation tasks, where the researcher uses complex prompts to induce a contrastive context on a particular constituent in order to assess the activity of the target focus-encoding strategy. The two examples below illustrate rich prompts, which are designed to explicitly contrast nominal constituents in juxtaposed sentences:[4]

(8) [*bahuë yohuë hueroye*] [*te'e guënarore huerohuë*]
bãã-wɨ **jo-wɨ** wero-je **te'-e**
NEG.AUX-N3S.PST.ASS canoe-CL:CONT buy-CL:GEN one-CL:GEN
gĩna-ro #(-re) wero-wɨ
metal-CL:RECIP-N.SBJ buy-N3S.PST.ASS
'[I didn't buy a BOAT], [(I) bought a COOKing pot].'
[VOL/SUG: 20230623ejepa001.013]

(9) [*yë' a'yëbi goameña nequëa'bi iño*] [*cuï'ne yë' yojeibi ye'yequëa'bi*]
jɨ'ɨ a'j-ɨ-bi goa-mahña nee-kɨ-a-'bi ihño kwĩ'ne
1SG older_sibling-CL:M-SBJ thing-DIM.PL make-CL:M-COP-3S.M here also
jɨ'ɨ johe-ɨ-bi je'je-kɨ-a-'bi
1SG younger_sibling-CL:M-SBJ study-CL:M-COP-3S.M
'[My OLDer brother works here], [but my YOUNGer brother studies].'
[VOL: 20230617elupa001.002]

CO-TEXT heuristics provide means to test the role of PROMINENT-marking outside of Q-A contexts, as discussed in Section 6.4. A combination of the three families of focus heuristics outlined above is employed is employed to establish the empirical backdrop for the analysis of Siona PROMINENT-marking strategy in this chapter.

6.2 Case-marking as a focus strategy

This section develops the analysis for how Siona PROMINENT-marking may be conceived as a (morphological) focus-encoding strategy. The following sections will determine to what extent the distribution of these markers is focus-sensitive. The domain of FOCUS REALIZATION refers to the range of strategies which languages may employ to express focal meaning — see Zimmermann and Onea 2011 for recent discussion. According to the fine-grained typology outlined in Büring (2009), languages may

[4]Per van der Wal (2016, §6), (8) and (9) diagnose different focal effects. The former concerns an *exclusive continuation* sequence (X ... but not Y); whereas the latter provides an instance of *overtly juxtaposed sequences* (X ..., and Y ...). In future work, particular co-textual diagnostics will be applied in a more careful way to test for exhaustivity effects.

recruit one or several prosody-, syntax-, and morphology-based strategies of focus realization. Focal case-marking is considered a morphological focus-marking language, belonging to the PARTICLE LANGUAGE-type proposed by Büring (2009, 22-24), where a special morpheme marks focus. This would be a purely morphological language since other devices such as word order and prosody are not at play. Many languages have dedicated focus particles, however PROMINENT case-marking strategy is fused with case-marking. A similar focal case-marking strategy is illustrated for Chickasaw in (10):[5]

(10) Chickasaw (*cic* — MUSKOGEAN)
[Gordon 2007, cited in Büring 2009, 22, ex. (44)]

a. *hat:ak-at koni(ã) pisa*
man-SBJ skunk sees
'The man sees the skunk.'

b. ***hat:ak-akot*** *koni(ã) pisa*
man-SBJ.FOC skunk sees
'The MAN sees the skunk.'

c. *hat:ak-at* ***koni-akõ*** *pisa*
man-SBJ skunk-OBJ.FOC sees
'The man sees the SKUNK.'

In languages displaying the focal case-marking strategy, the realization of the focus is restricted to nominal hosts. Following the terminology proposed in Aannestad (2021), focal case-marking is a type of ARGUMENT-ASSOCIATED FOCUS REALIZATION STRATEGY; as compared to SYMMETRICAL FOCUS SYSTEMS, which permit focus to be realized on hosts of various syntactic categories. The former type is exemplified for Rendille in (11), where separate strategies mark argument-level and predicate-level focus. The latter type is shown for Gùrùntùm in (12), where a single marker arises on nominal and verbal hosts, inducing different focal interpretations respectively. In neither language is case-marking realized to encode focus:

(11) Rendille (*rel* — CUSHITIC)
[Oomen 1978, 48-49, as cited in Aannestad 2021, 29]

a. ***ínam=é*** *y-imi*
boy=FOC 3SG-come:PST
'The BOY came.'

b. *ínam* ***á-y-imi***
boy PRED.FOC-3SG-come:PST
'The boy CAME.'

[5]Compared to the Siona PROMINENT-marking strategy, the Chickasaw pattern in (10) seems to be more transparent. In the latter language, focal case-markers are more complex than plain case-markers on morphological grounds, such that focal case-markers contain the sequence *-ko-*. This morpheme does not arise independently from case-marking. Setting aside the degree of opacity, Siona and Chickasaw are taken to display compatible focal case-marking strategies.

(12) Gùrùntùm-Mbaaru (*grd* — WESTERN-CHADIC)[6]
[Hartmann and Zimmermann 2009, in Assmann et al. 2023, 1352-1353]

a. ***á*** ***fúrmáyò*** *bà* *wúm* *y-imi*
FOC Fulani PROG chew colanut
'The FuLAni is chewing colanut.'

b. *tí* *bà* ***ròmb-á*** *g^{w}éì*
3SG PROG gather-FOC seed
'He is gathering the SEED(S).'

Upon further scrutiny, ARGUMENT-ASSOCIATED FOCUS REALIZATION STRATEGIES, as exemplified in (11), are commonplace in the Tukanoan language family.[7] For instance, the Desano [*des*] and Barasano [*bsn*] sentences shown in (13) and (14) illustrate uses of exclusively-nominal focus-encoding morphology:

(13) Desano (*des* — EASTERN-TUKANOAN)[8]
[Silva 2012, 162, ex. (96); 164, ex. (103)]

a. ***gʉa-pʉ*** ***weka-pʉ*** *~bobe-~da-ka*
1PL.EXCL-CONTR above-CONTR work-EXRT-EVID:REAS
'WE (EXCL) work up THERE (upstream).'

b. ***si-bu-pʉ-de*** *koa-du* *api-a-bʉ*
DEM.DIST-CL:BASKET-CONTR-N.SBJ gourd-CL:CONC leave-COMP-N3.PST
yʉʉ
1SG
'I put the gourd in THAT one (another basket).'

(14) Barasano (*bsn* — EASTERN-TUKANOAN)
[Jones and Jones 1991, 175, ex. (563)]
to-re *yá-ka-hʉ* *wʉhʉ* ***yoe-se-bã̀***
there-REF be-RMT.PST-3 reed(species) long-NMLZ-FOC
'There was *wʉhʉ* reed, the LONG stuff (not another material).'

Many Eastern Tukanoan languages display the ARGUMENT-ASSOCIATED FOCUS REALIZATION type; however, the FOCUS PARTICLE is not fused here with case-marking. This is clearly demonstrated in (13b) for Desano, where the case-marker *-re* and the focus-marker *-pʉ* are produced in sequence. Conversely, Siona exhibits a pattern where the presence of (certain) case-markers induces an information focus or contrastive reading in the sentence. Moving forward this is referred to as the focal case-marking strategy.

To summarize: Siona makes use of a case-marking strategy, which is restricted to arising on nominal hosts, whose distribution aligns with certain focus contexts.

[6]The authors indicate that the focus-marking found in (12b) is also used to achieve narrow verbal focus — i.e., GAthering vs. another activity concerning seeds, and broad predicative focus. A discussion of other scopal ambiguities arising with focus-encoding strategies is undertaken in Section 6.3.

[7]A similar nominal-specific focus clitic is described for A'ingae [*con*] (Fischer and Hengeveld 2023, 92, ex. (93)). This languages is in intimate contact with Siona+ languages — see Chapter 4.

[8](13a) illustrates an instance where the focus-marker, *-pʉ* arises on two nominals. In Siona, there is resistance to PROMINENT-marking on multiple core arguments of a clause; although, admittedly, patterns of multiple foci, including argumental and spatial nouns remain to be checked. I return to this point briefly in the following section of this chapter.

Concretely, in the relevant contexts, focus is expressed via the selection of the PROMINENT case-marker, as determined by the grammatical relation and the animacy class of the nominal host. On this basis, a leveled paradigm of PLAIN and PROMINENT case-markers emerges, as shown in Table 6.1 (adapted from Table 3.9 in Chapter 3):

	INANIMATE		ANIMATE ($\wedge$ $^{\uparrow}$INAN)	
	PLAIN	PROMINENT	PLAIN	PROMINENT
Argumental DCM:				
SUBJECT (S)	-∅	*-bi*	-∅	*-bi*
DIRECT OBJECT (P)	-∅	*-re*	*-re*	*-ni*
INDIRECT OBJECT, e.g., (R)	*-re/-na*		*-re*	*-ni*
Spatial DCM:	INANIMATE		$^{\uparrow}$INANIMATE	
LOCATION (L)	-∅	*-re*	*-re*	*-ni*
GOAL (G)	-∅	*-na*	*-na*	

Table 6.1: Siona DCM paradigm, PLAIN- vs. PROMINENT-marker mapping

An immediate consequence of the analysis proposed in Table 6.1 pertains to what is labeled the SPLIT *-re* ANALYSIS in Chapter 3. In short, the case-marker *-re* is not a PROMINENT-marker per se; rather its PROMINENT-marking status depends upon the noun class of the argument at hand. The two sentences provided in (15) are a case in point: The PLAIN *-re* in (15a) is required with the animate P for grammatical reasons, whereas PROMINENT *-re* on the inanimate P in (15b) evokes the emphatic contrast reading, as demonstrated by the context suggested by the speaker:

(15) a. *yë' ñicuë ai yëquëa'bi bi̱'a̱re*
ji'i̱ ñihkw-i̱ ai ji̱i̱-ki̱-a-'bi **bĩ'ã *(-re)**
1SG grandparent-CL:M INTENS like-CL:M-COP-3S.M bird-N.SBJ
'My grandfather loves birds.' [VOL: 20240918ejabi001.004]

b. *yë'ë ocore u̱cusi'i cayë*
ji'i̱ **ohko #(-re)** ũhku-si'-i kaa-ji̱
1SG water-N.SBJ drink-FUT-N3S say-N3S.PRS.ASS
'I want to drink WAter.' ⇒ {I want to drink coffee, to drink tea,... }
[**Context suggested by speaker:** A boy comes home to find coffee and tea on the table.] [VOL/SUG: 20230623ejabi001.024-025]

Ultimately, the differences between morphological focus languages with dedicated (nominal) focus-markers on the one hand, and with fusional, focal case-marking strategies is a matter of morphology. The underlying syntax for the two types of languages is presumed to be identical. I propose the outline for a syntactic analysis of (certain) case-markers as focus particles in Section 6.5. Inasmuch as this discussion points to a plausible formal analysis of Siona information structure-driven DCM as focus-marking, the following sections scrutinize the distribution of PROMINENT-marking to ascertain the degree of focus-sensitivity exhibited by this case-marking category. This evidence is subsequently evaluated to determine if Siona DCM is focus-encoding or whether it is simply focus-related (Matić and Wedgwood 2013).

6.3 Focal case-marking: The size question

This section is concerned with the semantic notion of *focus size*, namely the amount of structure which may be focalized by the presence of the PROMINENT case-marker. The previous section establishes that this strategy is limited to being expressed on nominal hosts; however, this section shows that this strategy is not limited to a narrow, argument focus interpretation. Firstly, Section 6.3.1 explores the availability of narrow and broad interpretations for the PROMINENT-marking strategy. Secondly, Section 6.3.2 explores how focal interpretations are established within complex nominal constituents. Two types of complex complements are considered in particular: (i) nominals with relative clause modifiers, and (ii) nominalized sentential complements to desire and modal predicates.

6.3.1 Broad and narrow interpretations for focal case-marking

Patterns where a particular form of focus-marking is recruited to focalize constituents of various sizes, are commonplace in the typological record (e.g., Arregi 2016; Samek-Lodovici 2016; Assmann et al. 2023). For instance, a minimal pair is reported for Buli (16), where each sentence accepts two focal interpretations accordingly: In the case of (16a), where the focus particle *alè* occurs, either narrow SUBJECT focus or sentential focus is available. On the other hand, in (16b), where another particle, *ká*, arises in the immediately post-verbal position, either narrow OBJECT focus or broad predicate focus is available. The available interpretations are ascertained based upon the questions which they felicitously answer (i.e., via Q-A heuristics):

(16) Buli (*bwu* — NIGER-CONGO)
[Assmann et al. 2023, 1356-1357, exx. (6)-(7)]

a. i. [Narrow SUBJECT focus: Who ate a mango?]
(**ká**) Àtìm **alè** dè mángó
FOC A FOC ate mango
'Atim ate a mango.'

ii. [Sentential focus: Why are you angry?]
(**ká**) Àtìm **alè** dè mángó
FOC A FOC ate mango
'Atim ate a mango.'

b. i. [Narrow OBJECT focus: What did Atim eat?]
wá dè **ká** mángó
3SG ate FOC mango
'He ate a mango.'

ii. [Broad predicate focus: What did Atim do?]
wá dè **ká** mángó
3SG ate FOC mango
'He ate a mango.'

The pattern where one form corresponds to multiple interpretations is often termed a *focus syncretism*, as shown for the Buli sentences (16a) and (16b) above; and in the

English examples in (4). On the one hand, the pattern of SUBJECT-sentential focus syncretism shown for Buli in (16a) does not arise in Siona grammar. However, on the other hand, a compatible pattern of narrow OBJECT-broad predicate focus syncretism is attested for the Siona focal case-marking strategy.

On the basis of the Q-A evidence shown in (17), the same answer is provided for the narrow-scope question in (17a) and for the broad-scope question in (17b), corresponding to different sizes of focal material in each answer:

(17) a. i. NARROW-SCOPE QUESTION: *quere nede'huaquë'ne ?*
ke-e-re nee + de'wa-ki̵-'ne
WH-CL:GEN-N.SBJ make + repair-2/3S.M.PRS.N.ASS-Q
'What are you (M) fixing?'

ii. ANSWER: *yë'ë ja̲ë̲rëre nede'huayë*
ji̵'i̵ **hãĩ̵-ri̵ #(-re)** nee + de'wa-ji̵
1SG hammock-CL:MAIZE-N.SBJ make + repair-N3S.PRS.ASS
'I am fixing a/the HAmmock.'

b. i. BROAD-SCOPE QUESTION: *quere yo'quë'ne ?*
ke-e-re jo'-ki̵-'ne
WH-CL:GEN-N.SBJ do-2/3S.M.PRS.N.ASS-Q
'What are you (M) doing?'

ii. ANSWER: *yë'ë ja̲ë̲rëre nede'huayë*
ji̵'i̵ **hãĩ̵-ri̵ #(-re)** nee + de'wa-ji̵
1SG hammock-CL:MAIZE-N.SBJ make + repair-N3S.PRS.ASS
'I am fixing a/the HAmmock.' [VOL/SUG: 20230617elupa001.017a-d]

Two critical insights emerge based upon the example presented in (17): Firstly, the focalizing effect of PROMINENT case-marker in Q-A contexts can scope beyond the nominal constituent which hosts it. This fact highlights a fundamental disconnect between the morphological realization of focal case-marking and the size of the interpretation that it achieves. Secondly, obligatory PROMINENT *-re* marking is observed in the answers in (17), as it is on the WH-item arising in each question, *kee-**re*** (what). These case matching effects are a component of Q-A congruence effects noted in the information question-answer construction.

As regards the second insight, this is not an instance of straightforward, surface case-matching. Instead, this may be re-framed as a PROMINENT case-matching requirement, such that the appropriate PROMINENT-marker is forced to arise on the focalized constituent and the WH-item in the question, even where surface mismatches obtain. Such discrepancies arise, for instance, where an animacy class mismatch arises across the focalized non-SUBJECT constituents. For instance, (18) adapts the answers to both the narrow- and broad-scoping questions in (17). In (18), the noun is fitted with a post-nominal relative clause modification, inducing the promotion procedure (detailed in Section 3.2.2). Per the paradigm in Table 6.1, the promoted noun displays the animate case-marking pattern, and PROMINENT *-ni* marking is required accordingly:

(18) a. [NARROW-SCOPE QUESTION: What are you (M) fixing?]
ANSWER′: *yë'ë ja̲ë̲rë yetesiconi nede'huayë*
ji̵'i̵ ↑[**hãĩ̵-ri̵** **jehte-sih-ko *(-ni)/(#-re)**] nee
1SG hammock-CL:MAIZE tear-PERF-CL:F-N.SBJ(2) make
+ de'wa-ji̵
+ repair-N3S.PRS.ASS
'I am fixing a/the frayed HAmmock.'

b. [BROAD-SCOPE QUESTION: What are you (M) doing?]
ANSWER′: *yë'ë ja̲ë̲rë yetesiconi nede'huayë*
ji'i̵ ↑[**hãĩ̵-ri̵** **jehte-sih-ko *(-ni)/(#-re)**] nee
1SG hammock-CL:MAIZE tear-PERF-CL:F-N.SBJ(2) make
+ de'wa-ji̵
+ repair-N3S.PRS.ASS
'I am fixing a/the frayed HAmmock.' [VOL/SUG: 20230617elupa001.017e-f]

In a similar vein, the broad-scoping question in (19a) accepts various answers, which contain nominals belonging to different grammatical relations. Where the answer contains an inanimate OBJECT-argument, as in (19b), the anticipated PROMINENT *-re* marker is required. However, PROMINENT *-na* marking is forced on the spatial GOAL-argument, as associated with the directed motion predicate *sai-* (go) in (19c):

(19) a. QUESTION: *quere ñamina'a̲ yo'ja̲'quëa'ne ?*
ke-e-re ñahmina'ã jo'-hã'-ki̵-a-'ne
WH-CL:GEN-N.SBJ tomorrow do-PRP-CL:M-COP-Q
'What are you (M) doing tomorrow?'

b. ANSWER: *go̲nore neja̲'quëa'ë*
gõno #(-re) nee-hã'-ki̵-a-'i̵
chicha-N.SBJ make-PRP-CL:M-COP-N3S
'(I) am going to make CHIcha.'

c. ANSWER′: *airona saija̲'quëa'ë*
ai-ro #(-na) sai-hã'-ki̵-a-'i̵
big-CL:LOC-GOAL go-PRP-CL:M-COP-N3S
'(I) am going to go to the FOrest (i.e., to go hunting).'
[VOL: 20241104ejabi001a-d]

The examples in (19) demonstrate that the appropriate PROMINENT-marking is forced to arise on the nominal within the focalized predicate in these instances. On this basis, the fact that PROMINENT-marking is rejected on the SUBJECT-argument in contexts of broad-scoping Q-A sequences, as shown in (20), is unsurprising. In this case, the nominal is part of the presupposition and not contained within the focalized constituent:

(20) a. QUESTION: *quere yo'co ?*
ke-e-re jo'-ko
WH-CL:GEN-N.SBJ do-2/3S.F.PRS.N.ASS
'What are you (F) doing?'

b. ANSWER: # *yë'bi ca̲i̲ñë yure*
ji̵'i̵ (# -bi) kãĩ̃-ñi̵ jure
1SG-SBJ sleep-N3S.PRS.ASS now
Intended: 'I̲ am sleeping right now.'
[**Speaker comment:** Using *-bi* is odd in this case. Something is missing and it seems that the speaker has more to say.]
[VOL/SUG: 20230622ejabi022a-c]

The examples above illustrate how PROMINENT-marking arises on a nominal host within the focalized constituent; corresponding to a narrow, argument-focus reading, or a broad, predicate-focus reading accordingly. The data presented here are ascertained on the basis of Q-A diagnostics. These findings are corroborated by naturalistic instances, such as (21) lifted from *The demon kwẽẽwahti* tale in the narrative sample analyzed in Chapter 5. In this excerpt, the P-token in (21b) bears PROMINENT *-re*, which signals a contrast between the *activity* performed by the wife character and that performed by the husband — i.e., an instance of broad focus:

(21) a. *nocabo jëosaiye ba̲'ë̲ña i̲ i̲o̲ ë̲jë̲maca ...*
nohka-bo hi̵o + sai-je
banana-CL:ENCLOS clear + go-CL:GEN
ba-∼'-i̵-ña [ĩ̃-i̵ ĩ̃-o
be-RMT.PST-2/3S.M.PST.N.ASS-REP 3PRO-CL:M 3PRO-CL:F
i̵hĩ̃-mahka]$_i$
husband-DIM.SG
'Her husband$_i$ went to tend to the banana plantation ...'

b. *... i̲o̲macacato a̲o̲re sëoco ñucoña*
ĩ̃-o-mahka=kato$_j$ **ãõ-re** si̵o-ko
3PRO-CL:F-DIM.SG=TOP cassava-N.SBJ fry-S.F.DEP
ñuh-ko-ña
be_seated-2/3S.F.PST.N.ASS-REP
'... (∼as for her,) she$_j$ was sitting, frying caSSAva.'
[NAT*: 20151112oespa001.164-165]

The empirical evidence considered in this section demonstrates that the Siona PROMINENT-marking strategy displays the typologically common pattern of narrow-broad focus syncretism in overt Q-A pairs (cf. Selkirk 1984; Lambrecht 1994; Rochemont 1998, 2013; Arregi 2016; Assmann et al. 2023 and references therein). PROMINENT case-marking arises on a nominal host which resides within the focalized constituent, such that the scope of the focus interpretation is underspecified. A series of related questions arise relating to the scopal properties of focus within complex noun phrases (e.g., as in the English distinction, 'the red CAR' vs. 'the RED car'). The following subsection explores focal case-marking within complex nominals.

6.3.2 Focal case-marking within complex nominals

The discussion thus far has considered (PROMINENT) case-marking on simplex nominals, which bear a single case-marker on their right edge. This subsection explores the expression and interpretation of PROMINENT-markers within complex nominals. In particular, two types of complex nominals are considered: (i) nouns with relative

clause modifiers, and (ii) nominalized clauses, with clause-internal arguments. Each of these configurations contain multiple nominal (sub-)constituents which may host case-marking morphology.

To begin with a consideration of relative clauses, Section 3.2.2 details how Siona displays parallel strategies for relative clause modification: (i) the pre-nominal modification strategy, and (ii) the post-nominal modification strategy. Speakers do not report any relevant interpretative differences between these alternatives. In the latter case, the modifier surfaces as a nominalization, which has implications for the process of promotion on inanimate singular nouns. (22) provides instances of pre- and post-nominal paraphrases for (18), repeated from the preceding section; illustrating obligatory PROMINENT-marking for each alternative:

(22) a. [QUESTION: What are you (M) fixing?]
ANSWER: *yë'ë yetesi jaë̲rëre nede'huayë*
jɨ'ɨ **[jehte-si hãĩ̲-rɨ #(-re)]** nee + de'wa-jɨ
1SG tear-PERF hammock-CL:MAIZE-N.SBJ make + repair-N3S.PRS.ASS
'I am fixing a/the frayed HAmmock.'

b. [QUESTION: What are you (M) fixing?]
ANSWER: *yë'ë jaë̲rë yetesiconi nede'huayë*
jɨ'ɨ ↑**[hãĩ̲-rɨ jehte-sih-ko *(-ni)/(#-re)]** nee
1SG hammock-CL:MAIZE tear-PERF-CL:F-N.SBJ2 make
+ de'wa-jɨ
+ repair-N3S.PRS.ASS
'I am fixing a/the frayed HAmmock.'
[VOL/SUG: 20230617elupa001.017d-g]

The nominal in (22) is modified by a straightforward relative clause, formed via a property-denoting predicate, *jehte-* (be torn). A single case-marker surfaces at the right edge in both instances. In more complex cases, the relative clause modifier may itself contain nominal constituents which are viable targets for PROMINENT case-marking. This is the case in (23), which displays the answer segment to a SUBJECT-oriented question where the focalized SUBJECT has a relative clause modifier. Two acceptable case-marking configurations are noted: In (23a), the canonical configuration arises where case-marking surfaces at the right edge of the noun phrase. In (23b), the embedded SUBJECT receives PROMINENT-marking, and the right edge is left unmarked:

(23) a. [QUESTION: What just went by?/What type of creature just went by?]
ANSWER: *yë' a'yë ñasi bi̲'a̲bi ja̲ saja'i*
[jɨ'ɨ a'j-ɨ ñaa-si **bĩ'ã̲ #(-bi)**] hã sa-ha'i
1SG older_sibling-CL:M see-COMP bird-SBJ just go-3S.M.PST.ASS
'The BIRD that my older brother saw just flew by.'

b. [QUESTION: What just went by?/What bird just went by?]
ANSWER′: *yë' a'yëbi ñasi bi̲'a̲ ja̲ saja'i*
[jɨ'ɨ a'j-ɨ #(-bi) ñaa-si bĩ'ã (*-bi)] hã sa-ha'i
1SG older_sibling-CL:M-SBJ see-COMP bird just go-3S.M.PST.ASS
'The bird that my older BROther saw just flew by.'
[**Context suggested by speaker:** There are several birds that everybody is taking about. Specifically, the one that your older brother and not, e.g., younger brother or father, saw flies by.]
[VOL/SUG: 20241104ejabi001.003-004]

The answers in (23) assert the same content, although they evoke different sets of alternatives, which are appropriate in different contexts. In fact, the answer suggested in (23a) accepts two possible readings. This answers the SUBJECT-oriented question directly. However, in another context, this PROMINENT-marking on the head of the noun phrase, *bĩ'ã-**bi*** (bird), contrasts this species of animal with others (e.g., bats, wasps). Such instances where a single focus morpheme corresponds to focalization of the entire constituent or of only the marked sub-constituent are considered a sub-type of DISJUNCTIVE FOCUS SYNCRETISM by Assmann et al. (2023, §4.2.1).

By way of contrast, the answer in (23a) focalizes the embedded SUBJECT-argument in the complex nominal constituent, presupposing that there are multiple salient birds in the context at hand.[9] The configuration in (23b) contributes a new finding to the description of Siona focal case-marking: The obligatory PROMINENT-marker in answer turns need not arise at the right edge of the noun phrase. A single instance of case-marking within the nominal constituent satisfies the Q-A congruence requirement, with concomitant effects on the interpretation.

A further refinement to the observation that a single case-marker satisfies the Q-A congruence requirement is in order: The requirement that PROMINENT-marking arises on one, but not all, eligible nominal hosts in the focalized constituent is separate from the overt case-marking requirement for animate non-SUBJECT arguments. (24) presents a similar instance to (23), which answers a question narrowly targeting the LOCATION-argument. The LOCATION in (24) contains a post-nominal relative clause, which triggers promotion accordingly. Two possible case-marking patterns are noted: Either a single PROMINENT *-ni* marker surfaces at the right edge of the nominal; or, the embedded SUBJECT receives PROMINENT *-bi* marking and the right edge of the promoted nominal bears PLAIN *-re* marking:

(24) a. [QUESTION: Where do you (PL) live?]
ANSWER: *yëquë ba'icua'ia'ë huë'e ñicuëba̲i̲ nesiconi*
jɨhkɨ ba'i-ko-wa'i-a-'ɨ ↑[wɨ'e ñihk^{w}-ɨ-bãĩ
1PL.EXCL live-CL:F-PL.AN-COP-N3S house grandparent-CL:M-PL.AN
nee-sih-ko *(-ni)/(#-re)]
make-PERF-CL:F-N.SBJ2
'We (EXCL) live in the HOUSE that (our) grandparents built.'

[9]The sentence in (23b) may serve as an answer to the same question as in (23a). It is generally acceptable to answer a question by addressing a more specific sub-question (e.g., Roberts 2012) — for this reason (23b) is shown to answer both the super- and the sub-question accordingly. This matches speaker judgments.

b. [QUESTION: Where do you (PL) live?/Which house do you (PL) live in?]
yëquë ba'icua'ia'ë huë'e ñicuëbaibi nesicore
jɨhkɨ ba'i-ko-wa'i-a-'ɨ ↑[wɨ'e **ñihk^{w}-ɨ-bãĩ #(-bi)**
1PL.EXCL live-CL:F-PL.AN-COP-N3S house grandparent-CL:M-PL.AN-SBJ
nee-sih-ko *(-re)]
make-PERF-CL:F-N.SBJ
'We (EXCL) live in the house that (our) GRANDparents built.'
[**Context suggested by speaker:** In the community there are many groups of houses, built over several generations. You live in a house built by your grandparents and not by, say, your parents and their siblings.]
[VOL/SUG: 20241202ejabi001.002a-c]

The interpretations noted in (24) for the LOCATION-argument align neatly with those noted for the SUBJECT-argument in (23a). Where the embedded SUBJECT is PROMINENT-marked, PROMINENT-marking is not accepted at the right edge of the complex nominal. However, due to the case-marking requirement for promoted non-SUBJECTs, the PLAIN *-re* marker is obligatory in that position. In this way PLAIN and PROMINENT case-markers execute different roles in the language, a testament to the multi-dimensionality of Siona DCM. Only the latter class of case-markers encode focus-related information.

A final pair of sentences containing a complex nominal with a nominalized modifier is demonstrated in (25). These examples complement (24), by separating the role of PLAIN *-re* marking and PLAIN *-ni* on animate non-SUBJECTs. These examples are not Q-A pairs. The assertive sentences below concern a discontinuous animate OBJECT-argument, where the nominalized relative clause modifier itself contains an embedded animate OBJECT. Each sub-constituent of the discontinuous noun phrase must bear PLAIN *-re* marking in order to produce a grammatical sentence — mirroring so-called 'co-case-marking' patterns described for Quechuan languages (e.g., Lefebvre and Muysken 1988; Muysken 1989; Hastings 2003) — whereas a single PROMINENT *-ni* marker affects the scope of the contrastive interpretation in the sentence:[10]

(25) a. *ye'yaquëbi te' joyaëre isibi betere be'tequëni*
je'ja-kɨ-bi [te'-e hoja-ɨ *(-re)] ĩhsi-bi
teach-CL:M-SBJ one-CL:GEN dog-CL:M-N.SBJ give-3S.M.PST.ASS
[behte *(-re) be'te-kɨ *(-re)]
duck-N.SBJ chase-CL:M-N.SBJ
'The teacher (M) gave (me) a dog (M) that chases ducks.'

[10]Although the speaker naturally produces *-bi* marking on the SUBJECT in (25), this is not a focal instances of PROMINENT-marking, as indicated by the suggested translation. Rather, overt case-marking is often utilized on deverbal nominalizations like these for the sake of communicative efficiency — i.e., conceived as a component of the disambiguating function of Siona DCM to be studied in greater detail in future research.

b. *ye'yaquëbi te' joyaëre i̱sibi beteni be'tequëre*
je'ja-ki̵-bi [te'-e hoja-i̵ *(-re)] ĩhsi-bi
teach-CL:M-SBJ one-CL:GEN dog-CL:M-N.SBJ give-3S.M.PST.ASS
[**behte *(-ni)/(#-re)** be'te-ki̵ *(-re)]
duck-N.SBJ2 chase-CL:M-N.SBJ(2)
'The teacher (M) gave (me) a dog (M) that chases DUCKS.'
[**Context suggested by speaker:** The teacher has several dogs and he lends one that chases ducks, for hunting, but not the other animals in the forest.] [VOL/SUG: 20241119ejabi002.006-008]

In (25a), PLAIN *-re* marking arises on both sub-constituents for the matrix OBJECT, and on the embedded OBJECT, such that no focal-marking is expressed within this complex noun phrase. Correspondingly, none of these (sub-)constituents are focalized for the sake of interpretation. Conversely, in (25b), PROMINENT-marking arises on the embedded OBJECT-argument, and this argument is focalized accordingly, as expected. As is the case in (24b), the presence of PROMINENT-marking on the embedded constituent does not affect the PLAIN-marking patterns which satisfy the case-marking requirement for animate non-SUBJECTs independently.

The above examples demonstrate how focal case-marking arises in complex nominals, formed with relative clause-modifiers. The PROMINENT-marking strategy arises equally within independent and dependent clauses on the basis of the facts considered thus far. Similar facts emerge where the clausal complements of certain predicates are taken into account, which are required to be event nominalizations.[11] Such event nominalizations, contrary to participant-denoting nominalizations, never undergo promotion, since they are derived via the generic classifier, *-je* (or nasal *-ñe*), as discussed in Bruil 2019. This discussion considers two classes of constructions which take nominalized non-finite complements: (i) desire constructions (e.g., *ji̵i̵-* (want/like); *goe-* (not want/like)), and (ii) impersonal modal constructions.[12]

To begin, (26) shows the possible PROMINENT-marking configuration available to the nominalized complement clause of the desire construction, *ji̵i̵-* (want/like).[13] These assertive sentences do not display obligatory PROMINENT-marking, such that where no such case-marking is produced, as in (26a), no contrastive (focus) effects arise. Nonetheless, where PROMINENT-marking is used, either on the event nominalization itself, *ñaañe* (see/read), as in (26b); or on the the embedded OBJECT, *bãĩkohka* (Siona language), as in (26c), a focus interpretation emerges narrowly on its nominal host. Speakers reject the sentence where PROMINENT-marking co-occurs on both the event nominalization and the embedded argument:

[11] Siona exhibits several complementation strategies. Some predicate types take finite complements (e.g., *report*, *perception* predicates), whereas others select for non-finite complements.

[12] Although the negative copula does take the main verb in generic nominalized form, speakers reject the presence of *-re* on the main predicate in any context: e.g., *bããñi̵ ba'ije(*-re) iño* (intended: I don't LIVE here). It remains to be explored whether further differences exist for these functions of the event nominalization.

[13] This is one common strategy to express 'to want/to like to X' in Siona. The transitive *desire construction* in (26) is used alongside another impersonal strategy with the desiderative applicative *-i̵a-* (glossed DES), and an idiomatic construction which is directly translated as 'X says X will do Y'. More specifically the latter strategy recruits the verb *kaa-* (say) with a co-referential S in the subordinate clause with the verb in the future tense. Neither of these strategies take nominalized complements of the pertinent type.

(26) a. *iye mo'se yë' ye'yaco yëco bãĩcoca ñañe*
i-je mo'se jɨ'ɨ je'ja-ko jɨɨ-ko [bãĩ
DEM.PROX-CL:GEN day 1SG teach-CL:F want-3S.F.PRS.ASS person.COL
+ kohka ñaa-ñe]
+ language see-CL:GEN
'Today my teacher (F) wants (us) to read Siona.'

b. *iye mo'se yë' ye'yaco yëco bãĩcoca ñañere*
i-je mo'se jɨ'ɨ je'ja-ko jɨɨ-ko [bãĩ
DEM.PROX-CL:GEN day 1SG teach-CL:F want-3S.F.PRS.ASS person.COL
+ kohka **ñaa-ñe #(-re)**]
+ language see-CL:GEN-N.SBJ
'Today my teacher (F) wants (us) to READ Siona.'
[**Context suggested by speaker:** In class the teacher wants to do reading excercises instead of other tasks — e.g., writing, singing.]

c. *iye mo'se yë' ye'yaco yëco bãĩcocare ñañe*
i-je mo'se jɨ'ɨ je'ja-ko jɨɨ-ko [**bãĩ**
DEM.PROX-CL:GEN day 1SG teach-CL:F want-3S.F.PRS.ASS person.COL
+ kohka #(-re) ñaa-ñe (*-re)]
+ language-N.SBJ see-CL:GEN
'Today my teacher (F) wants (us) to read SIOna.'
[**Context suggested by speaker:** At school there are materials written in several languages, but the teacher chooses the Siona reading.]
[VOL/SUG: 20230614erepa001.005a-d]

The examples in (26) demonstrate that Siona does permit for narrow verbal focus interpretation in constructions where the verbal element takes the form of an event nominalization. The derived nominal is a viable host for case-marking, such that the presence of PROMINENT-marking focalizes the event nominal itself — i.e., as shown in (26b). In order to focalize the embedded OBJECT, this argument must be PROMINENT-marked. On this basis, within nominalized complement clauses, the DISJUNCTIVE FOCUS SYNCRETISM effects observed with complex nominals with relative clause modifiers in (23) are blocked. The implications of this discrepancy for the syntax of relative clauses as compared with event nominalizations as in (26) remain to be fully unpacked in future research.

Event nominalizations are also selected in impersonal modal constructions (recently discussed by Jeretič (2021)). Two modal constructions are of interest to the present discussion: (i) the strong modal *ba'i-*, and (ii) the weak modal *de'o-*.[14] The predicate arises with default agreement (third-person masculine, singular), and the Experiencer-argument displays the dominant animate non-SUBJECT case-marking alternation: i.e., PLAIN *-re* vs. PROMINENT *-ni*. As anticipated, obligatory PROMINENT *-ni* marking arises on both the WH-item and the corresponding focalized Experiencer in the Q-A pair in (27):

[14] I neglect the negative counterpart *beo-* for this discussion, although it seems to behave in the same relevant ways noted for the affirmative modals.

(27) a. QUESTION: *queini hua'i moñe de'oquë?*
ke-ɨ *(-ni) wa'i moo-ñe de'o-kɨ
WH-CL:M-N.SBJ2 meat fish-CL:GEN be_good-2/3S.M.PRS.N.ASS
'Who is allowed to go fishing? (lit. To whom is fishing good?)'

b. ANSWER: *jaëni hua'i moñe de'oji*
hã-ɨ *(-ni)/(#-re) wa'i moo-ñe de'o-hi
DEM.MED-CL:M-N.SBJ2 meat fish-CL:GEN be_good-3S.M.PRS.ASS
'He is allowed to go fishing (lit. Fishing is good to him).'
[VOL: 20230614ejabi001.013a-b]

In the same elicitation session, the speaker is asked a follow-up to (27). The speaker is asked to produce a similar instance, except where the activity of fishing is contrasted with other activities. The resulting sentence is reported in (28), along with the conjured context, where PROMINENT *-re* marks the event nominalization, ***mooñe-re*** (to fish), producing a similar effect to that noted in (26b):

(28) *. . . më'recato hua'i moñere de'oji*
mɨ'ɨ̃-re=kato [wa'i **moo-ñe #(-re)**] de'o-hi
2SG-N.SBJ=TOP meat fish-CL:GEN-N.SBJ be_good-3S.M.PRS.ASS
'As for you, you can go FISHing.'
[**Context suggested by speaker:** At an event where several other activities are planned, you can go fishing, but not participate in other planned activities.] ⇒ {You can go hunting, You can go canoeing,. . . }
[VOL: 20230614ejabi001.014]

As demonstrated in (26) for complement clauses in desire constructions, the placement of PROMINENT-marking within an event nominalization in modal constructions has implications for the focal interpretation. (29) reports a pair of instances elicited via the CONTEXT CONJURING technique where PROMINENT *-re* is placed on different constituents within the dependent clause:

(29) [**Context suggested by speaker:** A mother is talking to her child about learning life skills. . .]

a. *duruhuëare isiye ye'yeye ba'iji*
[**duhru-wɨ-ã #(-re)** ĩhsi-je je'je-je]
basket-CL:CONT-PL.INAN-N.SBJ give-CL:GEN learn-CL:GEN
ba'i-hi
be-3S.M.PRS.ASS
'(You) must learn to sell BASKETS.' ⇒ {. . . sell bags, . . . sell clothes}

b. *duruhuëa isiyere ye'yeye ba'iji*
[duhru-wɨ-ã **ĩhsi-je #(-re)** je'je-je]
basket-CL:CONT-PL.INAN give-CL:GEN-N.SBJ learn-CL:GEN
ba'i-hi
be-3S.M.PRS.ASS
'(You) must learn to SELL baskets.'
⇒ {. . . weave baskets, . . . sort baskets} [VOL: 20230617elupa001.007a-c]

Together examples (26) through (29) demonstrate that narrow verb focus is attainable in Siona grammar, yet only under conditions where they arise as event nominalizations. Unlike the patterns of scopal underspecification noted in main clauses in Section

6.3.1, the scope of focalization is fixed in the case of event nominalizations and their internal nominal constituents. I am unaware of a grammaticalized strategy for narrow verb focus in the language outside of the embedded contexts considered here.

In order to round off the description of the distribution of PROMINENT case-marking on embedded event nominalizations, a final example is provided in (30). This example considers the Q-A heuristic, as applied to the nominalized complement clause to the desire construction, *jɨ̈ɨ-* (want/like). Whereas the WH-item in the question in (30a) demonstrates obligatory PROMINENT-marking, as anticipated, *kee-**re*** (what); the focalized event nominalization in the answer in (30b) resists PROMINENT-marking, as does the embedded OBJECT. This finding merits further investigation:

(30) a. QUESTION: *quere yo'ye yëquë ?*
ke-e-re jo'-je jɨ̈ɨ-kɨ
WH-CL:GEN-N.SBJ do-CL:GEN want-2/3S.M.PRS.N.ASS
'What do you (M) want to do?'

b. ANSWER: *baicoca ye'yeye yëyë*
[**bãĩ̃ + kohka je'je-je**] jɨ̈ɨ-jɨ
person.COL + language learn-CL:GEN-N.SBJ want-N3S.PRS.ASS
'(I) want to learn/practice SIONA.' [VOL: 20241104ejabi001.016a-b]

The discussion in this section demonstrates that the *size* question regarding focal effects for Siona DCM is not as straightforward as it might appear. In addition to a handful of focus syncretisms, the choice of host for PROMINENT-marking on a sub-constituent of complex nominals has interpretative consequences. The distribution of PROMINENT-marking is largely driven by the expression of information focus in interrogative contexts across all DCM relations, and by the encoding of contrast (or contrastive focus) in non-interrogative contexts. The majority of elicited data considered thus far concerns information focus uses of the PROMINENT case-marking strategy. The following section delves into the distribution of PROMINENT-marking across a range of (non-)interrogative focus contexts in order to better characterize its semantico-pragmatic function.

6.4 Distribution check: Siona DCM and focus

Whereas the previous section analyzes the scopal properties of PROMINENT-marking, the present section addresses the range of focal constructions where this strategy is attested. It is well-established in the literature that the precise distribution of focus-encoding strategies varies on language-by-language basis (e.g., Zimmermann and Onea 2011; Matić and Wedgwood 2013; van der Wal 2011, van der Wal 2016, 2022; Kratzer and Selkirk 2020). Ultimately this discussion contributes to the determination of whether the Siona PROMINENT-marking strategy is a bona fide focus-*encoding* strategy, or whether a weaker position is warranted — i.e., Siona DCM has a focus-sensitive distribution.

The discussion in this section is broken into two parts: Section 6.4.1 assesses the activity of Siona PROMINENT case-marking across a handful of constructions where focus strategies are implicated cross-linguistically: including, different types of questions, corrective focus constructions, similarity and opposition constructions. Section 6.4.2 describes the interaction of PROMINENT case-marking with nominal discourse

markers — i.e., -*se'e* (only), -*hẽ* (also), -tã'ã (even)[15], and the contrastive topic clitic, =*kato*. Together these sections reveal the distribution of the PROMINENT-marking strategy, which in turn sheds light on its semantico-pragmatic nature.

6.4.1 Siona DCM patterns across focal construction types

The data considered in the preceding sections of this chapter have demonstrated that focal case-marker is implicated in both interrogative contexts and to encode the emphatic contrast reading in plain assertions — i.e., information focus and contrastive focus (Dik et al. 1981; Dik 1997; Kiss 1998; Repp 2016; etc.). The former type displays obligatory, construction-level case matching effects, whereas the latter does not. This section explores both interrogative and assertive instances of Siona focal case-marking, each of which displays a more limited distribution than might be anticipated based upon the foregoing discussion. The discussion is designed moreso as a contribution to the empirical description of Siona DCM than as a full-fledged theoretical account of the patterns.

As a point of departure, (31) enumerates a set of constructions, where focus strategies prototypically occur on a cross-linguistic basis (inspired by Becker and Schneider-Blum 2020). Siona DCM is implicated in a subset of these constructions (identified with ✓), and is absent in others (identified with x). This section retains the shorthand labels in (31) to categorize examples in the discussion below as a matter of convenience:

(31) a. i. [x] ALTERNATIVE QUESTION constructions (henceforth Q-ALT):
e.g., Do you want [coffee]$_{A1}$ or [tea]$_{A2}$?

ii. [x] Answer to (31a-i) (henceforth SEL-ALT):
e.g., (I want) [coffee]$_{A1}$.

b. i. [✓] INFORMATION QUESTION constructions (henceforth Q-WH):
e.g., [What]$_{\{A\}}$ do you want?

ii. [✓] Answer to (31b-i) (henceforth SEL-WH):
e.g., (I want) [coffee]$_{A1}$.

c. [x] SIMILARITY constructions (henceforth SIM):
e.g., I made [the muffins]$_{A1}$ and I also made [the cake]$_{A2}$.

d. [✓] OPPOSITION constructions (henceforth OPP):
e.g., [[John]$_{A1'}$ is mowing the lawn]$_{A1}$, but [[Pete]$_{A2'}$ was pruning the roses]$_{A2}$.

e. [x] CORRECTION constructions (henceforth CORR):[16]
e.g., [Tom]$_{A1}$ did not make dinner, [Lea]$_{A2}$ made dinner.

[15]The morpheme -*tã'ã* is the COUNTER-EXPECTATIVE marker. Its precise semantic contribution is difficult to unpack, so for the present purposes 'even' is suggested as a basic English translation based upon glosses used elsewhere in the literature for markers with similar meanings.

[16]On the one hand, the bi-clausal CORR construction shown in (31e) often recruits PROMINENT-marking along the lines of the OPP construction archetype. However, in short corrective sentences speakers tend to reject PROMINENT-marking in the absence of explicit alternatives — e.g., 'John made dinner' → 'No, MARY made dinner'. In such instances a periphrastic strategy is selected in Siona.

Per (31), Siona DCM is implicated in information questions (Q-WH) and their answers (SEL-WH), and in the so-called opposition construction (OPP). Upon closer inspection, the distribution of PROMINENT-marking within these construction types is limited. If Siona DCM were to be a full-fledged focus-encoding strategy, it would be expected to arise in a larger subset of focus constructions. Before delving into the details of these particular construction types, examples are provided for the construction types where PROMINENT-marking does not occur.

In many languages, the same focus-encoding strategy arises with information questions as with alternative questions, and for their corresponding answers. According to (31), Siona does not recruit the PROMINENT case-marking strategy for the latter class of interrogative constructions. (32) presents an alternative question and its answer, where no PROMINENT-markers arise in either turn.[17] When asked explicitly, speakers indicate that case-marking is odd in both (32a) and (32b):

(32) a. Q-ALT: *huerore tsia banije noca ?*[18]
wero-re **tsia** bãnihẽ **nohka**
buy-N2/3S.PST.N.ASS egg or banana
'Did (you (PL)) buy EGGS or baNAnas?'

b. SEL-ALT: *tsiase'e huerohuë*
tsia-se'e wero-wɨ
egg-EXCL buy-N3S.PST.ASS
'(We) bought only EGGS.'
⇒ *#tsiase'**-re** huerohuë* **Intended:**(32b)
[VOL/SUG: 20230617eyopa001.019-20]

Despite an ample amount of empirical evidence for PROMINENT-marking in interrogative contexts, (32) demonstrates that this usage does not extend to alternative questions. Especially the lack of PROMINENT-marking in answer turn in (32b) is surprising, given that the selected alternative *tsia-se'e* (only eggs) is contrasted with the excluded alternative *nohka* (banana). The lack of PROMINENT-marking in (32b) is illuminating for another reason: PROMINENT-marking does not naturally arise with the exclusive discourse marker, *-se'e* (only) in such instances. On most accounts, this element contributes a focus-sensitive meaning (e.g., Rooth 1985, 1992; König 1991; Krifka 2001, 2012). A detailed discussion regarding PROMINENT-marking and *-se'e* is reserved for Section 6.4.2.

PROMINENT-marking is similarly averted in alternative question-answer sequences where the SUBJECT-argument is targeted, as shown in (33):

[17] The non-assertive clause-typing verbal morphology in the Q-ALT turn indicates that this is a question (Bruil 2014, §6.2.1). An optional question particle, *-'ne* is sometimes added to the conjugated verb to indicate unequivocally that the clause is interrogative.

[18] In the same elicitation session, the speaker confirms that the same judgments apply to this sentence where the conjunct *kwĩ'ne* (also, and) is used in place of the disjunctive marker, *banihẽ* (or). Speakers also produce the Spanish disjunctive marker *o* in place of the native one to the same effect.

(33) a. Q-ALT: *tsoa ëjao bąnijẹ më' gajeo goame necoa'ne iñona ?*
tsoa + ɨha-o bãnihẽ **mɨ'ɨ̃ gahe-o** goa-me nee-ko-a-'ne
tall + foreign-CL:F or 2SG friend-CL:F task make-CL:F-COP-Q
ihño-na
here-GOAL
'Does the tall FOReigner (F) or your FRIEND (F) work here?'

b. SEL-ALT: *tsoa ëjao goame necoa*[19]
tsoa + ɨha-o (#-bi) goa-me nee-ko-a
tall + foreign-CL:F-SBJ task make-CL:F-COP.3S.F
'The FOReigner (F) works (here).' [VOL: 20230617eyopa001.033]

Just as alternative question-answer sequences do not implicate DCM, neither is it implicated in polar question constructions (Q-POL and SEL-POL examples). As a case in point, (34) shows a polar question-answer sequence lifted from a telephone call role-play task. The speaker is calling a friend who works in a repair shop, and asks if he is fixing something particular, *tojasaiwɨ̃ã* (desks), in (34a). His friend responds in the negative and corrects the item which he is repairing in (34b), where he does not naturally produce PROMINENT *-re* marking:

(34) a. Q-WH.POL: *nede'huaquë toyasaihuëą ?*
nee + de'wa-kɨ **toja + sai-wɨ-ã**
make + repair-2/3S.M.PRS.N.ASS write + furniture-CL:CONT-PL.INAN
'Are you (M) fixing desks?'

b. SEL-WH.POL (≃CORR): *bąñë − jąërë nede'huayë*
bãã-ñɨ **hãɨ̃-rɨ** (#-re) nee
NEG.AUX-N3S.PRS.ASS hammock-CL:MAIZE-N.SBJ make
+ de'wa-jɨ
+ repair-N3S.PRS.ASS
'No, I'm fixing a HAmmock.' [VOL: 20241015ejabi001.017a-e]

Although the original Q-POL turn in (34a), volunteered by the speaker, does not contain PROMINENT-marking; upon further scrutiny, the speakers accepts this case-marking when asked. The speaker conjures the context laid out in (35a) where the expected contrastive interpretation emerges in that sentence. Nonetheless PROMINENT case-marking is still dispreferred on the focalized constituent in the Q-POL turn in (35b):

(35) a. Q-WH.POL′: *nede'huaquë toyasaihuëąre ?*
nee + de'wa-kɨ **toja**
make + repair-2/3S.M.PRS.N.ASS write
+ sai-wɨ-ã #(-re)
+ furniture-CL:CONT-PL.INAN-N.SBJ
'Are you (M) fixing DESKS?' ⇒ {. . . fixing bicycles, . . . fixing chairs,. . . }
[**Context suggested by speaker:** The friend he is calling owns a repair shop where they fix everyday items.]

[19]The speaker comments that, if *-bi* marking is used in (33b), it seems like there is more to say about the foreign woman. They deem this to be an odd answer to (33a). To my view, this judgment is similar to topicality effects described for related DCM patterns in Koreguaje [*coe*] in Cook and Levinsohn 1985. Overt case-marking signals that the referent is important to the immediately ensuing discourse.

b. SEL-WH.POL (≃CORR): *bą̲ñë – ja̲ë̲rë nede'huayë*
bãã-ñɨ **hãĩ-rɨ** nee + de'wa-jɨ
NEG.AUX-N3S.PRS.ASS hammock-CL:MAIZE make + repair-N3S.PRS.ASS
'No, I'm fixing a HAmmock.' [VOL: 20241015ejabi001.017a-e]

The examples in (34) and (35) show that PROMINENT-marking is not implicated in polar question constructions, although these also seem to have a corrective flavour. Corrections are a common focus construction cross-linguistically (i.e., Dik et al. 1981, 1997); which is not surprising, given that they negate a particular overt alternative, and suggest another. Outside of interrogative contexts, sentences such as (36) demonstrate that PROMINENT-marking is not recruited in more canonical corrective constructions either, as indicated in (31):

(36) CORR: *curare cua'coye ba'iji ba̲ji noca*
tsia kwa'ko-je ba'i-hi bãã-hi **nohka**
egg cook.TRS-CL:GEN be-3S.M.PRS.ASS NEG.AUX-3S.M.PRS.ASS plantain
'(You) have to cook EGGS, not planTAIN!'
[**Context suggested by speaker:** One offers to help their mother with cooking dinner. They grab a plantain to cook. . .] [VOL: 20230625ejabi001.007]

The final focus construction, which is claimed in (31) not to implicate the focal case-marking strategy, concerns so-called similarity constructions (SIM examples). This meaning is achieved by the additive suffix, *-hẽ* (also),[20] as shown in (37). This marker is compatible with PROMINENT-marking, as demonstrated in Section 6.4.2; however, the construction does not force the presence of the PROMINENT-marker:

(37) a. SIM: *gucubi so̲quë̲tenere̲je̲ tsiayare go'eo*[21]
guhku-bi **sõhkɨtene *(-re)-hẽ** tsia-ja-re
owl-SBJ catfish(spec)-N.SBJ-ADD river-CL:RIV-N.SBJ
go'e-o
look_for-3S.F.PST.ASS
'Owl (nickname) was also looking for CATfish (nickname) at the river.'

b. SIM′: *gucuje̲ so̲quë̲tenere tsiayare go'eo*
guhku (#-bi)-hẽ sõhkɨtene *(-re) tsia-ja-re
owl-SBJ-ADD catfish (spec)-N.SBJ river-CL:RIV-N.SBJ
go'e-o
look_for-3S.F.PST.ASS
'Also OWL (nickname) was looking for Catfish (nickname) at the river.'
[VOL: 20230627elupa001.019a-c]

The general pattern is such that PROMINENT case-markers are often avoided where the nominal is marked with focus-sensitve enclitics, such as *-hẽ* (also) and *-se'e* (only).

[20] This is a multi-functional morpheme, which also shows a related simulative function — e.g., *jɨ'ɨ ha'kɨ-hẽ ba'ɨ ñaagoañoñɨ* (I look like my father [VOL: 20241010ejabi001.003b]). This same marker exhibits interesting behaviour under negation and quantification, which merits further investigation — Martine Bruil, p.c.

[21] The speaker indicates that *-ni* marking is grammatically possible, *sõhkɨtene-**ni**-hẽ* (CATfish too), but it is odd in the case of (37a).

Nonetheless, this incompatibility is not a morphological co-occurrence restriction on case-marking, given instances such as (37a) where *-hẽ* occurs with the obligatory PLAIN *-re* marker on the animate non-SUBJECT argument as expected. Further details are considered in Section 6.4.2.

Up to this point this discussion has considered focus constructions where the PROMINENT-marking strategy is inactive. The remainder of this section explores the distribution of this strategy in those constructions where it is active: namely, information question-answer sequences, and opposition constructions. In fact, this strategy is only implicated in a subset of each construction type, further militating against the strong analysis of PROMINENT-marking as a focus-encoding strategy. These construction types are explored in greater detail in Section 6.4.1.1 and Section 6.4.1.2 respectively.

6.4.1.1 Case-marking and information-seeking questions

QUESTION-ANSWER congruence phenomena are central to the conception of Siona DCM patterns as PLAIN-PROMINENT pairs, introduced in the *preliminary description* in Chapter 3. More concretely, the Q-A diagnostic is applied to ascertain the PROMINENT-status of a given case-marker based upon their obligatory usage in information question-answer sequences. A subset of information questions demonstrate congruence effects, and PROMINENT-marking in the answer turn is only obligatory where it is also required in the corresponding question turn. To adequately establish the connection between PROMINENT-marking and particular information questions, a brief description of Ecuadorian Siona WH-item lexicon is in order.

Firstly, independently of their case-marking properties, all WH-items are fronted to the left clausal edge. Since only constituent WH-items concern this discussion, the high frequency interrogative particle, *me* (how, why), is set aside.[22] Siona constituent WH-items fall into two series, formed around two interrogative roots: i.e., *ke-* and *he-*.[23] Table 6.2 enumerates the WH-item inventory in Siona, where items identified with X^{PROM} display obligatory PROMINENT-marking:[24]

[22]Schwarz (2014, §2.3) provides a compatible description of the cognate item, *me* (how, why) as found in Ecuadorian Sekoya.

[23]Interestingly, the series *he* is not attested in Colombian Siona (Wheeler 1970, discussed in Schwarz 2014, 5-6). However, unlike Ecuadorian Sekoya, I have not come across human-denoting uses of the *he*-series, described in Schwarz (2014, 6-7) in Ecuadorian Siona. This variety is a middle ground.

[24]I have also come across one instance of *ke-o* (lit. what (F)) which was naturally produced where the answer is presupposed to be a de-verbal nominalization — e.g., ↑*kohka kaako* (cellphone), discussed in Chapter 3: (*keo-**ni** ko'eko?* [What (F) are you (F) looking for?]). This same strategy may be used with an animate OBJECT-seeking question, where the referent is presupposed to be feminine. Further work is needed to fully unpack the precise distribution of this item in both of these contexts.

Root	Classifier+	Meaning(s)
ke	*-ɨ* (CL:M)$^{\text{PROM}}$	who (M, *default*); what (ANIM)
ke	*-o* (CL:F)$^{\text{PROM}}$	who (F)
ke	*-je* (CL:GEN)$^{\text{PROM}}$	what (INAN)
ke	+ NP — e.g., *je'jekɨ*	what student (M)
he	*-ro* (CL:LOC)$^{\text{PROM}}$	where
he	*-sõ* + NP	how many (ANIM/INAN)
he	*-ja'ye* + NP	how X, how many (INAN)
he	*-jɨwa*	when
he	+ NP — e.g., *wɨ'e*	which house

Table 6.2: Constituent WH-item inventory, X$^{\text{PROM}}$ must bear PROMINENT-marking

The first take-away from Table 6.2 is that PROMINENT-marking is only required in question-answer sequences with simplex WH-items. Conversely, complex WH-items, such as *ke je'jekɨ* (which student (M)), do not require PROMINENT-marker. Contrast the near-synonymous information question-answer sequences in (38) and (39):

(38) a. Q-WH: *queibi ba'iquë iye huë'e ?*
ke-ɨ-bi$^{<\text{PROM}>}$ ba'i-kɨ ɨ-je wɨ'e
WH-CL:M-SBJ live-2/3S.M.PRS.N.ASS DEM.PROX-CL:GEN house
'Who lives in this house?'

b. SEL-WH: *yëquë tsëcabëbi ba'iyë*
jɨhkɨ tsɨhka-bɨ #(-bi)$^{<\text{PROM}>}$ ba'i-jɨ
1PL.EXCL family-CL:COL-SBJ live-N3S.PRS.ASS
'Our FAmily is living (there).' [VOL: 20220617ecebi001.016a-c]

(39) a. Q-WH: *que tsëcabë icore ba'iye ?*
[**ke tsɨhka-bɨ**] ih-ko-re ba'i-je
WH family-CL:COL DEM.PROX-CL:F-N.SBJ live-N2/3S.PRS.N.ASS
'Which family lives here?'

b. SEL-WH: *yëquë tsëcabë ba'iyë*
jɨhkɨ tsɨhka-bɨ (#-bi) ba'i-jɨ
1PL.EXCL family-CL:COL-SBJ live-N3S.PRS.ASS
'OUR family is living (there).'[25] [VOL: 20220617ecebi001.014a-b]

A similar pattern is reported in (40), where the complex WH-item designating the OBJECT in the question turn does not bear *-re* marking. The speaker did not produce PROMINENT-marking on the focalized constituent in the answer turn in (40b), but did not reject it when asked. For the current purposes, this instance of PROMINENT-marking is reported as optional. This is crucially dissimilar to the PROMINENCE-matching pattern noted with simplex WH questions:

[25]The speaker provides a similar judgment to that noted in (33b), namely that there is more to say about the referent, our family, as the conversation continues. This usage does not align with the focal case-marking patterns discussed in the present discussion.

(40) a. Q-WH: *je ocohuë baco jaǫ ?*
he ohko-wɨ baa-ko hã-o
WH water-CL:CONT have-2/3S.F.PRS.N.ASS DEM.MED-CL:F
'Which water bottle does she have?'

b. SEL-WH: *baco yë' ocohuë*
baa-ko **jɨ'ɨ ohko-wɨ(-re)**
have-3S.F.PRS.ASS 1SG water-CL:CONT-N.SBJ
'(She) has MY water bottle.' [VOL/SUG: 20220620eeupa001.012]

Per Table 6.2, quantity WH-expressions formed with *hesõ* (how many) do not bear PROMINENT-marking. This is demonstrated in (41), where PROMINENT *-ni* marking does not arise on the animate WH-expression, or in the corresponding SEL-WH turn in (41b). Instead PLAIN *-re* marking arises on the animate OBJECT as expected:

(41) a. Q-WH: *jesǫ yaëre airo ñaë ?*
[**hesõ ja-ɨ *(-re)**] ai-ro ñaa-ɨ
how_many jaguar-CL:M-N.SBJ big-CL:LOC see-2/3S.M.PST.N.ASS
'How many jaguars did (you (M)) see in the forest?'

b. SEL-WH: *si'ahua'i yaëre ñahuë*
[**si'a-wa'i ja-ɨ *(-re)**] ñaa-wɨ
all-CL.ANIM jaguar-CL:M-N.SBJ see-N3S.PST.ASS
'(I) saw ALL of the jaguars.' [VOL/SUG: 20230619elupa002.007a-b]

In a follow-up to (41b), the speaker accepted the suggested PROMINENT *-ni* marking on the OBJECT-argument, and conjured the context reported in (42). In this instance, the head noun of the OBJECT nominal receives the emphatic contrast interpretation:

(42) SEL-WH′: *si'ahua'i yaëni ñahuë*
[**si'a-wa'i ja-ɨ #(-ni)**] ñaa-wɨ
all-CL.ANIM jaguar-CL:M-N.SBJ2 see-N3S.PST.ASS
'(I) saw all of the JAGuars.'
[**Context suggested by speaker:** You are out surveying the hunting paths in the forest and come across many species of animals. There are many jaguars, but also tapyrs and peccaries, and other creatures.]
[VOL/SUG: 20230619elupa002.007a-b]

The quantity question-answer sequence shown in (41b), together with (42), demonstrates PROMINENT-marking is still available with WH-expressions that do not force the marking. A similar pattern is shown in (43), where the presence of PROMINENT *-re* on the inanimate OBJECT achieves an emphatic contrast reading:

(43) a. Q-WH: *jesǫ jachohuë̱a baquë ?*
[**hesõ hacho-wɨ-ã̱**] baa-kɨ
how_many shoot-CL:CONT-PL.INAN have-2/3S.M.PRS.N.ASS
'How many rifles do (you (M)) have?'

b. SEL-WH: *jachohuë̱a bayë cayaye*
[**hahcho-wɨ-ã̱**] baa-jɨ [**kaja-je**]
shoot-CL.CONT-PL.INAN have-N3S.PRS.ASS two-CL:GEN
'(I) have TWO rifles.'

c. SEL-WH′: *jachohuëgare bayë cayaye*[26]
[**hahcho-wɨ-ã-re**] baa-jɨ [**kaja-je**]
shoot-CL.CONT-PL.INAN-N.SBJ have-N3S.PRS.ASS two-CL:GEN
'(I) have two RIFLES...'
[**Context suggested by speaker:** You own various weapons and other implements for hunting, which you can lend out, but only two rifles.]
[VOL/SUG: 20220617ecebi001 .016a-e]

The examples shown in this section have demonstrated that Q-A congruence effects strictly obtain with a handful of simplex WH-items identified in Table 6.2. In other information question-answer sequences, PROMINENT-marking is not required on either turn in the information question construction. In these contexts, PROMINENT-marking exhibits the same contrastive focal effects noted in non-interrogative contexts. This fact in and of itself favours the weaker analysis proposed above, where Q-A case-matching effects are treated separately from contrastive (focus) effects. The following section explores the distribution of PROMINENT case-marking in a certain family of non-interrogative focus constructions, namely opposition constructions.

6.4.1.2 DCM and opposition constructions

Opposition constructions (i.e., OPP) belong to the umbrella category of PARALLEL FOCUS contexts identified by Dik (1981, 1997, etc.). On the basis of co-textual diagnostics (outlined in Section 6.1), rich contexts are established to elicit sentences where overt alternatives are explicitly contrasted in juxtaposed sentences. PROMINENT-marking is produced on one or both contrasted constituents. The one-marker pattern is demonstrated for animate OBJECT-arguments in (44), whereas the double-marking pattern is shown for spatial GOAL-arguments in (45) respectively:[27,28]

(44) OPP: [*ñeñebëre mohuë te'e yohuë*] cuï'ne [*suarani a'rimaña tsẽahuë*]
ñeñebɨ-re moo-wɨ te'-e jo-wɨ kwĩ'ne
pícolo-N.SBJ fish-N3S.PST.ASS one-CL:GEN canoe-CL:CONT also
swara-ni a'rimahña tsẽã-wɨ
bocachico-N.SBJ2 few catch-N3S.PST.ASS
'[I caught ∼one canoe-full of PÍcalo (fish species)], and [(just) a few bocaCHIcos (fish species)].' [VOL: 20240918ejabi001.006]

(45) OPP: [*yë'ë tsiayana gajeyë*] [*yë' yojei saiji huë'ena*]
jɨ'ɨ **tsia-ja #(-na)** gahe-jɨ jɨ'ɨ johe-ɨ
1SG river-CL:RIV-GOAL go_down-N3S.PRS.ASS 1SG younger_sibling-CL:M
sai-hi **wɨ'e #(-na)**
go-3S.M.PRS.ASS house-GOAL
'[I am going to the RIver], (whereas) [my little brother is going HOME].'
[VOL: 20230617eyopa001.014-015]

[26] When asked, the consultant found it odd to place *-re* on the numeral element in (43c).

[27] In the suggested English translation in (44), it is natural to produce a connector, e.g., *whereas*, although, for Siona speakers, a prosodic pause suffices to separate the opposed clauses in the construction.

[28] The pattern where only one overtly contrasted noun phrase is PROMINENT-marked may account for certain instances of PLAIN-marking on contrastive nominal tokens in the quantitative study in Chapter 5. In many instances a single focal case-marker suffices to establish the contrast.

On the one hand, in the case of (44), the speaker produces PROMINENT-marking in the first clause, but PLAIN *-re* marking in the second clause. When asked, PROMINENT-marking on both constituents is possible, but deemed to be slightly odd in this context. On the other hand, in (45), the speaker rejects PLAIN zero-marking on either spatial GOAL constituent. Although a degree of optionality regarding the selection of the one- or double-marker configuration is noted, a crucial generalization emerges: At least one PROMINENT-marker must be present on the focalized constituent in the opposition construction. Some descriptive nuance is added to this generalization in what follows based on a consideration of certain sub-types of opposition constructions.

There is a sub-class of opposition constructions where the one-marker pattern arises predictably. I label this the *polar* opposition construction. This construction juxtaposes one affirmative and one negative clause, sometimes containing the element *jo'okɨtã'ã* (but),[29] which does not the PROMINENT-marking pattern. The single PROMINENT-marker arises invariably within the affirmative clause, whereas PLAIN-marking is selected within the negative clause.

Three examples illustrate the polar opposition constructions in (46) through (48) below. (46) and (47) demonstrate instances of this construction targeting inanimate and animate OBJECT respectively. (48) provides a comparable instance targeting the Experiencer argument as found with the impersonal predicate, *sɨhsɨ-* (be cold). PROMINENT-marking is confined to the affirmative clause, and the ordering of negative and affirmative clauses has no bearing on case-marking:

(46) OPP: [*bahuë yohuë hueroye* NEG] [*te'e guënarore huerohuë* AFF]
bãã-wɨ **jo-wɨ** wero-je **te'-e**
NEG.AUX-N3S.PST.ASS canoe-CL:CONT buy-CL:GEN one-CL:GEN
gĩna-ro #(-re) wero-wɨ
metal-CL:RECIP-N.SBJ buy-N3S.PST.ASS
'[I didn't buy a BOAT NEG], [(I) bought a COOKing pot AFF].'
[VOL: 20230623ejepa001.013]

(47) OPP: *go'ye mo'se* [*tsihua'ore ñañe bahuë* NEG] [*tsihuaëni ñahuë* AFF]
go'je mo'se **tsi-wa'-o-re** ñaa-ñe bãã-wɨ
before day child-SG-CL:F-N.SBJ see-CL:GEN NEG.AUX-N3S.PST.ASS
tsi-wa'-ɨ *(-ni)/(#-re) ñaa-wɨ
child-SG-CL:M-N.SBJ2 see-N3S.PST.ASS
'Yesterday [(I) didn't see a GIRL NEG], [(I) saw a BOY AFF].'
[VOL: 20230617elupa001.019]

(48) OPP: [*Palomani tsëtsëji* AFF] *yo'oquëta'a* [*maire baji tsëtsëye* NEG]
Paloma *(-ni)/(#-re) sɨhsɨ-hi jo'okɨtã'ã **mai-re**
P-N.SBJ2 be_cold-3S.M.PRS.ASS but.M 1PL.INCL-N.SBJ
bãã-hi sɨhsɨ-je
NEG.AUX-3S.M.PRS.ASS be_cold-CL:GEN
'[PALOma is cold (lit. it's cold to PALOma) AFF], but [WE (INCL) aren't cold (lit. it's not cold to US) NEG].' [VOL: 20230617elupa001.021]

A point of clarification is in order: There is no general restriction on PROMINENT-marking within negative clauses; rather this is best viewed as a construction-specific

[29]The conjunction *jo'okɨtã'ã* has several gender-based alternatives chosen based on the gender of the speaker: i.e., *jo'okotã'ã* (F), *jo'okwa'itã'ã* (PL).

property. In independent negative clauses, Siona DCM behaves in the expected way. For instance, (49) is a volunteered negative sentence where PROMINENT-marking triggers the anticipated contrastive focal reading:

(49) *airo bahuë baini ñañe*
ai-ro bãã-wɨ **bãĩ *(-ni)/(#-re)** ñaa-ñe
big-CL:LOC NEG.AUX-N3S.PST.ASS person.COL-N.SBJ2 see-CL:GEN
'(I) didn't see PEOple in the forest.'
[**Context suggested by speaker:** You went to look for whoever was using the forest paths frequented by the Sionas. You didn't find them, but saw something else... e.g., creatures, trash.] [VOL: 20230615elupa001.007d]

This section has confirmed that the opposition construction, like certain Q-A constructions, implicates Siona DCM in a principled way. Concretely, PROMINENT-marking encodes local contrasts between overt constituents. The marked constituent in the polar opposition construction is always in the *affirmative* clause. It remains to be determined whether this is a rigid syntactic pattern, or whether this generalization has information structural implications. Klein and de Swart (2011) recognize similar certain 'construction effects' in better-studied DCM languages — e.g., Turkish OBJECT-oriented DCM is conditioned by embeddedness, Romanian *-pe* is obligatory on the OBJECT in the comparative construction (*id*, 11-12).

The instances of obligatory PROMINENT case-marking discussed here for opposition constructions are similar to many naturalistic instances identified in the narrative sample in Chapter 5. However, the examples elicited via co-textual techniques here always contain two overt constituents. PROMINENT case-marking is forced on at least one nominal where both nominals are overtly expressed, although it is unclear what effect that argument-dropping in naturalistic speech has on these patterns. Argumental noun phrases are more often covert than they are overt in the narrative sample — as shown in Table B.9 in Appendix B. In any case, these findings are taken as corroborating evidence for the contrastive focal effects in Siona DCM.

Up to this point it has been shown that focus-sensitive case-marking is not recruited in *all* contexts where it might be expected based on focus semantics alone; rather the PROMINENT-marking has a more narrow distribution. To round off this discussion, the following section considers the connection (or lack thereof) between PROMINENT-marking and nominal discourse markers, such as *-se'e* and *-hẽ* introduced above.

6.4.2 Compatibility of focal case- and discourse-marking

In addition to PROMINENT case-marking as a focus-encoding strategy, Siona exhibits a series of nominal discourse markers. This set comprises four morphemes, which arise at the right edge of the noun phrases alongside case-marking morphology in a particular order. This dissertation adopts the labels, and the glossing conventions regarding the suffix vs. enclitic status of the discourse markers, utilized in Bruil (2014, *et seq.*), as enumerated in (50):

(50) a. The EXCLUSIVE marker: e.g., *ohko-**se'e*** (only water).

b. The ADDITIVE marker: e.g., *Guhku-**hẽ*** (Owl (nickname) too).[30]

c. The COUNTER-EXPECTATIVE marker: e.g., *hãmu-**tã'ã*** (even the armadillo).

d. The TOPIC marker: e.g., *jɨ'ɨ**=kato*** (as for me ...).[31]

Any nominal word which may host case-marking may also bear the nominal morphemes outlined in (50). Each of these morphemes has a dedicated position at the right edge of the nominal constituent, typically arising as phrasal suffixes in the order schematized below in the form of a nominal template (Good 2011, 2016):

	CL	PL	DIM	INTENS	EXCL	CASE	ADD	TOP	CNT.EXP
ROOT	1	2	3	4	5	**6**	7	8	9
e.g.,	*-kɨ*	*-wa'i*	*-mahka*	*-reba*	*-se'e*	*-re...*	*-hẽ*	*=kato*	*-tã'ã*

The above nominal template comprises two major suffixal sequences — i.e., the inner positions 1 through 4,[32] which arise closer to the root, and peripheral positions 5 through 9. Whereas the former comprises several morphemes which are sensitive to the φ-properties of the nominal, the latter grouping concerns morphemes which are sensitive to the discourse status of the nominal. The case-marker occupies an intermediate position (position 6), displaying sensitivities to both properties.

The internal structure of the Siona nominal word aligns with that described for the Eastern Tukanoan language Koitiria: "Suffixes that are closest to the stem tend to be lexical, more closely linked to the core meaning of the word; suffixes that are farther from the stem mark the word for its role in the clause, and then for its role in discourse" (Stenzel 2013c, 168). Returning to Siona, the naturalistic sentence in (51) contains various nominals fitted with peripheral discourse markers, illustrating their respective interpretative functions. Most (if not all) of these markers interact with focus semantics in one way or another — see Krifka and Musan 2012 for an overview:

(51) *...yë'cato, iñno yë'se'e yë' huaredohuëmañaja're ba'iyë yë'je*
jɨ'ɨ=kato ĩhño **jɨ'ɨ-se'e** jɨ'ɨ ware-dowɨ-mahña-ha're *ba'i-jɨ*
1SG=TOP here 1SG-EXCL 1SG child.COL-PL.AN-DIM.PL-COM live-N3S.PRS.ASS
jɨ'ɨ-hẽ
1SG-ADD
'As for me$^{(TOP)}$, I also$^{(ADD)}$ live here, just me$^{(EXCL)}$, with my children.'
[NAT: 20140816saupi001.080]

Although both classes of nominal morphology are focus-sensitive in one way or another, perhaps surprisingly, the bulk of nominals bear either a discourse marker *or* a

[30]This marker also functions as a SIMILARITY marker — e.g., *ha'kɨ-**hẽ*** (like his father). It is interesting to note that this is considered to be a case function in certain descriptive traditions: including the 'essive case' in the Finnish tradition (Karlsson 2015, 148-150), and the 'semblative case' in the Australian tradition — e.g., Evans 1985 for Kayardild. Of course, this dissertation argues against *-hẽ* as a case-marker in Siona.

[31]I follow Bruil 2014 (and elsewhere) in treating the contrastive topic marker *=kato* as a enclitic rather than a suffix. This move is motivated by phonological criteria.

[32]For the present purposes, the intensifier morpheme, *-reba*, is not treated as having a discourse-related function. A common usage of this morpheme is *ihño-**reba*** (right here) — this marker contributes a meaning such as 'precisely', or 'the very X'.

case-marker (or neither) in Siona speech. This subsection explores the co-occurrence patterns of PROMINENT-marking and nominal discourse particles as a contribution to the description of Siona DCM — i.e., inspired by the conception of (IN)COMPATIBILITY DIAGNOSTICS for focus-encoding strategies in van der Wal (2015, 2016, etc.).

An initial observation is in order: Discourse markers and case-markers are not in complementary distribution. Instances where both types of markers co-occur on the same nominal constituent arise in elicitations and in texts. (52) lifts several tokens from the narrative sample analyzed in Chapter 5 where discourse marking and case-marking co-occur — see Table B.10 in Appendix B for a complete overview of discourse markers in the sample:

(52) a. [Context: The children of Baina are eating various jungle creatures.]
ja̲ë̲reje̲ a̲ni je̲o̲ni . . .
hã-ɨ-re-hẽ̃ ãni + hẽõ-ni
DEM.MED-CL:M-OBJ-ADD eat + leave_behind-SS
'(They) ate him too (and moved on) . . . '
[NAT*: 20150811sfryi001.549]

b. *i̲o̲ neato ñami ti̲'a̲ni a̲i̲sicobi a̲ni chaoni i̲recato a̲coña . . .*
ĩ-o nea-to ñahmi tĩ'ã-ni ãĩ-sih-ko-bi ãni + chao-ni
3PRO-CL:F dark-COND night arrive-SS eat-COMP-CL:F-SBJ eat + finish-SS
ĩ-ɨ-re=kato ãh-ko-ña
3PRO-CL:M-N.SBJ=TOP eat-2/3S.F.PST.N.ASS-REP
'. . . she came later in the night, the (demon (F)) who had been eating (him) came back to finish eating him up.'
[NAT*: 20151112oespa001.025]

c. [Context: The only female survivor of the global flood event comes to the spot where her house once stood.]
. . . goa yëiñë ai jaiñëse'bi baquëña
goa [**jɨi-ñɨ ai hai-ñɨ-se'e-bi**]
CNT.EXP cotton-CL:TREE very big-CL:TREE-EXCL-SBJ
bah-kɨ-ña
be-2/3S.M.PST.N.ASS-REP
'. . . but (to her dismay), there was only a big COTton tree there.'
[NAT*: 20151112oespa001.025]

The presence of both morphemes may have semantic implications, or not. For instance, in the naturalistic sentence in (53), the complex promoted LOCATION-argument comprises a head noun with a post-nominal relative clause nominalization — i.e., of the type discussed in Section 6.3.2. There are two viable hosts for case-marking and, by proxy, for discourse markers: On the one hand, PLAIN *-re* marking flags nominalized modifier, *tɨmɨsihkoreba**-re*** (. . . which was very full), at the right edge of the noun phrase, which strictly satisfies the syntactic case-marking condition for promoted non-SUBJECT nominals. On the other hand, the head nominal is not case-marked, but rather marked with the topic marker *=kato*, which evokes a contrastive topic reading of the spatial constituent, *wɨ'ereba**=kato*** (that very house — i.e., not other houses which have been mentioned). A similar effect for *=kato* is noted in (52b) above:

(53) *. . . jai huë'erebacato aireba tëmësicorebare bareña*
↑**[hai + wɨ'e-reba=kato ai-reba tɨmɨ-sih-ko-reba *(-re)]**
big + house-INTENS=TOP very-INTENS be_full-COMP-CL:F-INTENS-N.SBJ
bah-te-ña
be-N2/3S.PST.N.ASS-REP
'(Everybody) was in that big house, which was very full (so they say).'
[NAT: 20140816saupi002.072]

At this stage a generalization emerges such that discourse markers are compatible with case-marking, although their co-occurrence is relatively rare. With this morphosyntactic generalization at hand, the question of how *focal* case-marking interacts with discourse markers may be adequately addressed. For instance, in (53), PLAIN *-re* marking does not make a semantic contribution per se; so semantic incompatibility is not at issue. Conversely, in (52c), *haiñɨ**-se'e-bi*** (only a big TREE), focal case-marking naturally arises with the exclusive discourse marker. As discussed in Section 6.4.1, this co-occurrence is not required, and only arises in rare instances. The remainder of this discussion considers co-occurrence patterns between PROMINENT case-markers and discourse markers.

In order to demonstrate that PROMINENT-marking and discourse marking are *morphologically* compatible, but not necessarily semantically compatible, this discussion shifts to a consideration of elicited Q-A data points. Comparing (54) and (55), the former reports a polar question-answer sequence — i.e., where PROMINENT-marking is not expected to arise; whereas the latter reports an information question-answer sequence. The latter pair implicates DCM in the form of Q-A congruence effects, such that the focalized animate OBJECT argument bears obligatory PROMINENT *-ni*:

(54) a. Q-WH.POL: *ñare tsɨ̃ɨ̃re ?*[33]
ñaa-re **tsɨ̃ɨ̃ *(-re)**
see-N2/3S.PST.N.ASS child.COL-N.SBJ
'Did (you (PL)) see the kids?'

b. Q-SEL.POL: *ñahuë tsihua'ose're*
ñaa-wɨ **tsi-wa'-o-se'e *(-re)/(#-ni)**
see-N3S.PST.ASS child-SG-CL:F-EXCL-N.SBJ
'(We) only saw the GIRL.' [VOL: 20230617eyopa001.017]

(55) a. Q-WH: *queini ñare ?*
ke-ɨ-ni ñaa-re
WH-CL:M-N.SBJ2 see-N2/3S.PST.N.ASS
'Who did (you (PL)) see?'

b. Q-SEL: *tsihua'ose'ni ñahuë*
tsi-wa'-o-se'e *(-ni)/(#-re) ñaa-wɨ
child-SG-CL:F-EXCL-N.SBJ2 see-N3S.PST.ASS
'(We) only saw the GIRL.'

[33]The noun produced by the speaker in (54a), *tsɨ̃ɨ̃* (children) is classified as belonging to the class of lexical generics for the coding of the Siona narrative dataset in Chapter 5. This is due to the fact that narrators often produce these items with no case-marking outside of SUBJECT contexts — such that they pattern with the inanimate noun class. It is curious that the speaker rejects zero-marking in the context of this elicitation task. Similar judgments are made by other speakers. More work is needed to properly unpack the case-marking facts for these few items.

[**Context suggested by speaker:** As a follow-up to (54) − A parent has two children, a boy and a girl. Both children leave the house, so the parent asks their neighbour if they saw them leave. The neighbour says (55b), indicating that they did not see the son leaving, but only the daughter.] [VOL: 20230617eyopa001.018]

On the basis of the above examples, Q-A congruence are not overridden by the presence of the exclusive discourse marker. Thus, although their co-occurrence is grammatically licit, speakers hesitate to produce instances where these co-occur outside of question-answer sequences. The only non-Q-A instances which arise in my field notes concern spatial DCM examples, such as (55):

(56) *iye ye'yahuë'ese'na goamaña neñë*
[i-je je'ja + wɨ'e-se'e-na] goa-mahña nee-ñɨ
DEM.PROX-CL:GEN teach + house-EXCL-GOAL task-DIM.PL make-N3S.PRS.ASS
'(I) am working (lit. doing little tasks) only at this school.'
[**Context suggested by speaker:** The speaker is a teacher who used to work at various local school. When asked about his current situation, he states (56).] [VOL: 20230626eyopa001.021]

As is found in the narrative sample in Chapter 5, on the basis of instances like (56), it appears that spatial nouns arise with PROMINENT-marking more readily than their argumental counterparts. More work is warranted to further unpack this morphological discrepancy and its possible semantico-pragmatic underpinnings. In a similar vein, further research remains to check whether the case-marking effects noted with the exclusive discourse marker can be replicated for the additive marker (*-hẽ*) and the counter-expectative marker (*-tã'ã*).

The remainder of this discussion concerns the contrastive topic clitic, *=kato*. This is the most frequently-utilized discourse marker in Siona speech.[34] The contrastive status of this marker is clear on the basis of naturalistic instances such as (57), and the lines in (58) lifted from *The demon kwẽẽwahti* tale analyzed in Chapter 5:

(57) [**Context:** The speaker's wife says she is from the Cuyabeno region...]
jaiyana aide'osiquëa'ë yë'cato
hai-ja-na ai + de'o-sih-kɨ-a'-ɨ **jɨ'ɨ=kato**
big-CL:RIV-GOAL big + become-COMP-CL:M-COP-N3S 1SG=TOP
'As for me, I grew up at the big river (i.e., Napo river).'
[NAT: 20150810sbopa001.004]

(58) a. *nocabo jë̲o̲saiye ba̲'ëña i̲ i̲o̲ ë̲jë̲maca ...*
nohka-bo hĩõ + sai-je
banana-CL:ENCLOS clean + go-CL:GEN
ba-∼'-ɨ-ña [ĩ-ɨ ĩ-o
be-RMT.PST-2/3S.M.PST.N.ASS-REP 3PRO-CL:M 3PRO-CL:F
ɨhɨ̃-mahka]$_i$
husband-DIM.SG
'Her husband$_i$ went to tend to the banana plantation ...'

[34]The full distribution of discourse markers and their co-occurrence with case-marking is reported in Table B.10 in Appendix B.

b. *. . . i̱o̱macacato a̱o̱re sëoco ñucoña*
ĩ-o-mahka=kato$_{ii}$ ãõ-re sɨo-ko
3PRO-CL:F-DIM.SG=TOP cassava-N.SBJ fry-S.F.DEP
ñuh-ko-ña
be_seated-2/3S.F.PST.N.ASS-REP
'. . . (as for her,) she$_{ii}$ was sitting (there), frying caSSAva.'
[NAT*: 20151112oespa001.164-165]

The fact that both the topic marker *=kato* and PROMINENT case-marking have contrastive semantics of some type is indicative of semantic incompatibility. This is expected on the conception of topical material as presupposed, compared with focal material which is novel information (Dik et al. 1981, 1997; Chafe 1987; and many more).

The anticipated semantic incompatibility between *=kato* and PROMINENT-marking is borne out as a grammatical co-occurrence restriction. This fact is already presumed for PROMINENT *-bi* marking on the SUBJECT-argument. For instance, *=kato* marking motivates the exclusion of 46 S-tokens in the preceding chapter, given that these PLAIN tokens are invariably zero-marked.[35]

In the case of non-SUBJECT arguments, regarding animate nominals, the presence of *=kato* does not satisfy the syntactic overt case-marking requirement. In these instances, *=kato* co-occurs with obligatory PLAIN *-re* marking.[36] Reconsider (59) from the discussion of event nominalization complements to model constructions in this chapter. In this sentence, the Experiencer argument must be marked with the PLAIN *-re* marker in order to form a grammatical sentence:

(59) *. . . më'recato hua'i moñere de'oji*
mɨ'ĩ *(-re)/(*-ni) =kato [wa'i moo-ñe #(-re)] de'o-hi
2SG-N.SBJ(2)=TOP meat fish-CL:GEN-N.SBJ be_good-3S.M.PRS.ASS
'As for you$^{\text{TOP}}$, you can go FISHing$^{\text{FOC}}$.'
[**Context suggested by speaker:** At an event where several other activities are planned, you can go fishing, but not participate in other planned activities. A parent says (59) to one of their children, expressing that they may only go fishing; but the other sibling can do as they please.]
[VOL: 20230614ejabi001.014]

In addition to representing a morphological restriction, (59) elucidates the contrastive topic interpretation which it induces on the Experiencer constituent. Whereas the focal PROMINENT-marker *-re* produces a strong contrastive reading for the activity denoted by the event nominalization; impressionistically, the topical *=kato* marker induces a weaker brand of contrast. The construction shown in (59), implicating both a contrastive topic and contrastive foci, resembles constructions described for

[35]It is conceivable that similar co-occurrence restrictions observed in Colombian Siona, regarding the cognate, *-ga*, underpin the analysis of this morpheme as a case-marker. Since multiple case-marking is not attested in Siona+ languages (see Chapter 4), *-ga* is tricky to tease apart from focal case-marking, like in Ecuadorian Siona.

[36]Technically speaking, the fact that PROMINENT *-ni* is unavailable to P-tokens rends these toke invariable. In follow-up studies, these tokens ought to be excluded. Since only 5 such tokens arise in the narrative sample in Chapter 5, this does not have a big effect on the statistical findings reported there.

languages such as German and Hungarian by Büring (2003, 2016). The corresponding Siona construction deserves more attention.

A final observation is in order regarding the semantic incompatibility of *=kato* and focal case-marking, based upon corroborating evidence from the corpus-based investigation in the previous chapter. Of the total set of 67 (non-temporal) nominal tokens bearing *=kato* marking in the Siona narrative dataset (see Table B.10), these *all* refer to established referents. None of these tokens coded in the sample pertain to first-mention tokens. This finding supports the initial analysis of *=kato* as a topic marker, in addition to inducing a contrast.

To summarize: Discourse particles resist co-occurrence with PROMINENT-marking to varying degrees, and their co-occurrence is often deemed odd. Nonetheless, this co-occurrence is not grammatically blocked, and the presence of most discourse markers in (information) question-answer sequences, for example, does not interfere with the regular patterns of Q-A congruence. The contrastive topic marker *=kato* is an exception to this generalization, and a grammatical co-occurrence restriction arises. In order to fully unpack the semantic and grammatical dimensions of the noted incompatibility, a more elaborate semantic study of the phenomenon is appropriate. Such a study is outlined in the conclusion in the following chapter of this dissertation. For the present purposes, this section provides indispensable insights into the distribution of PROMINENT case-marking in Siona DCM.

6.4.3 In summary: Focus-sensitive Siona DCM

This chapter began with an analytical question: Can the obligatory case-marker congruence patterns be unified with the emphatic contrast patterns be unified under a single analysis — i.e., Siona PROMINENT case-marking is a bona fide focus-encoding strategy? The alternative is that information focus related to (certain) interrogative constructions is to be treated separately from the contrastivity effects discussed for Siona DCM in this chapter and elsewhere in this dissertation.

The discussion here has demonstrated that PROMINENT case-marking has a relatively limited distribution, strictly arising in a subset of the most common focus-related constructions. On the one hand, besides particular simplex WH-questions and their answers, and opposition constructions (X is doing Y, but Z is doing A), the PROMINENT-marking strategy is typically associated with the contrastive (focus) reading. Otherwise, as shown in (31), PROMINENT-marking is not implicated in alternative Q-A constructions, in similarity constructions (X did Y, X also did Z), or in correction constructions (X didn't do Y, Z did Y). On the basis of this limited distribution, an analysis of PROMINENT-marking as a bona fide focus-encoding strategy is untenable.

Taken together the data presented in this dissertation points to two separate information structure-related TRIGGERS, which largely determine the presence of PROMINENT case-marking across all principled DCM patterns in Siona grammar. On this basis, Siona DCM is better viewed as focus-sensitive, rather than focus-encoding per se (in the sense of Matić and Wedgwood 2013). On the one hand, under the appropriate Q-A contexts, case-marking is predictable. This is best treated as a component of simplex WH-question constructions themselves — i.e., a construction-level effect (Klein and de Swart 2011, 11-13). On the other hand, all else being equal, the dominant semantico-pragmatic effect of PROMINENT-marking is what I have labeled the emphatic contrast reading. This is even the effect which arises in interrogative contexts besides the relevant simplex WH-question constructions, which override this main

effect, namely where the referent of the relevant constituent is contrasted with salient alternatives. Of course, this is highly similar to the finding that contrastivity status is the strongest factor in the corpus analysis presented in Chapter 5.

6.5 Towards a theory of discourse-driven DCM

This chapter assembles a range of data which refines the characterization of information structure and the distribution of Siona DCM. Under the appropriate contexts, information focused or contrastive nominal constituents are marked with overt PROMINENT case-markers in a predictable way. This chapter argues that contrastivity is the dominant function of PROMINENT-marking, whereas certain Q-A sequences display PROMINENT case-matching patterns on a construction-specific basis. The same class of PROMINENT-markers are selected in each of these contexts. This empirical backdrop lays the groundwork for the development of a formal theory which is largely set aside for future research. The present section closes this chapter by considering the theoretical implications of the information structure effects described for Siona PROMINENT-marking above.

The first set of theoretical implications which concerns this discussion pertain to questions of focus realization, as examined in Section 6.2. On the one hand, Siona DCM is compatible with the FOCUS PARTICLE strategy (Büring 2009, 22-24), where a morpheme is responsible for realizing an underlying contrastive (focus) feature. Following the logic in Kratzer and Selkirk (2020), this is ought to be viewed as an interface feature, e.g., [±Contrast], whose distribution varies from language-to-language based upon the precise contrastive meaning that the structure conveys. In Siona grammar, this feature is strictly grammaticalized within the nominal domain, which accounts for its relatively limited distribution.

On the other hand, Siona does not make available focus-specific morphology, rather the putative [±Contrast]-feature factors into the realization of case-marking morphology on the noun phrase. Therefore, this feature has a similar status to the [±Animacy]-feature, which codifies the overt case-marking requirement on animate non-SUBJECT arguments. As such, the primary difference between these two nominal features is that [±Contrast] is active across all DCM relations, whereas the [±Animacy]-feature does not apply to the SUBJECT-argument. This analysis also makes the assumption that the simplex WH-construction supersedes the general usage of [±Contrast], and forces this feature to arise on the WH-item and on the focalized element in the corresponding answer turn.

The conspiracy of these features is most evident as case-marking patterns on the DIRECT OBJECT — i.e., where the case-marker *-re* is the PROMINENT-marking for the inanimate noun class, and the PLAIN-marking for animates.[37] This is labeled the SPLIT *-re* ANALYSIS in Chapter 3. Following conventions of Distributed Morphology (Halle and Marantz 1993, 121-124), (60) posits a featural decomposition[38] for case-marker morphemes inserted at Position 6 in the nominal template shown above. This position may correspond to an underlying K head in the extended nominal projection.

[37]The promoted inanimate noun class must acquire an [±Anim]-feature at some point in the derivation in order to capture the DCM patterns.

[38]The suggested decomposition provides privative feature representations by convention. Not much hinges on the binary or privative status of these features for the present purposes from what I gather.

However, it is not immediately clear how to handle instances of multiple case-marking (case concord effects) or what the status of the other nominal discourse markers would be.The feature [N] represents nominal category, and the [Acc] is an abstract Case category, presumably assigned to the nominal before Spell-Out occurs:[39]

(60) a. x-*re* ↔ [N, Acc, Anim]
b. x-*re* ↔ [N, Acc, Foc]
c. x-*ni* ↔ [N, Acc, Anim, Foc]

Per (60), the marker -*re* is selected on the basis of a featural disjunction: i.e., either it has the [±Animacy]-feature, as in (60a), or it has the [±Contrast]-feature, as in (60b), but never both. Conversely, (60c) represents the featural make-up for the marker -*ni*, which realizes the conjunction of these features. Following the definitions established in this dissertation, only those instances which realize an underlying [±Contrast]-feature constitute PROMINENT-marking. Non-focal, inanimate P does not receive overt case-marking because there simply is no overt morpheme which corresponds to the minimal entry [N, Acc]. On the other hand, no [±Animacy]-plus-[±Contrast] featural conjunction is implicated in the calculation of SUBJECT case-marking. To capture PROMINENT -*bi* marking, a more straightforward entry such as (61) is required, where [Nom] is presumed to be assigned by the T-head:

(61) x-*bi* ↔ [N, Nom, Foc]

There are two possible analyses for (PLAIN) zero-marking: (i) There simply is no morphological entry for these case-marking, (ii) Features are deleted via an Impoverishment rule (e.g., Glushan 2010; Keine 2010; Keine and Müller 2015; Bárány 2018). This rule suppresses the case-marker terminal and eliminates case-marking morphology accordingly. One analysis accounts for certain instances, and the other analysis for other instances where morphological case-marking is absent. An Impoverishment-type approach has some merit since non-focal arguments may still arise with case-marking, as shown to be the case in the quantitative study in Chapter 5. In fact, there may even be several types of non-focal PROMINENT-marking — i.e., (i) Those driven by WEAK TRIGGERS, such as specificity, topicality, and word order; (ii) Those driven by the disambiguation function of DCM (e.g., de Swart 2007; Malchukov 2008; Klein and de Swart 2011). To my view, this is not a language-specific issue, but rather an analytical puzzle concerning multi-dimensional case-marking systems and the analysis of WEAK TRIGGERS for DCM more broadly.

A final comment is in order regarding the connection between [±Contrast] and the empirical evidence considered in this chapter. In light of the prevalence of rigid Q-A congruence patterns in Siona grammar, PROMINENT case-marking appears to be compatible with the family of 'question-under-discussion' (qud) approaches (Velleman and Beaver 2016 for recent overview). This framework is developed on the basis of the

[39]I remain non-committal regarding the processes of (accusative) Case-assignment in Siona for the sake of this brief discussion. Siona, like many (DCM) languages displays similar case-marking facts for (IN)DIRECT OBJECTS, and it is not clear whether the same mechanism assigns the same Case value in both instances. For that matter, assuming that DIRECT OBJECTS are formally syncretic with LOCATIONS regarding case-marking, it is conceivable that a single mechanism 'assigns' the same Case feature to these spatial nominals as well. An adequate theory ought to unify the case-marking licensing mechanism across these grammatical functions in an informative way.

notion that natural discourse may be modeled as question-answer sequences, where questions may be explicit or implicit.

At first glance, the QUD approach is an appealing analytical direction since it permits focal case-marking in interrogative and assertive contexts to be unified under a single analysis. For example, the non-interrogative focal instance in (62), produced in the context of a CONTEXT CONJURING task, aligns formally with the corresponding Q-A example in (63) on the basis of PROMINENT-marking patterns:

(62) [Implicit Q-WH: What do you want to drink?]
(SEL-WH) *yë'ë ocore u̲cusi'i cayë*
jɨ'ɨ **ohko #(-re)** ũhku-si'-i kaa-jɨ
1SG water-N.SBJ drink-FUT-N3S say-N3S.PRS.ASS
'I want to drink WAter.' ⇒ {I want to drink coffee, I want to drink tea,...}
[**Context suggested by speaker:** A boy comes home to find coffee and tea on the table.] [VOL/SUG: 20230623ejabi001.024-025]

(63) a. Q-WH: *quere u̲cusi' caquë ?*
ke-e-re ũhku-si'-i kaa-kɨ
WH-CL:GEN-N.SBJ drink-FUT-N3S say-2/3S.M.PRS.N.ASS
'What do you (M) want to drink?'

b. SEL-WH: *yë'ë ocore u̲cusi'i cayë*
jɨ'ɨ **ohko #(-re)** ũhku-si'-i kaa-jɨ
1SG water-N.SBJ drink-FUT-N3S say-N3S.PRS.ASS
'I want to drink WAter.' [VOL/SUG: 20230627ejabi001.012-014]

Although the appeal of the QUD approach is evident, it runs into some analytical problems for the Siona data. Obligatory PROMINENT-marking only arises for a subset of information Q-A sequences, as discussed in Section 6.4.1.1. Although all WH-items are fronted, only a subset require PROMINENT-marking, and force it to arise on the focal element in the answer. In these explicit question-answer sequences, no contrastive context is required per se; however, outside of these constructions, a contrastive context is required for the felicitous usage of PROMINENT-marking. (This excludes certain uses of overt case-marking for ambiguity avoidance or communicative efficiency purposes.)

On this basis, the distribution and the function of PROMINENT-marking in non-contrastive explicit question contexts does not align with those found in implicit question contexts. The latter contexts are necessarily contrastive. An analysis conflating implicit and explicit question contexts would predict an even more limited distribution of PROMINENT-marking than is observed in practice. For instance, such marking would be expected to be incompatible with contrasting quantities, since *jesõ* (how many) does implicate PROMINENT case-marking, or with any complex questions such as *ke mamakɨ* (which boy) for the same reason. However, contrastive focus uses of PROMINENT-marking have been demonstrated for each of these contexts in the discussion above. Ultimately, interrogative focus effects, arising with certain WH-items, ought to be treated as a separate, secondary effects in Siona DCM; whereas the dominant information structural driver of overt PROMINENT-marking is to induce the emphatic contrast reading, highlighting particular referents in discourse. These observations corroborate the statistical evidence for contrast-driven usage of PROMINENT-marking in Siona narratives in Chapter 5.

CHAPTER 7

Conclusions

This dissertation set out to refine the characterization of Siona DCM patterns, as established in previous research (Bruil 2014, §4.4; Case and Jeretič 2021).[1] In particular, the research questions in (1) are addressed, as reiterated from (8) in Chapter 1:

(1) a. Which DCM patterns can be identified in Siona? What properties do these DCM patterns have in common, and how do they differ?

b. Which factors determine the selection of case marking alternatives in a given context — such as animacy status, focus, specificity? A combination thereof? How do the active factors interact to determine the DCM patterns in usage?

c. How does Siona DCM compare to other DCM systems described in the Tukanoan literature? In the northwestern Amazonia area? In the broader DCM typology?

In order to achieve a holistic description of DCM in the language, a broad array of data types are assembled via the implementation of different methodologies. This diverse research program integrates both controlled elicitation-based and naturalistic text-based data so as to establish corroborating evidence for a strong empirical backdrop for the description and analysis purported here (Matthewson 2004, 2022; Tonhauser and Matthewson 2016; Davidson 2020; etc.), laying the groundwork for future research.

Chapter 2 lays out definitions for DCM, and related phenomena, in addition to establishing the pillars for this holistic description aspired to in this dissertation. An adequate characterization of any DCM phenomenon touches upon three descriptive

[1] In Bruil & Case, *forthcoming*, we undertake a corpus-based analys of P-oriented DCM patterns, which will also contribute to this broader enterprise.

dimensions — i.e., as schematized in Table 2.1: (i) The *formal* set of case-marking alternatives implicated in a given alternation; (ii) The *functional* dimension, pertaining to distribution of DCM patterns across grammatical relations; and (iii) The identification and ranking of TRIGGERS, which drive the observed DCM patterns. Each of these dimensions is addressed in this dissertation regarding Siona DCM, although the latter dimension receives the lion share of attention.

The *preliminary description* in Chapter 3 largely satisfies the first two dimensions, and sets the stage for the targeted exploration of the third. By mapping the six case-markers in the language to the grammatical relations that they encode, the *preliminary description* lays out a novel conception of principled DCM patterns in the language as comprising PLAIN-PROMINENT pairs. Principled argumental DCM patterns define case-marking for core arguments – i.e., SUBJECT and (IN)DIRECT OBJECT;[2] and strikingly similar patterns characterize the case-marking of certain spatial arguments – i.e., STATIC LOCATION and GOAL. Table 7.1 summarizes the distribution of Siona case-markers in paradigmatic fashion:

	INANIMATE		ANIMATE ($\wedge$ $^{\uparrow}$INAN)	
	PLAIN	PROMINENT	PLAIN	PROMINENT
Argumental DCM:				
SUBJECT (S)	-∅	*-bi*	-∅	*-bi*
DIRECT OBJECT (P)	-∅	*-re*	*-re*	*-ni*
INDIRECT OBJECT (R)	*-re* (*-na*)		*-re*	*-ni*
Obligatory case-marking:	*-bi* (Instrument); *-hã're* (Companion)			
Spatial DCM:	INANIMATE		$^{\uparrow}$INANIMATE	
LOCATION (L)	-∅	*-re*	*-re*	*-ni*
[*some L] GOAL (G)	-∅	*-na*	*-na*	
Obligatory case-marking:	*-bi* (Source); *-hã'ã* (Path, Limit)			

Table 7.1: Siona case-inventory, incl. DCM patterns across argumental and spatial domains

The paradigm laid out in Table 7.1 establishes the foundation for the analysis of active TRIGGERS for these DCM patterns. The highest level TRIGGER concerns a rigid animacy-based class split, which forms a wedge between S-oriented DCM on the one hand, and *all* non-SUBJECT DCM patterns in the language. Once the phenomenon of promotion (identified with $\uparrow$) is taken into account, a pattern emerges such that the PLAIN- and PROMINENT-markers for inanimate non-SUBJECTs *never* align with those found with their animate (or promoted) counterparts; which extends equally to spatial DCM. Animacy has a unique role in determining the distribution of formal alternatives, whereas other TRIGGERS are different in kind.

[2] **Section 3.2.4** demonstrates that the PLAIN *-re* and PROMINENT *-ni* pattern for animate non-SUBJECT DCM also obtains for the coding of non-core relations, such as the Experiencer and the Beneficiary argument. Similar patterns are attested throughout the Tukanoan family.

Contrastivity is shown to be a STRONG TRIGGER, which characterizes all principled DCM alternations, as it is the best predictor for the selection of PROMINENT-marking. The *preliminary description* banks on stable patterns of QUESTION-ANSWER congruence, where PROMINENT-marking is forced on the focal element in the ANSWER turn − i.e., as demonstrated in (2) through (4) below. In each instance the appropriate PROMINENT-marking is found, such that, for both non-SUBJECT arguments (i.e., shown in (3) and (4)), the noun class split comes to the fore:

(2) [ANSWERS THE QUESTION: *Who wants to go (somewhere)?*]
yë'ë gajeobi yëco saiye
ji'i gahe-o #(-bi) jii-ko sai-je
1SG friend-CL:F-SBJ want-3S.F.PRS.ASS go-CL:GEN
'My FRIEND (F) wants to go.' [VOL: 20220623emipa002.008-009]

(3) [ANSWERS TO THE QUESTION: *What did you buy yesterday?*]

a. *yohuëre huerohuë*
jo-wi #(-re) wero-wi
canoe-CL:CONT-N.SBJ buy-N3S.PST.ASS
'(I) bought a caNOE.'

b. *yohuë huajëconi huerohuë*
↑[**jo-wi wahi-ko *(-ni)/(#-re)**] wero-wi
canoe-CL:CONT new-CL:F-N.SBJ(2) buy-N3S.PST.ASS
'(I) bought a new caNOE.'

c. *huequëni huerohuë*
wẽki *(-ni)/(#-re) wero-wi
cow-N.SBJ(2) buy-N3S.PST.ASS
'(I) bought (a) COW/COWS.' [VOL: 20230617eyopa001.003-005]

(4) [ANSWERS THE QUESTION: *Where do you (PL) live?*]

a. *ba'iyë iye huë'ere*
ba'i-ji **i-je wi'e #(-re)**
go-N3S.PRS.ASS learn + house-GOAL
'(We) live in THIS house.'

b. *ba'iyë yë'ë ñicuëbai huë'e nesiconi*
ba'i-ji ↑[**ji'i ñihkw-i-bãĩ wi'e**
live-N3S.PRS.ASS 1SG grandparent-CL:M-PL.AN house
nee-sih-ko *(-ni)/(#-re)]
make-PERF-CL:F-N.SBJ2
'(We) live in the HOUSE that my grandparents built.'
[VOL: 20230619elupa002.006a-g]

The question-answer sequences shown above demonstrate obligatory uses of PROMINENT-marking; although these are not the only contexts where PROMINENT case-markers are attested in Siona. In other words, several questions emerge as to how PROMINENT-marking functions in other, non-obligatory contexts − i.e., corresponding to the notorious 'gray area' described for most, if not all, DCM systems described in the literature (Moravcsik 1978; Bossong 1984, 1985; Aissen 1999, 2003; Haspelmath 2021; etc.): Which WEAK TRIGGERS are active in the attested multi-dimensional DCM

patterns in Siona (as conceived in Chapter 2)? On the basis of which other TRIGGERS do principled DCM patterns in Table 7.1 converge? In what ways do they diverge?

The remaining chapters in this dissertation take the *preliminary description* and bolster it in various respects. As such, the ultimate goal is to further the characterization of particular aspects of the REALIZATIONAL DOMAIN of Siona DCM. Section 7.1 summarizes the contributions of each individual chapter. Section 7.2 closes this conclusion by outlining future directions of research to further refine the description and analysis of Siona DCM patterns.

7.1 Summary of findings

The content arising in Chapter 4 through Chapter 6 comprises independent studies, where I put on a different metaphorical *hat* and implement diverse analytical techniques to build upon the *preliminary description*. This section is concerned with summarizing the key evidence purported in each content chapter and with bridging these findings into broader discussion at hand. This section is organized as follows: Section 7.1.1 summarizes the findings from Chapter 4. Chapter 5 is summarized in Section 7.1.2. Finally, Chapter 6 is summarized in Section 7.1.3.

7.1.1 A typological investigation of Siona case-marking

For Chapter 4, I put on my *comparativist* hat. This chapter aims to situate various properties of Siona case-marking and DCM, laid out in the *preliminary description* in Chapter 3, in their typological context. Via the development and implementation of the structural questionnaire technique (Haspelmath 2005, 2010; Krasnoukhova 2012; Birchall 2014), the NWA case-marking dataset in Appendix A positions the Siona facts within a representative sample of Tukanoan and non-Tukanoan languages spoken in the putative NWA area. Besides a handful of typical Tukanoan traits, various Siona case-marking properties, including the set of complex, discourse-driven DCM patterns, diverge sharply from those described for Eastern Tukanoan languages. Some evidence is garnered which suggests that certain Siona traits may have arisen under the pressures of contact-induced convergence. As such, this chapter affords certain insights into the diachronic development of Siona DCM, while highlighting fundamental synchronic discrepancies with Eastern Tukanoan languages.

To begin with the prototypical Tukanoan traits exhibited by Siona, these concern the case-marker *-re*. All Tukanoan languages possess this marker in their respective inventories and aspects of its distribution are shared across the family (Barnes 1999, 2006; Stenzel 2008, 2013d; Ramirez 2019[1997]; etc.). Every Tukanoan language in the sample recruits *-re* marking to encode both DIRECT and INDIRECT OBJECTS, where zero-marking is blocked on the latter. Additionally, besides the above-noted argumental uses of *-re* marking, all Tukanoan languages extend *-re* marking to the encoding of spatial arguments in one way or another. Besides these points of overlap, Siona case-marking diverges markedly from that noted for Eastern Tukanoan languages in the sample (and for the Western Tukanoan outlier, Máíhɨ̃kì [*ore*]).

On the one hand, Siona and its closest relatives − conceived as the Siona+ language complex (incl. Colombian Siona [^{c}*snn*], Ecuadorian Sekoya [*sey*], and Koreguaje [*coe*]),[3] share a set of four 'primary' case-markers, which, in addition to *-re*, comprises

[3]The Peruvian variety of Sekoya has received some attention in recent work − Vallejos

cognates to the marker -*ni*; the GOAL-marker -*na*; and the multi-functional marker -*bi*. The latter marker encodes the SUBJECT, INSTRUMENT, and spatial SOURCE in each of these languages. Each of the 'primary' markers is implicated in the complex Siona+ DCM patterns, exemplified by the Ecuadorian Siona paradigm in Table 7.1 above. The identification of cognate candidates in the Eastern Tukanoan languages Tanimuka [*tnc*] and Kubeo [*cub*] for -*bi* and -*na* points to an earlier, genetic origin. Otherwise cognate markers are not attested in the family.

On the other hand, besides the formal case-marker inventory, Siona+ languages diverge from Eastern Tukanoan languages on the basis of TRIGGERS for the attested DCM patterns. Eastern Tukanoan languages are primarily driven by referentiality (i.e., specificity or definiteness), which is, at best, a WEAK TRIGGER in Siona (as found in Chapter 5). Conversely, Siona+ languages display discourse-driven DCM patterns, which extend equally to SUBJECT-marking, aligning closely to the facts gathered regarding Murui [*huu*] and Tariana [*tae*], spoken to the east. Additionally, as regards non-SUBJECT DCM in Siona+ languages, a STRONG animacy-related TRIGGER determines morphological case-marking splits for particular grammatical relations (also in *Maíhĩkì* [*ore*]), which conspires with discourse-related TRIGGERS to drive the selection of particular alternatives in their discursive context. A unique brand of multi-dimensional DCM emerges, unlike that attested for Eastern Tukanoan in this sample.

Ultimately, the typology developed in Chapter 4 indicates that Siona displays an admixture of Tukanoan and non-Tukanoan case-marking traits. As regards the latter, on account of the geographic position of Siona+ languages at the north-westernmost edge of Amazonia, candidates for grammatical convergence are noted with languages spoken in the Andean foothills region to the west, and with the languages of the *People of the centre* spoken to the east (Echeverri 1997; Epps and Michael 2017; etc.). On the one hand, the presence of a dedicated GOAL-marker, and its exceptional usage with certain non-prototypical predicates, overlaps notably with similar marking patterns described for Quechuan languages spoken to the west. Siona+ languages also display certain FUNCTIONAL GAPS – i.e., they have certain plausible CASE FUNCTIONS which are not encoded via case-marking, including to encode temporal nouns or POSSESSORS. These same gaps are described for languages spoken to the east, such as Murui [*huu*] and Bora [*boa*]. A possible easterly origin for S-oriented DCM is also suggested above. Ultimately, various Siona case-marking traits, including aspects of the particular brand of DCM patterns attested in the language, likely developed due to the unique linguistic context at the northwestern frontier between Amazonia and the Andean foothills.

7.1.2 A corpus-based investigation of Siona DCM

For Chapter 5 (supplemented by Appendix B), I put on my *variationist* hat and develop the modest Siona narrative dataset (8078 word forms; 1914 nominal tokens). This chapter provides a welcome contribution of quantitative evidence to the description of Siona DCM. On the one hand, this chapter measures the rate of PROMINENT-marking in a naturalistic sample: i.e., for SUBJECT (S), DIRECT OBJECT (P), and LOCATION (L) tokens accordingly.

On the other hand, it ascertains statistical evidence, which is utilized to determine the relative impact of a handful of factors (TRIGGER CANDIDATES) on the proportion

and Schwarz 2016; Vallejos 2021; Vallejos and Brown 2021. Although a complete description is lacking at present.

of PROMINENT-marking. The results of this analysis provide statistical support for the generalization laid out in Chapter 3, that discourse-related TRIGGERS have primacy in Siona DCM, particularly the notion of contrast. Furthermore, this chapter sheds light on the relevance of various WEAK TRIGGERS, whose subtle effects are notoriously tricky to address via elicitation-based methodologies. What emerges is a ranking of statistically significant TRIGGERS for each grammatical relation, corresponding to the notion of TRIGGER STRENGTH – as conceived in Section 2.2.3, which is at the core of the analysis of the REALIZATIONAL DOMAIN of Siona DCM in this dissertation.

The results are summarized in Table 7.2, reiterated from Section 5.4:

TRIGGER CANDIDATE [GLOBAL MARKING RATE]	S-marking [21.18%]	P-marking [24.27%]	L-marking [52.63%]
		Sentence-level factors	
WORD ORDER	MODERATE EFFECT$_3$	SMALL EFFECT$_4$	!NO EFFECT
THEMATIC STATUS	!NO EFFECT	!NO EFFECT	!NO EFFECT
		Referent-level factors	
ANIMACY STATUS	!NO EFFECT	SMALL EFFECT$_5$	N/A
SPECIFICITY STATUS	MODERATE EFFECT$_2$	MODERATE EFFECT$_3$	MODERATE EFFECT$_2$
		Discourse-level factors	
CONTRASTIVITY STATUS	LARGE EFFECT$_1$	LARGE EFFECT$_1$	LARGE EFFECT$_1$
LOOK-AHEAD TOPICALITY STATUS	!NO EFFECT	MODERATE EFFECT$_2$	!INVARIABLE
LOOK-BACK TOPICALITY STATUS	!NO EFFECT	!NO EFFECT	!NO EFFECT

Table 7.2: Summary of proportional and statistical evidence, Chapter 5

As shown in Table 7.2, the *varbrul* results (reported in full in Appendix B) indicate that CONTRASTIVITY STATUS is the strongest predictor for PROMINENT-marking for each grammatical relation tested in the study. No other large effects are detected in this analysis. This finding corroborates the descriptive generalization that Siona DCM is primarily driven by a contrast-related STRONG TRIGGER. For instance, this pattern aligns with the fact that the emphatic contrast reading is typically associated with PROMINENT-marking in elicited contexts, described in Chapter 3 and Chapter 6.

Regarding other significant factors, smaller effects are noted for a handful of TRIGGER CANDIDATES in this analysis. Firstly, the discourse-related TOPICALITY STATUS is correlated with PROMINENT-marking for non-SUBJECT DCM patterns in the sample. This analysis applied two topicality metrics (see Section 5.2), and the significant effect concerns LOOK-AHEAD TOPICALITY, as conceived here: Categorizing tokens based upon whether their referent(s) persist(s) in immediately ensuing narration. This finding, along with the above-noted contrastivity effect, confirm the primacy of discourse-related TRIGGERS for Siona DCM established throughout this dissertation.

Besides discourse-related TRIGGERS, this chapter finds relatively smaller effects for SPECIFICITY STATUS across all tested DCM patterns. Specific full nouns favour PROMINENT-marking as compared with their non-specific and pronominal counterparts across-the-board. Another smaller effect is detected regarding WORD ORDER; although this effect does not obtain in the L-marking sub-sample. The general pattern is such that pre-verbal arguments display a propensity for PROMINENT-marking, such that sentential word order and case-marking may represent interrelated strate-

gies for indicating the pragmatic status of noun phrases in natural discourse. This pattern is not noted in previous research, and is difficult to ascertain on the basis of elicitation techniques, given that speakers generally judge word order, and sometimes case-marking, to be flexible components of Siona grammar. Ultimately, these findings shed light on the activity of certain WEAK TRIGGERS in Siona DCM in actual usage.

Setting aside some grammatical relation-particular quirks, the analysis largely detects stable patterns regarding the role of particular factors. It is clear that discourse-related factors outrank others, and that SPECIFICITY STATUS is ranked next highest for all DCM patterns. Of course, by extension, this chapter also finds statistical evidence *against* the activity of certain TRIGGERS: i.e., THEMATIC STATUS (*agentivity*, *affectedness*, and *subcategorization status*) is not relevant to Siona DCM. Similarly, GLOBAL TOPICALITY STATUS, pertaining to the recency of previous mentions of a given referent are non-consequential for Siona DCM.

A final observation is in order regarding the smallest significant effect detected in the P-marking sub-sample, concerning ANIMACY STATUS. On the one hand, this finding confirms that PROMINENT *-re* on inanimate P and PROMINENT *-ni* on animate P display compatible distributions in the sample. On the other hand, it also inspires another hypothesis. Given that inanimate *-re* marking alternates with zero-marking, which is not the case for animate *-ni*, this subtle difference in PROMINENT-marking may be due to the fact that *-re* is selected in certain instances to avoid zero-marking – i.e., for the sake of ambiguity avoidance. This dissertation side-steps the question of disambiguation in Siona DCM (see Section 2.2.3.2), although a more elaborate investigation of ambiguity avoidance in a naturalistic sample is in order to more thoroughly test for the extent of such effects on case-marker selection in usage.

7.1.3 An elicitation-driven investigation: Focus-sensitive case-marking

The final content chapter of this dissertation, Chapter 6, develops an analysis of PROMINENT case-marking as a contrastive encoding strategy, sensitive to certain focus contexts. More specifically, overt PROMINENT-markers are analyzed as a brand of argument-associated FOCUS PARTICLES (Büring 2009, 22-24; Aannestad 2021). This analysis complements the description of emphatic contrast and QUESTION-ANSWER congruence effects laid out in the *preliminary description*, and the statistical effects for CONTRASTIVITY STATUS in Chapter 5, by implementing an elicitation-based research program to properly *diagnose* the focal effects of PROMINENT-marking (van der Wal 2015, 2016, etc.).

The example in (5), for instance, is elicited via the CONJURED CONTEXT technique (van der Wal 2016, §2). The speaker provides careful, introspective judgments regarding the use of PROMINENT-marking, which are often judged as 'optional' on the basis of straightforward translation tasks:

(5) *yë'ë ocore u̲cusi'i cayë*
ji'ɨ **ohko #(-re)** ũhku-si'-i kaa-jɨ
1SG water-N.SBJ drink-FUT-N3S say-N3S.PRS.ASS
'I want to drink wAter.' ⇒ {I want to drink coffee, I want to drink tea,... }
[**Suggested context:** A boy comes home to find coffee and tea on the table.]
[VOL/SUG: 20230623ejabi001.024-025]

The careful application of elicitation-based techniques, including CONJURED CONTEXT tasks, QUESTION-ANSWER (Q-A) tests, and CO-TEXT diagnostics injects a degree of researcher control over variables and context, which refine the empirical basis for the analysis of the dominant TRIGGER for PROMINENT-marking in Siona DCM. With this methodological arsenal at hand, the focus-sensitive distribution of the PROMINENT case-marking strategy are elucidated.

This chapter establishes two sets of findings regarding PROMINENT-marking, which align with properties of cross-linguistically common focus constructions (Büring 2009): On the one hand, Section 6.3 unveils certain focal syncretisms, as emerge via the application of Q-A and CO-TEXT diagnostics. This includes the cross-linguistically robust broad-narrow focus syncretism (e.g., Dik et al. 1981, 1997; Rooth 1985, 1992; Krifka and Musan 2012). For instance, (6) demonstrates how the PROMINENT *-re* marker is obligatory in both broad- and narrow-scope questions, and PROMINENT-marking arises obligatorily on the answer turn in either case, shown in (6b). This happens to surface as *-re* in the case of (6b-i), but as *-ni* on the promoted argument in (6b-ii):

(6) a. i. NARROW-SCOPE QUESTION: *quere nede'huaquë'ne ?*
ke-e *(-re) nee + de'wa-kɨ-'ne
WH-CL:GEN-N.SBJ make + repair-2/3S.M.PRS.N.ASS-Q
'What are you (M) fixing?'

ii. BROAD-SCOPE QUESTION: *quere yo'quë'ne ?*
ke-e *(-re) jo'-kɨ-'ne
WH-CL:GEN-N.SBJ do-2/3S.M.PRS.N.ASS-Q
'What are you (M) doing?'

b. i. ANSWER′: *yë'ë jaërëre nede'huayë*
jɨ'ɨ **hãĩ-rɨ #(-re)** nee + de'wa-jɨ
1SG hammock-CL:MAIZE-N.SBJ make + repair-N3S.PRS.ASS
'I am [fixing [a/the HAmmock]].'

ii. ANSWER″: *yë'ë jaërë yetesiconi nede'huayë*
jɨ'ɨ ↑[**hãĩ-rɨ jehte-sih-ko-ni/(# -re)**] nee
1SG hammock-CL:MAIZE tear-PERF-CL:F-N.SBJ(2) make
+ de'wa-jɨ
+ repair-N3S.PRS.ASS
'I am [fixing [a/the frayed HAmmock]].' or,
'I am [fixing [a/the FRAYed hammock]].'
[VOL/SUG: 20230617elupa001.017a-e]

In fact, the multiple interpretations possible within the complex noun phrase in (6b-ii) reveal a second type of syncretism — i.e., (i) where focus is interpreted narrowly on the head noun 'hammock', or (ii) where the modifier, 'frayed', receives focus. Similar patterns are described for various FOCUS PARTICLE languages (Aannestad 2021, §4.2.1). Together these focus size-related patterns point to the fact that Siona case-marking is not restricted to expressing WH focus on the full nominal constituent which is case-marked, despite the 'particle' being morphologically restricted to realization of nominal hosts.

On the other hand, in addition to size-related facts, Section 6.4 illustrates certain focus contexts where the PROMINENT-marking strategy, which go beyond QUESTION-ANSWER pairs and the emphatic contrast reading. In particular, Siona PROMINENT-marking is implicated in so-called OPPOSITION constructions, which overtly contrast

multiple nominal constituents. For instance, it is deemed infelicitous to leave either of the GOAL-arguments in (7) zero-marked:

(7) *yë'ë tsiayana gajeyë – yë' yojei saiji huë'ena*
ji̵'i̵ **tsia-ja #(-na)** gahe-ji̵ ji̵'i̵ johe-i̵
1SG river-CL:RIV-GOAL go_down-N3S.PRS.ASS 1SG younger_sibling-CL:M
sai-hi **wi̵'e #(-na)**
go-3S.M.PRS.ASS house-GOAL
'I am going to the RIver, (whereas) my little brother is going HOME.'
[VOL: 20230617eyopa001.014-015]

Section 6.4 refines the analysis of the PROMINENT-marking strategy by demonstrating that it is *not* recruited for other prototypical functions: i.e., CORRECTIVE FOCUS constructions, SIMILARITY constructions. Additionally, this strategy is absent for various classes of QUESTION-ANSWER pairs, concerning alternative questions or even certain information questions where complex WH-items are employed – e.g., *ke tsi̵hkabi̵* (which family), on either the QUESTION or the ANSWER turn. The incompatibility with certain discourse suffixes, such as *-se'e* (only) and *-hẽ* (also), equally indicates a relatively limited distribution of the PROMINENT-marking strategy compared to other focus-encoding strategies described in the literature. There is no monolithic conception of contrast (or contrastive focus) which applies equally to all languages (Zimmermann and Onea 2011; Matić and Wedgwood 2013; Kratzer and Selkirk 2020; etc.). As such, the relatively limited distribution of PROMINENT-marking in Siona, compared to focus-encoding strategies in better-studied languages, points to PROMINENT-marking strategy as a focus-sensitive case-marking pattern, rather than a bona fide focus-encoding strategy. The dominant information structural effect related to overt PROMINENT-marking is contrastivity, complementing findings in the other chapters of this dissertation. Ultimately, these findings bolster the language-particular description aspired to in this dissertation and delimit the shape of a technical focus-based account of Siona DCM in future work.

The content chapters of this dissertation contribute to the characterization of Siona DCM laid out in the *preliminary description* in several respects. Corroborating corpus-based and elicitation-based evidence are assembled, which demonstrate that PROMINENT case-markers encode contrastive focus for all grammatical relations displaying principled DCM patterns in the language. In this way, once animacy- and promotion-based marking splits are accounted for, the distribution of case-markers is predictable in the various focal contexts where the strategy is recruited. Despite certain points of micro-variation, compatible facts are identified across the Siona+ complex analyzed in Chapter 4.

The information structural uses of PROMINENT-marking uncovered in this dissertation do not exhaustively account for selection of PLAIN and PROMINENT markers, however. Put simply, not all PROMINENT-marking is based upon information structure, since Siona grammar strictly recruits PROMINENT-marking in a small subset of focal contexts (as discussed in Chapter 6). The text-based investigation in Chapter 5 provides some further, often subtle, insights into the WEAK TRIGGERS that also factor into Siona DCM. The following section outlines certain directions for expanding this investigation in future research.

7.2 Future research directions

This dissertation addresses various aspects of Siona DCM via comparativist, variationist, and elicitation-driven techniques in Chapters 4 through Chapter 6, and follow-up studies are postulated at various points throughout these chapters. The opening of this section provides a summary of the future investigations proposed for the continued implementation of these techniques. Other appealing techniques exist in the literature, which are not discussed here.

Firstly, the typological-comparative dataset developed in Appendix A to feed Chapter 4 may be elaborated upon in several ways. The modest typological sample (N=21) can be expanded to include certain neglected Eastern Tukanoan languages, and to have geographical coverage beyond the restricted NWA region scrutinized in this chapter. More and better grammatical descriptions are constantly being released which may be integrated into the study accordingly. Since several components of Appendix A are designed specifically to contextualize the Siona facts − such as the selection of tested CASE-MARKING CORRESPONDENCES; less language-specific sub-parts may be expanded upon more easily than others.

Besides simply expanding the case-marking typology in Appendix A, Chapter 4 also calls for a more targeted typologization of (nominal) discourse-markers. A comparative study of discourse markers in the region is lacking at present, despite the fact that most grammars consulted in this dissertation have such a class of morphemes. Such a study is particularly appealing given the presence of various cognate forms for the Siona marker *=kato*, noted in the region (see Section 4.3).

It is relevant to mention one final follow-up study to Chapter 4: This chapter highlights various cognate candidates and plausible points of grammatical convergence concerning Siona and other (non-)Tukanoan languages in the typological sample in Appendix A. A more appropriate technique ought to be applied in order to further elucidate the diachronic pathways for the case-markers found in Siona+ languages, and this chapter lays some preliminary groundwork for such a study.

Turning to the corpus-based study reported in Chapter 5, there is a clear need for a follow-up to expand the narrative sample for the sake of ascertaining more viable statistics for spatial DCM patterns. This ought to be a first order of business regarding follow-up textual research.

A second aspect of Siona DCM which warrants its own investigation concerns disambiguation effects. A text-based analysis is conceivable, which analyzes certain plausible instances of overt case-marking which do not fall out naturally from discourse-related uses of PROMINENT-marking. Variationist techniques may be appropriate to account for the effect of ambiguity avoidance for certain case-marking discrepancies, such as that noted between inanimate and animate P-marking in Chapter 5. In general, disambiguation effects require further attention and further elicitation-based fieldwork is also in order.

Finally, Chapter 6 applied elicitation-based technique to bolster the empirical coverage for focal case-marking patterns. The evidence sets the descriptive backdrop for a formal analysis, incorporating negative data points and collecting introspective speaker judgments. The next step is to flesh out the formal analysis. Section 7.3 closes the present discussion by laying out the groundwork for a formal Minimalist analysis of Siona DCM, where focus is treated as a morphosyntactic feature (e.g., Kratzer and Selkirk 2020).

7.3 Towards a formal analysis of Siona DCM

The general approach to a formal analysis suggested here assumes all empirical evidence presented in Chapter 3 and Chapter 6. An adequate formal theory of Siona case-marking (including DCM patterns) must satisfy the following desiderata:

- Animacy effects arise such that *all* non-SUBJECT, animate arguments must bear overt case-marking. The ideal analysis accounts for the class of promoted nominals, which inherit this animacy status for the sake of case-marking.
  - This case-marking requirement is separate from the obligatory case-marking on the R-argument, and found with pseudo-spatial *-na* marking.
- The animacy-based STRONG TRIGGER for non-SUBJECT DCM must outrank the focus-related STRONG TRIGGER, which is active in all principled DCM patterns.
- The principled interaction of these STRONG TRIGGERS must outrank other plausible WEAK TRIGGERS, which factor into the distribution of case-marking alternatives in practice.

The set of desiderata laid out above certainly implicate a multi-dimensional analysis of Siona DCM (Klein and de Swart 2011). Additionally, based upon the rigidity of the animacy and focus effects outlined in this chapter, there is ample evidence that two separate features (or feature sets) ought to be independently active in the syntax: e.g., (i) animacy as φ [anim], and (ii) focus as δ [foc] (Miyagawa 2009, *et seq.*).[4] This analysis presumes that focus may be codified as a feature, implicated in the calculus of case-marking, and that the language does not make available focus-specific morphology.

The conspiracy of these putative features can be taken to underpin the proposed SPLIT *-re* ANALYSIS, where two featural decompositions can be posited for obligatory *-re* marking on the P-argument, as in (8). By extension, the featural decomposition is conceived as in (8b):[5]

(8) a. $\textit{-re}_{[\varphi]}$: PLAIN *-re* on animate P ;
b. $\textit{-re}_{[\delta]}$: PROMINENT *-re* on focal inanimate P ;
c. $\textit{-ni}_{[\varphi \wedge \delta]}$: PROMINENT *-ni* on focal animate P

Several questions persist as to how these features are projected in the syntax. For instance, how do singular, inanimate nouns acquire the feature φ under particular modification contexts? It is also unclear which model can account for *-re* as realizing an underlying featural disjunction (i.e., $\varphi \vee \delta$), compared with *-ni* as a featural conjunction, as in (8c). An adequate theory would capture these generalizations based upon the empirical facts laid out in this dissertation.

The remainder of this section outlines various families of analyses which have been postulated in the literature for DCM to date, although most analyses are restricted to *differential object marking*. According to Kalin (2014, 2017, 2018), in the Minimalist

[4] Perhaps the latter feature is similar to the [±Foc] feature suggested in Kratzer and Selkirk 2020. The authors suggest that the distribution of this feature is sensitive to language-particular factors, relating to the active focus contexts where the focus-encoding strategy is recruited.

[5] I am currently working with Paloma Jeretič on an analysis along these lines.

literature, there are three primary families of syntactic analyses: (i) movement-based accounts, (ii) size-based accounts, and (iii) licensing-based accounts. Below, I demonstrate how the final type of account is likely the best avenue for analyzing Siona DCM.

Firstly, as concerns movement-based approaches to DCM, many authors analyze case-marking (on the DIRECT OBJECT) as arising due to A-movement of the argument from its base-generated position as the verbal complement (e.g., de Hoop 1996; Rodríguez-Mondoñedo 2007; Baker and Vinokurova 2010; López 2012; Ormazabal and Romero 2013a, 2013b; Baker and Kramer 2014; Baker 2015). This movement permits interaction with higher Case-assigners, and is often taken to induce a specific or definite interpretation (e.g., Diesing 1992); which is considered a primary driver of DCM alternations cross-linguistically. To my view, this approach falls flat for Siona for several reasons. On the one hand, position-based diagnostics are untenable in Siona on the basis of rampant scrambling, both in semi-spontaneous speech and in controlled elicitations. Other researchers working with DCM languages reject this class of analysis for this reason: e.g., Kalin (2018) regarding Neo-Aramaic, Levin (2019) regarding Palauan. On the other hand, it is not clear that this approach can be extended to account for DCM patterns besides P-oriented DCM in Siona.

Turning to size-based approaches to DCM, the general analysis is such that the lack of case-marking is associated with (pseudo-)incorporation, whereas case-marked noun phrases have a more elaborate internal structure (e.g., Massam 2001; Danon 2006, 2011; Dayal 2011; Lyutikoval and Pereltsvaig 2023). Whereas, Siona does exhibit certain fringe constructions that resemble incorporation (discussed in Bruil & Case, *ms*, *forthcoming*), this approach is certainly not tenable as a general analysis for Siona DCM. It is likely that Siona is one of several languages which exhibit both pseudo-incorporation and DCM as independent components of its grammar (Driemel 2023). Besides positional constraints, there is no semantic or syntactic reason to consider unmarked arguments in Siona as incorporated (i.e., unlike what is claimed for other Tukanoan languages, like Koitiria, as analyzed by Stenzel 2008). I follow Kalin (2017, 2018) and others in acknowledging conceptual reasons to eschew the Split DP-hypothesis, for the purposes of DCM or otherwise.

Finally, licensing-based Minimalist approaches to DCM have been recently cultivated in the works of Kalin (2014, 2017, 2018; and since employed in Mursell 2018; Levin 2019; Murphy and Meyase 2022; van der Wal 2022, and many more). These approaches offer a promising direction for analyzing Siona DCM. Oversimplifying considerably, case-marking arises based on the interaction of two principles: (i) only certain classes of nominals *require* licensing (e.g., animates); and (ii) certain nominal-licensers are always active, like the S-licenser, whereas others are optional, and only arise as a last-resort mechanism to check the Case-feature on a nominal which requires it. This is not altogether different the classical conception of 'of'-insertion in English in the absence of a structural Case-assigner — e.g., inside a nominalization: *the descruction* *(*of*) *the city* (Chosmky 2004[1970], 149).

The licensing-based model addresses some of the desiderata for an analysis of Siona DCM outlined above − i.e., the SUBJECT vs. non-SUBJECT divide is accounted for based upon the different licensing mechanisms which define these grammatical relations. It is even conceivable that the SPLIT *-re* HYPOTHESIS may be accounted for if both animacy and focus are encoded by such features (e.g., φ and δ accordingly). However, it is not immediately clear how PROMINENT S-marking can be modeled in this way. Future work will navigate these issues as part of an account of Siona DCM.

Appendices

APPENDIX A

Northwest Amazonia case-marking dataset

This Appendix presents the Northwest Amazonia (NWA) case-marking dataset, underpinning the typological discussion in Chapter 4 of this dissertation. Section A.1 outlines the sampling procedures and presents the structural questionnaire administered to derive this dataset. Section A.2 presents the resulting NWA case-marking dataset in full, and lays out the data gathered on a Question-by-Question basis.

A.1 Methodology matters

This section of the Appendix outlines the procedures followed to produce the NWA sample, which is utilized to contextualize Siona case-marking patterns in the discussion in Chapter 4. Section A.1.1 lays out the sampling protocols, including inclusion criteria for the survey, and enumerates the languages comprising the dataset. Section A.1.2 presents the NWA case-marking questionnaire which is administered to the language sample in order to develop the NWA dataset. The operational definitions and precise techniques implemented to build the dataset are spelled out for each Question accordingly.

A.1.1 Presenting the NWA sample

Sampling is a central component to establishing any comparative linguistic survey, since it is necessary to consult a range of descriptive sources in order to derive an informative comparative dataset. There are two broad families of sampling techniques: (i) bottom-up sampling and (ii) top-down sampling (cf. Muysken et al. 2015; Miestamo et al. 2016). Whereas the former is primarily concerned with gathering data without a prioristic assumptions; the latter determines the appropriateness of languages based upon a predetermined set of traits, with the aim of testing and refining

the typology. Given that the NWA dataset is designed to situate several Ecuadorian Siona case-marking traits, established in Chapter 3 in this dissertation, the latter approach is deemed appropriate. This section outlines the inclusion criteria followed in assembling the sample, and identifies the languages comprising the NWA sample.

The top-down design of the NWA sample assembled in this Appendix adheres to four primary inclusion criteria (Velupillai 2012, 50-59):

- The INFORMATIVITY CRITERION simply requires that adequate descriptive materials be available for a particular language to be included in the present study. The consulted sources must provide sufficient information for the NWA questionnaire to be administered in an informative fashion.
- The GENETIC CRITERION reflects the fact that the NWA sample is designed to represent the Western and Eastern branches of the Tukanoan family in a balanced way. Nonetheless, the Eastern branch is larger, and comprises several internal groupings — see discussion in Section A.1.2 in the introductory chapter of this dissertation. The NWA dataset assembles a modest, representative sample of this larger branch, and matches this to the number of tested Western Tukanoan languages, for the present purposes.
- The AREAL CRITERION defines the geographical coverage of the NWA sample to the putative NWA region, subsuming all Tukanoan languages and the various languages spoken in their vicinity. This region corresponds roughly to the region surrounding the Napo-Putumayo-Vaupés watersheds. The areal coverage of this study extends slightly to the west, into the adjacent foothills region, to include certain languages, which have been suggested to display contact-induced convergence with Western Tukanoan groups (e.g., Quechuan languages, Barbacoan languages, Shiwiar) — cf. Bruil 2015a; Piispanen 2021; van Gijn et al. 2023.
- There are two TYPOLOGICAL CRITERIA, which constrain the language sample included in the NWA dataset. These typological criteria ensure that only those languages whose grammars are compatible with the notions of DCM, espoused in this dissertation, are included in the comparative dataset:
  - Firstly, the *core case-marking baseline* requires that a language recruits overt case-marking to encode core grammatical functions. For instance, the isolate Waorani [*auc*], spoken nearby to Western Tukanoan languages, does not display core case-marking (Fawcett 2023).
  - Secondly, the *accusative-alignment criterion* blocks inclusion of the few ergative-aligned languages spoken in the region: e.g., the Kawapanan languages of northern Peru (cf. Valenzuela 2011, 2015; Rojas-Berscia and Bourdeau 2017; Rojas-Berscia et al. 2023).

On the basis of the aforementioned inclusion criteria, the NWA dataset gathers data from a modest sample (N=21) in order to contextualize the Ecuadorian Siona case-marking facts put forth in Chapter 3. Table A.1 enumerates the sampled languages and indicates bibliographical information for the descriptive sources consulted for the sake of data collection. In this Appendix, and in the corresponding discussion in Chapter 4, languages are identified via their ISO codes, indicated in the table below:

Language [ISO]	**Genetics**	**Area**	**Primary source(s)**
*Ecuadorian Siona [^{E}*snn*]	W-TUKANOAN	Foothills	Bruil 2014 [B14] ; *personal fieldwork*
Colombian Siona [^{C}*snn*]	W-TUKANOAN	Foothills	Wheeler 1970 [W70], 1987 [W87],
Ecuadorian Sekoya [^{E}*sey*]	W-TUKANOAN	Foothills	Johnson and Levinsohn 1990 [JL90]
Koreguaje [*coe*]	W-TUKANOAN	Foothills	Cook and Levinsohn 1985 [CL85] ; Cook and Criswell 1993 [CC93]
Máíhɨ̄kì[*ore*]	W-TUKANOAN	Centre	Farmer 2015 [F15] ; Velie 2008[1975] [V08]
Kubeo [*cub*]	E-TUKANOAN	Vaupés	Morse and Maxwell 1999 [MM99] ; Chacon and Genetti 2019 [CG19]
Tukano [*tuc*]	E-TUKANOAN	Vaupés	Ramirez 2019[1997] [R19] ; Duarte and de Oliveira Lopes 2020 [DL20]
Wanano [*gvc*]	E-TUKANOAN	Vaupés	Stenzel 2008 [S08], 2013c [S13]
Desano [*des*]	E-TUKANOAN	Vaupés	Miller 1999 [M99] ; Silva 2012 [DS12]
Barasano [*bsn*]	E-TUKANOAN	Vaupés	Jones and Jones 1991 [JJ91]
Tanimuka [*tnc*]	E-TUKANOAN	Vaupés	Eraso 2015 [E15]
A'ingae [*con*]	*isolate*	Foothills	Fischer and Hengeveld 2023 [FH23]
Shiwiar [*acu*]	CHICHAM	Foothills	Kohlberger 2020 [K20]
Tena K [*quw*]	QUECHUAN	Foothills	Grzech 2016 [G16]
Imbabura K [*qvi*]	QUECHUAN	Highlands	Cole 1982 [C82] ; Jake 1983 [J83]
Awa Pit [*kwi*]	BARBACOAN	Foothills	Curnow 1997 [C97]
Murui [*huu*]	WITOTOAN	Centre	Wojtylak 2021 [W21]
Bora [*boa*]	BORAN	Centre	Thiesen and Weber 2012 [TW12]
Tariana [*tae*]	ARAWAKAN	Vaupés	Aikhenvald 2003 [A03]
Kakua [*cbv*]	KAKUA-NUKAK	Vaupés	Bolaños 2016 [B16]
Hup [*jup*]	NADAHUPAN	Vaupés	Epps 2008 [E08]
Yuhup [*yab*]	NADAHUPAN	Vaupés	Ospina Bozzi 2002 [O02]

Table A.1: Language specimens in NWA dataset and consulted primary sources

In addition to ISO short-hands, this Appendix and Chapter 4 utilize a series of secondary identifiers to facilitate the comparison of Tukanoan and non-Tukanoan languages included in the sample. Western Tukanoan languages are identified by the symbol ♦, and Eastern Tukanoan by ◊. Ecuadorian Siona, the standard of comparison for this typology, is additionally identified with an asterisk (*). Non-Tukanoan languages spoken within the Vaupés region (see Section 1.3.2) are identified in-text via the short-hand V. Figure A.1 plots the sampled languages geographically in a map diagram, where the Vaupés region is loosely traced with a dotted gray line. This map visually delineates the putative NWA area covered by this dataset:

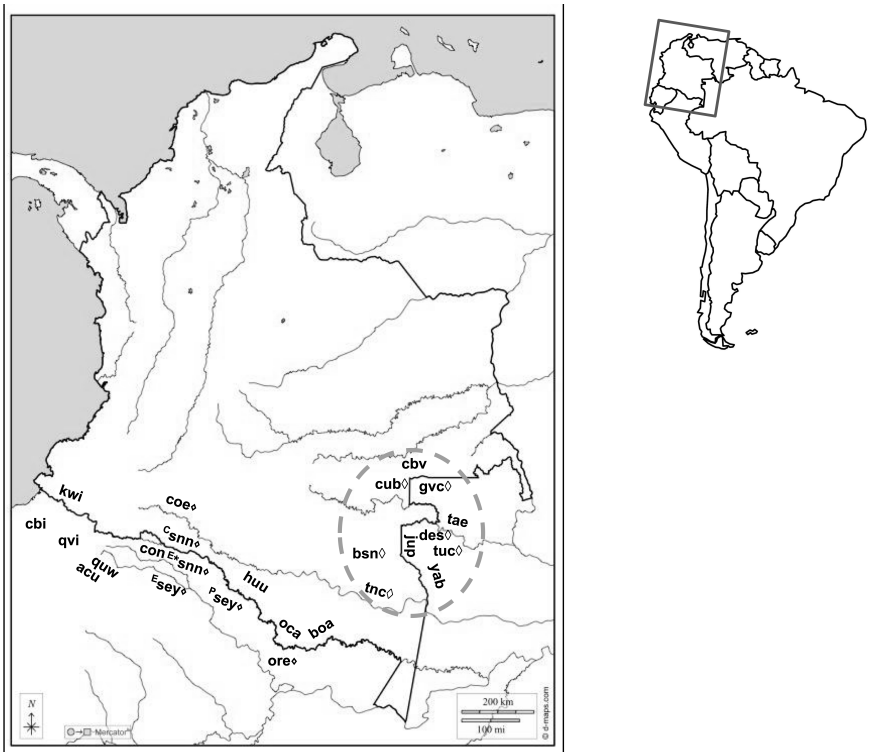


Figure A.1: Geographical distribution of languages . Map template downloaded from *https://d-maps.com* (Colombia borders, including rivers).

A.1.2 Formulating and administering the questionnaire

This section presents the NWA case-marking questionnaire, which is administered to the sample outlined in Table A.1, in order to develop the NWA dataset. The NWA case-marking questionnaire is a structural questionnaire, which is a tool often implemented in linguistic typological research (e.g., Haspelmath 2008; Krasnoukhova 2012; Birchall 2014; Janic and Haspelmath 2023). Structural questionnaires are generally organized into three levels (Birchall 2014, §2.2.2): DOMAIN > QUESTION[1] > VALUE (i.e., answer to QUESTION). DOMAINS refer to bundles of thematically-related QUESTIONS, and the researcher determines the set of suitable VALUES, which may be attributed to a particular language per QUESTION, in order to facilitate cross-linguistic comparisons.

The NWA case-marking questionnaire is developed around three DOMAINS, as inspired by the three pillars of DCM description, proposed in Chapter 2:

1. The FORMAL DOMAIN (QUESTIONS 1.1 - 1.2)
2. The FUNCTIONAL DOMAIN (QUESTIONS 2.1 - 2.7)
3. The DCM DOMAIN (QUESTIONS 3.1 - 3.5)

[1]Birchall 2014 utilizes the term FEATURE, rather than QUESTION; however, I avoid this terminology given the range of meaning that *feature* has in linguistic research.

(1) spells out the NWA case-marking questionnaire in full. Accepted VALUE-types for each question are provided in gray text. The precise instructions for deriving these VALUES are laid out in the remainder of this section on a DOMAIN-by-DOMAIN basis:

(1)
a. *Q1.1: How many case markers does the language have?* {1,2,3,...,n}
b. *Q1.2: Is multiple case-marking permitted?* {0,1}
c. *Q2.1: What type of **Subject-Instrument** correspondence is observed?* {a=identity, b=differentiation, c=mixed}
d. *Q2.2: What type of **Object-Recipient** correspondence is observed?* {a,b,c}
e. *Q2.3: What type of **Companion-Instrument** correspondence is observed?* {a,b,c}
f. *Q2.4: What type of **Object-Location** correspondence is observed?* {a,b,c}
g. *Q2.5: What type of **General Location** correspondence is observed?* {a,b,c}
h. *Q2.6: Is there case-marking for **Possessor**?* {0,1}
i. *Q2.7: Is there case-marking for **Temporal adjuncts**?* {0,1}
j. *Q3.1: Are DCM patterns observed on S?* {0,1}
   i. *Q3.1.1: Which formal DCM pattern is observed?* {a=optional (∅), b=alternating (+), c=hybrid (+∅)}
   ii. *Q3.1.2: What triggers drive the observed DCM pattern?* {a=animacy (a), b=referentiality (r), c=discourse (d), d=other (o)}
k. *Q3.2: Are DCM patterns observed on P?* {0,1}
   i. *Q3.2.1: Which formal DCM pattern is observed?*
   ii. *Q3.2.2: What triggers drive the observed DCM pattern?*
l. *Q3.3: Are DCM patterns observed on R?* {0,1}
   i. *Q3.2.1: Which formal DCM pattern is observed?*
   ii. *Q3.2.2: What triggers drive the observed DCM pattern?*
m. *Q3.4: Are DCM patterns observed on L?* {0,1}
   i. *Q3.2.1: Which formal DCM pattern is observed?*
   ii. *Q3.2.2: What triggers drive the observed DCM pattern?*
n. *Q3.5: Are DCM patterns observed on G?* {0,1}
   i. *Q3.2.1: Which formal DCM pattern is observed?*
   ii. *Q3.2.2: What triggers drive the observed DCM pattern?*

A.1.2.1 Coding the Formal domain of the questionnaire

The FORMAL DOMAIN comprises two QUESTIONS, which pertain to the expression of morphological case-marking. The first question concerns the size of the case-marker

inventory (*Q1.1*), adopting the methodology for identifying the number of overt case-marking morphemes in *WALS (49): Number of cases* (Iggesen 2013).[2]

Per the working definition in Iggesen (2013), bona fide case-markers are morphemes arising on the nominal dependent, which indicate the grammatical function or the thematic role of a given nominal. Case-markers must be phonologically integrated into the nominal word, whereas independent morphemes are disregarded. This definition permits the identification of several case-markers to encode a particular FUNCTION; although, crucially, allophonic variants are not tallied separately. The determination of case-marker inventory size is exemplified for Trumai in (2):

(2) Trumai [*tpy* — *isolate*, Brazil]: **5-member inventory** [Iggesen 2013, ex. (4)]
Absolutive: *axos* [*axos* (child)]
Ergative: *axos-**ak***$_1$
Dative: *axos-**atl***$_2$, *axos-**ki***$_3$
Genitive: *axos-**kate***$_4$
Locative: *esak-**en***$_5$ [*esak* (hammock)]

Like Iggesen (2013), this survey counts POSSESSOR-marking in the inventory tally, corresponding to the Genitive function identified in (2). This survey excludes various controversial categories, such as the vocative, or focus- and topic-markers, from the determination of case-marker inventory size. The result is a numerical VALUE.[3]

The second QUESTION in the FORMAL DOMAIN (*Q1.2*), classifies languages based on whether the MULTIPLE CASE-MARKING pattern is attested or not. As such, this QUESTION accepts a binary VALUE. A given language is determined to display this pattern where more than one case-marking morpheme may co-occur on a nominal without incurring a grammatical violation. This pattern is demonstrated in (3) for Warlpiri,[4] and a comparable instance from Tukano is provided in (4):

(3) Warlpiri (*wvp* — PAMA-NYUNGAN)
[Hale 1982, cited in Sadler and Nordlinger 2006, 463, ex.(3)]
*ngarrka-ngku ka-rlajinta yankirri-ki luwa-rni **ngapa-ngka-ku***
man-ERG PRS-3SG.DAT emu-DAT shoot-N.PST water-LOC-DAT
'The man is shooting at the emu at the watering hole.'

[2]Other specialists have developed alternative approaches to measuring case-marker inventory size: e.g., the *University of Texas Hunter-Gatherer linguistic database* (https://huntergatherer.la.utexas.edu/grammar) and the *South American Indigenous Language Structures database* (SAILS — https://sails.clld.org/) directed by Pieter Muysken, both particular to South America.

[3]The coding of *Q1.1* for the NWA dataset departs from the classification in Iggesen (2013) in that the present study disregards zero-marking. Rather, the QUESTION developed here is strictly concerned with *overt* case-marking morphemes; whereas Iggesen (2013) includes zero-marking in his tally. On this basis, a language with only one overt case-marker is typed as a displaying a two-member inventory (see example with Mapudungun in Iggesen 2013). In order for the results of the present questionnaire to be compatible with the *WALS (49)* dataset, totals ought to be considered as $N + 1$.

[4]The formal pattern of multiple case-marking is commonplace in the languages of Australia (cf. Dench and Evans 1988, Nordlinger 1997, *a.o.*). It is not immediately clear whether Eastern Tukanoan-type multiple case-marking is the same morphosyntactic phenomenon found in case-stacking configurations in Australian languages, East Asian languages, etc.

(4) Tukano$^{\diamond}$ (*tuc* — Eastern-Tukanoan)[5] [R97:169, ex. (340)]
toó-pi-de *peêdu* *sĩ'dí-wɨ*
ana.loc-loc-obj caxiri_drink drink-rep.pst.vis
'I will drink *caxiri* <u>there</u> (an established location).'

The Formal domain in this questionnaire comprises two Questions, although several other formal properties of case-marking have been subject to typological inquiry. For instance, the affix vs. clitic status of case-markers has received attention (e.g., *WALS (51): Position of case affixes* (Dryer 2013)). However, given the non-standardized treatment of such phenomena in the descriptive sources, these properties are excluded from the nwa dataset, and reserved for future research.

A.1.2.2 Coding the Functional domain of the questionnaire

The coding of the Functional domain comprises seven Questions, which fall into two categories: The first five Questions (*Q2.1*-*Q2.5*) concern case-marking correspondence, adapting *WALS (52): Comitatives and instrumentals* (Stolz et al. 2013), as outlined in what follows; whereas, the final two Questions (*Q2.6* and *Q2.7*) determine whether overt case-marking encodes particular Functions: i.e., the Possessor and Time functions, corresponding to the Functional gaps discussed in Section 4.2.2.2 in Chapter 4. Before outlining the instructions for addressing each of these Questions below, a word is in order regarding the establishment of Case functions, which guided data collection.

The aim of establishing a set of Case functions is to avoid gathering data in an ad hoc way for the Functional domain, as much as possible. This is common practice in typological research, where conventions are established for defining and operationalizing comparative concepts (e.g., Bickel 2007, 2009; Croft 2001, 2003, 2009; Haspelmath 2005, 2007, 2008, 2010; Newmeyer 2002). For the present purposes, the aim is to establish a set of Case functions, like those described for Ecuadorian Siona in Chapter 3 in this dissertation, which are compatible with other tested languages for the sake of this comparative exercise. Although it is a notoriously supple task to determine cross-linguistically valid definitions for grammatical relations, such as Subject or Object, other thematic roles are easier to identify across languages, e.g., the Instrument-argument.

The nwa dataset aligns with most authors listed above, by prioritizing prototypical uses of arguments with generalizable predicates to define roles as comparable Case functions: For instance, this approach is taken to define the case-marking function, *Dative*, as a comparative concept par excellence, based upon the encoding of the Recipient role as its prototypical thematic correlate, in the work of Haspelmath (2005, 2011), and others. The accusative-alignment condition on inclusion in the current sample allows for a similar tact to be taken to define core Case functions in a valid way. For example, *the man*$^{(S)}$ *killed the deer*$^{(P)}$, where the Subject-argument is a volitional Agent and the Direct object-argument is an undergoer Patient. Generally, the consulted descriptive sources already postulate generalizations based on such prototypical instances, which facilitates this comparative exercise considerably.

[5] Ramirez (2019[1997]) glosses the case-markers *-de* (cognate with Siona+ *-re*) and *-pi*, ref and foc respectively. These are semantic labels which reflect what the author determines to be the dominant triggers characterizing their usage. Alternative functional labels are used in (4) for convenience.

For the sake of coding the NWA dataset, a set of ten such CASE FUNCTIONS is established, as commonly attested in the typological record. (5) enumerates these FUNCTIONS (and short-hand labels, where relevant), along with prototypical sentences which guide the determination of which case-markers are available to each FUNCTION across the NWA sample:

(5)
a. SUBJECT (S) — e.g., *the man*$^{(S)}$ *is sleeping*; *the man*$^{(S)}$ *killed the deer*
b. DIRECT OBJECT (P) — e.g., *the man killed the deer*$^{(P)}$; *he found the house*$^{(P)}$
c. INDIRECT OBJECT (R) — e.g., *send the man*$^{(R)}$ *a card*; *show me*$^{(R)}$ *the money*
d. INSTRUMENT (INST) — e.g., *the man attacked his enemy with a weapon* $^{(INST)}$
e. COMPANION (COMP) — e.g., *the man came to the party with a friend* $^{(COMP)}$
f. POSSESSOR (POSS) — e.g., *the man's* $^{(POSS)}$ *friend works here*
g. LOCATION (L) — e.g., *my house is in the town*$^{(L)}$; *he is sleeping in his canoe*$^{(L)}$
h. GOAL (G) — e.g., *he is going to the party*$^{(G)}$; *she brought rice to the kitchen*$^{(G)}$
i. SOURCE (SRC) — e.g., *my friend came from the jungle* $^{(SRC)}$
j. TIME (TEMP) — e.g., *he arrives on Monday* $^{(TEMP)}$; *in summer* $^{(TEMP)}$ *it's hot*

Next, based upon the available descriptions, case-markers are mapped onto these ten CASE FUNCTIONS, as demonstrated for Ecuadorian Siona in Table A.2, based upon the description in Chapter 3:

CASE FUNCTIONS	AVAILABLE CASE-MARKERS
SUBJECT (S)	*-bi*
DIRECT OBJECT (P)	*-re*, *-ni*
INDIRECT OBJECT (R)	*-re*, *-ni*, *-na*
INSTRUMENT (INST)	*-bi*
COMPANION (COMP)	*-hã're*
POSSESSOR (POSS)	N/A
LOCATION (L)	*-re*, *-ni*, *-na*
GOAL (G)	*-na*
SOURCE (SRC)	*-bi*
TIME (TEMP)	N/A

Table A.2: Ecuadorian Siona case-markers mapped onto CASE FUNCTIONS

A similar mapping exercise is completed for each included language in the NWA sample, permitting the QUESTIONS in the FUNCTIONAL DOMAIN may be answered post hoc.

The first five QUESTIONS in the FUNCTIONAL DOMAIN concern so-called case-marking correspondences, as developed in *WALS (52): Comitatives and instrumentals* (Stolz et al. 2013). Put simply, the authors distinguish three 'correspondence types' based upon the case-markers attested across these two CASE FUNCTIONS. The first two types are straightforward: The IDENTITY-type is where the exact same set of case-markers encodes both the INSTRUMENT and the COMPANION functions, as exemplified for Estonian in (6). The DIFFERENTIATION-type is where different marker(s) are attested across the concerned CASE FUNCTIONS, as shown for Finnish in (7):

(6) Estonian (*ekk* — Uralic) [Lavotha 1973, cited in Stolz et al. 2013, ex. (2)]

a. *ma kirjuta-n* ***sule-ga***
1sg write-1sg pen-inst
instrument: 'I am writing with a pen.'

b. *Villem jaluta-b* ***isa-ga***
V go_for_walk-3sg father-com
companion: 'Villem is going for a walk with his father.'

(7) Finnish (*fin* — Uralic) [Karlsson 1978, cited in Stolz et al. 2013, ex. (3)]

a. *hän kirjoittaa* ***kynä-llä***
3sg write.3sg pen-inst
instrument: 'S/he is writing with a pen.'

b. *läsnä oli* *V.V.* ***vaimo-ine-en***
near be.pst.3sg V.V. wife-com-poss.3
companion: 'V.V. was present with his wife.'

The third type concerns the Mixed-type, shown in (8) for Hungarian. This refers to the pattern where there is some overlap regarding the sets of available case-markers across these Case functions, but where these sets are not identical. In Hungarian, the same case-marker *-val*,[6] encodes both the Companion and Instrument functions — i.e., shown in (8a) and (8b). However, the so-called *associative* marker may also be used to encode the Companion function, as in (8c):

(8) Hungarian (*hun* — Uralic) [Bánhidi et al. 1975, cited in Stolz et al. 2013, ex. (4)]

a. ***tol-lal*** *ír-ok*
pen-with write-1sg
instrument: '(I) am writing with a pen.'

b. *Jan is megjelenik* ***barát-já-val***
J also appear.3sg friend-poss.3sg-with
companion[1]: 'Jan too shows up with his friend.'

c. *csónak-há'-ak sport-és játszóter-ek vár-ják* ***család-ostul***
boat-house-pl sport-adj playground-pl wait-3pl family-com.assoc
gyerek-estül *az ember-ek-et*
child-com.assoc def man-pl-acc
companion[2]: 'Boathouses and sports grounds are waiting for the people with family and children.'

The first five Questions in the Functional domain adapt the typological methodology in Stolz et al. (2013), and apply this three-way distinction to a handful of case-marking correspondences based upon the ten Case functions laid out in (5). The tested correspondences in the nwa qustionnaire, outlined below, are selected based upon a subset of those noted in Ecuadorian Siona grammar, per Table A.2. The nwa dataset contextualizes the Ecuadorian Siona correspondences enumerated below:

[6]The morpheme *-val* is realized as the allomorph *-lal* in (8a).

- *Q2.1*: Subject-Instrument [Siona: Identity-type.]
- *Q2.2*: Object-Recipient [Siona: Mixed-type.]
- *Q2.3*: Companion-Instrument [Siona: Differentiation-type.]
- *Q2.4*: Object-Location [Siona: Mixed-type.]
- *Q2.5*: Location-Goal-Source [Siona: Differentiation-type.]

The remaining two Questions in the Functional domain consider whether overt case-marking is attested to encode certain Case functions: i.e., the Possessor function (*Q2.6*) and the Time function (*Q2.7*). According to Table 4.3, Ecuadorian Siona does not recruit case-marking to encode either of these Case functions — i.e., Functional gaps in Ecuadorian Siona. However, these Functions are readily encoded via case-marking in other nwa languages, and in the broader typology. Accordingly, all Questions in the Functional domain are designed to typologize the Ecuadorian Siona patterns in particular, in keeping with the aims of Chapter 4.

A.1.2.3 Coding the dcm domain of the questionnaire

The final domain of the nwa questionnaire is most directly concerned with comparing dcm patterns. First and foremost, this domain is designed around five Questions (*Q3.1-Q3.5*), which test whether the five attested dcm patterns, described for Ecuadorian Siona in Chapter 3, arise in the languages included in the nwa sample:

(9) a. *Q3.1: Are DCM patterns observed on S?* {0,1}
b. *Q3.2: Are DCM patterns observed on P?* {0,1}
c. *Q3.3: Are DCM patterns observed on R?* {0,1}
d. *Q3.4: Are DCM patterns observed on L?* {0,1}
e. *Q3.5: Are DCM patterns observed on G?* {0,1}

Wherever a dcm pattern is observed, two Sub-questions are addressed to flesh out the comparative typology, as in (10). The first Sub-question determines the formal-type of the observed dcm pattern, following terminology and definitions established in Section 2.2.1 in Chapter 2 of this dissertation. The second concerns the set of active Macro-triggers (i.e., Section 2.2.3), which drive the observed alternation:

(10) *Q3.1: Are DCM patterns observed on S?* {0,1}
a. *Q3.1.1: Which formal DCM pattern is observed?*
{a=optional, b=alternating, c=hybrid}
b. *Q3.1.2: What triggers drive the observed DCM pattern?*
{a=animacy (*a*), b=referentiality (*r*), c=discourse (*d*), d=other (*o*)}

As pertains to the first Sub-question, (10a), this dissertation recognizes three formal dcm types. The first two are well-established in the literature: i.e., the Optional-type, where an overt case-marker alternates with zero-marking, shown in (11); and the Alternating-type, where more than one case-marker may encode a particular Funciton, as shown in (12):

(11) Persian (*fas* — INDO-IRANIAN)
[Lazard 1994, cited in Chappell and Verstraete 2019, 5, ex.(5)]

a. ***ketâb-râ*** *xând-am*
book-ACC read.PST-1SG
'I read the book.'

b. ***ketâb*** *xând-am*
book read.PST-1SG
'I read a book/books.'

(12) Evenki (*evn* — TUNGUSIC)[7]
[Nedjalkov 1997, cited in Chappell and Verstraete 2019, 9, ex.(11)]

a. ***oron-mo*** *java-kal*
reindeer-ACC1 take-PRS.IMP.2SG
'Catch that reindeer.'

b. *min-du* ***ulle-ye*** ***kolobo-yo*** *by:-kel*
1SG-DAT meat-ACC2 bread-ACC2 give-PRS.IMP.2SG
'Give me (some) meat and (some) bread.'

The third type recognized in this dissertation, and the NWA questionnaire, is the HYBRID-type, proposed in Section 2.2.1. This refers to the pattern where multiple case-markers alternate with zero-marking, as demonstrated for Colombian Siona SUBJECT-marking in (13):

(13) Colombian Siona♦ (*snn* — WESTERN-TUKANOAN)[8]
[W70:47]

a. ***yɨ'ɨ-ga*** *saí-yɨ*
1SG-SBJ2 go-N3S.PRS.ASS
'And I, I am going.'

b. ***ča'kádi-wa'-na-bi*** *bá-gɨ-na* *gahe-ni*
fire_ant-ANIM-PL-SBJ1 PRO.HUM-CL:M-GOAL go_down-SS
sɨ'á-wɨ
attach-N3S.PST.ASS
'The fire ants came down to him and latched on.'

c. ***'okó*** *ka'ní-hi*
rain gather-3S.M.PRS.ASS
'The rain is gathering.'

In addition to comparing formal DCM types, the second SUB-QUESTION, (10b), identifies which TRIGGERS are active for each attested DCM pattern. The tested families of TRIGGERS align with those laid out in Section 2.2.3. Drawing inspiration from the *differential object-marking* survey in Sinnemäki (2014), this questionnaire types DCM patterns on the basis of MACRO-TRIGGERS, collapsing families of related TRIGGERS,

[7]Note that the author utilizes the glosses DEF.ACC for ACC1, and INDEF.ACC for ACC2 respectively in the original work. Note that for the present purposes, this sub-question is strictly formal in nature.

[8]As in the case of the Evenki example in (12), the suggested glosses here opt for a numbering system: e.g., SBJ1, SBJ2. However, Wheeler uses the gloss SBJ for the case-marker *-bi*, and the gloss SBJ.EMPH for *-ga* — i.e., the emphatic subject marker per his analysis.

which are commonly attested in the literature. Table A.3 enumerates the MACRO-TRIGGER labels used in this study and indicates some corresponding descriptive notions, which are commonly found in the consulted sources:

MACRO-TRIGGER LABEL	Descriptive notions
[*a*]	**animacy**: also, e.g., *humanness*, *empathy*
[*r*]	**referentiality**: also, e.g., *definiteness*, *specificity*, *pronoun*
[*d*]	**discourse**: also, e.g., *topic*, *focus*, *emphasis*
[*o*]	**other**: e.g., *predicate/clausal effects*, *proximity to speaker*

Table A.3: Macro-trigger labels used in this comparative survey

Given the fact that many DCM patterns are multi-dimensional (cf. Klein and de Swart 2011), several patterns exhibit multiple MACRO-TRIGGERS for a given DCM pattern. A determination is made, where possible, as to whether a particular TRIGGER is a STRONG TRIGGER — e.g., animates are marked one way, and inanimates another; or as a WEAK TRIGGER, which more reflects a marking tendency. WEAK TRIGGERS are identified within parentheses as a matter of convenience.

The relevant Ecuadorian Siona DCM facts, which are typologized in the DCM DOMAIN of the NWA dataset, are outlined in Table A.4:

CASE FUNCTIONS	FORMAL TYPE	AVAILABLE FORMS	MACRO-TRIGGERS
SUBJECT (S)	OPTIONAL	*-bi* (*-∅*)	[*d*]
DIRECT OBJECT (P)	HYBRID	*-re*, *-ni* (*-∅*)	[*a*, *d*, (*r*)]
INDIRECT OBJECT (R)	ALTERNATING	*-re*, *-ni*, *-na*	[*a*, *d*]
LOCATION (L)	HYBRID	*-re*, *-ni*, *-na* (*-∅*)	[*d*, (*r*), o_{PRED}]
GOAL (G)	OPTIONAL	*-na* (*-∅*)	[*d*, (*r*)]

Table A.4: Ecuadorian Siona DCM patterns

In summary, the DCM DOMAIN in the questionnaire is designed to typologize three aspects of Ecuadorian Siona DCM. Firstly, it checks which languages exhibit which of the five DCM patterns described in Chapter 3. Secondly, it sorts each of the attested DCM patterns on the basis of formal types. Finally, it compares the set of active TRIGGERS found across these DCM patterns. This methodology facilitates the comparison of various DCM facts across the NWA sample.

A.2 Presenting the NWA dataset

Table A.5 spells out the full NWA case-marking dataset. This presentation of the full dataset is followed by a detailed presentation of the VALUES extracted across the sample as divided based on domains. The FORMAL DOMAIN is explored in A.2.1, followed by a discussion of the FUNCTIONAL DOMAIN results in A.2.2, and finally by the DCM DOMAIN results in A.2.3:

	(*N*)	multiple	S=Inst	P=R	Comp=Inst	P=L	GenLoc	Poss	Temp	S	P	R	L	G	
*Siona (E)♦	**(6)**	x	a	c	b	c	b	x	x	∅✓[d]	+∅✓[a,r,d]	+✓[a,d]	+∅✓[d(r,o)]	∅✓[d(r)]	[E*snn*]♦
Siona (C)♦	**(6)**	x	a	c	b	c	b	x	x	+∅✓[d]	+∅✓[a,r,d]	+✓[d(a)]	+∅✓[d(r,o)]	∅✓[r]	[C*snn*]♦
Koreguaje♦	**(4)**	x	a	a	b	c	b	x	x	∅✓[d]	+∅✓[a,r,d]	+✓[d(a)]	+∅✓[d,r(o)]	∅✓[r,d]	[*coe*]♦
Máíhɨ̄kì♦	**(2)**	x	b	a	a	b	b	x	x	-∅	∅✓[a,r]	N/A	-∅	-∅	[*ore*]♦
Kubeo◊	**(5)**	✓	b	a	a	c	a	✓	✓	-∅	∅✓[r,d(a,o)]	N/A	+✓[r(d)]	+✓[r(d)]	[*cub*]◊
Tanimuka◊	**(7)**	x	b	a	b	c	c	✓	x	∅✓[a]	∅✓[a(d)]	N/A	N/A	N/A	[*tnc*]◊
Tukano◊	**(4)**	✓	b	a	a	c	a	✓	✓	-∅	∅✓[r(d)]	N/A	+∅✓[r(d)]	+∅✓[r(d)]	[*tuc*]◊
Desano◊	**(4)**	✓	b	a	a	c	a	✓	✓	-∅	∅✓[r,d]	N/A	+∅✓[r(d)]	+∅✓[r(d)]	[*des*]◊
Barasano◊	**(4)**	x	b	a	a	c	a	✓	✓	-∅	∅✓[r,d]	N/A	∅✓[r(d)]	∅✓[r(d)]	[*bsn*]◊
Wanano◊	**(4)**	✓	b	a	a	c	a	x	✓	-∅	∅✓[r,d]	N/A	+✓[r(o)]	+✓[r]	[*gvc*]◊
A'ingae	**(9)**	x	b	b	a	b	c	x	x	-∅	+✓[o]	N/A	N/A	N/A	[*con*]
Murui	**(6)**	x	b	b	a	b	c	x	✓	∅✓[d]	∅✓[r,d]	∅✓[d]	+∅✓[o]	+∅✓[o]	[*huu*]
Bora	**(5)**	x	b	c	b	b	b	x	x	-∅	+∅✓[a(o)]	N/A	N/A	N/A	[*boa*]
T-Kichwa	**(8)**	x	b	c	a	b	b	✓	✓	-∅	∅✓[o]	+✓[o]	+✓[o]	N/A	[*quw*]
I-Kichwa	**(7)**	x	b	c	a	b	b	✓	✓	-∅	N/A	N/A	N/A	+✓[o]	[*qvi*]
Shiwiar	**(5)**	x	b	a	a	b	c	✓	✓	-∅	∅✓[r]	N/A	+✓[r,o]	+✓[r,o]	[*acu*]
Awa Pit	**(9)**	x	b	a	a	c	c	✓	✓	-∅	∅✓[a,r]	N/A	+✓[o]	+✓[o]	[*kwi*]
Kakuaᵛ	**(4)**	✓	b	a	a	b	a	✓	x	-∅	∅✓[a,r]	N/A	N/A	N/A	[*cbv*]ᵛ
Tarianaᵛ	**(5)**	✓	a	a	a	b	a	x	✓	∅✓[d]	+∅✓[r,d]	+✓[r]	∅✓[o]	∅✓[o]	[*tae*]ᵛ
Hupᵛ	**(4)**	x	b	a	c	c	c	✓	✓	-∅	∅✓[a,r]	N/A	+✓[d(o)]	N/A	[*jup*]ᵛ
Yuhupᵛ	**(7)**	x	b	a	b	b	a	✓	x	-∅	∅✓[r(a)]	N/A	+✓[o]	+✓[o]	[*yab*]ᵛ
	Formal		Functional							dcm					
			a (ident) — b (diff) — c (mixed)							-∅ (no attested case-marking); n/a (obligatory case-marking)					

Table A.5: Results: nwa case-marking dataset

♦(West-Tukanoan); ◊(East-Tukanoan); v (Vaupés sub-area)

A.2.1 The dataset: The Formal domain

Table A.6 identifies all case-markers used to determine the case-marker inventory size (*Q1.1*) along with the glosses suggested in the original sources. All included, bona fide case-markers are presented in black text, following the operational definition laid out in Section A.1.2.1. Morphemes which do not meet these criteria, but which execute a case-like function, are labelled 'pseudo-case-markers', and are excluded from the calculation of the numerical VALUE for this question. 'Pseudo-case-markers' are identified with gray text, and with an asterisk (*), in this table.

A few instances of attested 'pseudo-case-markers' in the sample are provided below. (14) illustrates the independent comitative particle *-jã'me* in Koreguaje [*coe*]. Similar particles are described for encoding the PATH and COMPANION FUNCTIONS in Colombian Siona, as shown in (15). Another example concerns the so-called *terminative*-marker in Kakua [*cbv*], expressing a LIMIT-like FUNCTION, which is encoded by a non-integrated preposition per Bolaños (2016, §6.4.3), as shown in (16):

(14) Koreguaje♦ (*coe* — WESTERN-TUKANOAN) [CC93:51]

a. *rañawahɨ ǰai nɨkʰa-mɨ* ***rɨ̃jo-na*** ***jã'me***
Siona healer be_standing-M.SG wife-GOAL COM
'The Siona healer is standing <u>with his wife</u>.'

b. ***i-na-re*** ***jã'me*** *pa-sa ǰɨi-mɨ ǰɨ'i*
this-PL.ANIM-N.SBJ COM live-DES.SG want-M.SG 1SG
'(I (M)) want to live <u>with them</u>.'

(15) Colombian Siona♦ (*c-snn* — WESTERN-TUKANOAN) [W70:76]

a. ***Pedro wɨ'e-se'e-ga*** ***tɨká*** *saí-yɨ*
P house-EXCL-GA PATH go-N3S.PRS.ASS
'(I) am going <u>only toward Pedro's house</u>.'

b. ***wa'í maú-si'-kɨ*** ***nakóni*** *koká ká-wɨ*
fish catch-COMP-CL:M COM word say-N3S.PST.ASS
'(I) conversed <u>with the one who had caught the fish</u>.'

(16) Kakua$^{\vee}$ (*cbv* — KAKUA-NUKAK) [B16:210, ex. (87)]
pɨ ***joà'-tak=jũb*** ***pu'ba***
TERM front_yard-middle=EMPH EMPHZ.TAG
'ã̃=t-hɨ̃w'=wɨt=hĩ
3SG=EVID-come=REP.EVID=RMT.PST
'(He) came <u>up right to the middle of the front yard</u> (it is said).'

Language [*iso*]	(*N*)	**Inventory** (original gloss)
*E-Siona♦ [*snn*]$_i$	**(6)**	***-re*** (N.SBJ); *-ni* (N.SBJ2); *-bi* (SBJ,INST,SRC); *-na* (GOAL); *-ja̲'a̲* (PATH); *-ja̲'re* (COM)
C-Siona♦ [*snn*]$_i$	**(5)**	***-re*** (OBJ.FOC); *-ni* (OBJ.EMPH.FOC); *-bi* (SBJ.FOC,INST,SRC); *-ga* (SBJ.EMPH.FOC); *-na* (GOAL(.EMPH).FOC); *-ja'an* (along) — **naconi* (COM); **tëca* (PATH) [W62,W70]
Koreguaje♦ [*coe*]	**(4)**	***-re*** (ESP≃specific); *-pi* (FUEN≃source); *-ni* (PAC≃patient); *-na* (META≃goal) — **ja̲'me* (COM); **jatëca* (LIM) [CC93]
Máíhɨ̃kì♦ [*ore*]	**(2)**	***-re*** (NON.SUBJ); *-jànù* (COM/INST) [F15]
Kubeo◊ [*cub*]	**(5)**	***-re*** (NNOM); *-i* (POS,LOC); *-rã* (FOC.LOC); *-ke* (INST/COM); *-ta* (PATH) [CG19, MM99]
Tukano◊ [*tuc*]	**(4)**	***-re*** (REF); *-pʉ* (FOC); +*bẽ'da* (COM); +*yaá* (POSS) [R97]
Wanano◊ [*gvc*]	**(4)**	**+*re*** (OBJ); +*i* (LOC1); *-pʉ* (LOC2) ; ≃*be're* (COM/INST) [S13]
Desano◊ [*des*]	**(4)**	***-re*** (REF); =*ya* (POSS); -˜*beda* (COM/INST); *-ge* (LOC) [DS12]
Barasano◊ [*bsn*]	**(4)**	***-re*** (O); *rãka* (*with*) ; *-hʉ* (*place,time*); *-ya* (GEN) [JJ91]
Tanimuka◊ [*tnc*]	**(7)**	***-té*** (DEIC1); *-re'ká* (LOC); *-ϕí* (INST,ABL); -˜*rá* (ALL); *-a'ká* (COM); *-ri'ká* (GEN); *-ro'sí* (BEN) [E15]
A'ingae [*con*]	**(9)**	=*ma* (ACC.REAL); =*ve*/=*me* (ACC.IRR); =*nga* (DAT); =*i'khû* (INS); =*ni* (LOC); =*ye*/=*ñe* (ELAT); ='*pi* (LIM); =*ngae* (MN/PTH (manner/path)); =*ne* (SO (source)); =*mbe* (BEN) [FH23]
Shiwiar [*acu*]	**(5)**	=*nʲa* (OBJ); =*hʲãĩ̃* (COM); =*nʲama*∼=*i* (LOC); =*ja* (ABL); =*tin* (TEMP) [K20]
T-Kichwa [*quw*]	**(8)**	*-ta* (ACC); *-ma* (DAT); *-manda* (ABL); *-wa* (INSTR); *-pa* (GEN); *-pi* (LOC); *-gama* (LAT (lative ≃PATH)); *-k* (BEN) [G16]
I-Kichwa [*qvi*]	**(7)**	*-ta* (ACC); *-man* (*to*); *-manda* (*from*); *-wan* (*with*); *-paj* (*possessive,benefactive*); *-pi* (*in*); *-kaman* (*up to*) [C82]
Awa Pit [*kwi*]	**(9)**	=*ta* (ACC,*in*); =*kasa* (*with*); =*pas* (*from*); =*pa* (*in*(*approx*)); =*mal* (LOC); =*ki* (*at*); =*kima* (*until*); =*pa*∼=*wa* (POSS); =*ma* (TEMP) [C97]
Murui [*huu*]	**(6)**	=*dɨ* (S/A.TOP); *-na* (N.S/A.TOP); *-mo* (LOC,TEMP); *-mona* (ABL); *-do* (INS); *-ri* (TRANS≃PATH) — **-nino* (PRV≃*without*) [W21]
Bora [*boa*]	**(5)**	*-kʰè* (OBJ.AN); *-βù* (GOAL); *-tʰù* (SOU(source)); *-rì* (OBL.IN); *-ma* (*with*) [TW12]
Tariana[V] [*tae*]	**(5)**	*-n(h)e* (FOC.A/S); *-na* (OBJ); *-nuku* (TOP.NON.A/S); *-se* (*on,to,from*); *-ne* (COM,INS) [A03]
Kakua[V] [*cbv*]	**(4)**	=*di'* (OBJ); =*bŭ* (LOC); =*hĩ'* (INST,COM); =*i'* (POSS) — **pɨ̌* (TERM≃LIM) [B16]
Hup[V] [*jup*]	**(4)**	*-ăn* (OBJ); *-an* (DIR); *-Vt* (OBL); *nɨ̌h* (POSS) [E08]
Yuhup[V] [*yab*]	**(7)**	*-dìh* (*accusatif*); *-Vt* (PRED); *-kûj* (*comitatif*); *-dèh* (*déterminatif de possession*); *-bâh* (*distance proche*); *-'áh* (*distance moyenne*); *'áhà* (*distance lointaine*) [O02]

Table A.6: Case-marker inventory sizes and members across NWA sample, including original glossing conventions from source material

Table A.7 sorts languages in the NWA sample on the basis of whether the multiple case-marking pattern is attested or not (*Q1.2*):

ATTESTED **(6)**	NOT ATTESTSED **(15)**
cub◇ *tuc*◇ *des*◇ *gvc*◇ *cbv*ᵛ *tae*ᵛ	**e-snn*♦ *c-snn*♦ *coe*♦ *ore*♦ *tnc*◇ *bsn*◇ *con huu boa quw qvi acu kwi jup*ᵛ *yab*ᵛ

Table A.7: Attestation types for multiple case-marking across sample

Where this pattern is attested, it typically arises within the spatial domain. For instance, Kubeo [*cub*] encodes the PATH relation via the obligatory combination of a locative marker, which may arise on its own, and a dedicated PATH-marker, as demonstrated in (17). Other instances of spatial multiple case-marking are shown for Desano [*des*] in (18), for Tukano [*tuc*] in (19), and for Tariana [*tae*] in (20). These instances typically implicate the co-occurrence of the non-SUBJECT-marker (e.g., Tukanoan *-re*) and a locative marker:

(17) Kubeo◇ (*cub* — EASTERN-TUKANOAN) (MM99:122, ex. (392))
ke te-Rĩ ape-kɨ ɨ̃bɨ̃ 'bA-kɨ-RE
like_that do-GER other-CL:M man be-NMLZ.N.FUT.M.SG-N.SBJ
xaro-RExa-Ibã-jA ***boa-I-Rõ-I-ta***
send-RMT.PST-3PL-REP kill-I-NMLZ.SG.INAN-LOC-PATH
'Doing it so, so they say, a long time ago they sent another man, who has died since then, up to that killing spot.'

(18) Desano◇ (*des* — EASTERN-TUKANOAN) [M99:32, ex. (119)]
ero-ge-re *gɨa árĩ-bɨ*
there-LOC-OBJ 1PL.EXCL be-N3.PST
'We were there, at that place.'

(19) Tukano◇ (*tuc* — EASTERN-TUKANOAN) [R97:169, ex. (340)]
toó-pɨ-de *peêdu sĩ'dí-wɨ*
ANA.LOC-FOC-OBJ caxiri_drink drink-REP.PST.VIS
'I will drink *caxiri* THERE (an established location).'

(20) Tariana ᵛ (*tae* — ARAWAKAN) [A03:159, ex. (7.78)]
nu-ñha nu-dia nhua ***nu-ya-dapana-se-nuku***
1SG-eat 1SG-return 1SG 1SG-POSS-CL:HAB-LOC-N.SBJ.TOP
'I'll go back to eat (my catch) in my very house.'

In certain Vaupés languages, multiple case-marking is also found on (non-spatial) oblique arguments, as demonstrated for Kakua [*cbv*] in (21):

(21) Kakuaᵛ (*cbv* — KAKUA-NUKAK) [B16:206, exx. (71)-(72)]
a. *webit* ***'ɨ'=hĩ'=di'*** *bû'jup=di'* *mâw=na=ka*
child blow_gun=INST=N.SBJ hummingbird=N.SBJ kill=DECL=ASS
'ĩ=tʃãhãp
3PL=PROG
'The children are killing hummingbirds with the blow gun.'

b. *hiw* *'ã=ñap=na=be* ***nin=bŭ=di'=hé'***
jaguar 3S.M=jump=DECL=REC.PST this=LOC=N.SBJ=INTENS
'The jaguar jumped right here.'

A.2.2 The dataset: The Functional domain

The first five QUESTIONS in the FUNCTIONAL DOMAIN type languages on the basis of three correspondence categories identified in Section A.1.2.2. The full typology for each correspondence is provided in what follows, along with illustrative examples for these types for each QUESTION in turn.
Table A.8 categorizes languages in the sample on the basis of SUBJECT-INSTRUMENT correspondence type (Q2.1):

IDENTITY **(3)**	DIFFERENTIATION **(17)**	MIXED **(1)**
e-snn♦ *coe*♦ *tae*ᵛ	*ore*♦ *cub*◇ *tnc*◇ *tuc*◇ *des*◇ *bsn*◇ *gvc*◇ *con huu boa quw qvi acu kwi cbv*ᵛ *jup*ᵛ *yab*ᵛ	*c-snn*♦

Table A.8: Correspondence types for SUBJECT-INSTRUMENT

The IDENTITY type is displayed for Koreguaje [*coe*] in (22) and Tariana [*tae*] in (23):

(22) Koreguaje♦ (*coe* — WESTERN-TUKANOAN) [CC93:47-48]

a. ***ǰɨ'ɨ-pi*** *ǰo'o-ra*
1SG-SBJ do-DES.SG
'I (M) (i.e., nobody else) will do (it).'

b. ***wa't^{h}i-pi*** *ǰo'o-mɨ* *ǰɨ'ɨ*
knife-INST do-M.SG 1SG-SBJ
'(I (M)) will do (it) with a knife.'

(23) Tariana^{v} (*tae* — ARAWAKAN) [A03:141, 152, exx. (7.4), (7.50)]

a. *kiya-ku* *hiniri* *di-wha-ka* *di-ka* *di-niwa*
strong-PERS ukuki_fruit 3S.N.FUT-fall-SUB 3S.N.FUT-see 3S.N.FUT-collect
di-wake-ta *di-yã-nhi-pidana* ***diha***
3S.N.FUT-join+CAUS1-CAUS2 3S.N.FUT-stay-ANT-RMT.PST.REP ART
nawiki-ne
man-SBJ.FOC
'The man saw that the ukukí fruit was falling down a lot (and) he was collecting (them) . . . '

b. *di-wapa* ***mawipi-ne***$_i$ ***diha pupawa di-sue-ta-nipe-ne***$_{ii}$
3S.F-wait hook-COM ART poison 3S.F-lay+CAUS1-CAUS2-NMLZ-COM
'He (the Makú man) was waiting (for the demon) with the stick$_i$ (and) with the poison which he had prepared$_{ii}$.'

The dominant type in the region is the DIFFERENTIATION-type, since most sampled languages lack overt S-marking altogether. However, two DIFFERENTIATION-type

languages do display overt S-marking, which does not coincide with INSTRUMENT-marking: i.e., Tanimuka [*tnc*] shown in (24), and Murui [*huu*] shown in (24b):

(24) Tanimuka$^{\diamond}$ (*tnc* — EASTERN-TUKANOAN) [O02: 230, 28: exx. (176b), (260a)]

a. ***ϕo'i∼baha-ré*** *jaí-á* *∼háa-re'ka*
people-DEIC1 jaguar-N kill-RMT.PST
'The people killed a jaguar.'

b. ***∼aúá-ϕí*** *biá* *ji-túté-ba'a-jú*
cassava-INST aji_pepper 1SG-SPREAD-EAT-PRS
'I eat aji, by dipping it with cassava.'

(25) Murui (*huu* — WITOTOAN) [W21: 247, 281: exx. (8.15), (12.9)]

a. [***nai-e*** ***jɨko=dɨ***] *jai* *une-gɨ* *bai-t-e*
ANA-CL:GEN dog=SBJ.TOP already bee-CL:OVAL.BIGGER find-LK-3
'That dog found a bee.'

b. *nai-mɨe* *da-ma* *abɨ* *fa-t-e* ***yoe-fai-do***
ANA-CL:M one-CL:M SELF kill-LK-3 metal-CL:SHORT.THICK-INST
'He killed himself with a machete.'

Finally, Colombian Siona [^{c}*snn*] is the only language coded as displaying the MIXED-type, given that S-marking is complex in the language, shown in (13) above. Conversely, the INSTRUMENT FUNCTION is strictly encoded with *-bi*, as in other Siona+ languages, such as Koreguaje [*coe*] in (22b).

Table A.9 spells out the distribution of DIRECT OBJECT-INDIRECT OBJECT (i.e., P-R) correspondence types (*Q2.2*) in the sample:

IDENTITY **(13)**	DIFFERENTIATION **(3)**	MIXED **(5)**
ore$^{\blacklozenge}$ *cub*$^{\diamond}$ *tnc*$^{\diamond}$ *tuc*$^{\diamond}$ *des*$^{\diamond}$ *bsn*$^{\diamond}$ *gvc*$^{\diamond}$ *acu* *kwi* *tae*$^{\vee}$ *cbv*$^{\vee}$ *jup*$^{\vee}$ *yab*$^{\vee}$	*con* *huu* *qvi*	*e-snn*$^{\blacklozenge}$ *c-snn*$^{\blacklozenge}$ *coe*$^{\blacklozenge}$ *quw* *boa*

Table A.9: Correspondence types for P-R functions

The dominant pattern is one where the DIRECT OBJECTS and INDIRECT OBJECTS of the ditransitive clause recruit the same overt case markers — i.e., IDENTITY is the most commonly attested type. The following examples from Wanano [*gvc*] (26), Tukano [*tuc*] (27), Kakua [*cbv*] (28), Hup [*jup*] (29), and Shiwiar [*acu*] (30) illustrate this type:[9]

[9]Birchall (2014, §8.3) arrives at a similar conclusion in his survey for the corresponding *Western Amazonia* region. In his methodology and terminology, he finds that these languages (incl. Tukanoan languages and other languages of the Vaupés) display *marked neutral case-marking*, where P and R receive the same case-marking.

(26) Wanano$^{\Diamond}$ (*gvc* — EASTERN-TUKANOAN) [S13:336, ex. (54)]
~bichá-ré tí-kó-ró ***hó-ré***$_{(P)}$ ***~bak-ɨ́-~dá-ká-ré***$_{(R)}$
today-N.SBJ ANA-CL:F-SG banana-N.SBJ child-CL:M-DIM-PL-N.SBJ
wá-ré
give-VIS.PERF.2/3
'Today she gave the little boys$_{(R)}$ banana(s)$_{(P)}$.'

(27) Tukano$^{\Diamond}$ (*tuc* — EASTERN-TUKANOAN)[10] [R97:166, ex. (330)]
mɨ'ɨ̃ ba'â-dí ***+ di'i-do-de***$_{(P)}$ *kã'dó* ***ã'dí-de***$_{(R)}$ *suâ*
2SG eat-NMLZ.INAN.SG + meat-CL:LOC-N.SBJ a_little 3PRO.M-N.SBJ peel.off
+ o'o-ya
+ give-IMPER
'Give a little of your piece of meat$_{(P)}$ to him$_{(R)}$!'

(28) Kakua^{V} (*cbv* — KAKUA-NUKAK) [B16: 189, ex. (15)]
Hãmu=di'$_{(R)}$ ***ma=nɨm=di'***$_{(P)}$ *ma=wɨ̌-ɨ́*
H=OBJ 2SG-daughter=OBJ 2SG=give-IMPER
'Give Hamu$_{(R)}$ your daughter$_{(P)}$!'

(29) Hup^{V} (*jup* — NADAHUPAN) [E08:170, ex. (21)]
núp ***hõp-ăn***$_{(P)}$ *'ăn* *no'-'ũh-'áy* *hám*
DEM.PROX fish-OBJ 1SG.OBJ give-APPL-VENT.IMPER go.IMPER
pawdína-ăn$_{(R)}$
P-OBJ
'Go give this fish$_{(P)}$ to Paulina$_{(R)}$ for me!'

(30) Shiwiar (*acu* — CHICHAM) [K20:250, ex. (7.169)]
ʧuú ***+ muukɨ̃=n***$_{(P)}$ ***nu*** ***nũwɨ̃=n***$_{(R)}$ *su-sa-r-i*
monkey + head=OBJ ANA wife=OBJ give-PFV-PL-3S+DECL
'They gave his wife$_{(R)}$ the head of the monkey$_{(P)}$.'

Only three tested languages display the DIFFERENTIATION-type in the NWA dataset. A'ingae [*con*] has a dedicated DATIVE-marker, shown in (31). Murui [*huu*] assigns the LOCATIVE-marker to encode the INDIRECT OBJECT-argument (R), as in (32), which is unavailable to the DIRECT OBJECT-argument (P):

(31) A'ingae (*con — Isolate*) [FH23:80, 103; ex. (42), (138)]
a. ***rande*** ***kuri-fi'ndi=ma=ngi***$_{(P)}$ ***ke=nga=ja***$_{(R)}$ *afe*
big gold-CL:BITS=ACC.REAL=1 2SG=DAT=CNTR.TOP give
'I gave you$_{(R)}$ a large bill$_{(P)}$ (lit. big money).'
b. *chhi~chhi=pa phiña=mba* ***api=nga si'nge=nga*** *utsian*
slice~REP=SS put=SS pot=DAT fire=DAT put.on
'Having sliced (it), (she) put (it) in the pot and set (it) on the fire.'

[10] For Ramirez (2019[1997], §8.12), the non-SUBJECT-marker, *-re*, is presented as *-de* in the underlying representation, but as *-re* in in the orthography. Similar claims are made for other Tukanoan languages, including regarding the Western Tukanoan language, Koreguaje, where a *-re* ∼ *-t^{h}e* alternation is noted, particularly among older speakers (Cook and Criswell 1993, 47). These are all cognates of the same *-re* marker, found in Ecuadorian Siona and all other Tukanoan languages.

(32) Murui (*huu* — WITOTOAN) [W21:270,273; exx. (8.99),(8.110)]

a. ***aros-na***$_{(P)}$ ***kue-mo***$_{(R)}$ *akata*
rice-N.SBJ.TOP 1SG-LOC show.IMPER
'Show ME$_{(R)}$ the rice$_{(P)}$!' (abrupt reading)

b. *bi-rui-yai-do* ***nofiko-mo*** *i-ti-kue*
this.CTS-CL:DAY-PL-INST La_Chorrera-LOC exist-LK-1SG
'Nowadays I live in La Chorrera.'

Five languages in the sample display the MIXED-type as defined here. In the Siona+ languages, the same markers are available to P as to R; however, the R-argument may also accept *-na* — see Section 4.2.2.1 in the discussion in Chapter 4 regarding points of micro-variation across the Siona+ grouping. A highly similar pattern is shown for Tena Kichwa [*quw*] in (33), and a similar pattern is also noted in Bora [*boa*] in (34):

(33) Tena Kichwa (*quw* — QUECHUAN) [G16:103, ex. (2.107)]
Pablo ***sisa-guna-ta***$_{(P)}$ *kuya-n* ***Maria-ma/Maria-ta***$_{(R)}$
P flower-PL-ACC give-3 M-DAT/ACC
'Pedro gave flowers$_{(P)}$ to Maria$_{(R)}$.'

(34) Bora (*boa* — BORAN)[11] [WT12:271,279; exx. (640),(663)]

a. *muha* ***díí-bye-ke*** *mé ájtyumí-hi*
1PL DEM.MED-M.S-OBJ SAP see-<*t*>
'We saw him.'

b. *tsiíju* ***í*** ***tsiíméne-ke***$_{(R)}$ *ajcu-ì* ***majchó-vu***$_{(P)}$
mother SELF child-OBJ give-<*t*> food-GOAL
'The mother gave food$_{(P)}$ to her baby$_{(R)}$.'

Table A.10 spells out the distribution of INSTRUMENT-COMPANION correspondence types (*Q2.3*) in the sample. This table includes the distribution of types in the WALS-52 sample (Stolz et al. 2013), spanning 322 language specimens from around the world, for context. No MIXED type languages are found in this sample:

	IDENTITY **(15)**	DIFFERENTIATION **(6)**	MIXED **(0)**
	ore$^{\blacklozenge}$ *cub*$^{\diamond}$ *tuc*$^{\diamond}$ *des*$^{\diamond}$ *bsn*$^{\diamond}$ *gvc*$^{\diamond}$ *con huu quw qvi acu kwi tae*^{V} *cbv*^{V} *jup*^{V}	*e-snn*$^{\blacklozenge}$ *c-snn*$^{\blacklozenge}$ *coe*$^{\blacklozenge}$ *tnc*$^{\diamond}$ *boa yab*^{V}	N/A
WALS (52)	*76* (23.6%)	*213* (66.15%)	*33* (10.25%)

Table A.10: Correspondence types for INSTRUMENT-COMPANION functions

The dominant pattern in the NWA sample is the IDENTITY-type. Examples are provided below for Kubeo [*cub*] (35), Barasano [*bsn*] (36), Tariana [*tae*] (37), Hup [*jup*] (38), Imbabura Kichwa [*qvi*] (39), and for Shiwiar [*acu*] in (40):

[11] Bora displays a pattern where *-ke* (glossed OBJAN in Thiesen and Weber 2012) is used to flag the (animate) P-argument in transitive contexts, but the R-argument in ditransitive ones. The P-argument in the double-object construction, however, takes the marker *-vu* (glossed THM by Thiesen and Weber 2012).

(35) Kubeo$^{\diamond}$ (*cub* — Eastern-Tukanoan) [MM99:119, exx. (383b)-(383c)]

a. *ape-kɨ̃* *ĩbĩ* *tota-ta-RExa-Abẽ* ***õpõ-jĩ-ke***
other-CL:M man hit-ITER-RMT.PST-3M.SG explode-CL:FUNNEL-INST
ʼĩ-RE
he-N.SBJ
'Another man hit him several times with a rifle.'

b. *aru dõ-I* *xi* *ki-E-RE* *toroxɨ̃-wɨ̃* ***jɨ̃***
and DEM.MED-LOC my be-NMLZ.PL.INAN-N.SBJ be_happy-N3S 1SG
ʼxi-bã-Rã-ke
my-child-PL.AN-INST
'And when I am there, I am (being) happy with my children.'

(36) Barasano$^{\diamond}$ (*bsn* — Western-Tukanoan) [JJ91:68-69, exx. (189)-(190)]

a. ***bĩsi-bedo-rãka*** *gate* *bʉ̃to* *asi* *bahi-ro-dẽ* *hʉ̃a-ko-aka-hʉ̃*
vine-ring-INST toast very_much hot be-NMLZ-SPCR crisp-FF-MOT-PST-3
ti
3PRO.PL
'Toasting (leaves) with a vine-ring, when they are very hot they become crisp.'

b. ***yʉ̃-rãka*** *wa-rʉ̃a-be-a-ti* *bʉ̃*
1SG-COM move-DES-NEG-PRS-Q 2SG
'Don't you (SG) want to come with me?'

(37) Tariana$^{\text{V}}$ (*tae* — Arawakan) [A03:152, exx. (7.49)-(7.50)]

a. *di-wapa* ***mawipi-ne***$_i$ ***diha pupawa di-sue-ta-nipe-ne***$_{ii}$
3S.F-wait hook-COM ART poison 3S.F-lay+CAUS1-CAUS2-NMLZ-COM
'He (the Makú man) was waiting (for the demon) with the stick$_i$ (and) with the poison which he had prepared$_{ii}$.'

b. *diha* *naha* ***na-ketʃi-ni-ne*** *di-ema-ka-pidana*
3PRO.M ART.PL 3PL-relative-CL:M-COM 3S.F-stand-SUB-RMT.PST.REP
'While he stood together with their relatives (at the port).'

(38) Hup$^{\text{V}}$ (*jup* — Nadahupan) [E08:185, 186, ex. (77), (81)]

a. ***mʼăc-ãt*** *pɨ́d* *hɨd* *bibʼ-ní-h* *dĕh=teg-éh*
mud-OBL DIST 3PRO.PL close-INFR2-DECL water=tree-DECL
'They would stop it up again with mud, the water tree.'

b. ***ʼãh=ʼíp-ít*** *ʼãh* *ni-ʼe'-ní-h*
1SG=father-OBL 1SG be-PERF-INFR2-DECL
'I lived with my father.'

(39) Imbabura Kichwa (*qvi* — Quechuan) [C82:114, exx. (448),(450)]

a. *pamba-pi* ***yunda-wan*** *yapu-ni*
field-LOC pair_of_oxen-INST plow-1SG
'(I) plow in the field with a pair of oxen.'

b. ***ñuka wawki-wan*** *kawsa-ni*
1SG brother-COM live-1SG
'(I) live <u>with my brother</u>.'

(40) Shiwiar (*acu* — CHICHAM) [K20:263-264, exx. (7.220),(7.223)]

a. *nukútsəí-rú* *wíi* ***tsuwír=hʲãĩ̃***
grandmother-1SG.P+VOC 1SG hot-COM
i-mʲa-i-hmĩ̃
CAUS-bather-PFV-1SG.S>2SG.O+DECL
'Grandmother, I just bathed you <u>with the hot (water)</u>.'

b. *wíi* *win-i-tʲát-ha-i* ***huríhuri=hʲãĩ̃***
1SG come-PFV-FUT-1SG.S-DECL cannibal-COM
'I will come <u>with the cannibals</u>.'

The DIFFERENTIATION-type is found in all remaining languages in the NWA sample — i.e., no MIXED-type is attested for this particular correspondence. The DIFFERENTIATION-type is attested in all Siona+ languages, as demonstrated for Koreguaje [*coe*] in (41), although this language uses a 'pseudo-case-marker' to encode the COMPANION function as in (41b). The fact that INSTRUMENT-marking, shown in (41a), is distinct from COMPANION-marking in (41b). Similar patterns are shown for Tanimuka [*tnc*] in (42) and for Yuhup [*yab*] in (43):

(41) Koreguaje♦ (*coe* — WESTERN-TUKANOAN) [CC93:47-48]

a. ***wa'tʰi-pi*** *ǰo'o-mɨ* *ǰɨ'ɨ*
knife-INST do-M.SG 1SG-SBJ
'(I (M)) will do (it) <u>with a knife</u>.'

b. ***i-na-re*** ***jãʼme*** *pa-sa* *ǰii-mɨ* *ǰɨ'i*
this-PL.ANIM-N.SBJ COM live-DES.SG want-M.SG 1SG
'(I (M)) want to live <u>with them</u>.'

(42) Tanimuka◊ (*tnc* — EASTERN-TUKANOAN) [E15:280,178, ex. (260),(91a)]

a. ***ãvá-ɸí*** *biá* *ɟi-túté-ba'a-ɟú*
cassava-INST aji_pepper 1SG-spread-eat-PRS
'I eat aji pepper, spreading it <u>with cassava</u>.'

b. *i'-kó* *wã-rũ-kó* ***Juan-'ka***
3PRO-CL:F study-FUT-F J-COM
'She is studying <u>with Juan</u>.'

(43) Yuhup▽ (*yab* — NADAHUPAN)[12] [O02:158,161, exx. (30b), (36b)]

a. ***pŏh-ót*** *jâp* *wédⁿí*
hand-PRED 3PRO eat
'He eats <u>with his hands</u>.'

[12]For the sake of coding the NWA dataset, I lift the gloss PRED (originally *prédicatif*) from Ospina Bozzi 2002. This refers to a morpheme which may arise on both nouns and verbs, and displaying a predicativizing function, and, crucially, a case-marking function, particularly on certain oblique CASE FUNCTIONS, including INSTRUMENT-marking as shown in (43a). For the present purposes, this meets all the criteria for being an overt, bona fide case-marker (*id*, 137, 159-161).

b. ***tɨ́h*** ***báb^{m}-kûj*** *jɨ̀dəh* *wɨ̀dnèd^{m}í*
3PRO companion-COM 3PRO.PL arrive
'They arrived with their companion.'

Table A.11 spells out the distribution of correspondence types regarding the DIRECT OBJECT (P) and the STATIC LOCATION (L) in the sample, i.e., *Q2.4*. No IDENTITY-type languages arise in the NWA sample:

IDENTITY **(0)**	DIFFERENTIATION **(10)**	MIXED **(11)**
N/A	*ore*$^{\blacklozenge}$ *con huu boa quw qvi acu cbv*$^{\vee}$ *tae*$^{\vee}$ *yab*$^{\vee}$	*e-snn*$^{\blacklozenge}$ *c-snn*$^{\blacklozenge}$ *coe*$^{\blacklozenge}$ *cub*$^{\diamond}$ *tnc*$^{\diamond}$ *tuc*$^{\diamond}$ *des*$^{\diamond}$ *bsn*$^{\diamond}$ *gvc*$^{\diamond}$ *kwi jup*$^{\vee}$

Table A.11: Correspondence types for P-L functions

Máíhɨ̃kì is the only Tukanoan language which does not display the MIXED-type since the L-argument is strictly zero-marked in the language, as shown in (44), compared with overt *re*-marking available to the P-argument (Farmer 2015, §3.3), shown in (44a):

(44) Máíhɨ̃kì$^{\blacklozenge}$ (*ore* — WESTERN-TUKANOAN)

a. *tóyá-tɨ̀kà* ***bɨoto*** *bai-jɨ*
write-CL:STICK bag be-3S.M.PRS
'The pencil is in the bag.' (Neveu 2012b, ex. (11b))

b. ***ñíò-re*** *hásó-bɨ*
3S.F.PRO-N.SBJ shoot-1PL.PST.DECL
'(We) shot her (the tapir).' [F15:30; ex. (13a)]

Other DIFFERENTIATION-type patterns are commonplace in the sample, as demonstrated for A'ingae [*con*] (45) and Tariana [*tae*] (46):

(45) A'ingae (*con* — *Isolate*) [FH23:93,107; exx. (95), (159)]

a. ***Santa Rosa=ni=ja*** *tsa='ka=mbi='ya*
SR=LOC=CNTR.TOP ANA=SIM=NEG=ASS
'It is not like that in SANTA ROSA.'

b. ***matichi=ve=ta=ti=ki*** *in'jan='fa*
machete=ACC.IRR=NEW.TOP=INT=2 want=SBJ.PL
'Do you want machetes?'

(46) Tariana$^{\vee}$ (*tae* — ARAWAKAN) [A03:143,149; ex. (7.9),(7.41)]

a. ***puale-se*** *nehpani-pida*
elsewhere-LOC 3PL+work-REP
'They worked in a different place.'

b. *nhuaniri* ***nu-na*** *di:ta* *di-eme-ta*
1SG+father 1SG-N.SBJ 3S.N.FUT+lock 3SG.N.FUT-stay+CAU1-CAUS2
hiperi-ne
parí-INST
'My father locked me with parí (a type of wood).'

Turning to the MIXED-type in the NWA sample, Table A.12 displays the distribution of case-markers across the P and the L FUNCTIONS accordingly. This is the pattern found in all Tukanoan languages in the sample, excepting Máíhɨ̃kì [*ore*], where *-re* (or cognates thereof) are distributed in both the argument and spatial domains:

LANGUAGE	P	L
*Ecuadorian Siona♦ [E*snn*]	***-re***, *-ni*	***-re***, *-ni*, *-na*
Colombian Siona♦ [C*snn*]	***-re***, *-ni*	***-re***, *-na*
Koreguaje♦ [*coe*]	***-re***, *-ni*	***-re***, *-na*
Kubeo◇ [*cub*]	***-re***	***-re***, *-i*, *-rã*
Tanimuka◇ [*tnc*]	***-re***	***-re***, *-re'ká*
Tukano◇ [*tuc*]	***-re***	***-re***, *-i*, *-pʉ*
Desano◇ [*des*]	***-re***	***-re***, *-ge*
Barasano◇ [*bsn*]	***-re***	***-re***, *-hʉ*
Wanano◇ [*gvc*]	***-re***	***-re***, *-i*, *-pʉ*
Awa Pit [*kwi*]	*=ta*	*=ta*, *=pa*, *=mal*, *=ki*
Hup▽ [*jup*]	*-ăn*	*-an*, *-V́t*, *-có*

Table A.12: Forms distributed to P and L functions, MIXED-type

The Siona+ pattern is discussed in various places in Chapter 4, including Section 4.2.2.1 and Section 4.2.3.4. As found in the case of the P-R correspondence in (*Q2.2*), the crucial factor is the presence of *-na* on certain L-arguments, but never on P. This is displayed for Colombian Siona [C*snn*] in (47). A similar pattern where L-marking is richer than P-marking, although there is some overlap, is shown in the series of Awa Pit [*kwi*] sentences in (48):

(47) Colombian Siona♦ (*snn* — WESTERN-TUKANOAN) [W70:62]

a. ***yógu-na*** *kãí-hɨ* *daí-ma-té*
canoe-GOAL sleep-PL.PRS.DEP come-RMT.PST-N2/3S.PST.N.ASS
'(We) came as (we) slept in the canoe.'

b. ***yógu-té*** *yũ'í-hi*
canoe-N.SBJ be_seated-3S.M.PRS.ASS
'(He) is seated in the canoe.'

c. ***wɨ'é-de*** *ba'í-hi*
house-N.SBJ be-3S.M.PRS.ASS
'(He) is in the house.'

(48) Awa Pit (*kwi* — Barbacoan)
[C97:73,135,137,139; exx. (65),(213),(225),(235)]

a. *Camilo=na* ***Santos=ta*** *pala* *kwin-ti-zi*
C=top S=acc plantain give-pst-non.locut
'Camilo gave Santos a plantain.'

b. *wisha* ***payl=ta=yɳ*** *ayna-t* *ku-m*
1pl.excl pot=in=rest cook-sv eat-adjzr
'(We) cooked and ate in clay pots.'

c. ***aɳ=pa*** *awa* *su* *paa-ma-ti*
here=in(approx) person earth become-comp-term
'Around here (this side of the river) it's become cultivated.'

d. *Dolores* ***Pueblo Viejo=mal*** *tu=ma* *ka* *ki ?*
D PV=loc be_in_place=inter be_permanently q
'Is Dolores in Pueblo Viejo?'

Table A.13 spells out the distribution of the 'generalized spatial-marking' pattern, discussed in Section 4.2.2.1, namely where L, G, and Source arguments are encoded in using the same set of case-markers (*Q2.5*). Given that Máíhɨ̃kì does not have spatial case-marking, this language is identified in gray text:

IDENTITY **(9)**	DIFFERENTIATION **(3)**	MIXED **(9)**
cub◇ *tuc*◇ *des*◇ *bsn*◇ *gvc*◇ *cbv*ᵛ *tae*ᵛ *yab*ᵛ *ore*♦	*tnc*◇ *boa* *qvi*	*e-snn*♦ *c-snn*♦ *coe*♦ *con* *huu* *acu* *quw* *kwi* *jup*ᵛ

Table A.13: Correspondence types for generalized spatial marking

The Identity-type is the dominant type found within the Vaupés region, which includes all Eastern Tukanoan languages (excl. Tanimuka [*tnc*]). The 'generalized spatial marking' pattern is shown for Desano [*des*] (49), for Kakua [*cbv*] (50), and for Tariana [*tae*] (51):

(49) Desano◇ (*des* — Eastern-Tukanoan) [M99:59-61, ex. (241),(245),(254)]

a. *gɨa* *yoa-ri-bohe* *árĩ-bɨ* ***ero-ge-re*** ***iri***
1pl.excl be_long-dbv-time be-n3.pst there-loc-obj dem.prox
bãkã-ge-re
town-loc-obj
'We were a long time there in this town.'

b. *gɨa* ***õ-re***$_i$ *era* *pɨ'rɨ* *bu'a-bɨ* *pare* ***finka-ge***$_{ii}$
1pl.excl here-n.sbj arrive after go.down-n3.pst finally farm-loc
'We arrived here$_i$, after which we went down to a farm$_{ii}$.'

c. ***gahi-rã-ge*** *ari-biri-bã*
other-an.p-loc come-neg-3pl
'(People) did not come from other places.'

(50) Kakua^{V} (*cbv* — Kakua-Nukak) [B16:248,206-207, ex. (27),(75),(76)]

a. ***victor-ĩ' mɨ̌=bŭ=hĩ*** *fĩ=hɨm-ɨp=hĩ*
V-poss house=loc=rmt.pst 1pl=exist-pst=rmt.pst
mik-wĕ'e=na *fĩ=tʃãhãp=hĩ*
refl-chat=decl 1pl=prog=rmt.pst
'We were at Victor's house.'

b. ***Mitú=bŭ*** *fĩ=beh-ep=hĩ*
M=loc 1pl=go-pst=rmt.pst
'We left to Mitú.'

c. ***tʃena=bŭ*** *kan=ná* *wăt-jù'-nɨ'-ɨp*
both=loc 3sg=decl big_flower_blossom-toss-stop-pst
'(It) blossomed from both sides.'

(51) Tariana^{V} (*tae* — Arawakan) [A03:148-149, ex. (7.41),(7.33),(7.36)]

a. ***puale-se*** *nehpani-pida*
elsewhere-loc 3pl+work-rep
'They worked in a different place.'

b. *na-pidana* ***uni-se***
3pl+go-rmt.pst.rep water-loc
'They (the women's hair) went into the water.'

c. *hĩ* ***wyaka-se*** *ka-nu-karu* *dhuma-naka*
dem.anim far-loc rel-come-pst.rel.f 3sg.f+hear-prs.vis
waku-nuku
1pl+speech-top.n.a/s
'She who came from far away understands our speech.'

Table A.14 lays out all available case-marking forms across the primary spatial functions for languages of the Identity-type:

Language	L	G	SRC
Kubeo$^{◊}$ [*cub*]	***-re***, *-i*, *-rã*	***-re***, *-i*, *-rã*	***-re***, *-i*, *-rã*
Tukano$^{◊}$ [*tuc*]	***-re***, *-i*, *-pʉ*	***-re***, *-i*, *-pʉ*	***-re***, *-i*, *-pʉ*
Desano$^{◊}$ [*des*]	***-re***, *-ge*	***-re***, *-ge*	***-re***, *-ge*
Barasano$^{◊}$ [*bsn*]	***-re***, *-hʉ*	***-re***, *-hʉ*	***-re***, *-hʉ*
Wanano$^{◊}$ [*gvc*]	***+re***, *+i*, *-pʉ*	***+re***, *+i*, *-pʉ*	***+re***, *+i*, *-pʉ*
Kakua^{V} [*cbv*]	*=bŭ*	*=bŭ*	*=bŭ*
Tariana^{V} [*tae*]	*-se*	*-se*	*-se*
Yuhup^{V} [*yab*]	*-bah*, *-'ãh*, *'áhà*	*-bah*, *-'ãh*, *'áhà*	*-bah*, *-'ãh*, *'áhà*

Table A.14: Forms across primary spatial functions, identity type (*Q2.5*)

Turning to the three languages which display the Differentiation-type, a similar table to the above is provided in Table A.15:

Language	L	G	SRC
Tanimuka$^{\Diamond}$ [*tnc*]	***-re***, ***-re'ká***	*-rã*	*-ϕi*
Bora [*boa*]	*-ri*	*-vu*	*-tu*
Imbabura Kichwa [*qvi*]	*-pi*	*-man*, *-ta*	*-manta*

Table A.15: Forms across primary spatial functions, Differentiation-type (*Q2.5*)

The minimal series presented in (52) displays the Differentiation-type as attested in Tanimuka [*tnc*]:

(52) Tanimuka$^{\Diamond}$ (*tnc* — Eastern-Tukanoan)
[E15:210-211, ex. (151a)-(151c)]

a. ***wi'í-a-re'ká*** *ji-ĩbé*
maloca_house-N-LOC 1SG-COP
'I am in the maloca house.'

b. ***wi'í-a-∼rá*** *ji-káka-jú*
maloca_house-N-GOAL 1SG-enter-PRS
'I am entering the maloca house.'

c. ***wi'í-a-ϕi*** *ji-ϕó-jú*
maloca_house-N-SRC 1SG-leave-PRS
'I am leaving from the maloca house.'

Finally, the Mixed-type accounts for the Siona+ languages, and a handful of sampled non-Tukanoan languages, as demonstrated in Table A.16:

Language	L	G	SRC
*Ecuadorian Siona$^{\blacklozenge}$ [^{E}*snn*]	***-re***, *-ni*, *-na*	*-na*	*-bi*
Colombian Siona$^{\blacklozenge}$ [^{C}*snn*]	***-re***, *-na*	*-na*	*-bi*
Koreguaje$^{\blacklozenge}$ [*coe*]	***-re***, *-na*	*-na*	*-pi*
A'ingae [*con*]	*=ni*	*=ni*	*=ne*
Murui [*huu*]	*-mo*	*-mo*	***-mo****na*
Shiwiar [*acu*]	*=nyuma*	*=nyuma*	*=njuja*
Tena Kichwa [*quw*]	*-pi*, *-ma*	*-ma*	*-manta*
Awa Pit [*kwi*]	*-ta*, *-pa*, *-mal*, *-ki*	*-ta*, *-pa*, *-mal*	***-ta****s*, ***-pa****s*
Hup^{V} [*jup*]	*-Vt*, *-an*, *có*	*-Vt*, *-an*	*-an*

Table A.16: Forms across primary spatial functions, mixed type (*Q2.5*)

The Siona+ Differentiation-type pattern is discussed in various places in Section 4.2.2.1 in Chapter 4. The basic pattern is one where the Goal-marker, *-na*, arises on the L-argument arising with certain predicates. The Koreguaje [*coe*] pattern in (53) illustrates the broader Siona+ pattern. Highly similar patterns are found in Tena Kichwa [*quw*], as reported in (54):

(53) Koreguaje (*coe* — WESTERN-TUKANOAN) [CC93:48-49]

a. ***wi'e-re*** *pa'i-mo repa-o*
house-N.SBJ be-F.SG that-F.SG
'She is in the house.' [*Ella está en la casa*]

b. ***wi'e-na*** *k^{h}ãi-me j̃ĩi*
house-GOAL sleep-PL children
'The children are sleeping
in the house.' [*Los niños duermen en la casa*]

(54) Tena Kichwa (*quw* — QUECHUAN) [G16:62-68, exx. (2.47),(2.28),(2.29),(2.30)]

a. *Maria kaspi-wa wajta-n Juan-ta* ***wasi-pi***
M stick-INST hit-3 J-ACC house-LOC
'Maria hits Juan with a stick in the house.'

b. *Maria* ***chagra-ma*** *traba-nga ra-w-n*
M field-DAT work-FUT make-PROG-3
'Maria will work in the field.'

c. *Pablo* ***ista-ma*** *shamu-n*
P party-DAT come-3
'Pablo came to the party.'

d. ... *kay awa* ***luma-manda***
DEM.PROX high hill-SRC
'...from here above, from the hill.'

The MIXED-type is also shown for Murui [*huu*] in (55), where the same case-marking for L and G FUNCTIONS, but separate marking is recruited for the SOURCE FUNCTION:

(55) Murui (*huu* — WITOTOAN) [W21:256 (§8.1.1.3)]

a. *bi-rui-yai-do* ***nofiko-mo*** *i-ti-kue*
this.CTS-CL:DAY-PL-INST La_Chorrera-LOC exist-LK-1SG
'Nowadays I live in La Chorrera.'

b. *nai-mie rii-tai-ya-no* ***erai-mo*** *jai jaai-d-e*
ANA-CL:PR.M angry-BECOME-E.NMLZ-SEQ estuary-LOC already go-LK-3
'After becoming angry, (he) left for El Encanto (lit. estuary).'

c. ***Ikato-mona*** *duaibi-ti-kue*
El_Encanto-SRC chew.coca.VENT-LK-1SG
'I came from El Encanto to chew coca.'

The remainder of this section considers certain FUNCTIONAL GAPS — i.e., FUNCTIONS, for which case-markers are not recruited in Ecuadorian Siona, although they are attested elsewhere in the sample. Firstly, Table A.17 types sampled languages based on the presence of overt POSSESSOR-marking (*Q2.6*) — i.e., whether something like a *genitive*-case exists in a given language. Where such a case-marker is attested, it is explicitly identified in the table:

ATTESTED **(12)**	NOT ATTESTED **(9)**
cub$^{\diamond}$ (*-i*)	**e-snn*$^{\blacklozenge}$
tnc$^{\diamond}$ (*-ri'ká*)	*c-snn*$^{\blacklozenge}$
tuc$^{\diamond}$ (*+yaá*)	*coe*$^{\blacklozenge}$
des$^{\diamond}$ (*=ya*)	*ore*$^{\blacklozenge}$
bsn$^{\diamond}$ (*-ya*)	*gvc*$^{\diamond}$
quw (*-pa*)	*con*
qvi (*-paj*)	*huu*
acu (*=n/=n*^{y}*a*)	*boa*
kwi (*=pa*)	$^{\text{V}}$*tae*
cbv$^{\text{V}}$ (*-i'*)	
jup$^{\text{V}}$ (*=nɨ̀h*)	
yab$^{\text{V}}$ (*-*$^{\sim}$*dèh*)	

Table A.17: Attestation of overt POSSESSOR-marking (*Q2.6*)

Overt POSSESSOR-marking is illustrated for Desano [*des*] (56), Tukano [*tuc*] (57), Hup [*jup*] (58), Yuhup [*yab*] (59), and for Awa Pit [*kwi*] in (60):

(56) Desano$^{\diamond}$ (*des* — EASTERN-TUKANOAN) [DS12:162, exx. (95)-(96)]

a. ***gɨa-ya*** *wi'i*
1PL.EXCL-POSS house'
'our (EXCL) house.'

b. ***yɨ'ɨ-ya*** *bãleta*
1SG-POSS suitcase'
'my suitcase.'

(57) Tukano$^{\diamond}$ (*tuc* — EASTERN-TUKANOAN) [R97:§12.4,242-244]

a. ***Pédudu-yaá*** *+ wi'i*
P-POSS + house
'Pedro's house.'

b. ***wiô-gɨ-yaá*** *+ wese*
chief-CL:M-POSS + farm
'The chief's farm.'

(58) Hup$^{\text{V}}$ (*jup* — NADAHUPAN) [E08:225-226, exx. (34a)-(34b)]

a. ***pedú=nɨ́h*** *cug'ɛ̆t*
P=POSS book
'Pedro's book.'

b. *cug'ɛ̆t* ***pedú=nɨ́h***
book P=POSS
'Pedro's book.'

(59) Yuhup$^{\text{V}}$ (*yab* — NADAHUPAN) [O02:243, ex. (101e)]
dɨ̀d^{n}-dɔ̀h *ûjh*
woman-POSS bag
'The woman's bag.'

(60) Awa Pit (*kwi* — Barbacoan) [Curnow 1997, 106, (131)]

a. ***katsa=wa*** *kwizha*
big=poss dog
'The big person's dog.'

b. ***Santos=pa*** *pimpul*
S=poss leg
'Santos's leg.'

According to the analysis in Stenzel (2013b, 374-375), the Wanano [*gvc*] marker *-ya-* is a head-marker, despite its cognacy with several Eastern Tukanoan languages in Table A.17. This does not meet the dependent-marking criterion adopted in this study, following Iggesen (2013), so this language is coded as not displaying Possessor-marking:

(61) Wanano$^{\Diamond}$ (*gvc* — Eastern-Tukanoan) [Stenzel 2013b, 374]
Koitiria *ya-~baka-ri*
K poss-village-pl.inan
'(the) Koitiria's villages.'

For all languages where overt Possessor-marking is not attested, a Possessor-Possessum juxtaposition strategy, in that order, is described. This is the pattern found in Ecuadorian Siona, shown in (62), and also shown for Bora [*boa*] in (63) and Murui [*huu*] in (64):

(62) Ecuadorian Siona$^{\blacklozenge}$ (*e-snn* — Western-Tukanoan)
yë'ë gaje-i *yo-huë*
1sg friend-cl:m canoe-cl:cont
'My friend's canoe.'

(63) Bora (*boa* — Boran) [Thiesen and Weber 2012, 253-254, exx. (58b)-(58c)]

a. ***kpà'áró*** *mɛ:ní-mɯ̀*
mother pig-pl.anim
'(my) mother's pigs.'

b. ***kpà'árò-mɯ́ts^{h}í mɛ:ní-mɯ̀*** *hà*
mother-dual.kin pig-pl.anim house
'(my) parent's pigs' house.'

(64) Murui (*huu* — Witotoan) [Stenzel 2013b, 8, ex. (1.7)]
Lusio *yoe-fai*
L metal-cl:short.thick
'Lucio's machete.'

Table A.18 categorizes languages on the basis of whether or not case-marking is recruited to encode the Time function (*Q2.7*). As described in Chapter 3, and confirmed in the corpus-based investigation in Chapter 5, Ecuadorian Siona does not employ case-markers to encode temporal nouns. Where such case-markers are attested, they are explicitly identified in the table:

ATTESTED **(12)**	NOT ATTESTED **(9)**
cub$^{\diamond}$ (*-i*, *-rã*, *-re*)	**e-snn*$^{\blacklozenge}$
tuc$^{\diamond}$ (*-pɨ*, *-re*)	*c-snn*$^{\blacklozenge}$
des$^{\diamond}$ (*-ge*, *-re*)	*coe*$^{\blacklozenge}$
bsn$^{\diamond}$ (*-hʉ*, *-re*)	*ore*$^{\blacklozenge}$
gvc$^{\diamond}$ (*-pʉ*, *-re*)	*tnc*$^{\diamond}$
huu (*-na*)	*con*
quw (*-pi*)	*boa*
qvi (*-ta*, *-pi*)	$^{\text{V}}$*cbv*
acu (*=tin*)	$^{\text{V}}$*yab*
kwi (*=ta*, *=pa*, *=mal*, *=ma*)	
tae$^{\text{V}}$ (*-se*)	
jup$^{\text{V}}$ (*-V́t*, *-có*)	

Table A.18: Attestation of overt TIME-marking (*Q2.7*)

Per Table A.18, most Eastern Tukanoan languages recruit spatial case-markers, including *-re*, to encode TIME nominals. This is demonstrated for Tukano [*tuc*] in (65) and for Kubeo [*cub*] in (66):

(65) Tukano$^{\diamond}$ (*tuc* — EASTERN-TUKANOAN) [R97:237, ex. (580)]
dĩ'ká-de *bu'ê-dã*
today-N.SBJ study-IMPER
'Let's study TODAY!'

(66) Kubeo$^{\diamond}$ (*cub* — EASTERN-TUKANOAN) [MM99:114,118; exx. (361b),(378)]

a. *xãrãwi-RE*
day-N.SBJ
'by day, in the daytime, during the day…'

b. *oko-Rĩbĩ-A-RE*
water-CL:TEMP-PL.INAN-N.SBJ
'in the rainy season…'

c. *ke* *te-Rĩ* ***ape-xãrãwɨ-I*** *dĩ-karã* …
like_that do-GER other-CL:DAY-LOC go-1PL.EXCL
'So we went another day…'

Overt TIME-marking is also exemplified for Murui [*huu*] in (67), for Awa Pit [*kwi*] in (68), and for Shiwiar [*acu*] in (69):

(67) Murui (*huu* — WITOTOAN) [W21:259, ex. (8.57)]
ua *jari-re-na* *kasi* ***naɨo-na*** *mei* *kaɨ* *zai-ta-d-e*
really quick-ATT-N.S/A.TOP almost night-N.S/A.TOP so 1PL step-CAUS-LK-3
'It was very quick, (they) almost stepped on us at night.'

(68) Awa Pit (*kwi* — BARBACOAN) [C97:139, ex. (234)]
verano=wa=na *ii* *ki-mtu-s*
summer=in(approx)=TOP be_hot be_hot-IMPF.LOCUT
'In summer it is hot (to us).'

(69) Shiwiar (*acu* — CHICHAM) [K20:265, ex. (7.225)]
uwɨ́h ikɨ́-tai^{N} *kint*^{j}*a=**tin***
hand sit-NS.NMLZ day=TEMP
'in five days...'

Other languages in the sample display zero-marked TIME nominals, like Ecuadorian Siona, as demonstrated for Yuhup [*yab*] in (70), and for A'ingae [*con*] in (71):

(70) Yuhup^{V} (*yab* — NADAHUPAN) [O02:159, ex. (31a)]
jɨ́htóh *ɨ̃* *kéjép* *tî'*
nowadays 1PL watch *yuruparí*
'Nowadays we are watching *yuruparí*.'

(71) A'ingae (*con* — *Isolate*) [FH23:96, ex. (103)]
tû'i *t*^{h}*û*∼*t*^{h}*û-'ngi-ye*
tomorrow fell∼ITER-AM:GO&DO-INF
'Tomorrow we'll come to fell (it).'

A.2.3 The dataset: The DCM domain

Table A.19 identifies the DCM patterns on the SUBJECT (S) attested in the NWA sample (*Q3.1*). The bulk of languages do not display these patterns (i.e., N=15), since they simply do not encode the S-argument via a case-marking strategy. For those languages which do display S-oriented DCM, all available case-marking alternatives, including zero-marking, are identified so that the formal DCM-type may be discerned. The described active MACRO-TRIGGERS are also identified in the table:

LANGUAGE		FORMS	MACRO-TRIGGERS			
			ANIM	REF	DISC	OTHER
*Ecuadorian Siona♦ [^{E}*snn*]	∅✓[d]	{-∅, -*bi*}			✓FOC	
Colombian Siona♦ [^{C}*snn*]	+∅✓[d]	{-∅, -*bi*, -*ga*}			✓FOC	
Koreguaje♦ [*coe*]	∅✓[d]	{-∅, -*pi*}			✓CONTR	
Tanimuka◇ [*tnc*]	∅✓[a]	{-∅, -*re*}	✓HUM			
Witotoan [*huu*]	∅✓[d]	{-∅, =*dɨ̃*}			✓TOP	
Tariana^{V} [*tae*]	∅✓[d]	{-∅, -*n*(*h*)*e*}			✓TOP	
NO DCM ON S (-∅):	Máíhɨ̃kì♦ [*ore*]; Kubeo◇ [*cub*]; Tukano◇ [*tuc*]; Desano◇ [*des*]; Barasano◇ [*bsn*]; Wanano◇ [*gvc*]; A'ingae [*con*]; Bora [*boa*]; Tena Kichwa [*quw*]; Imbabura Kichwa [*qvi*]; Shiwiar [*acu*]; Awa Pit [*kwi*]; Kakua^{V} [*cbv*]; Hup^{V} [*jup*]; Yuhup^{V} [*yab*]					

Table A.19: S-oriented DCM patterns (*Q3.1*)

Focusing firstly on the formal typology of the six attested S-oriented DCM patterns, no ALTERNATING-type is attested. The dominant pattern is the OPTIONAL-type (∅), described for Ecuadorian Siona in Chapter 3, as illustrated for Murui [*huu*] in (72):

(72) Murui (*huu* — WITOTOAN) [W21:249 (§8.1.1.2)]

a. ***Kata=dɨ*** *bi-ya*
K=SBJ.TOP come-E.NMLZ
'Kata (not anybody else) came back.'

b. ***Kata*** *jaai-ya*
K go-E.NMLZ
'Kata went away.'

The only HYBRID-type (+∅) attested in the sample concerns Colombian Siona [^{c}*snn*], as demonstrated in the sentences in (73). This language permits zero-marking, in alternation with two overt case-markers, *-bi* and *-ga* (Wheeler 1970, 173-176; 1987, 131):

(73) Colombian Siona♦ (*snn* — WESTERN-TUKANOAN) [W70:47]

a. ***yɨ'ɨ-ga*** *saí-yɨ*
1SG-SBJ.EMPH go-N3S.PRS.ASS
'And I, I am going.'

b. ***ča'kádi-wa'-na-bi*** *bá-gɨ-na* *gahe-ni*
fire_ant-ANIM-PL-SBJ PRO.HUM-CL:M-GOAL go_down-SS
sɨ'á-wɨ
attach-N3S.PST.ASS
'The fire ants came down to him and latched on.'

c. ***'okó*** *ka'ní-hi*
rain gather-3S.M.PRS.ASS
'The rain is gathering.'

Turning now to the set of attested MACRO-TRIGGERS, per Table A.19, all languages, except for Tanimuka [*tnc*], display discourse-triggered DCM patterns. These patterns are illustrated for Koreguaje [*coe*] in (74) and for Tariana [*tae*] in (75):

(74) Koreguaje♦ (*coe* — WESTERN-TUKANOAN) [CC93:47]
jɨ'ɨ-pi *jo'o-ra*
1SG-SBJ do-DES.SG
'I (M) (i.e., not anyone else) will do (it).' [*Yo (en contraste con otros) lo haré*]

(75) Tariana (*tae* — ARAWAKAN) [A03:141, ex. (7.4)]
[The S-argument is contrasted with another referent in this narrative episode.]
kiya-ku *hiniri* *di-wha-ka* *di-ka* *di-niwa*
strong-PERS ukuki_fruit 3S.N.FUT-fall-SUB 3S.N.FUT-see 3S.N.FUT-collect
di-wake-ta *di-yã-nhi-pidana* ***diha***
3S.N.FUT-join+CAUS1-CAUS2 3S.N.FUT-stay-ANT-RMT.PST.REP ART
nawiki-ne
man-SBJ.FOC
'The man saw that the ukukí fruit was falling down a lot (and) he was collecting (them) (while the evil spirit was trying to steal the fruit).'

The Tanimuka [*tnc*] pattern demonstrated in (76) is exceptional, since it is the only animacy-triggered S-oriented DCM pattern in the sample. Here the human S-argument in (76a) is obligatorily marked with-*re*, whereas other S-arguments are zero-marked across-the-board, e.g., the animal S-argument in (76b):

(76) Tanimuka$^{\Diamond}$ (*tnc* — EASTERN-TUKANOAN) [E15:230,233; exx. (176b),(183a)]

a. *ϕo'i~* ***baha-ré*** *ɲaí-á* *~háa-re'ka*
people-N.SBJ jaguar-N kill-PST.RMT
'The people killed a jaguar.'

b. ***jaí-à*** *kára-ká* *~háá-ko'ó*
jaguar-N chicken-N kill-PST.REC
'The jaguar killed the chicken.'

Table A.20 details the P-oriented DCM patterns in the NWA sample. Every sampled language specimen exhibits such a pattern, except for Imbabura Kichwa [*qvi*]. Many of the attested patterns implicate multi-dimensional DCM. All formal types and semantic TRIGGERS are encountered at least once in the dataset:

Language		Forms	Triggers			
			anim	ref	disc	other
*Ecuadorian Siona♦ [E*snn*]	+∅✓[a,r,d]	{-∅, **-re**, -*ni*}	✓anim	✓pro,(spec)	✓foc,(top)	✓coord
Colombian Siona♦ [C*snn*]	+∅✓[a,r,d]	{-∅, **-re**, -*ni*}	✓(anim)	✓pro,(spec)	✓foc,(top)	
Koreguaje♦ [*coe*]	+∅✓[a,r,d]	{-∅, **-re**, -*ni*}	✓(anim)	✓pro,(spec)	✓foc,(top)	
Máíhɨ̌kì♦ [*ore*]	∅✓[a,r]	{-∅, **-re**}	✓anim	✓(spec)		
Kubeo◊ [*cub*]	∅✓[r,d(a,o)]	{-∅, **-re**}	✓(anim)	✓pro,(ref)	✓(top)	✓coord
Tanimuka◊ [*tnc*]	∅✓[a(d)]	{-∅, **-re**}	✓hum		✓(contr)	
Tukano◊ [*tuc*]	∅✓[r(a)]	{-∅, **-re**}	✓(anim)	✓pro,def		
Desano◊ [*des*]	∅✓[r,d]	{-∅, **-re**}		✓pro,spec		
Barasano◊ [*bsn*]	∅✓[r,d]	{-∅, **-re**}		✓pro,def		
Wanano◊ [*gvc*]	∅✓[r,o]	{-∅, **-re**}	✓(anim)	✓pro,ref		✓position
A'ingae [*con*]	+✓[o]	{=*ma*, =*ve*∼=*me*}				✓realis
Murui [*huu*]	∅✓[r,d]	{-∅, -*na*}		✓(spec)	✓top	✓position
Bora [*boa*]	∅✓[a(o)]	{-∅, -*ke*}	✓anim			
Tena Kichwa [*quw*]	∅✓[o]	{-∅, -*ta*}				✓tbd
Shiwiar [*acu*]	∅✓[r]	{-∅, -$n^y a$}		✓(ref)		✓glob
Awa Pit [*kwi*]	∅✓[a,r]	{-∅, =*ta*}	✓hum	✓pro,(ref)		
Kakuaᵛ [*cbv*]	∅✓[a,r]	{-∅, =*dɨ̀h*}	✓hum	✓(def)	✓(disc)	
Tarianaᵛ [*tae*]	+∅✓[r,d]	{-∅, -*na*, -*nuku*}		✓pro	✓top	
Hupᵛ [*jup*]	∅✓[a,r]	{-∅, -*ăn*}	✓hum	✓pro,(ref)		✓pl
Yuhupᵛ [*yab*]	∅✓[r(a)]	{-∅, -*dɨ̀h*}	✓hum	✓(def)		✓glob
No dcm on p: Imbabura Kichwa [*qvi*] (-*ta*)						

Table A.20: p-oriented dcm patterns (*Q3.2*)

Beginning with the distribution of formal types, the dominant P-oriented DCM type in the sample is the OPTIONAL-type (∅). This pattern is demonstrated in the pair of Máíhɨ̃kì [*ore*] sentences presented in (77), reflecting the zero-marking vs. *re*-marking pattern also described for all Eastern Tukanoan languages:

(77) Máíhɨ̃kì (*ore* — WESTERN-TUKANOAN) [F15:30,198; ex. (13a), (225)]

a. ***ñíò-re*** *hásó-bɨ́*
3PRO.F-N.SBJ shoot-1PL.PST.DECL
'(We) shot her (the tapir).'

b. ***hòyà*** *báá-má-yi*
domestic_animals have-NEG-1SG.PRS.DECL
'(I) don't have domestic animals.'

One tested language displays the ALTERNATING-type (+), i.e., A'ingae [*con*]. One of two overt P-markers is selected as based upon the reality status of the clause (Fischer and Hengeveld 2023) — i.e., the irrealis accusative marker *=ve* (with predictable allomorph *=me*), shown in (78a), alternates with the realis marker *=ma* in (78b):

(78) A'ingae (*con* — *Isolate*) [FH23:93, exx. (95),(97)]

a. ***matichi=ve=ta=ti=ki*** *in'jan='fa*
machete=ACC.IRR=NEW.TOP=INT=2 want=SBJ.PL
'Do you want machetes?'

b. *sumbu-en=jan* ***ain-fa='u=ma***
emerge-CAUS-IMP dog-CL:LAT=AUG=ACC.REAL
'Get the dog out.'

The HYBRID-type (+∅), described for Ecuadorian Siona in Chapter 3, is formally identical to those described for other Siona+ languages. The corresponding Colombian Siona pattern is illustrated in (79). The Siona+ case-marker *-ni* is strictly available to animate P-arguments — e.g., (79c). A similar pattern is described for Tariana [*tae*], shown in (80), where the marker *-na* is strictly available to pronominal P-arguments, and a zero-marking vs. *-nuku* marking alternation characterizes the pattern for full nouns:

(79) Colombian Siona (*snn* — WESTERN-TUKANOAN) [W70:43-44]

a. ***jo'ó*** *yo'ó-wɨ́*
work do_work-N3S.PST.ASS
'(I) did the work.'

b. ***'ãõ-de*** *k^{w}a'kú-yɨ́*
food-N.SBJ cook.TRS-N3S.PRS.ASS
'(I) am cooking food.'

c. ***ya'ó-ni*** *hu-ɨ́'ɨ́* 'aídu-na
peccary-N.SBJ2 kill-N3S.PST.ASS forest-GOAL
'(I) shot a peccary in the woods.'

(80) Tariana^{V} (*tae* — Arawakan) [A03:143-146 ex. (7.9), (7.19), (7.23)]

a. ***a:si*** *na:* *na-ñha* *du-na*
pepper 3pl+give 3pl-eat 3s.f-n.sbj
'They giver her pepper to eat.'

b. *di-hẽ-ta-pidana* ***diha*** ***pa:ku-nuku***
3s.n.fut-see+caus1-caus2-rmt.pst.rep art gold-n.sbj.top
'He (the Makú gold miner) showed the gold (to the master about to kill him).'

c. *nhuaniri* ***nu-na*** *di:ta* *di-eme-ta*
1sg+father 1sg-n.sbj 3s.n.fut+lock 3sg.n.fut-stay+cau1-caus2
hiperi-ne
parí-inst
'My father locked me with parí (a type of wood).'

Turning to the active Micro-triggers shown in Table A.20, there is a considerable amount of variation within the sample. Most languages display multi-dimensionality to a certain extent. Firstly, considering animacy-related Triggers, several languages are described as having obligatory-marking on human-denoting p-arguments, e.g., Kakua [*cbv*] (Bolaños 2016, 193-196), as shown in (81). Similar patterns are shown for Máíhɨ̃kì [*ore*], except where humans and high animals require *re*-marking, but low animals and inanimates do not (Farmer 2015, 97) illustrated in (82). These are still animacy-driven dcm patterns, although with different 'cut-off points' in the sense of Bossong (1985), Aissen (1999, 2003), and others:

(81) Kakua^{V} (*cbv* — Kakua-Nukak)
[B16:195-196,200-202; exx. (33),(38),(49),(57),(60)]

a. ***kãk-wã=di'*** *''ĩ=t-maw-hĕm'-ep=wɨt=hĩ*
person-pl=n.sbj 3pl=evid-kill-eat-pst=rep.evid=rmt.pst
'(They) killed and ate people (it is said).'

b. ***'ĩ=nɨm=di'*** *wˇẽj* *kĕt* *beh-wˇẽj*
3pl=daughter=n.sbj coddle 3pro.pl go-coddle
'They coddle their daughter.'

c. i. *webit* ***mɨhiw=di'*** *'ã=mâw=na=ka*
child dog=n.sbj 3s.m=kill-decl=ass
'The child hits the dog.'

ii. *webit* ***mɨhiw*** *'ã=mâw=na=ka*
child dog 3s.m=kill-decl=ass
'The child hits the dog.'

d. i. ***tɨ̂=di'*** *mâw-a=ka*
firewood=n.sbj hit-imper=ass
'Pick up the firewood.'

ii. *kăn'* ***tɨ̂*** *bûd=na=ka* mi=tʃãhap
3pro.f firewood cut=decl=ass 3s.f=prog
'She is cutting firewood.'

(82) Máíhɨ̄kɨ̀♦ (*ore* — Western-Tukanoan)[13]
[F15:94-95; exx. (105b),(103a-b),(106b),(109)]

a. ***násó-rè*** *áí-yí* *násó-rè*
wolly_monkey-N.SBJ eat-1PL.PRS.DECL wolly_monkey-N.SBJ
'(We (EXCL)) would eat wolly monkey.'

b. i. ***gɨ̀ɨ̀-re*** *hɨ́á-má-yí*
louse-N.SBJ locate-1SG.PST.DECL
'I found a louse.'

ii. ***gɨ̀ɨ̀*** *báá-mà-kò* *hɨ́tò ?*
louse have-NEG-3S.F.PRS.INTERR niece
'Don't you have lice, niece?'

c. i. *ɨ̂ɨ̂* ***wèè-rè*** *gósá-rè* sá-kɨ̀
3SG.M.PRO house-N.SBJ think-SS.SEQ go.PST.NI-3SG.M.PST.DECL
'He thought of his house and went (there).'

ii. ***yóù*** *íí-bɨ́*
canoe borrow-1PL.PST.DECL
'We borrowed a canoe.'

A similar pattern is described for Hup [*jup*], which requires case-marking on human-denoting P-arguments. This language also displays the typologically rare STRONG TRIGGER of plurality (Epps 2008, 177-178) — i.e., classified as an OTHER type TRIGGER, where *all* plurals require overt case-marking via the overt form, *=n'ăn*, shown in (83d):

(83) Hup[v] (*jup* — Nadahupan) [E08:174-178, exx. (36), (41), (43), (47), (51)]

a. ***'ayŭp=ɨ̃h-ăn*** *'ãh kéy-éy j'ŭg-ăn*
one=CL:M-N.SBJ 1SG see-DYNM forest-N.SBJ
'I saw a man in the forest.'

b. i. ***hohóh=mah*** *tɨh* *'ey-yohóy-óh*
frog=REP 3PRO call-search-DECL
'He was calling and searching for the frog.'

ii. ***tɨnɨ̀h*** ***cápu-ăn=yɨ́'*** *tɨh* *'éy-cud'ũhnɨ́y*
3.SG.POSS frog(Pt)-N.SBJ=TEL 3PRO call-INFR.2.maybe
'He's apparently calling for his frog.'

c. *yɨkán* ***mŏy*** *hɨd* *bɨ'-pɨ́d-ɨ́h* *póg*
over_there house 3PRO.PL make-DIST-DECL big
'There they built a house, (it was) big!'

d. *'ãh* ***cug'ǽt=n'ăn*** *pũhũt-d'əh-hi-yɨ́'-ɨ́y*
1SG leaf=N.SBJ.PL blow-send-descend-TEL-DYNM
'I blew the papers down.'

Other languages display a WEAK ANIMACY TRIGGER for P-oriented DCM, as is shown for Kubeo [*cub*] in (84) — "animate nouns are more often marked by *-re* than inanimate ones" (Chacon and Genetti 2019, 412; cf. Morse and Maxwell 1999, 111):

[13]Note that Farmer (2015, 95) uses the scientific, Latinate name for the wolly monkey in (82a), *Lagothrix lagothricha*.

(84) Kubeo (*cub* — Eastern-Tukanoan) [MM99:111-113, exx. (350), (351)]

a. *'xã-bɨ* ***tibĩ-re***
see-M.SG otter-OBJ
'(He) recently saw <u>an otter</u>.'

b. *dõ-I* *ea-karã* ***kũĩ'dã-kɨ*** ***ãjã*** ...
DEM.MED-LOC find-1PL.EXCL one-M.SG snake
'There we found a snake...'

c. *'ke-Rõ-RA* *'kari* ***bue-I-jãbĩ-A***
like_that-NMLZ.INAN.SG-EXCL IN teach-I-CL:BUILD-PL.INAN
ja-RExa-Ibã
make-RMT.PST-3PL
'(They) built <u>schools</u> a long time ago now.'

Turning to the referentiality Macro-trigger, note that a similar claim is made for Kubeo [*cub*], as suggested regarding the animacy Trigger — i.e., "[d]efinite and referential P [...] are marked by *-re* more often than those that are indefinite, generic, or non-referential" (Chacon and Genetti 2019, 412), as exemplified by the following near-minimal pair:

(85) Kubeo$^{\diamond}$ (*cub* — Eastern-Tukanoan) [CG19: 412, ex. (18)]

a. ***oko*** *ũkũ-rĩ* bɨaha-yɨ
water drink-CVB finish-NMLZ.MSC
'After you are done drinking <u>some (unspecified amount of) water</u>.'

b. ***oko-re*** *ũkũ-rĩ* hebe-kobe
water-NNOM drink-CVB finish-INFRR.3FEM
'She drank <u>all the water</u> (i.e., a specified amount of water).'

In other Eastern Tukanoan languages, referentiality-related Strong triggers are described. For instance, the Barasano [*gvc*] pattern displayed in (86) demonstrates how the use of *-re* is unavailable on generic P-arguments. Similar rigid marking patterns are shown for Tukano [*tuc*] in (87):

(86) Barasano (*bsn* — Eastern-Tukanoan) [JJ91:65-66, exx. (175)-(176)]

a. ***bʉ̃-re*** *sĩa-gʉ-bĩ* *yai*
2SG-N.SBJ kill-CL:M-3S.M wildcat
'The wildcat will probably kill <u>you</u>.'

b. ***rase-a*** *sĩa-to* *bãdi*
toucan-PL.AN kill-HORT 1PL.INCL
'Let's kill <u>toucans</u>!'

(87) Tukano$^{\diamond}$ (*tuc* — Eastern-Tukanoan) [DL20:30-31, exx. (33)-(34),(38)-(40)]

a. i. ***ni'ká*** ***wi'i*** *da'rê-gɨ'* wee-mí
one house make-SS.M.SG do-PRS.VIS.3S.M
'(He) is making <u>a house</u>.'

ii. ****ni'ká*** ***wi'i-re*** *da'rê-gɨ'* wee-mí
one house-N.SBJ make-SS.M.SG do-PRS.VIS.3S.M
Intended: '(He) is making <u>a house</u>.'

b. i. ***ohô*** *ba'â-ya*
banana eat-IMPER
'Eat (some) banana!'

ii. ***ohô-poro-re*** *ba'â-ya*
banana-CL:OBLS-N.SBJ eat-IMPER
'Eat the banana!'

iii. ****ohô-re*** *tɨá-yɨ*
banana-N.SBJ eat-IMPER
Intended: 'Eat some/the banana!'

Besides the Eastern Tukanoan languages, referentiality-triggered P-oriented DCM is relatively rare in the sample. One such pattern is described for Kakua [*cbv*], where definiteness is described as a secondary factor — i.e., a WEAK TRIGGER, exmplified in (88):

(88) Kakua^{V} (*cbv* — KAKUA-NUKAK) [B16:202, exx. (60), (57)]

a. *kăn'* ***tɨ́*** *bûd=na=ka* mi=tʃãhap
3SG.F firewood cut=DECL=ASS 3SG.F=PROG
'She is cutting firewood.'

b. ***tɨ́-di'*** *mâw-a=ka*
firewood=OBJ hit-IMP=ASS
'Pick up the firewood!'

Another relevant referentiality-related TRIGGER concerns pronominality. All Tukanoan languages require case-marking, typically *-re*, on pronominal P-arguments (Barnes 1999, 2006; Zúñiga 2007; Stenzel 2008, 2013d). This is demonstrated with personal pronouns in Barasano [*bsn*] in (89) and in Kotiria, a variety which is highly related to Wanano [*gvc*], in (90):

(89) Barasano$^{\Diamond}$ (*bsn* — EASTERN-TUKANOAN) [JJ91:65, ex. (175)]
bʉ̃-re *sĩa-gʉ-bĩ* *yai*
2SG-N.SBJ kill-CL:M-3S.M wildcat
'The wildcat will probably kill you.'

(90) Kotiria$^{\Diamond}$ (*gvc* — EASTERN-TUKANOAN) [S08:156, ex. (4)]
∼bʉ'ʉ ***∼sa-re*** *∼ba'yo-re*
2SG 1PL.EXCL-OBJ lie-VIS.PERF.2/3
'You (SG) lied to us (EXCL).'

A similar pattern is described for certain Vaupés languages, e.g., Hup [*jup*], Epps (2008, 170), shown in (91), and Yuhup [*yab*], as shown in (92):

(91) Hup^{V} (*jup* — NADAHUPAN) [E08: 167, ex. (6)]
hɨ́d-ăn *g'əç-tuk-yó'=mah*
3PL-OBJ bite-want-SEQ=REP
'Having tried to bite them, it's said. . . .'

(92) Yuhup^{v} (*yab* — NADAHUPAN) [O02: 143-144, ex. (12c), (10c)]

a. ***jɨ̀də̃h-˜dɨ̀h*** *tɨ́h* ˜ihkéjí
3PL-OBJ 3SG greet
'He greets <u>them</u>.'

b. ***tɨ̀jí'*** *˜úb^{m} í ˜jàb^{m}bĕh*
man hit dog
'The man hits <u>the dog</u>.'

Turning now to the discourse MACRO-TRIGGER, these TRIGGERS are commonly attested for P-oriented patterns in the NWA sample. Compatible patterns with discourse-related STRONG TRIGGERS are described for all Siona+ languages, for Murui [*huu*] (Wojtylak 2021, 245-258), and for Tariana [*tae*] (Aikhenvald 2003, 145).[14] The Kakua [*cbv*] pattern is described such that discourse-related properties have a secondary status (Bolaños 2016, 203-205) — i.e., coded. as a WEAK TRIGGER for the present purposes.

The Murui [*huu*] instance presented in (93) aligns neatly with contrastive DCM in Koreguaje [*coe*], shown in (94), and in Siona, as shown in (95), lifted from Chapter 6 in this dissertation:

(93) Murui (*huu* — WITOTOAN) [W21:256,267 (§8.1.1.3)]
aros *atɨ-ñe-itɨ-o* ***asukar-na*** *atɨ-itɨ-o*
rice bring-NEG-FUT.LK-2SG sugar-N.SBJ.TOP bring-FUT.LK-2SG
'(You) won't bring <u>rice</u>. (You) will bring <u>the SUGAR</u>.'

(94) Koreguaje$^{\blacklozenge}$ (*coe* — WESTERN-TUKANOAN) [CC93:49, ex. (201)]
k^{h}ura-wa'-ɨ-ni *k^{h}u'e-mo* *ǰɨ'ɨ*
chickens-creature-M.SG-N.SBJ2 look.for-F.SG 1SG
'I (F) am looking for <u>a chicken</u> (a particular one, not just any).'

(95) Ecuadorian Siona$^{\blacklozenge}$ (^{e}*snn* — WESTERN-TUKANOAN)
b<u>a</u>huë yohuë hueroye te'e gu<u>ë</u>narore huerohuë
bãã-wɨ **jo-wɨ** wero-je **te'-e**
NEG.AUX-N3S.PST.ASS canoe-CL:CONT buy-CL:GEN one-CL:GEN
gĩna-ro-re wero-wɨ
metal-CL:RECIP-N.SBJ buy-N3S.PST.ASS
'I didn't buy <u>a BOAT</u>, (I) bought <u>a COOKing pot</u>.'
[VOL: 20230623ejepa001.013]

Besides the MACRO-TRIGGERS outlined above, certain OTHER-type TRIGGERS are noted.[15] In addition to plurality-driven DCM in Hup [*jup*], shown in (83), two OTHER

[14] Following Aikhenvald (2003), Wojtylak (2021) utilizes the term *topical* to refer to a range of discourse-related notions: "The term *topical* refers to [S] and [P] arguments that are topical in discourse; to be marked with a topical marker, the constituent has to: (i) be the topic of the narrative; (ii) be referential (that is, treated as existing within the universe of discourse); and (iii) be important (but not necessarily contrastive)".

[15] I set aside Tena Kichwa [*quw*], given that the active TRIGGERS remain an open question at this time. As Grzech (2016, 61) puts it: "In elicitation context and in careful speech, [P is] always marked with the *accusative* suffix [*-ta*]. However, in more spontaneous speech the marker is often omitted. Preliminary observation suggests that these omissions are not associated with a given semantic type of verb, or properties of direct objects, including number, animacy, or semantic role in a given clause."

TRIGGERS are described in the relevant sources: namely positional effects and coordination effects.

Turning firstly to positional effects, Stenzel (2008) demonstrates in (96) that DCM is restricted to P-arguments arising in their canonical, immediately pre-verbal position in Wanano [*gvc*]. She goes onto claim that similar restrictions are found in Tukano [*tuc*], Barasano [*bsn*] and Kubeo [*cub*] (Stenzel 2013c, §10.3). Other observations regarding postional constraints are made in passing regarding Murui [*huu*] (Wojtylak, 2021, 258):

(96) Wanano$^{\Diamond}$ (*gvc* — EASTERN-TUKANOAN) [S08:8-9, exx. (23a)-(23b)]

a. *hi-piti-ro* ***chʉa*** *~da-ta-ra*
COP-COL-SG food get-come-VIS.IMPERF.2/3
'Everyone brings food.'

b. *ti-~da ~da-~sa'a* ***chʉa-re*** *chʉ yoa-ra*
ANA-PL get-MOV.inside food-OBJ eat make-VIS.IMPERF.2/3
'They take the food inside and eat (it).'

Another pattern is described for Kubeo [*cub*], such that *-re* may (optionally) arise on the first internal constituent in a coordinated P (Morse and Maxwell 1999, 111-113). This pattern is shown in (97). A similar restriction is noted for Ecuadorian Siona, as demonstrated in (98), lifted from the discussion in Chapter 5, where the regular requirement for animate P to be case-marked is suspended under coordination:

(97) Kubeo$^{\Diamond}$ (*cub* — EASTERN-TUKANOAN) [MM99:112-113, ex. (353)]
boa-'bI ***jupari-Ri-re*** *ape-ko* ***dẽĩdoki-ko*** ***warĩ-ko***
kill-3S.M rabo_colorado-CL:3D-OBJ other-CL:F ñacundá-CL:F jacha-CL:F
pidubã-ko *boa-'bI ĩ̃ bi̇ki̇-ki̇*
tucunaré-CL:F kill-3S.M 3PRO.S.M old-CL:M
'The elderly man recently caught a *rabo colorado*, (and also) a *ñacundá*, a *jacha*, and a *tucunaré* (guan).'

(98) Ecuadorian Siona$^{\blacklozenge}$ (*snn* — WESTERN-TUKANOAN)
. . . yo'oquë sai uye naso huasaiya baquë . . .
jo'o-ki̇ sai-i̇ **uje** **nahso** wa + sai-a
do-S.M.PRS.DEP go-S.M.PRS.DEP turkey wolly_monkey kill + go-NEG
bah-ki̇
be-N2/3S.PST.N.ASS
'. . . doing (this), (he) went out to hunt turkey, monkey(, etc.).'
[NAT*: 20151023orocr001.324]

Table A.21 outlines the DCM patterns for the INDIRECT OBJECT (R) in the NWA sample (*Q3.3*). A small subset of languages display these patterns (N=6), whereas the majority of the sample displays obligatory R-marking. The corresponding case-markers are identified in the table accordingly:

Language		Forms	Macro-triggers			
			anim	ref	disc	other
*Ecuadorian Siona♦ [Esnn]	$_{+}\checkmark^{[a,d]}$	{**-re**, -*ni*, (-*na*)}	✓ANIM		✓FOC	
Colombian Siona♦ [Csnn]	$_{+}\checkmark^{[d(a)]}$	{**-re**, -*ni*, -*na*}	✓(ANIM)		✓FOC	
Koreguaje♦ [*coe*]	$_{+}\checkmark^{[d(a)]}$	{**-re**, -*ni*}	✓(ANIM)		✓CONTR	
Murui [*huu*]	$_{\varnothing}\checkmark^{[d]}$	{-∅, -*mo*}		✓PRO	✓CONTR	
Tena Kichwa [*quw*]	$_{+}\checkmark^{[o]}$	{-*ta*, -*ma*}				✓TBD
Tariana◊ [*tae*]	$_{+}\checkmark^{[r]}$	{-*na*, -*nuku*}		✓PRO		

Obligatory marking on r: Máíhɨ̄kì♦ [*ore*] (-*re*); Kubeo◊ [*cub*] (-*re*); Tanimuka◊ [*tnc*] (-*re*); Tukano◊ [*tuc*] (-*re*); Desano◊ [*des*] (-*re*); Barasano◊ [*bsn*] (-*re*); Wanano◊ [*gvc*] (-*re*); A'ingae [*con*] (=*nga*); Imbabura Kichwa [*qvi*] (-*man*); Bora [*boa*] (-*vu*); Shiwiar [*acu*] (=$n^y a$); Awa Pit [*kwi*] (=*ta*); Kakua◊ [*cbv*] (=*di'*); Hup◊ [*jup*] (=*ăn*) Yuhup◊ [*yab*] (-*dìh*)

Table A.21: r-oriented dcm patterns (*Q3.3*)

Most languages in the sample display obligatory case-marking on the R-argument, including all Eastern Tukanoan languages, and Máíhɨ̃kì(Barnes 1999, 2006; Stenzel 2013d). Among those languages which do exhibit R-oriented DCM, the dominant pattern is the ALTERNATING-type (+), where various overt case-markers are accepted. This is demonstrated for the included Quechuan languages, where Cole (1982, 104-105) claims that "[t]he use of dative [*-man*] for [R] is obligatory in [Imbabura Kichwa [*qvi*], unlike other Quechuan languages]", as shown in (99). In contrast, (100) displays the dominant Quechuan pattern from Tena Kichwa [*quw*] where both *-ta* (glossed ACC) and *-ma* (glossed ALL) are acceptable on the R-argument (Grzech, 2016, 103-104):

(99) Imbabura Kichwa (*qvi* — QUECHUAN) [C82:104, ex. (418b),(419)]
mama-ka ***Juzi-man/*Juzi-ta***$^{(R)}$ *muti-ta*$^{(P)}$ *kara-rka*
mother-TOP J-DAT/ACC boiled_corn-ACC give-3.PST
'Mother served boiled corn$^{(P)}$ to José$^{(R)}$.'

(100) Tena Kichwa (*quw* — QUECHUAN) [G16:103, ex. (2.107)]
Pablo sisa-guna-ta$^{(P)}$ *kuya-n* ***Maria-ma/Maria-ta***$^{(R)}$
P flower-PL-ACC give-3 M-DAT/ACC
'Pedro gave flowers$^{(P)}$ to Maria$^{(R)}$.'

Similar ALTERNATING formal type is also attested across the Siona+ continuum — R-oriented DCM is described in detail in Section 3.2.3 in the *preliminary description* (Chapter 3). Although *-na* is strictly available on inanimate R in Ecuadorian Siona; this marker is available to encode all types of R-arguments in Colombian Siona [^{C}*snn*], as in (101):

(101) Colombian Siona$^{\blacklozenge}$ (*c-snn* — WESTERN-TUKANOAN) [W87:127-128]

a. ***ja'yë-na*** *do'rohuë-re insi-jë'ën*
older_brother-GOAL basket-OBJ give-IMPER
'Give (your) older brother the basket!'

b. ***yë'ë ma'yë-re*** *coca quëa-huë*
1SG older_brother-OBJ word tell-1SG.PST
'I gave the information to my older brother.'

The only language displaying the OPTIONAL-type (+) for R-oriented DCM is Murui [*huu*], as shown in (102). In this language, personal names and pronouns permit zero-marking, in alternation with the *locative*-marker, *-mo*, which is otherwise obligatory on the R-argument:

(102) Murui (*huu* — WITOTOAN) [W21:269, exx. (8.95),(8.96)]

a. *dɨo-kaɨ* ***kue*** *ine*
tobacco-CL:STEM 1SG give.IMPER
'Give me a cigarette! (normal reading)'

b. *dɨo-kaɨ* ***kue-mo*** *ine*
tobacco-CL:STEM 1SG-LOC give.IMPER
'Give ME a cigarette! (abrupt reading, brusque)'

As concerns the attested MACRO-TRIGGERS for R-oriented DCM in the NWA, these fall into two categories. On the one hand, there are those TRIGGERS, which restrict the distribution of particular alternatives: e.g., *-na* is restricted to inanimate R in Ecuadorian Siona, whereas only animates can be marked with *-ni* in all Siona+ languages. A similar pattern is noted in Tariana [*tae*], where *-na* is restricted to pronouns, and in Murui [*huu*] where *-mo* is only optional on personal names and pronouns. Setting aside these class-based restrictions, the dominant pattern is discourse-triggered, mirroring the facts presented for S-oriented DCM addressed in *Q3.1*.

Table A.22 collapses the presentation of L- and G-oriented DCM patterns in the NWA sample (*Q3.4* and *Q3.5*). Where languages display independent DCM patterns for each spatial function, separate rows are provided in the table, otherwise these are presented as a single row: [16]

[16] Note that the animacy-related MACRO-TRIGGER is eliminated from this discussion given that this dimension is generally semantically incompatible with spatial functions. Some authors do describe such fringe spatial arguments — i.e., see Stolz et al. 2014 and Haspelmath 2019 in particular. However, no comparable generalizations can be gleaned from the NWA sources at this time.

			Form(s)	Macro-triggers		
				REF	DISC	OTHER
*Ecuadorian Siona♦ [^{E}*snn*]	L	$+_{\varnothing}\checkmark^{[d(r,o)]}$	{-∅, **-re**, *-ni*, *-na*}	✓(PRO, SPEC)	✓FOC	✓(PRED)
	G	$_{\varnothing}\checkmark^{[d(r)]}$	{-∅, *-na*}	✓(PRO, SPEC)	✓FOC	
Colombian Siona♦ [^{C}*snn*]	L	$+_{\varnothing}\checkmark^{[d(r,o)]}$	{-∅, **-re**, *-na*}	✓PRO, (SPEC)	✓TOP	✓(PRED)
	G	$_{\varnothing}\checkmark^{[r]}$	{-∅, *-na*}	✓PRO, (SPEC)		
Koreguaje♦ [*coe*]	L	$+_{\varnothing}\checkmark^{[d,r(o)]}$	{-∅, **-re**, *-na*}	✓SPEC	✓TOP	✓(PRED)
	G	$_{\varnothing}\checkmark^{[r,d]}$	{-∅, *-na*}	✓SPEC	✓TOP	
Kubeo◇ [*cub*]	L=G^{+}	$+\checkmark^{[r(d)]}$	{**-re**, *-i*, *-rã*}	✓SPEC	✓(FOC)	
Tukano◇ [*tuc*]	L=G^{+}	$+_{\varnothing}\checkmark^{[r(d)]}$	{-∅, **-re**, *-pʉ*}	✓SPEC	✓(TOP)	
Desano◇ [*des*]	L=G^{+}	$+_{\varnothing}\checkmark^{[r(d)]}$	{-∅, **-re**, *-ge*}	✓SPEC	✓(FOC)	✓(PROTO)
Barasano◇ [*bsn*]	L=G^{+}	$+_{\varnothing}\checkmark^{[r(d)]}$	{-∅, *-hʉ*, (**-re**)}	✓(SPEC)	✓(FOC)	
Wanano◇ [*gvc*]	L	$+\checkmark^{[r(o)]}$	{**-re**, *-i*, *-pʉ*}	✓REF		✓(DIST)
	G^{+}	$+\checkmark^{[r]}$	{**-re**, *-pʉ*}	✓(SPEC)		
Murui [*huu*]	L=G	$+_{\varnothing}\checkmark^{[(r,o)]}$	{-∅, *-mo*, *-na*}	✓(PRO)		✓(PROTO)
Tena Kichwa [*quw*]	L	$+\checkmark^{[(o)]}$	{*-pi*, *-ma*}			✓(PRED)
Imbabura Kichwa [*qvi*]	G	$+\checkmark^{[(o)]}$	{*-man*, *-ta*}			✓TBD
Shiwiar [*acu*]	L=G	$+_{\varnothing}\checkmark^{[r,o]}$	{-∅, *=nam*, *=i*}	✓PRO		✓PROTO
Awa Pit [*kwi*]	L	$+\checkmark^{[o]}$	{*=ta*, *=pi*, *=mal*, *=ki*}			✓DIST
	G	$+\checkmark^{[o]}$	{*=ta*, *=pi*, *=mal*}			✓DIST
Tariana^{V} [*tae*]	L=G	$_{\varnothing}\checkmark^{[(o)]}$	{-∅, *-se*}			✓(PROTO)
Hup^{V} [*jup*]	L	$+\checkmark^{[d(o)]}$	{*-an*, -V*t*, *-có*}		✓CONTR	✓(PRED)
Yuhup^{V} [*yab*]	L=G^{+}	$+\checkmark^{[o]}$	{*-bâh*, *-'áh*, *-'áhà*}			✓DIST

No spatial case-markers: Máíhɨ̃kì♦ [*ore*] (-$\varnothing_{(L/G)}$)

Obligatory case-markers: Tanimuka◇ [*tnc*] (*-re'ká*$_{(L)}$; *-rã*$_{(G)}$); A'ingae [*con*] (*=ni*$_{(L/G)}$); Bora [*boa*] (*-ri*$_{(L)}$; *-vu*$_{(G)}$); Tena Kichwa [*quw*] (*-ma*$_{(G)}$); Imbabura Kichwa [*qvi*] (*-pi*$_{(L)}$); Kakua^{V} [*cbv*] (*=bŭ*$_{(L/G)}$); Hup^{V} [*jup*] (*-an*$_{(G)}$)

Table A.22: Spatial DCM patterns (*Q3.4* and *Q3.5*)

Firstly, considering the formal types of spatial DCM in the sample, certain languages are described as displaying ALTERNATING DCM (+). This is demonstrated in (103) for Kubeo [*cub*] and for Desano [*des*] in (104), each of which displays the 'generalized spatial-marking' pattern (*Q2.5*):

(103) Kubeo$^{\diamond}$ (*cub* — EASTERN-TUKANOAN) [CG19:413, exx. (21a)-(21c)]

a. ***kĩrãmi tãibɨ-rã*** *eda-rĩ* *õ-re* *upa-rĩ* wãi-ne
house yard-F.LOC arrive-CVB she-N.SBJ dance-N.SBJ spin-N.SBJ
da-ima=da
make-TAEM.II.3PL=REP
'They arrived <u>in the house yard</u> and made her dance, spinning.'

b. ***yo-i*** ***kĩrãmi-i*** *eda-ma*
here-LOC house-LOC arrive-3PL.ANIM
'They arrived <u>at the house</u>.'

c. *yãrãdawi* *obe-be=wɨ* ***yo-re*** eda-rãma
white_man be_many-NEG=CL:COL here-N.SBJ arrive-INFRR.3PL.ANIM
'A few whites have arrived <u>here</u>.'

(104) Desano$^{\diamond}$ (*des* — EASTERN-TUKANOAN) [M99:60, ex. (245)]
gɨa ***õ-re*** *era* *pɨ'rɨ* *bu'a-bɨ* *pare* ***finka-ge***
1PL.EXCL here-N.SBJ arrive after go_down-N3.PST finally farm-LOC
'We arrived <u>here</u>, after which (we) went down to <u>a farm</u> (our final destination).'

A similar spatial DCM pattern is found in Tukano [*tuc*], shown in (105), which also permits zero-marking — i.e., demonstrating the HYBRID-type (+∅):

(105) Tukano$^{\diamond}$ (*tuc* — EASTERN-TUKANOAN)
[R97:168-170,237; exx. (339),(580),(340)]

a. ***a'tó*** *kãrigɨ́* *dĩĩ-á-pɨ*
here sleep-NOM.M.SG be-REC.PST-1SG
'I usually sleep <u>here</u>.'

b. ***a'tó-de*** *dõ'ó-pɨ* *kãdí-gɨ́* *dĩĩ-a-ti*
here-N.SBJ where-FOC sleep-NOM.M.SG be-REC.PST-Q
'<u>Here</u>, where did you (M) sleep? (here in the city, where did you sleep?)'

c. ***a'tó-pɨ*** *a'tî-a*
here-FOC come-IMPER
'Come <u>HERE</u>!'

d. ***toó-pɨ-de*** *peêdu* *sĩ'dí-wɨ*
ANA.LOC-FOC-N.SBJ caxiri_drink drink-REP.PST.VIS
'I will drink *caxiri* <u>THERE</u> (an established location).'

Spatial DCM patterns are slightly more complex in Siona+ languages, which encodes the L-argument distinctly from the G-argument. For instance, Siona+ languages display OPTIONAL-type G-oriented DCM, where *na*-marking alternates with zero-marking, as shown for Koreguaje [*coe*] in (106):

(106) Koreguaje♦ (*coe* — WESTERN-TUKANOAN)
[CL85:111, ex. (45i)-(45iii)]

a. ***kee*** ***ja'o-cha*** *co~~ɨ~~* *sai-na'-me* *ch~~ɨ~~k~~ɨ~~na*
DEM.MED leaf-CL:RIV turtle go-PRP-N3S 1PL.EXCL
'We (EXCL) go <u>to the Caquetá River</u> for turtles . . . '

b. ***meja-hu~~ɨ~~-na*** *co~~ɨ~~* *mai-na'-me*
sand-CL:CONT-GOAL turtle come-PRP-N3S
'. . . the turtles come up <u>on the beach</u> . . . '

c. *jain~~ɨ~~ko* ***meja-hu~~ɨ~~*** *co~~ɨ~~* *mai-me*
many.PL.ANIM sand-CL:CONT turtle come-N3S
'. . . many are the turtles which come up <u>on the beach</u>.'

Turning to the TRIGGERS which drive the observed spatial DCM patterns, Tukanoan languages tend to display overlap with argumental MACRO-TRIGGERS — i.e., referentiality-related TRIGGERS are especially common in Eastern Tukanoan, and discourse-related TRIGGERS take precedence in Western Tukanoan. An instance of the former is demonstrated in the Tukano [*tuc*] examples in (105), where, for instance, the usage of spatial *-re* is similar to that found for its argumental functions. These examples also incorporate the latter category of TRIGGERS. In the Koreguaje [*coe*] sentences displayed in (106), Cook and Levinsohn (1985) claim that the usage of *-na* is constrained by the topicality status of the location in the unfolding discourse.

Considering the non-Tukanoan sample, spatial DCM is generally not triggered by the same MACRO-TRIGGERS as noted in the argumental domain. One dominant pattern across non-Tukanoan languages in the sample concerned the restriction of spatial DCM to prototypical spatial arguments — i.e., labeled PROTO in Table A.22. This pattern is described for Murui [*huu*] in (107) and for Shiwiar [*acu*] in (108):

(107) Murui (*huu* — WITOTOAN)
[W21:273,276, exx. (8.110)-(8.117),(8.123)-(8.124)]

a. i. *bi-rui-ya~~ɨ~~-do* ***nofi~~ɨ~~ko-mo*** *i-t~~ɨ~~-kue*
this.CTS-CL:DAY-PL-INST La_Chorrera-LOC exist-LK-1SG
'Nowadays, I live in <u>La Chorrera</u>.'

ii. *kue* *a~~ɨ~~* ***i~~ɨ~~*** *jaai-d-e*
1SG wife jungle_garden go-LK-3
'My wife went <u>to the jungle garden</u>.'

b. i. ***beno-mo*** *i-t~~ɨ~~-ka~~ɨ~~*
here-LOC exist-LK-1PL
'We live <u>here</u> (in this specific place, e.g., in this village).'

ii. ***beno-na*** *i-t~~ɨ~~-ka~~ɨ~~*
here-N.SBJ.TOP exist-LK-1PL
'We are <u>here</u> (passing through this place, e.g., in the jungle).'

iii. *jaai-ño-kai-ñe-no* *bu-e* ***beno*** *ñee∼ñee-d~~ɨ~~-o*
go-?-RAPID-NEG-SEQ Q1-CL:GEN here do∼RED-LK-2SG
'You didn't go, what are you doing (and doing) <u>here</u>?'

(108) Shiwiar (*acu* — CHICHAM) [K20:258, exx. (7.201)-(8.124)]

a. *wí=nʲa nukú-r=ka* ***hujuinʲtsá=nam***
1SG=OBJ mother-1SG.P=TOP J=LOC
puhú=itʲ-a-i
live+S.NMLZ=COP-3SG-DECL
'My mother lives <u>in Juyuintsa</u>.'

b. ***ikʲám*** *miŋka-ká-mia-ji*
forest disappear-PFV-DIST.PST-3SG+DECL
'He disappeared <u>in the forest</u>.' NB: non-spatial *íkʲam* (forest)

c. *turá amí=nʲa nukú-utsi-rmɨ=s* ***tu=í*** *puh-á-wa*
so 2SG=OBJ mother-DIM-2SG.P=FOC where-LOC live-IMIPFV-3SG
'So <u>where</u> does your mother live?'

Another commonly-described spatial-specific DCM TRIGGER found in the NWA sample concerns deictic distance (i.e., DIST). In some languages, case-marking distinctions encode distance from the speaker, in the domain of L- and G-oriented DCM patterns. This is shown for Yuhup [*yab*] in (109), and for Awa Pit [*kwi*], as illustrated in (110):

(109) Yuhup^{V} (*yab* — NADAHUPAN) [O02:160-161, exx. (34a)-(34c), (36a)]

a. *jàbᵐ bĕh óhí* ***tèjgⁿhôd-bâh***
dog sleep entryway-LOC.PROX
'The dog sleeps <u>close to the entryway</u>.'

b. ***tèjgⁿhôd-'áh*** *bŏk díí*
entryway-LOC.MED cooking_pot be
'The cooking pot is <u>beside the entryway</u> (i.e., moderately distant).'

c. *dó'jâp pébᵐí* ***tèjgⁿhôd-'áhà***
child be_seated entryway-LOC.DIST
'The child is sitting <u>beside the entryway</u> (i.e., distant from speaker).'

(110) Awa Pit (*kwi* — BARBACOAN) [O02:135-140, exx. (215),(225),(235),(239)]

a. ***Nulpe Media=ta*** *tu-y* *profesor=na*
NM=in be_in_place-NON.LOCUT teacher=TOP
'The teacher is <u>in Nulpe Medio</u>.'

b. ***aŋ=pa*** *awa su paa-ma-tɨ* ***uŋ=pa***
here=in(APPROX) person earth become-COMP-TERM there=in(APPROX)
ɨnkal i
mountain be.NON.LOCUT
'<u>Around here</u> (this side of the river) it's become cultivated. <u>Over there</u> (the other side of the river) it's bush.'

c. *Dolores* ***Pueblo Viejo=mal*** *tu=ma ka ki*
D PV=LOC be_in_place=INTER be_permanently Q
'Is Dolores <u>in Pueblo Viejo</u>?'

d. *paas* [***pala*** ***kwal=ki***] *pana-y*
two plantain trunk=at be_standing-NON.LOCUT
'The two [people] are standing
<u>at (beside) the trunk of the plantain tree</u>.'

(111) Hup^{V} (*jup* — NADAHUPAN)
[E08:174,187,189,367; exx. (36),(85),(95a),(108)]

a. *hohtĕg* ***dĕh-an*** *tɨh* *j'ɨ́d-íy*
canoe water-LOC 3PRO wash-DYNM
'He washes the canoe <u>at the water</u> (i.e., the port).'

b. *'ãh* *yamhidó'-óh* ***cãw-yucé-ét***
1SG sing-DECL São_José-OBL
'I sang <u>at São José village</u>.'

c. *cãw-ăn* *yæ̃wæ̃c-yɨ́'* ***n'ɨ́-có'=b'ay*** ***tŏk-có***
other-N.SBJ meet-TEL that-LOC=AGAIN belly-LOC
'(She) had already gotten another (child), <u>there</u>, <u>in the belly</u>.'

The final strictly spatial TRIGGER concerns certain patterns, determined by the predicate (i.e., PRED). For instance, this pattern is found in Siona+ languages, where *-na* is found on certain L-arguments, in place of the dominant L-marker, *-re* (see details in Section 3.3). The pair of Koreguaje [*coe*] sentences presented in (112) presents the general pattern. In this case, *-na* is selected with the predicate *k^{h}ãi-* (sleep), as it is in Siona and Sekoya varieties:

(112) Koreguaje$^{♦}$ (*coe* — WESTERN-TUKANOAN) [CC93:48-49]

a. ***wɨ'e-re*** *pa'i-mo* *repa-o*
house-N.SBJ be-F.SG that-F.SG
'She is <u>in the house</u>.'

b. ***wɨ'e-na*** *k^{h}ãi-me* *j̃ĩi*
house-GOAL sleep-PL children
'The children are sleeping <u>in the house</u>.'

This same pattern is demonstrated in the Tena Kichwa [*quw*] examples in (113). Here the *dative*-marker, *-ma*, is selected with the predicate *traba-* (work). The related Ecuadorian Siona predicate, *goamahña nee-* (do tasks, work), also requires *na*-marking where the location of work is expressed. Some similar instances have been described for Hup [*jup*] (Epps 2008, 367-369):

(113) Tena Kichwa (*quw* — QUECHUAN) [G16:68,62; exx. (2.47),(2.29)]

a. *Maria* *kaspi-wa* *wajta-n* *Juan-ta* ***wasi-pi***
M stick-INST hit-3 J-ACC house-LOC
'Maria hits John with a stick <u>in the house</u>.'

b. *Maria* ***chagra-ma*** *traba-nga* *ra-w-n*
M field-DAT work-FUT make-PROG-3
'Maria will work <u>in the field</u>.'

(114) Ecuadorian Siona$^{♦}$ (*snn* — WESTERN-TUKANOAN)
iye *ye'yahuë'ese'na* *goamaña* *neñë*
i-je ***je'ja + wɨ'e-se'e-na*** *goa-mahña* *nee-ñɨ*
DEM.PROX-CL:GEN teach + house-EXCL-GOAL task-DIM.PL make-N3S.PRS.ASS
'(I) work (lit. do little tasks)
<u>only at this school</u>.' [VOL: 20230626eyopa001.021]
[**Speaker comment:** The case-marker *-re* cannot be used on *school* in this sentence.]

The data presented above, in Section A.2 of the this Appendix comprises the NWA dataset. This dataset underpins the typological discussion put forth in Chapter 4 of this dissertation. Many of the data points laid out above merit further and more targeted investigation; however, this dataset is sufficient to facilitate the typologization of the Ecuadorian Siona facts, characterized in Chapter 3, on the basis of other patterns in the Tukanoan language family, and related patterns throughout the putative NWA area. Although this dataset is the result of a comparative study, outlined in Section A.1.2, several insights herein lay the groundwork for inquiry into diachronic aspects of Siona case-marking.

APPENDIX B

Siona DCM narrative dataset (supplement)

This Appendix complements the analysis of the Siona narrative dataset discussed in Chapter 5 of this dissertation. The dataset is composed by the folktale narrations identified in Table B.1, which corresponds to Table 5.1 in the content chapter:[1]

STORY TITLE (Story code)	TIME (HH:MM:SS)	WORD COUNT	NOMINAL TOKENS *N* (%)
Siona Genesis story (20151023orocr001)	21:33	1714	376 (21.94%)
The demon kwẽẽwahti (20151112oespa001)	15:48	1291	327 (25.33%)
The mother of jaguars (20151001oolpi001)	14:12	981	251 (25.59%)
The mother of the moon (20151112orapi001)	16:16	1350	331 (24.52%)
The children of Baina (20150811sfryi001/2)	33:25	2742	629 (22.94%)
POOLED TOTAL	**1:41:04**	**8078**	**1914 (23.69%)**

Table B.1: Composition of the Siona narrative dataset

[1] Consulted materials are all gathered from Bruil 2012 (Project ID IPF0211SG0067 [https://www.elararchive.org/dk0184]); archived as multimedia bundles, analyzed via the ELAN (EUDICO linguistic annotator) tool: ELAN (Version 6.9) [Computer software]. (2024). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from https://archive.mpi.nl/tla/elan.

In parallel fashion to the discussion in Chapter 5, English titles are attributed to each story in the dataset as a matter of convenience. The statistical models developed in this analysis strictly concern the full set of tokens in the pooled sample. Section 5.1 provides further details regarding the creation of the corpus for this analysis.

This Appendix is broken into two parts: Section B.1 spells out the annotation guidelines used to code the nominal tokens extracted from the dataset for the purposes of statistical analysis. Section B.2 provides full reports for the *varbrul* models produced in this quantitative analysis — i.e., for the S-marking sub-sample, the P-marking sub-sample, and for the L-marking sub-sample respectively. Each statistical model is accompanied by a series of illustrative examples for each tested grammatical relation.

B.1 Token-by-token annotation protocols

Section 5.2 of Chapter 5 presents in broad strokes the variationist methodology undertaken in order to tokenize the noun phrases arising in the Siona narrative dataset in Table B.1. The ultimate goal is to establish a token-based dataset against which the *varbrul* statistical technique may be implemented — see Section 5.2.2.

On the one hand, Chapter 5 indicates how the *dependent variable* is ascertained for nominal tokens (i.e., PROMINENT-marking status), which coincides with the descriptive facts laid out in the *preliminary description* in Chapter 3. Section 5.2.1 also motivates the set of *independent variables* selected for this analysis (i.e., TRIGGER CANDIDATES), based upon plausible TRIGGER for DCM identified in the literature.[2] As in Chapter 5, *independent variables* as sorted conceptually into FACTOR GROUPS by convention: (i) sentence-level factors, (ii) referential token-level factors, and (iii) information structural token-level factors. In what follows, the annotation protocols are made explicit for each FACTOR GROUP in turn.

B.1.1 Annotating sentence-level factors

This analysis tests for two sentence-level FACTORS: WORD ORDER and THEMATIC STATUS. They pertain to properties of nominal tokens which are determined by other elements in the sentence, rather than by properties of the nominal token or its referent per se. Section 2.2.3 briefly surveys this class of TRIGGERS as noted in the broader DCM literature.

The coding of WORD ORDER consists in identifying overt nominal tokens and the main verbal word in a given clause. The main verbal word is defined as the word within the clause which is inflected for tense, for clause type, and for SUBJECT-agreement (Bruil 2014, §5). A binary opposition is ascertained as to whether a given nominal token arises before the main verbal word (i.e., coded as *pre-verbal*), or after it (i.e., coded as *post-verbal*). (1) illustrates the application of this coding procedure:

[2] In variationist parlance, the crucial step of systematic annotation of FACTORS on a token-based dataset is often referred to as '*circumscribing* the variables' (Tagliamonte 2006, 2012).

(1) a. *i̱cato sa'nahuëbi etaë ?* [SENTENTIAL WORD ORDER: SXV]
ĩ-ɨ=kato$^{<\text{ pre-verbal S }>}$ **sa'niwɨ-bi**$^{<\text{ pre-verbal SOURCE }>}$
PRO3-CL:M=TOP inside-SRC
ehta-ɨ
go_out-2/3S.M.PST.N.ASS
'Did he come out from inside (there)?' [NAT*: 20151023orocr001.039]

b. *bacoña ja̱re guënahuëre debao* [SENTENTIAL WORD ORDER: VXS]
bah-ko-ña hãre **gɨna-wɨ-re**$^{<\text{ post-verbal L }>}$
live-3S.F.PST.N.ASS-REP like_that hard-CL:CONT-N.SBJ
debao$^{<\text{ post-verbal S }>}$
turkey_vulture
'The Debao (vulture *spec*) was in the sky.'
[NAT*: 20151112orapi001.243]

Although this technique permits the determination of sentential word order for each clause in the narrations, as shown in (1); it should be borne in mind that it is the coding of the WORD ORDER status of the nominal tokens themselves which concerns the present analysis. Accordingly, both tokens in (1a) are coded as *pre-verbal*, and those in (1b) as *post-verbal*. This design masks distinctions such as SOV vs. OSV in the resulting dataset.[3]

The second sentence-level FACTOR coded in this analysis concerns THEMATIC STATUS. The coding protocols adopted here are inspired by the corpus-based investigation of affectedness effects in Spanish *differential object marking* by von Heusinger and Kaiser (2011). The technique is adapted in this study such that it may be applied to the S-, P-, and L-tokens accordingly: As concerns S-marking, the *agentivity status* of a given token is ascertained; for P-marking, *affectedness status* is determined; whereas, for L-marking, the *subcategorization status* of each token is determined — i.e., classifying L-tokens by whether they are selectionally implicated by their predicate, or adjunctive in nature. Each of these types of THEMATIC STATUS is coded as a binary FACTOR, as laid out below.

Firstly, the predicate is recorded alongside the tokens for which it defines the semantic role. In this way, the THEMATIC STATUS of a given nominal token may be 'read off' the corresponding predicate in post hoc fashion. Following von Heusinger and Kaiser (2011), predicates are set against the typology proposed in Tsunoda (1985, 388), where certain predicate types select for, e.g., *agentive* or *non-agentive* S-arguments accordingly. The predicate typology assumed in this analysis is displayed in Table B.2, which is utilized to determined THEMATIC STATUS for S- and for P-tokens — note that [±affected] indicates that this is determined on a case-by-case basis:

[3]Table B.9 in the following section assembles baseline information about sentential word order in the sample, which is referenced at certain points in the discussion in Chapter 5 accordingly.

Predicate type	(e.g., Siona predicate)	S-status	P-status
0 (Copula, posture)	*ba'i-* (be,live)	[-agent]	
1 (Activity, motion)	*sai-* (go), *kohko-* (cough)	[±agent]	
1a (Effective action, +result)	*wai-* (kill)	[+agent]	[+affected]
1b (Effective action, -result)	*hahcho-* (shoot)	[+agent]	[±affected]
2 (Perception, ±attained)	*ahcha-* (listen to,hear)	[+agent]	[-affected]
3 (Knowledge, conception)	*gwahcha-* (know, think)	[-agent]	[-affected]
4 (Experience, undergoer)	*hũ̃-* (be sick,die)	[-agent]	[-affected]
5 (Possessive, stative)	*baa-* (have), *soa* (tall)	[-agent]	[-affected]

Table B.2: Predicate classes for corpus study and coding thematic status

A similar technique is applied to derive a binary variable for the *subcategorization status* for L-tokens. It is well-established that certain spatial arguments are naturally implicated by certain predicates, but are strictly adjunctive in other instances. For instance, the English sentences in (2) distinguish what Andrews (2007, 9) labels the *inner* and *outer* location respectively:

(2) a. Kangaroos inhabit these lands [Inner location → < Subcategorized >]
b. The boy was born in Scotland [Outer location → < Non-Subcategorized >]

In this study, *subcategorized* L-tokens are found with the basic spatial predicate, *ba'i-* (be, live), and with posture predicates, e.g., *ñu'i-* (be seated);[4] whereas other predicate types are classified as recruiting *non-subcategorized* L-arguments.

B.1.2 Annotating referential token-level factors

This analysis codes for two referent-level properties, which correspond to the best-attested DCM TRIGGERS in the literature (Aissen 2003; Sinnemäki 2014 — see extensive discussion in Chapter 2): namely, (i) SPECIFICITY STATUS, and (ii) ANIMACY STATUS. For the coding of each of these TRIGGER CANDIDATES, a two-part coding procedure is undertaken: Firstly, a more fine-grained categorial classification is performed. Secondly, these categories are converted into broader, FACTOR categories is performed for the sake of statistical analysis. This technique is intended to avoid ad hoc classifications and to promote systematicity and replicability of coding.

Beginning with SPECIFICITY STATUS, Table B.3 enumerates the available categories for coding tokens as a matter of corpus annotation, and the corresponding categories for the sake of statistical analysis:

[4] Although it is certainly possible to express posture without expressing location. However, given that this is a property of the L-token itself, this distinction is irrelevant to the present study. Postural predicates strictly reject *-na* marking, whereas *change-of-posture* predicates — e.g., *ha'ru-* (sit down), are coded as G-arguments. This is the original inspiration for this coding practice.

Annotated specificity status (Category attributed to token)	**Statistical specificity status** (Valuation used in *varbrul*)
< Definite > < Specific >	[+specific]
< Non-specific > < Generic >	[-specific]
< Pronoun >	

Table B.3: Specificity status taxonomy for sorting noun tokens in corpus study

At this stage a point of clarification is in order: The *varbrul* technique does not strictly accept binary variables, such that categorial variables are freely admitted (Tagliamonte 2006, 2012). Nonetheless best practices dictate that the smallest number of informative, and non-redundant, categories ought to be used as input for the application of the *varbrul* statistical technique.

Per Table B.3, pronominal tokens are separated from full noun tokens. Pronominal tokens refer to classes of independent reduced nominal expressions — e.g., *ĩõ* (she), *hãõ* (that (F), there), *to* (there). Full noun tokens are subsequently categorized on the basis of a binary *specific* vs. *non-specific* divide.

SPECIFICITY STATUS on full noun tokens is a context-sensitive property, which may shift from one mention to another for a given referent. (3) illustrates this fact with the two G-tokens lifted from the *Siona Genesis tale*. Although these tokens are surface-indistinguishable, *bẽhẽtubɨ-na* (fallen trunk); the first-mention token in (3a) is coded as *non-specific*, whereas the later mention in (3a) is coded as *specific*:

(3) a. *bejetubëna mënituiquë ba'quëña*
bẽhẽ-tubɨ-na$^{<\text{ Non-Specific }>}$ mɨni + tui-kɨ
be_fallen-CL:STICK-GOAL go_up + sit_on_something-CL:M
ba-∼'-kɨ-ña
be-RMT.PST-2/3S.M.PST.N.ASS-REP
'(He) climbed onto a fallen trunk and
sat.' [NAT*: 20151023orocr001.021]

b. *. . . bejetubëna ya'o jenaquë ba'quëña*
bẽhẽ-tubɨ-na$^{<\text{ Specific }>}$ ja'o hẽna-kɨ
be_fallen-CL:STICK-GOAL mud stick-CL:M
ba-∼'-kɨ-ña
be-RMT.PST-2/3S.M.PST.N.ASS-REP
'(He) stuck mud
to the fallen tree trunk.' [NAT*: 20151023orocr001.061]

The (non-)specificity of a tokens as in (3) is determined by a careful reading of the tales and, where applicable, by specificity cues in the translations (originally in Spanish). Since Siona grammar does not have (in)definite articles, often a value judgment must be made on a mention-by-mention basis.

Despite the lack of grammaticalized articles in Siona, this study recognizes a *definite* category for nominal tokens, borrowing a convention from the corpus-based investigation Paraguayan Guaraní *differential object marking* (Shain and Tonhauser

2010). The *definite* category is identified based upon the presence of certain elements within the noun phrase, associated with definite semantics: i.e., demonstratives, overt Possessors (shown in (4a)), and appositive pronoun-noun configurations (shown in (4b)):

(4) a. [**Context:** After the demon *kwẽẽwahti* escapes, she leaves her ax behind.]
go'ini i̲o̲ tsë'bo baëña – baquëna i̲ dëjo̲ bacoña
go'i-ni **ĩ-o** **tsɨ'bo**$^{<\text{Definite}>}$ baa-ɨ-ña
return-SS 3PRO-CL:F ax have-2/3S.M.PST.N.ASS-REP
baa-ki-na **ĩ-ɨ** **dĩhõ**$^{<\text{Definite}>}$ bah-ko-ña
have-M.PRS.DEP-DS 3PRO-CL:M wife be-2/3S.F.PST.N.ASS-REP
'He went back and had her ax. Having (it), his wife was (there) (so they say).'
[NAT*: 20151112oespa001.159]

b. [**Context:** The mother goes to ask the grandmother if she has seen her kids.]
cani ja̲maca i̲o̲ tsi̲ bë'cacobi saniñacona ...
kaa-ni hãmahka [**ĩ-o** **tsĩĩ** **+ bɨ'ka-ko-bi**]$^{<\text{Definite}>}$ sani
say-SS then 3PRO-CL:F child + parent-CL:F-SBJ go
+ ñaa-ko-na
+ see-S.F.PRS.DEP-DS
'(She) said (this), and then she, the mother, left to go and look (for them) (so they say).' [NAT*: 20151001oolpi001.042]

A final non-pronominal *generic* category is recognized in Table B.3, which subsumes noun tokens that refer to masses or substances. Such an instance is shown in (5) from *The mother of jaguars* tale. These tokens are treated as a sub-type of non-specific token for the present purposes:

(5) *...ñataquëna ñocua oyaja̲'coa'ë cacona jo̲tsi̲ ga̲'yojë bani bëajë'ë̲ caoña*
ñahta-ki-na **ñohkwa**$^{<\text{Generic}>}$ oja-hã'-ko-a'-ɨ
dawn-S.M.PRS.DEP-DS chambira press-PRP-CL:F-COP-N3S
kaa-ko-na hotsĩĩ gã'jo-hɨ bani + bɨa-hɨ̃'ɨ̃
say-S.F.PRS.DEP-DS grandchild.COL play-PL.PRS.DEP be + stay-IMPER
kaa-o-ña
say-2/3S.F.PST.N.ASS-REP
'When (the day) dawns, (she) wants to press chambira. (The grandmother) said, "grandkids, stay here and play!" ' [NAT*: 20151001oolpi001.055]

A similar two-step procedure is implemented for the sake of coding the ANIMACY STATUS of nominal tokens. Unlike SPECIFICITY STATUS, this FACTOR is context-independent, and may be 'read off' of the nominal token once its referent is identified. It is crucial to note that this is *not* the same as *grammatical* animacy status, in light of the grammatical process of promotion, outlined in Section 3.2.2 of the *preliminary description*.[5] Table B.4 enumerates the annotated animacy categories and indicates their conversion into statistical categories:

[5]Such a discrepancy between notional ANIMACY STATUS and grammatical animacy status arises with certain lexical items, such as celestial bodies: e.g., *ñañɨ* (moon).

Annotated animacy status (Category attributed to token)	**Binary animacy status** (Valuation used in *varbrul*)
< Human >	[+animate]
< High animal >	
< Low animal >	
< Inanimate (count) >	[-animate]
< Inanimate (mass) >	
< Abstract >	
< Location noun >	
< Temporal noun >	

Table B.4: Animacy status taxonomy for sorting noun tokens in corpus study

Firstly, a higher-level classification is performed, where spatiotemporal nominal tokens are distinguished from argumental nominal tokens. Notional animacy is conceptually incompatible with the former class, and elimiated as a FACTOR for the L-marking *varbrul* model reported in Section B.2.3. Secondly, the initial fine-grained classification of argumental tokens is undertaken as in Table B.4. Unlike specificity, *notional* animacy status may be determined based upon the referent of each noun token. Therefore, the class of promoted nominals, described in Section 3.2.2, is coded as *inanimate* tokens for the present purposes. The sentences lifted in (6) demonstrate instances of each of the fine-grained annotated animacy categories enumerated in Table B.4:

(6) a. *jɑ̲macarebacato baquëña te'i*
hãmahka-reba=kato$^{<\text{Time}>}$ bah-ki-ña **te'-ɨ**$^{<\text{Human}>}$
then-INTENS=TOP be-2/3S.M.PST.N.ASS-REP one-CL:M
'Once upon a time, there was a man (lit. there was one (M)).'

b. *. . . yaje u̲cuquë ba'iquëbi ñaëña*
jahe$^{<\text{Mass}>}$ ũhku-kɨ **ba'i-kɨ-bi**$^{<\text{Human}>}$ ñaa-ɨ-ña
ayahuasca drink-CL:M live-CL:M-SBJ see-2/3S.M.PST.N.ASS-REP
'(He) lived on (lit. the one who lived (M)), drinking jahe and having visions.' [NAT*: 20151112oespa001.003-004]

c. *cani jɑ̲maca nasobi mëco'ë tsi̲ gajaco i̲oje̲ . . .*
kaa-ni hãmahka **nahso-bi**$^{<\text{High animal}>}$ mɨh-ko'ɨ
say-SS then wolly_monkey-SBJ climb-3S.F.PST.ASS
tsĩĩ$^{<\text{Human}>}$ gaha-ko **ĩ-o-hẽ**$^{<\text{Human}>}$
child put_down-3S.F.PRS.ASS PRO3.F-CL:F-ADD
'(She) said "then the monkey climbed (up there)", then she put down her child too.' [NAT*: 20151112oespa001.102]

d. *jɑ̲ro utibëbi to̲meji huare jɑ̲ utibi to'teji caoña*
hã-ro$^{<\text{Location}>}$ **uhti-bɨ-bi**$^{<\text{Count}>}$ tõme-hi ware hã
DEM.MED-CL:LOC wasp-CL:RND-SBJ fall-3S.M.PRS.ASS child DISC.PART
uhti-bi$^{<\text{Low animal}>}$ to'te-hi kaa-o-ña
wasp(spec)-SBJ sting-3S.M.PRS.ASS say-2/3S.F.PST.N.ASS-REP
' "The wasp nest is falling there. Careful, child, (this) wasp stings!", (she) said.' [NAT*: 20151112oespa001.132]

No meaningful grammar differences are expected across these sub-categories on the basis of the *preliminary description* in Chapter 3. The initial fine-grained classification is preformed to ensure systematicity in the annotations.

B.1.3 Annotating information structural properties

This analysis codes for three information structural token-level properties: (i) CONTRASTIVITY STATUS, (ii) LOOK-BACK TOPICALITY STATUS, and (iii) LOOK-AHEAD TOPICALITY STATUS. The annotation protocols are outlined for each of these in turn.

Several techniques exist for coding CONTRASTIVITY STATUS in corpus-based research — see Vallduví (2016, §23.5) for a recent discussion. This analysis follows the operational definition of *contrastivity* established Myhill and Xing (1996, as implemented more recently in Vallejos Yopán 2009). For the authors, a nominal token is contrastive where the referent (or the concept) that it denotes belongs to an explicit or an implicit discourse set, comprising alternative entities (or concepts).[6,7] The relevant set of alternatives must be interpreted as activated in the common ground, at the point of the mention at hand.

Consider the S-tokens identified in (7), lifted from the *Mother of the moon* tale, where two referents are explicitly contrasted. An instance of implicit contrast is exemplified by the L-token in (8) from *The demon kwẽẽwahti* narrative:

(7) [**Context:** Two personified birds are out hunting and checking their traps for food.][8]

a. *pë'përi ba'iji iñore mëabecato ba'i ñabi − yo'ni jɑ̱maca i̱ɔ̱maca ba'iona ...*
pɨ̱'pɨ̱ri$_i$ ba'i-hi [ihño-re mɨ̱abe=kato] ba'i-ɨ̱
vulture be-3S.M.PRS.ASS here-N.SBJ above=TOP be-S.M.PRS.DEP
ñaa-bi − jo'-ni hãmahka **ĩ-o-mahka**$_i$ ba'i-o-na
see-3S.M.PST.ASS − do-SS then 3PRO-CL:F-DIM.SG be-S.F.PST.DEP-DS
'The VULture$_i$ was up HERE, looking out. Doing this, as she$_i$ was (there)...'

b. *...jɑ̱re tacarobi ña baco − ñani nëcadojaiyo*
hãre **tahkaro-bi**$_j$ ñaa-a bah-ko − ñaa-ni
like_that black_vulture-SBJ see-NEG be-2/3S.F.PST.N.ASS − look-SS
nɨhka + dohai-o
be_standing + wander-S.F.PST.DEP
'...likewise the black VULture$_j$ was looking out. She wandered on foot (i.e., on the ground) to go on the look out ...'
[NAT*: 20151112orapi001.174-176]

[6] Myhill and Xing (1996, 310-312) identify seven different types of *discourse* sets: (i) complementary, (ii) organizational, (iii) proximate, (iv) hierarchical, (v) rhetorical, (vi) conjoined, (vii) analogical; while also recognizing other types of conventionalized or implicit sets. This study disregards their parameter of *lists*, which does not seem to be applicable to the data considered here.

[7] Note that by *concept* contrast I refer to kind-level readings of the type "X wants FISH" (i.e., not another kind of food). Such semantic type-flexibility is often tacitly assumed in these works focused on matters of implementation.

[8] In (7), the first token, *pɨ̱hpɨ̱ri* (vulture), is treated as non-contrastive; whereas the other two token, *ĩ̃ɔ̃mahka* (she (DIM)) and *tahkaro-**bi*** (black vulture), are coded as contrastive with respect to one another. Therefore, only the *closest* mention to its referential competitor is coded as contrastive.

(8) [**Context:** The demon *kwẽ̃wahti* is chasing the protagonist, and he goes from tree to tree to find a new place to hide. After sitting up on a branch, it breaks. . .]
. . . jëyequë neni ja'obi pë'pëni ja'ruña − ja'runi ja̲rore tuquëña
hɨje-kɨ̵ nee-ni ha'o-bi pɨ̵∼pɨ̵-ni
break.INTRS-S.M.PRS.DEP make-SS leaf-INST pile∼REDUP-SS
ha'ru-ɨ̵-ña ha'ru-ni **hã-ro-re**
sit_down-2/3S.M.PST.N.ASS-REP sit_down-SS DEM.MED-CL:LOC-N.SBJ
tuh-kɨ̵-ña
sit_atop-2/3S.M.PST.N.ASS-REP
'. . . (it) broke and so (he) took leaves and covered (himself) and sat down. (He) sat down and was sitting at THAT spot.'
[NAT*: 20151112oespa001.097]

To complement the instances *contrastive* noun tokens identified above, (29) illustrates two instances of *non-contrastive* inanimate P-tokens. In the case of (9a), the token *gõ'no* (chicha)[9] displays PLAIN zero-marking; whereas the token, *biada'ca-**re*** (soup, lit. ají pepper broth) bears PROMINENT *-re* marking in (9b). These demonstrate how the determination for CONTRASTIVITY STATUS is made independently from case-marking:

(9) a. [**Context:** God sends the squirrel-person to search for active watering holes. He is given chicha to sustain himself. This drink is not contrasted with alternatives.]
. . . saiquëbi ja̲re tres to'to sani ja'runi i̲ beoje̲o̲ña go̲'no u̲cuchaoquë ba̲'quëña
sai-kɨ̵-bi hãre tres to'to sani + ha'ru-ni ĩ-ɨ̵
go-CL:M-SBJ like_that three watering_hole go + sit_down-SS 3PRO-CL:M
beo + hẽ̃õ-ɨ̵-ña **gõ'no** ũhku + chao-kɨ̵
NEG.EXIS + leave_behind-3S.M.PST.N.ASS-REP chicha drink + finish-CL:M
ba-∼'-kɨ̵-ña
be-RMT.PST-3S.M.PST.N.ASS-REP
'. . . (he) (lit. the going one (M)) went on to three watering holes and sat down and had none left. He finished drinking the chicha.'
[NAT*: 20151023orocr001.280]

b. [**Context:** The man wants to eat the stew, which he does not realize contains the remains of his children. This meal is not contrasted with another.]
. . . ñajëna yë'ë biada'care cusi'i caëña
ñaa-hi-na jɨ̵'ɨ̵ **bia** **+ da'ka-re** ũhku-si'-ɨ
see-PL.PRS.DEP-DS 1S aji_pepper + liquid-N.SBJ drink-FUT-N3S
kaa-ɨ̵-ña
say-2/3S.M.PST.N.ASS-REP
'. . . (they) watched (him) as (he) said "I will take (some) soup".'
[NAT*: 20150811sfryi001.264]

In general, this analysis adopts a conservative coding approach for CONTRASTIVITY STATUS, where only instances with a clear, local contrast meet the coding threshold.

[9] *Chicha* is a commonly consumed fermented beverage, which is typical of various indigenous communities across the NWA area.

Borderline instances are coded as *non-contrastive* by default, so as to avoid over-representation of controversial tokens and to focus on the case-marking patterns found on clearly contrastive tokens.

Turning to the coding protocols for LOOK-BACK TOPICALITY STATUS, the basic idea is that tokens which refer to entities which are more recently mentioned are more topical. This analysis adopts the referential distance metric, as applied by Shain and Tonhauser (2010) in their study of topicality effets in *differential object marking* in Paraguayan Guaraní. This metric is calculated specifically for the sample under scrutiny. As the stories are coded, each referent is given an arbitrary referent ID number, which is attributed to referential nominal tokens accordingly.[10] Each main clause in the narration is given a clause number in sequence, such that the number of intervening clauses between mentions may be tallied post hoc — i.e., its *referential distance* value.

A binary *look-back topical* vs. *look-back non-topical* distinction is ascertained via a two-step process: Firstly, the mean referential distance is calculated in the sample, thereby fixing the topicality threshold.[11] Secondly, nominal tokens are categorized as *look-back topical* if its referential distance is at or below the topicality threshold, or as *look-back non-topical* where its referential distance surpasses the threshold. Table B.5 presents the topicality thresholds developed for S- and P-marking sub-samples in the Siona narrative dataset, formatted as *mean referential distance* (token count):[12]

STORY TITLE	S-DISTANCE	P-DISTANCE
Siona Genesis story	*7.78* (141)	*7.028* (36)
The demon kwẽẽwahti	*5.393* (117)	*8.5* (42)
The mother of jaguars	*5.169* (77)	*2.967* (30)
The mother of the moon	*6.546* (99)	*6.897* (39)
Children of Baina	*5.431* (223)	*5.221* (131)
Pooled narrations	***6.065* (657)**	***5.942* (278)**
Excl. non-referential tokens	80	204

Table B.5: Referential distance thresholds for look-back topicality status

Based upon Table B.5, a threshold of 6 clauses is maintained for all DCM relations in the sample. Accordingly, tokens are coded as follows: (i) *look-back topical* ≤ 6 intervening clauses between mentions, vs. (ii) *look-back non-topical* > 6 intervening clauses.[13]

Turning to what is labelled LOOK-AHEAD TOPICALITY STATUS, on conceptual grounds, this FACTORS refers to the persistence of a nominal token's referent(s) in unfolding discourse.[14] This study adapts the methodology employed by Chiriacescu

[10]Note that a referent index is not appropriate for non-referential noun tokens, such as those coded as *abstract* (e.g., *ohko* (water)). These are coded as *look-back non-topical* by default.

[11]The term *topicality threshold* does not arise in Shain and Tonhauser (2010).

[12]Whereas Shain and Tonhauser (2010) observe a mean referential distance for objects at 15 contiguous main clauses, the mean distance was lower in the Siona data.

[13]An example of the implementation of this technique is provided in (25) in what follows.

[14]Several languages are argued to exhibit more local topicality effects in their DCM patterns: e.g., in *differential object marking* effects in Romanian (Chiriacescu and von Heusinger 2010);

and von Heusinger (2010) to categorize *look-ahead topical* and *look-ahead non-topical* nominal tokens in the sample. The former value is attributed to tokens whose referent (or the concept that they denote) is maintained as an at-issue argument in the ensuing two main clauses. This referent need not be overtly expressed in these continuations.

The narrative episode in (10) is from *The mother of the moon* tale. Both of the noun tokens in (10a) are coded as *look-ahead topical*, given that their referents (i.e., identified as *i* and *k* for convenience) are implicated in the clause following its mention, although the *k*-token is not fitted with an overt noun phrase in (10b). Conversely, the P-token, *jahe* (ayahuasca), in (10b), is coded as *look-ahead non-topical* since it does not persist in the immediately following lines of narration:

(10) a. *bacoña io debaocato – jamacarebacato ëjëre baoña bacona ...*
bah-ko-ña [**ĩ-o dẽbao=kato**]$_i$ < Look-ahead topical >
be-2/3S.F.PST.N.ASS-REP 3PRO-CL:F vulture(spec)=TOP
hãmahka-reba=kato **ĩhĩ-re**$_k$ < Look-ahead topical >
then-INTENS=TOP husband-N.SBJ
baa-o-ña bah-ko-na
have-2/3S.F.PST.N.ASS-REP live-S.F.PST.DEP-DS
'There was her$_i$, the Dẽbao$_i$. Then, (she$_i$) lived on and had a husband$_k$ ...'

b. *... yaje cuquë si'ahua'ire ñaquë ba'quëña: se'se, yahuë, huequë, si'abaire ñaquë*
jahe < Local non-topical > ũhku-kɨ si'a-wa'i-re ñaa-kɨ
yaje drink-S.M.PRS.DEP all-PL.AN-N.SBJ know-CL:M
ba-∼'-kɨ-ña sẽ'se yawɨ wẽhkɨ
be-RMT.PST-2/3S.M.PST.N.ASS-REP boar collared_peccary tapir
si'a-bai-re ñaa-kɨ
every-PL.ANIM-N.SBJ know-CL:M
'... (he$_k$) drank yaje and got to know everybody – (he$_k$) knew the boars, the peccaries, the tapirs, everybody ...'
[NAT*: 20151112orapi001.005-009]

The same coding protocol is applied to spatial tokens.

B.2 Results: Statistics and illustrative data

This section supplements the presentation of the quantiative analysis in Section 5.3, by supplying two types of data. Firstly, illustrative examples of tokens are lifted from the Siona narrative dataset, to accompany this more targeted discussion in Chapter 5. Secondly, full statistical reports are spelled out for the *varbrul* models developed for each fully tested sub-sample in this analysis — including the S-marking sub-sample (B.2.1), the P-marking sub-sample (B.2.2), and the L-marking sub-sample (B.2.3) respectively. Illustrative examples for coded categories are organized on a grammatical

in, e.g., Evenki (Dalrymple and Nikolaeva 2011); and many other (unrelated) languages (Iemmolo 2010); and in the domain of *differential agent marking* in a number of Australian, Austronesian and Tibeto-Burman languages (e.g., Chappell and Verstraete 2019, Dunn and Meakins 2023, Riesberg 2018).

relation-by-relation basis in what follows. Section B.2.4 assembles miscellaneous examples from the sample, which pertain to patterns that do not concern these three fully tested grammatical relations.

B.2.1 The results for S-oriented DCM

This section supplements the discussion of DCM patterns in the S-marking sub-sample in Section 5.3.2. Table B.6 reports the full results for the *varbrul* analysis, which represents a very good fit for the dataset (i.e., Input = 0.393);[15] on this basis, strong statistical evidence is produced for each of the effects reported in this table:

[15]In statistical parlance the *goodness of fit* metric determines how informative the significance and factor weight measures are for a given linear regression model (i.e., *varbrul*-analysis). Both the number of included tokens and the degree of variation determine the goodness of fit for resulting statistical model. This metric produces a number between 0 and 1, and a value closer to 0 than to 1 is indicative of a strong fit for the dataset.

N = 727			* p = < 0.001
Input: 0.393			! p = > 0.05
	Factor weight	% Prominent	***N***
***Contrastivity**			
Contrastive S	0.93	95.18%	83
Non-contrastive S	0.07	11.6%	644
Range	*86*		
***Specificity**			
Pronominal S	0.7	17.6%	346
Specific non-pro S	0.42	26.4%	273
Non-specific S	0.37	19.4%	108
Range	*33*		
***Word order**			
Post-verbal S	0.62	12.1%	173
Pre-verbal S	0.38	24%	554
Range	*24*		
		Non-significant Factors	
! Look-ahead topic (persistence)			
Look-ahead topical S		22.67%	375
Look-ahead topical S		19.6%	352
! Agentivity			
Agentive S		23.88%	448
Non-agentive S		16.85%	279
! Animacy			
Animate S		21.7%	599
Inanimate S		18.75%	128
! Look-back topic (distance)			
Look-back topical S		20.74%	351
Look-back topical S		21.18%	376

Table B.6: Full *varbrul* report: S-marking sub-sample

The remainder of this section lifts illustrative tokens for each tested factors in the analysis, presented in the order in which they arise in Section 5.3.2. An illustrative instance of a *contrastive* S-token is provided in (7) above.

ANIMACY STATUS does not have a significant effect on the PROMINENT-marking rate S-marking in the sample. (11) provides an instance of super-token, where both an *animate* and an *inanimate* S-token arises in the same line with PROMINENT *-bi*:

(11) *ja̲ro utibëbi to̲meji huare ja̲ utibi to'teji caoña*
hã-ro **uhti-bi-bi**$^{<\text{Inanimate}>}$ tõme-hi ware hã
DEM.MED-CL:LOC wasp-CL:RND-SBJ fall-3S.M.PRS.ASS child DISC.PART
uhti-bi$^{<\text{Animate}>}$ to'te-hi kaa-o-ña
wasp (spec)-SBJ sting-3S.M.PRS.ASS say-2/3S.F.PST.N.ASS-REP
' "The wasp nest is falling there. Watch out, child, (this) wasp stings!", she said.' [NAT*: 20151112oespa001.132]

Likewise AGENTIVITY STATUS does not impact the selection of PROMINENT-marking in the S-marking sub-sample. (12) illustrates instances of both PROMINENT- and PLAIN-marking arising on *agentive* S-tokens. (13) lifts instances of PROMINENT- and PLAIN-marking arising on *non-agentive* S-tokens respectively:

(12) a. *ja añabi cujų'į caëña*
hã **ãña-bi** kũ-hu'i kaa-i-ña
DISC.PART snake-SBJ bite-3S.M.PST.ASS say-2/3S.M.PST.N.ASS-REP
' "Oh no, a snake bit (me)!" (he) said.'
[NAT*: 20151112orapi001.234]

b. [**Context:** A tapir character is helping the children of Baina get fruit down from a tree. He tells them to go find cover.]
. . . yo'jëna iñona titouña į hueqëjẹ
jo'-hi-na ihño-na tihto-i-ña **ĩ-i̵**
do-PL.PRS.DEP-DS here-GOAL kick-2/3S.M.PST.N.ASS-REP 3PRO-CL:M
wẽhki̵-hẽ
tapir-ADD
'. . . (they) did so and the tapir kicked (the tree) right here.'
[NAT*: 20150811sfryi001.479-480]

(13) a. *saina u'ubi baquëña*
sai-i-na **u'u-bi** bah-ki-ña
go-S.M.PST.DP-DS sloth-SBJ be-2/3S.M.PST.N.ASS-REP
'(He) left and there was a sloth.' [NAT*: 20150811sfryi001.307-308]

b. [**Context:** After the children of Baina kill their mother, a jaguar, they take her teeth to a shaman.]
yai cujį gachoquëjẹ ju'įñe bą'ëña
jai kũhĩ gahcho-ki̵-hẽ hũ'ĩ-ñe
jaguar tooth remove-CL:M-ADD die-CL:GEN
ba-∼'-i-ña
be-RMT.PST-2/3S.M.PST.N.ASS-REP
'The one who removed the jaguar's tooth died.'
[NAT*: 20150811sfryi001.479-480]

Concerning WORD ORDER, (14) demonstrates instances of PROMINENT *-bi* marking and PLAIN zero-marking arising on *pre-verbal* S-tokens. (15) lifts instances containing *post-verbal* PROMINENT *-bi* marking, which is relatively rare, and *post-verbal* PLAIN zero-marking:

(14) a. *. . . huaëna tsįmacabi gọẹọ gọẹọ gọẹọ caëña*
wai-i-na **tsĩĩ-mahka-bi** kwẽ kwẽ kwẽ
kill-3S.M.PRS.DEP-DS child-DIM.SG-SBJ " "
kaa-i-ña
say-2/3S.M.PST.N.ASS-REP
'(He) killed (him), and the child screamed "kwẽ kwẽ kwẽ!" '
[NAT*: 20151001oolpi001.022-023]

b. [**Context:** The demon *cu̲e̲huati* only found one of the two men hiding in the village.]
i̲o̲ ñañe ba̲o̲
ĩ-o ñaa-ñe bãã-o
3PRO-CL:F see-CL:GEN NEG.AUX-3S.F.PST.ASS
'She didn't see (him).' [NAT*: 20151112oespa001.029]

(15) a. *dëmëquë ba̲quëña ja̲mubi*
dɨmɨ-kɨ ba-∼-kɨ-ña **hãmu-bi**
emerge_from_water-CL:M be-RMT.PST-3S.M.PST.N.ASS-REP armadillo-SBJ
'An armadillo emerged.' [NAT*: 20151023orocr001.038]

b. *bacoña ja̲re guënahuëre debao*
bah-ko-ña hãre gɨna-wɨ-re **debao**
live-3S.F.PST.N.ASS-REP like_that hard-CL:CONT-N.SBJ turkey_vulture
'The Debao (vulture *spec*) was in the sky.'
[NAT*: 20151112orapi001.243]

Turning to instances of the various SPECIFICITY STATUS categories, the instances in (16) serve to demonstrate instances of PROMINENT *-bi* and PLAIN zero-marking on *pronominal* S-tokens. Corresponding PROMINENT-marking and PLAIN-marking S-tokens are provided for *specific* tokens in (17), and *non-specific* tokens in (18):

(16) a. [**Context:** The demon *kwẽ̃ẽwahti* maliciously imitates a human man.]
i̲o̲ ëjë̲ cañeje̲ caoña i̲o̲bi
ĩ-o ɨ̃hɨ̃ kaa-je-hẽ kaa-o-ña
3PRO-CL:F husband say-CL:GEN-ADD say-2/3S.F.PST.N.ASS-REP
ĩ-o-bi
3PRO-CL:F-SBJ
'She (i.e., the *cu̲e̲huati* demon) spoke like her husband speaks.'
[NAT*: 20151112oespa001.170]

b. [**Context:** The children of Baina are speaking to the father from inside the womb of their mother, telling him to follow them.]
. . . cajë saijëna be'teña i̲
kaa-hɨ sai-hɨ-na be'te-ɨ-ña
say-PL.PRS.DEP say-PL.PRS.DEP-DS follow-2/3S.M.PST.N.ASS-REP
ĩ-ɨ
3PRO-CL:M
'. . . (they) said this and went on and he followed (them).'
[NAT*: 20151023orocr001.108]

(17) a. [**Context:** The mother of the attacked children asks the grandmother if she has seen the children, she says 'no'.]
. . . cani ja̲maca i̲o̲ tsi̲ bë'cacobi sani ña̲o̲ña
kaa-ni hãmahka [**ĩ-o tsĩ̃ bɨ'ka-ko-bi**] sani
say-SS then 3PRO-CL:F child parent-CL:F-SBJ go
+ ñaa-o-ña
+ see-2/3S.F.PRS.N.ASS-REP
'. . . (she) said (this) then the mother went out to see.'
[NAT*: 20151001oolpi001.042]

b. [**Context:** Shaman protagonist scares away the demon *kwẽ̃wahti* with his newfound magical abilities. She leaves behind an ax.]
. . . go'ini i̲o̲ tsë'bo baëña − baquëna i̲ dë̲j̲o̲ bacoña
go'i-ni **[ĩ-o tsɨ'bo]** baa-ɨ-ña
return-SS 3PRO-CL:F ax have-2/3S.M.PST.N.ASS-REP
baa-kɨ-na ĩ-ɨ dɨ̃hõ bah-ko-ña
have-S.M.PRS.DEP-DS 3PRO-CL:M wife be-2/3S.F.PST.N.ASS-REP
'He went back and had her ax − he had it when his wife was (there, at the house).' [NAT*: 20151112oespa001.159]

(18) a. *dëmëquë ba̲quëña ja̲mubi*
dɨmɨ-kɨ ba-∼-kɨ-ña **hãmu-bi**
emerge_from_water-CL:M be-RMT.PST-3S.M.PST.N.ASS-REP armadillo-SBJ
'An armadillo emerged.' [NAT*: 20151023orocr001.038]

b. [**Context:** one of Baina's children kills the man who is taking car of them. Then moves onto their next victim.]
. . . huani saquëña − a'ritsiaya baquëña
wani + sah-kɨ-ña **a'ri + tsia-ja**
kill + go-2/3S.M.PST.N.ASS-REP small + river-CL:RIV
bah-kɨ-ña
be-2/3S.M.PST.N.ASS-REP
'. . . (he) killed (him) and left − there was a stream.' [NAT*: 20150811sfryi001.418-419]

Other coded categories of S-tokens are not clearly illustrated by lifted tokens as the above-noted categories.

B.2.2 The results for P-oriented DCM

This section supplements the discussion of DCM patterns in the P-marking sub-sample in Section 5.3.3. Table B.7 reports the full results for the *varbrul* analysis, which represents a moderate fit for the dataset (i.e., Input = 0.606); on this basis, fairly strong statistical evidence is produced for each of the effects reported in this table:

N = 478			* p = < 0.001
Input: 0.606			! p = > 0.05
	Factor weight	**% prominent**	***N***
***Contrastivity**			
Contrastive P	0.86	87.9%	33
Non-contrastive P	0.14	19.6%	445
Range	*72*		
***Look-ahead topic (persistence)**			
Look-ahead topical P	0.67	51.1%	45
Local non-topical P	0.33	21.5%	433
Range	*34*		
***Specificity**			
Specific P	0.64	32.7%	165
Pronominal P	0.46	20.2%	109
Non-specific P	0.40	19.6%	204
Range	*24*		
Word order			
Pre-verbal P	0.60	26%	382
Post-verbal P	0.40	19.7%	86
Range	*20*		
Animacy			
Inanimate P	0.58	27.2%	290
Animate P	0.42	19.7%	188
Range	*16*		
		Non-significant Factors	
! Affectedness			
Affected P		25.09%	291
Non-affected P		22.99%	157
! Look-back topic (distance)			
Look-back topical P		23.65%	148
Non-agentive S		24.55%	330

Table B.7: Full *varbrul* report: P-marking sub-sample

The remainder of this section lifts illustrative tokens for each tested factors in the analysis, presented in the order in which they arise in Section 5.3.3.

As is established in Section 5.3.3, this analysis is based upon the PLAIN-PROMINENT distinction, which corresponds to PLAIN zero-marking and PROMINENT *-re* marking for inanimate P-tokens, and to PLAIN *-re* and PROMINENT *-ni* marking for promoted and animate P-tokens. In what follows, illustrative examples are provided for each of these sub-patterns for the sake of reference, although surface case-marking is irrelevant for the statistics reported in Table B.7. To open the discussion, canonical instances of PROMINENT and PLAIN marking alternatives are provided for *animate* and *inanimate*

P in the sub-sample in (19) and (20) respectively:

(19) a. *ja̲ja̲e̲ yo'co a̲co'ë ts̲i̲dohuëre yureta̲'a̲ quëaja̲'cua'ia'ë ja'quëre careña*
hãhẽ jo'-ko ãh-ko'ɨ **tsı̄ı̄-dowɨ-re**
like_that do-S.F.PRS.DEP eat-3S.F.PST.ASS child.COL-PL.AN-N.SBJ
jure-tã'ã kɨa-hã'-ko-wa'i-a'-ɨ ha'-kɨ-re
now-CNT.EXP tell-PRP-CL:F-PL.AN-COP-N3S parent-CL:M-N.SBJ
kaa-re-ña
say-N2/3S.PST.N.ASS-REP
‘ “(She) did this and ate the (other) children, now we will tell father!”, (the remaining, hiding, children) said.’ [NAT*: 20151001oolpi001.067]

b. [**Context:** A caiman had bit and killed one of the children of Baina, so his sibling avenges him and then goes and kills the rest of the caimans in the river.]
. . . yo'quë yecua'ini cu̲cu̲joa ba̲'quëña
jo'-kɨ **jeh-ko-wa'i-ni** kũh∼kũ + hoo-a
do-S.M.PRS.DEP other-CL:F-PL.AN-N.SBJ2 bite∼REDUP + cut-NEG
bah-kɨ
be-3S.M.PST.N.ASS
‘. . . (he) did (this) and tore up the Others (∼with his teeth).’
[NAT*: 20150811sfryi001.438]

(20) a. *nocabo bahuë yë'ë ja̲o̲je̲ beo . . .*
nohka-bo baa-wɨ jɨ'ɨ hã-o-hẽ
banana-CL:ENCLOS have-N3S.PST.N.ASS 1S DEM.MED-CL:F-ADD
beo-ɨ
NEG.EXIS-S.M.PRS.DEP
‘ “I had a banana plantation (and now) it is not (there).” ’
[NAT*: 20151001oolpi001.067]

b. [**Context:** The children of Baina have killed the mother of the household and put her in the stew, but the father is not yet aware.]
. . . ñajëna yë'ë biada'care cusi'i caëña
ñaa-hɨ-na jɨ'ɨ **bia** **+ da'ka-re** ũhku-si'-i
see-PL.PRS.DEP-DS 1S aji_pepper + liquid-N.SBJ drink-FUT-N3S
kaa-ɨ-ña
say-2/3S.M.PST.N.ASS-REP
‘. . . (they) watched (him) as (he) said “I will take (some) soup.” ’
[NAT*: 20150811sfryi001.264]

The analysis reported in Table B.7 detects no statistical effect on the basis of AFFECTEDNESS STATUS. Firstly, strictly concerning animate P-tokens, PLAIN and PROMINENT-marking is freely found on both *affected* P-token, lifted in (21), and *non-affected* P-tokens, as shown in (22):

(21) a. *jaje yo'co aco'ë tsidohuëre yureta'a quëaja'cua'ia'ë ja'quëre careña*
hãhẽ jo'-ko ãh-ko'ɨ **tsĩĩ-dowi-re**
like_that do-S.F.PRS.DEP eat-3S.F.PST.ASS child.COL-PL.AN-N.SBJ
jure-tã'ã kɨa-hã'-ko-wa'i-a'-ɨ ha'-kɨ-re
now-CNT.EXP tell-PRP-CL:F-PL.AN-COP-N3S parent-CL:M-N.SBJ
kaa-re-ña
say-N2/3S.PST.N.ASS-REP
' "(She) did this and ate the children, now we will tell father", (they) said.' [NAT*: 20151001oolpi001.067]

b. *... turini cha'cani huasi'i cani yë'ë yo'huë jeguaye jo'na yë're quëaye caoña*
turi-ni cha'ka-ni wah-si'-i kaa-ni jɨ'ɨ jo'-huë hewaye
mouse-N.SBJ2 jump-SS kill-FUT-N3S say-SS 1S do-N3S.PST.ASS when
jo'-o-na jɨ'ɨ-re kɨa-je kaa-o-ña
do-S.F.PST.DEP-DS 1S-N.SBJ tell-CL:GEN say-2/3S.F.PST.N.ASS-REP
' "I want to pounce on the rats and kill (them). I didn't do all of this so (they) tell on me," (she) said.' [NAT*: 20151001oolpi001.070]

(22) a. [**Context:** Speaking of the husband of the *Mother of the moon* who has just died at the beginning of the tale.]
ba'iquëbi juquëna yureta'a mamajëre cayabare baëña ëmë
ba'i-kɨ-bi hũh-kɨ-na jure-tã'ã **mama-hɨ-re**
live-CL:M-SBJ die-S.M.PST.DEP-DS now-CNT-EXP child-CL:COL-N.SBJ
kaja-ba-re baa-ɨ-ña ɨ̃mɨ
two-AN-N.SBJ have-2/3S.M.PST.N.ASS-REP man
'(He) lived on and died, at that time (he) had two children, sons...'
[NAT*: 20151112orapi001.011-013]

b. [**Context:** The mother of the children of Baina has a final meal before having the eggs which will contain his children.]
...anichaojëna tsiare baoña io
ãni + chao-hɨ-na **tsia-re** baa-o-ña
eat + finish-PL.PRS.DEP-DS egg-N.SBJ have-2/3S.F.PST.N.ASS-REP
ĩ-o
3PRO-CL:F
'Once (they) finished eating, she had eggs.'
[NAT*: 20150811sfryi001.112]

In a similar vein to the above, both PLAIN zero-marking and PROMINENT *-re* marking are readily observed on *affected* and *non-affected* inanimate P-tokens in the sample. The former category is exemplified in (23), and the latter category in (24):

(23) a. *ñacona yureta'a io ñocuare oyaco ñu'icobi huare aijë'ë hua'i bani caoña*
ñaa-ko-na jure-tã'ã ĩ-o **ñohkwa-re**
see-S.F.PRS.DEP-DS now-CNT.EXP 3PRO-CL:F chambira-N.SBJ
oja-ko ñu'ĩ-ko-bi ware ãĩ-hɨ̃'ɨ̃ wa'i baa-ni
press-S.F.PRS.DEP be_seated-CL:F-SBJ child.VOC eat-IMPER meat have-SS
kaa-o-ña
say-2/3S.F.PST.N.ASS-REP
'(She) saw that she was sitting there, pressing chambira (lit. the sitting one (F) pressed chambira) – "(My) child, eat!", (she) held meat and said (this).' [NAT*: 20151001oolpi001.043]

b. *. . . ñataquëna ñocua oyaja'coa'ë cacona jotsi ga'yojë bani bëajë'ë caoña*
ñahta-kɨ-na **ñohkwa** oja-hã'-ko-a'-ɨ
dawn-S.M.PRS.DEP-DS chambira press-PRP-CL:F-COP-N3S
kaa-ko-na hotsĩ gã'jo-hɨ bani + bɨa-hɨ̃'ɨ̃
say-S.F.PRS.DEP-DS grandchild.COL play-PL.PRS.DEP be + stay-IMPER
kaa-o-ña
say-2/3S.F.PST.N.ASS-REP
'When (the day) dawns, (she) wants to press chambira. (The grandmother) said, "grandkids, stay here and play!" '
[NAT*: 20151001oolpi001.055]

(24) a. [**Context:** God has saved one man after the flood and makes his destroyed house reappear from thin air.]
ñajë'ë më'ë huë'e cani ñaquëna jaoje huë'e
ñaa-hɨ̃'ɨ̃ **mɨ'ɨ̃ wɨ'e** kaa-ni ñaa-kɨ-na hã-o-hẽ
see-IMPER 2SG house say-SS see-S.M.PRS.DEP-DS DEM.MED-CL:F-ADD
wɨ'e
house
' "See your house," (he) said and (he) look and the house (was there) too.' [NAT*: 20151023orocr001.133]

b. [**Context:** The mother of the children of Baina has just had eggs and the goes into the trees to find fruit. Somebody is approaching to take the eggs so she goes to talk to them (in their eggs).]
të'caëna yureta'a gachani yëquëni ñaja'co cani yureta'a io co'rere
tɨ'ka-ɨ-na jure-tã'ã gahcha-ni **jɨhkɨ-ni**
dry_up-S.M.PST.DEP-DS now-CNT.EXP bring_down-SS 1PL.EXCL-N.SBJ2
ñaa-hã'-ko kaa-ni jure-tã'ã ĩ-o ko're-re
see-PRP-CL:F say-SS now-CNT.EXP 3PRO-CL:F louse-N.SBJ
'Once (it) dried up, then he took (them) down – "(she) will see us", she said to the lice.' [NAT*: 20150811sfryi001.089-090]

The coded property of LOOK-BACK TOPICALITY, based upon the referential distance metric outlined in Section B.1.3, is not visible on the token. Contrast the *non-topical* instance in (25a), which displays 33 intervening main clauses between it and its previous mention in the tale, and the *topical* coreferential token, which arises in the following line in (25b) — i.e., a single intervening clause makes this highly topical on the basis of this metric:

(25) a. *i̱ ja̱mu ti ja'o dasiquë ja̱re sa'nahuëna bëaë ?*
[**ĩ-ɨ̵ hãmu ti ha'o daa-sih-kɨ̵**] hãre sa'niwɨ̵-na
3PRO-CL:M armadillo ANA mud bring-COMP-CL:M like_that inside-GOAL
bɨa-ɨ̵
stay-2/3S.M.PST.N.ASS
'Did the armadillo that brought the mud stay inside (there)?'
[NAT*: 20151023orocr001.098]

b. *ja̱mucato ja̱re beouña – nehuesëña*[16]
hãmu=kato hãre beo-ɨ̵-ña
armadillo=TOP like_that not_be-2/3S.M.PST.N.ASS-REP
neewesɨ̵-ɨ̵-ña
do_forever-2/3S.M.PST.N.ASS-REP
'As for the armadillo, (he) was not around, he went away.'
[NAT*: 20151023orocr001.099]

Next, turning to CONTRSATIVITY STATUS, (26) provides instances of *contrastive* animate and inanimate P-tokens. Instances containing *non-contrastive* animate and inanimate P-tokens are laid out in (27) and (29) respectively:

(26) a. [**Context:** A recently deceased husband transforms into a wasp to bring his family to the Upper World to join him in the afterlife.]
. . . ja̱ ñacona – ja̱ i̱oni jëasi'i cani huajëconi jëasi'i cani yo'quëna . . .
ñaa-ko-na hã **ĩ-o-ni** hɨa-si'-i
see-S.F.PRS.DEP-DS DISC.PART 3PRO-CL:F-N.SBJ2 bring_across-FUT-N3S
kaa-ni **wahɨ-ko-ni** hɨa-si'-i kaa-ni yo'-kɨ-na
say-SS alive-CL:F-N.SBJ2 bring_across-FUT-N3S say-SS do-S.M.PRS.DEP-DS
'. . . when (she) saw (this), oh, (he) wanted to bring HER,
the living woman, (to the other side). (He) did (this) and then. . . '
[NAT*: 20151112orapi001.081-082]

b. [**Context:** In order to ascend and rejoin their father in the sky, the children of Baina must drink yage.]
. . . hueo yë'cato so̱quëtene hua'ire a̱i̱coa'ë caco ja̱ë̱hua'ini . . . i̱o̱cato goeye ba̱co
wee-o jɨ'ɨ=kato **sõhkɨ̵tene + wa'i-re**
lie_in_hammock-S.F.PST.DEP 1S=TOP catfish(spec) + meat-N.SBJ
ãĩ-ko-a'-ɨ̵ kaa-ko hã-ɨ̵-wa'i-ni
eat-PRP-CL:F-COP.3S.F say-S.F.PRS.DEP DEM.MED-CL:M-PL.AN-N.SBJ2
ĩ-o=kato gwee-je bãã-ko
3PRO-CL:F=TOP not_want-CL:GEN NEG.AUX-S.F.PRS.DEP
'. . . (she) lay (there, in the hammock) and said to them, "As for ME, I eat (only) CATfish meat". She didn't want (i.e., yage).'
[NAT*: 20151001oolpi001.065]

[16]It remains to be seen whether the presence of the contrastive topic marker *=kato* is conditioned by the high degree of (global) topicality associated with the P-argument in (25b).

(27) a. [**Context:** The jaguar demon talks to herself as she watches the grandchildren flee. She uses the term *turi* (mouse) to refer to the human children.]
. . . turini cha'cani huasi'i cani yë'ë yo'huë jeguaye jo'na yë're quëaye caoña
turi-ni cha'ka-ni wah-si'-i kaa-ni jɨ'ɨ jo'-huë hewaye
mouse-N.SBJ2 jump-SS kill-FUT-N3S say-SS 1S do-N3S.PST.ASS when
jo'-o-na jɨ'ɨ-re kɨa-je kaa-o-ña
do-S.F.PST.DEP-DS 1S-N.SBJ tell-CL:GEN say-2/3S.F.PST.N.ASS-REP
‘ “I want to pounce on the mice and kill (them). I didn't do all of this so (they) tell on me,” (she) said.’ [NAT*: 20151001oolpi001.070]

b. [**Context:** The children of Baina were born from the saliva of Baina on the ground in the form of lice. He sees a human woman and warns them.]
. . . të'caëna yureta'a gachani yëquëni ñaja'coa cani yureta'a i co're rere . . .
tɨ'ka-ɨ-na jure-tã'ã gahcha-ni **yɨhkɨ-ni**
dry_up-S.M.PST.DEP-DS now-CNT.EXP go_down-SS 1PL.EXCL-SS
ñaa-hã'-ko-a kaa-ni jure-tã'ã ĩ-ɨ ko're-re
see-PRP-CL:F-COP.3S.F say-SS now-CNT.EXP 3PRO-CL:M louse-N.SBJ
‘. . . once (his saliva) dried up, (he) came down and told the lice, “(she) will see us”. . . ’ [NAT*: 20150811sfryi001.089-090]

(28) [**Context:** The demon *kwẽẽwahti* tracks the shaman protagonist in a tree. She calls him *nahso* (wolly monkey), as many demons call humans in Siona narratives.]
ñu'ɨjë'ë nasore cueni aija'cua'ia'ë caoña mamaquëre
ñu'ĩ-hɨ̃'ɨ̃ **nahso-re** kwẽẽ-ni
be_seated-IMPER wolly_monkey-N.SBJ get_down-SS
ãĩ-hã'-ko-wa'i-a'-ɨ kaa-o-ña mama-kɨ-re
eat-PRP-CL:F-PL.AN-COP-N3S say-2/3S.F.PST.N.ASS-REP child-CL:M-n.sbj
‘ “(Stay) sitting down! (We) will get the monkey down and eat (him)”, (she) said to (her) son.’ [NAT*: 20151112oespa001.106-107]

(29) a. [**Context:** God sends the squirrel-person to search for active watering holes in the newly formed land. He is given chicha to sustain himself.]
. . . saiquëbi jare tres to'to sani ja'runi i beojeoña go'no ucuchaoquë ba'quëña
sai-kɨ-bi hãre tres to'to sani + ha'ru-ni ĩ-ɨ
go-CL:M-SBJ like_that three watering_hole go + sit_down-SS 3PRO-CL:M
beo + hẽõ-ɨ-ña **gõ'no** ũhku + chao-kɨ
NEG.EXIS + leave_behind-3S.M.PST.N.ASS-REP chicha drink + finish-CL:M
ba-∼'-kɨ-ña
be-RMT.PST-3S.M.PST.N.ASS-REP
‘. . . (he) (lit. the going one (M)) went on to three watering holes and sat down and had none left. He finished drinking the chicha.’
[NAT*: 20151023orocr001.280]

b. [**Context:** The man wants to eat the stew, which he does not realize contains the remains of his children. This meal is not contrasted with another.]
. . . ñajëna yë'ë biada'care cusi'i caëña
ñaa-hi̵-na ji̵'i̵ **bia** **+ da'ka-re** ũhku-si'-i
see-PL.PRS.DEP-DS 1S aji_pepper + liquid-N.SBJ drink-FUT-N3S
kaa-i̵-ña
say-2/3S.M.PST.N.ASS-REP
'. . . (they) watched (him) as (he) said "I will take (some) soup".'
[NAT*: 20150811sfryi001.264]

As concerns LOOK-AHEAD TOPICALITY (persistence or (re)activation), two PLAIN-marked tokens are shown in (10) above. This example demonstrates an instance of PLAIN *-re* marking on an animate *topical* token, and PLAIN zero-marking on an inanimate *non-topical* accordingly. Another clear instance of the latter is shown in (30) below:

(30) a. *huajoturubë sasiquëbi huajo dutani jëyoni te'ja'a wani jeosi'i cani huahueaëña . . .*
waho + turu-bë saa-sih-ki̵-bi **waho** duhta-ni hi̵jo-ni
arrow + bag-CL:RND take-COMP-CL:M-SBJ arrow remove-SS break.TRS-SS
te'e-hã'ã wani + hẽõ-si'-i kaa-ni
one-PATH kill + leave_behind-FUT-N3S say-SS
+ wẽã-i̵-ña
+ leave_behind.TRS-2/3s.m.pst.n.ass-rep
'Since he was carrying a quiver of arrows − he took out one arrow and broke (it). He wanted to kill (the animal with it), so he killed (it).'

b. *. . . huaëna tsimacabi goeo goeo goeo caëña*
wai-i̵-na tsĩ̃-mahka-bi kwẽ kwẽ kwẽ
kill-3S.M.PRS.DEP-DS child-DIM.SG-SBJ " "
kaa-i̵-ña
say-2/3S.M.PST.N.ASS-REP
'(He) killed (him), and the (animal) child screamed "kwẽ kwẽ kwẽ!" '
[NAT*: 20151001oolpi001.022-023]

A corresponding instance of PLAIN *-re* marking arises on the *non-topical* animate P-token lifted in (31):

(31) [**Context:** In *Mother of the moon* the narrator mention how the protagonist's husband dies, but he has sons, then the narrator's attention turns back to the wife.]
a. ba'iquëbi juquëna yureta'a mamajëre cayabare baëña ẽmẽ
ba'i-ki̵-bi hũh-ki̵-na jure-tã'ã **mama-hi̵-re**
live-CL:M-SBJ die-S.M.PST.DEP-DS now-CNT-EXP child-CL:COL-N.SBJ
caya-ba-re baa-i̵-ña ĩmi̵
two-AN-N.SBJ have-2/3S.M.PST.N.ASS-REP man
'(He) lived on and died, at that time he had two children, sons . . . '

b. *. . . daoco i̱ airo daoquë hua'i nesi ma'a̱ja̱'a̱ onidojaioña*
dao-ko ĩ-ɨ ai-ro dao-kɨ wa'i
wander-S.F.PRS.DEP 3PRO-CL:M big-CL:LOC wander-S.M.PRS.DEP meat
nee-si + ma'ã-hã'ã oni + dohai-o-ña
make-COMP + path-PATH cry + wander-2/3S.F.PST.N.ASS-REP
'. . . (she) along the path where he would go to hunt, crying as (she) wandered.' [NAT*: 20151112orapi001.011-015]

Unlike the instances above, PROMINENT-marking is also attested on *non-topical* P-tokens, although at a relatively lower frequency. (32) demonstrates an instance of PROMINENT *-re* marking on the *non-topical* inanimate P. A similar instance with PROMINENT *-ni* marking on a *non-topical* animate P-token is provided in (33):

(32) a. *ñacona yureta̱'a̱ i̱o̱ ñocuare oyaco ñu'i̱cobi – huare a̱ijë'ë hua'i bani caoña*
ñaa-ko-na jure-tã'ã ĩ-o **ñohkwa-re**
see-S.F.PRS.DEP-DS now-CNT.EXP 3PRO-CL:F chambira-N.SBJ
oja-ko ñu'ĩ-ko-bi ware ãĩ-hɨ̃'ɨ̃ wa'i baa-ni
press-S.F.PRS.DEP be_seated-CL:F-SBJ child.VOC eat-IMPER meat have-SS
kaa-o-ña
say-2/3S.F.PST.N.ASS-REP
'(She) saw that she was sitting there, pressing chambira (lit. the sitting one (F) pressed chambira) – "(My) child, eat!", (she) held meat and said (this).'

b. *. . . caona hua'i mani a̱si'i cani co̱meco ñacona i̱o̱ aibë ba'idë̱ so̱quëperubë nesicobi*
kaa-o-na wa'i maa-ni ãh-si'-i kaa-ni kõme-ko
say-S.F.PST.DEP-DS meat take-SS eat-FUT-N3S say-SS serve-S.F.PRS.DEP
ñaa-ko-na ĩ-o ai-bɨ ba'i-dɨ̃ ↑[sõhkɨ
see-S.F.PRS.DEP-DS 3PRO-CL:F ancestor-CL:COL be-CL:TEMP wood
+ peru-bɨ nee-sih-ko-bi]
+ instrument-CL:RND make-COMP-CL:F-INST
'. . . when (she) said (this), she (i.e., the mother) wanted to serve herself from soup using a wooden ladel, the type that the ancestors used at that time, as (the other) watched.' [NAT*: 20151001oolpi001.043-044]

(33) a. *ja̱o̱ni huatoje̱o̱ni sani yo'ni saquëña i̱. . .*
hã-o-ni wahto + hẽõ-ni sai-ɨ jo'-ni
DEM.MED-CL:F-NSBJ2 separate + leave_behind-SS go-S.M.PRS.DEP do-SS
sah-kɨ-ña ĩ-ɨ
go-2/3S.M.PST.N.ASS-REP 3PRO-CL:M
'He left her behind and went on . . . '

b. *. . . sai hua'i sasëyotëoquë saquëña jobo airona*
sai-ɨ wa'i saa + sɨjo + tɨo-kɨ
go-S.M.PRS.DEP meat bring + smoke + put-S.M.PRS.DEP
sah-kɨ-ña hobo ai-ro-na
go-2/3S.M.PST.N.ASS-REP middle big-CL:LOC-GOAL
'. . . he went on and brought smoked meat and put (it) in the middle of the forest.' [NAT*: 20150811sfryi001.030-032)]

In order to round off the discussion of LOOK-AHEAD TOPICALITY, (34) provides an instance of PLAIN zero-marking arising on a *topical* inanimate P-token:

(34) a. [**Context:** The children of Baina (jaguar-people) have killed their human mother and are bringing her teeth to be enchanted by a shaman.]
a'ri cujimaña isiña isijë'ë cani saquëña yaije ...
a'ri kũhĩ-mahña ĩhsi-hɨ̃'ɨ̃ kaa-ni sah-kɨ-ña
small tooth-DIM.PL give-IMPER say-SS go-2/3S.M.PST.N.ASS-REP
jai-hẽ
jaguar-ADD
'(He) handed over the little teeth. (He) said "give (them)!" and then the jaguar left ...'

b. *... me caquë sëquë duñera achasi'i caëña jare i tëijani achaquëna huahuaro cujimañare sëquë sëquë sëquë baiñobi caco uiño*
mee kaa-kɨ sii-kɨ duñe-reba
how say-2/3S.M.PRS.N.ASS enchant-2/3S.M.PRS.N.ASS better-INTENS
ahcha-si'-i kaa-ɨ-ña hãre ĩ-ɨ
understand-FUT-N3S say-2/3S.M.PST.N.ASS-REP like_that 3PRO-CL:M
tɨiha-ni ahcha-kɨ-na wawaro + kũhĩ-mahña-re
turn_over.TRS-SS listen-S.M.PRS.DEP-DS cockroach + tooth-DIM.PL-N.SBJ
sɨhkɨ sɨhkɨ bãĩ-o-bi kaa-ko ũĩ-o
" " person-CL:F-SBJ say-S.F.PRS.DEP be_lying-S.F.PRS.DEP
'... (he) wanted better understand how to do this enchantment. So (he) turned around and listened as the woman lying there said *sëquë sëquë sëquë* over the cockroach teeth.' [NAT*: 20150811sfryi001.316-324]

The next factor, which is discussed in Section 5.3.3 concerns SPECIFICITY STATUS. Firstly, as concerns pronominal P-tokens, all anticipated PLAIN and PROMINENT case-markers are attested. (35) demonstrates instances of PROMINENT *-ni* marking and PLAIN *-re* marking on promoted inanimate pronominal tokens. Similarly marked animate personal pronoun tokens are lifted in (36):

(35) a. [**Context:** The mother of the children of Baina is starting to become ill, when Baina himself flies away to the Upper World.]
... yo'quëna io bë'cacoje caoña iconi guijë'ë caoña – neo garabëre
jo'-kɨ-na ĩ-o bɨ'ka-ko-hẽ
do-S.M.PRS.DEP-DS 3PRO-CL:F parent-CL:F-ADD
kaa-o-ña ↑**ih-ko-ni** gũĩ-hɨ̃'ɨ̃
say-2/3S.F.PST.N.ASS-REP DEM.PROX-CL:F-N.SBJ2 bite-IMPER
kaa-o-ña nee-o gara-bɨ-re
say-2/3S.F.PST.N.ASS-REP make-3S.F.PST.ASS charcoal-CL:RND-N.SBJ
'When (he) did this, her mother said, "bite into this!", she had made charcoal.' [NAT*: 20150811sfryi001.092-093]

b. [**Context:** All of the animals in the forest are coming to take down a great tree.]
icore cueni jeoñu'u cueni jeouna jaobi taitota'a tsiaya ba'ija'coa
↑**ih-ko-re** kwẽẽ-ni hẽõ-ñu'ũ kwẽẽ-ni
DEM.PROX-CL:F-N.SBJ take_down-SS leave_behind-HORT take_down-SS
hẽõ-ɨ-na ↑hã-o-bi tãi-to-tã'ã
leave_behind-S.M.PST.DEP-DS DEM.MED-CL:F-SBJ fall-COND-CNT.EXP
tsia-ja ba'i-hã'-ko-a'
river-CL:RIV be-PRP-CL:F-COP.3S.F
‘ “Let's take it (i.e., the tree) down, when (it) comes down, we will leave (it) and it (lit. she) will become the river.” ’
[NAT*: 20151023orocr001.251-253]

(36) a. [**Context:** The husband of a lady from Earth dies and wants to bring his family with him to the Upper World.]
. . . cani daijẽ'ẽ cani sañu'u mësaruni caquë daajẽ'ẽ si'a tsëcabëre
kaa-ni dai-hɨ̃'ɨ̃ kaa-ni saa-ñu'ũ **mɨhsaru-ni** kaa-kɨ
say-SS come-IMPER say-SS bring-HORT 2PL-N.SBJ2 say-S.M.PRS.DEP
daa-hɨ̃'ɨ̃ si'a tsɨhka-bɨ-re
bring-IMPER all family-CL:COL-N.SBJ
‘. . . (he) said “come! (I) want to bring you (PL). The whole family!” ’
[NAT*: 20151112orapi001.028-029]

b. [**Context:** The jaguar people, children of Baina, have just killed their human step-mother, now the want to kill the rest of their human family.]
sahuë'huë mëareña yaibaicato iohua'ire cucujoñu'u cani . . .
sawɨ'wɨ + mɨa-re-ña jai + bãĩ=kato
fly + go_up-N2/3S.PST.N.ASS-REP jaguar + person.COL=TOP
ĩ-o-wa'i-re kũh∼kũ + hõã-ñu'ũ kaa-ni
3PRO-CL:F-PL.AN-N.SBJ bite∼REDUP + slash-HORT say-SS
‘As for the jaguar people, they flew up above and wanted to attack them (lit. (they) said “let's bite and slash them”).’
[NAT*: 20151112orapi001.028-029]

Regarding non-pronominal tokens, PROMINENT *-re* marking is displayed both for a specific P-token in (37), whereas an instance with PLAIN zero-marking is shown in (5) above. A PROMINENT *-re* marked non-specific token in found in (38):

(37) *i coriyore dutani baquëbi aide'huani i dëjore. . .*
[**ĩ-ɨ** **kori-jo-re**] duhta-ni baa-kɨ-bi aide'wa-ni
3PRO-CL:M rib-CL:LONG-N.SBJ remove-SS have-CL:M-SBJ make_grow-SS
[**ĩ-ɨ** **dɨ̃hõ-re**]
3PRO-CL:M wife-N.SBJ
‘. . . (he) took out his rib and, having (it), created (lit. grew) his wife.’
[NAT*: 20150811sfryi001.004]

(38) [**Context:** The protagonist is in the Upper World and some cricket creatures find her and bring her to the *Debao*. They have something for her to drink.]
. . . ja̲ë̲hua'ibi go̲nosëbë ti̲'tese're dahuë.
hã-ɨ-wa'i-bi [**gõno + sɨhbɨ tĩ'te-sih-je-re**]
DEM.MED-CL:M-PL.AN-SBJ chicha + flask bring_along-COMP-CL:GEN-N.SBJ
daa-wɨ
bring-N3S.PST.ASS
'. . . they brought (her) some flasks of chicha they had brought along.'
[NAT*: 20151112orapi001.106-109]

Similar patterns to the above are shown for animate P-tokens. An instance of PROMINENT *-ni* marking on a specific animate P-token in (39). Instances of PLAIN *-re* marking and PROMINENT *-ni* marking on animate P-tokens are illustrated in (40):

(39) Example of PROMINENT-marked definite animate P:
jerorana saquë ja'që cani yureta̲'a̲ bë'caquëni co'ehuëoni sareña
he-ro-reba-na sah-kɨ ha'-kɨ kaa-ni
WH-CL:LOC-INTENS-GOAL go-2/3S.M.PST.N.ASS parent-CL:M say-SS
jure-tã'ã **bɨ'ka-kɨ-ni** ko'e + wɨo-ni
now-CNT.EXP parent-CL:M-N.SBJ2 look_for + begin-SS
saa-re-ña
bring-N2/3S.PST.N.ASS-REP
' "Where exactly did father go?" they said, and then they brought (the others) to start looking for (their) father.' [NAT*: 20150811sfryi001.444]

(40) a. *bacoña i̲o̲ de̲baocato – ja̲macarebacato ë̲jë̲re baoña. . .*
bah-ko-ña ĩ-o dẽbao=kato –
be-2/3S.F.PST.N.ASS-REP 3PRO-CL:F vulture(spec)=TOP –
hãmahka-reba=kato **ɨhĩ-re** baa-o-ña
then-INTENS=TOP husband-N.SBJ have-2/3S.F.PST.N.ASS-REP
'There was *the Dẽbao*. At that time (she) lived on and had a husband. . . '
[NAT*: 20151112orapi001.005-006]

b. [**Context:** The protagonist goes to catch an electric eel, his son's favourite meat, but he can't find one.]
. . . yo'jujani hua̲ñumini huarej̲e̲ caquëni da baquë
jo'-huha-ni **wañumi-ni** ware-hẽ kaa-kɨ-ni daa-a
do-FRUST-SS anaconda-N.SBJ2 child-ADD say-CL:M-N.SBJ2 bring-NEG
bah-kɨ
be-2/3S.M.PST.N.ASS
'. . . (he) couldn't do (that), so (he) brought an anaCONda to (his) son, who asked him (to bring the eel).' [NAT*: 20150811sfryi001.444]

As concerns WORD ORDER, the dominant pattern is such that P-tokens arise in *pre-verbal* position. Both of the sentences in (40) above display *pre-verbal* P-tokens. (41) displays instances of PLAIN *-re* and PROMINENT *-ni* on *post-verbal* animate P-tokens:

(41) a. *yureta'a iohua'i ga'yojë quëohuë nesiconi chojëna ti'a baquë ...*
jure-tã'ã ĩ-o-wa'i gã'jo-hɨ ↑[**kɨo-wɨ**
now-CNT.EXP 3PRO-CL:F-PL.AN play-PL.PRS.DEP crush-CL:CONT
nee-sih-ko-ni] cho-hɨ-na tĩ'ã-a
make-COMP-CL:F-N.SBJ2 laugh-PL.PRS.DEP-DS arive-NEG
bah-kɨ
be-2/3S.M.PST.N.ASS
'Now they were playing (with) the trap they had made (lit. the crusher) and laughing when (he) arrived ...' [NAT*: 20150811sfryi001.368]

b. *... tsoe ja quëojëyouna jaona huaquëña te'ire*
tsoe hã kɨo + hɨjo-ɨ-na
already DISC.PART crush + break.INTRS-S.M.PST.DEP-DS
hã-o-na wah-kɨ-ña **te'-ɨ-re**
DEM.MED-CL:F-GOAL kill-2/3S.M.PST.N.ASS-REP one-CL:M-N.SBJ
'(The trap) already crushed (him) and it broke and he killed one (M) (of them) right there.' [NAT*: 20150811sfryi001.380-381]

B.2.3 The results for L-oriented DCM

Table B.8 reports the results of the full *varbrul* model for the L-oriented sub-sample. The FACTOR of ANIMACY STATUS is eliminated from the coding of this spatial DCM pattern. The eight coded *look-ahead topical* L-token invariably display PROMINENT-marking. In addition, on the basis of the small set of L-tokens, and, most likely, the generally higher proportion of PROMINENT-marking, this model is not a strong fit for the data (i.e., Input = 0.738) — just as is disclaimed in Chapter 5, the statistical findings reported below are to be taken with a grain of salt:

N = 152 Input: 0.738			* p = < 0.01 ! p = > 0.05
	Factor weight	**% prominent**	***N***
***Contrastivity**			
Contrastive L	0.79	90%	20
Non-contrastive L	0.21	47%	132
Range	*58*		
Specificity			
Specific non-pro L	0.66	62.7%	67
Pronoun L	0.43	46.2%	65
Non-specific L	0.40	40%	20
Range	*26*		
		Non-significant Factors	
◊**Look-ahead topic (persistence)**		◊INVARIABLE	
Look-ahead topical L		**100%**	8
Local non-topical L		**50%**	**144**
! Word order			
Pre-verbal L		54.95%	111
Post-verbal L		46.34%	91
! Subcategorization			
Subcategorized L		55.68%	88
Non-subcategorized L		48.44%	64
! Look-back topic (distance)			
Look-back topical L		64.52%	31
Global non-topical L		49.59%	121

Table B.8: Varbrul results: factors conditioning PROMINENT L-marking

In light of the relevance of promotion to the determination of PLAIN- and PROMINENT-marking forms, the PROMINENT-marking status is established on a token-by-token basis. (42) demonstrates an instance of promoted L-tokens as arise with PLAIN *-re* marking, and with PROMINENT *-ni* marking accordingly:

(42) a. [**Context:** The children see fruit up in the trees so they go up to get it.]
. . . ba'ina ja̲o̲re saniti̲'a̲ni abita mëni o̲cuereña
ba'i-i-na ↑**hã-o-re** sani + tĩ'ã-ni abita
be-S.M.PST.DEP-DS DEM.MED-CL:F-N.SBJ go + find-SS abita_fruit
mɨni + õhkwe-re-ña
go_up + slurp-N2/3S.PST.N.ASS-REP
'. . . since there was (fruit), (they) climbed up there to find and to eat (it) up.' [NAT*: 20150811sfryi001.450-452]

b. *tsiaya ba'isiconi yureta̲'a̲ ja̲ tuina ti tubëje̲ yureta̲'a̲ jëatubë de'ouña*
↑[**tsia-ja ba'i-sih-ko-ni**] jure-tã'ã hã
river-CL:RIV be-COMP-CL:F-N.SBJ2 now-CNT.EXP DISC.PART
tui-i̵-na [ti tu-bi̵-hẽ] jure-tã'ã
sit_on_something-S.M.PRS.DEP-DS ANA trunk-CL:RND-ADD now-CNT.EXP
hi̵a + tu-bi̵ de'o-i̵-ña
hard + trunk-CL:RND become-2/3S.M.PST.N.ASS-REP
'(He) sat <u>where there was a river</u>. That tree trunk had now hardened too.' [NAT*: 20151023orocr001.092]

Section 5.3.4 considers an additional formal complexity, concerning the usage of *-na* with the L-tokens arising with particular predicates. On the basis, under conditions of promotion, *-na* marking is obligatory, and the corresponding invariable tokens are excluded. (43) lifts two such instances from the sample:

(43) a. *... sa'nihuë jai cua'coro tsi̲u̲sicona aide'ouña*
↑[**sa'niwi̵ hai kwa'ko-ro tsĩũ-sih-ko-na**] ai
inside big cook-CL:RECEP cover-COMP-CL:F-GOAL big
+ de'o-i̵-ña
+ become-2/3S.M.PST.N.ASS-REP
'... (they) grew <u>inside the big covered pot</u>.' [NAT*: 20150811sfryi001.147]

b. *... go'ini caquë ba̲'quëña baja'i tsiaya – ja̲o̲na ja̲ë̲hua'i* guya-yë
go'i-ni kaa-ki̵ ba-∼'-ki̵-ña
return-SS say-S.M.PRS.DEP be-RMT.PST-2/3S.M.PST.N.ASS-REP
ba-ha'i tsia-ja ↑**hã-o-na**
be-3S.M.PST.ASS river-CL:RIV DEM.MED-CL:F-GOAL
hã-i̵-wa'i guja-ji̵
DEM.MED-CL:M-PL.AN bathe-N3S.PRS.ASS
'... (he) came back and said, "There is a river. They are bathing <u>there</u>." ' [NAT*: 20151023orocr001.240-242]

The role of CONTRASTIVITY STATUS is evident on the basis of examples such as (44), where PROMINENT *-re* marking arises:

(44) [**Context:** The demon *kwẽ̃ẽwahti* is chasing the protagonist, and he goes from tree to tree to find a new place to hide. After sitting up on a branch, it breaks...]
... jëyequë neni ja'obi pë'pëni ja'ruña – ja'runi ja̲rore tuquëña
hi̵je-ki̵ nee-ni ha'o-bi pi̵∼pi̵-ni
break.INTRS-S.M.PRS.DEP make-SS leaf-INST pile∼REDUP-SS
ha'ru-i̵-ña ha'ru-ni **hã-ro-re**
sit_down-2/3S.M.PST.N.ASS-REP sit_down-SS DEM.MED-CL:LOC-N.SBJ
tuh-ki̵-ña
sit_atop-2/3S.M.PST.N.ASS-REP
'... (it) broke and so (he) took leaves and covered (himself) and sat down. (He) sat down and was sitting <u>at THAT spot</u>.' [NAT*: 20151112oespa001.097]

As regards the WORD ORDER, this does not appear to factor into L-marking. (45) lifts two typical instances of PROMINENT *-re* on both *pre-verbal* and *post-verbal* L-tokens:

(45) a. *iñore ba'ico cani sanisaëña . . .*
ihño-re ba'i-ko kaa-ni sani + saa-ɨ-ña
here-N.SBJ live-3S.F.PRS.ASS say-SS go + bring-2/3S.M.PST.N.ASS-REP
' "(She) lives HERE" (he) said and (he) brought (her there). . . '
[NAT*: 20150811sfryi001.450]

b. [**Context:** Baina helps his pregnant wife retrieve arari fruit from the trees]
. . . ñacona arari darëse'e baquëña ma'are
ñaa-ko-na arari da'rɨ-si-je
see-S.F.PRS.DEP-DS arari_fruit fall_to_ground-COMP-CL:GEN
bah-kɨ-ña **ma'ã-re**
be-2/3S.M.PST.N.ASS-REP path-N.SBJ
'. . . (she) saw the arari fruits that had fallen, (they) were on the path.'
[NAT*: 20150811sfryi001.073]

As regards SPECIFICITY STATUS, PROMINENT-marking freely arises on both *specific* and *non-specific* L-tokens. Instances PLAIN zero-marking and PROMINENT *-re* marking are shown in (46):

(46) a. [**Context:** After her husband passes away, the protagonist searches for his spirit in the forest.]
. . . caona achani ja beohue'ña ba'isiquëta'a
kaa-o-na ahcha-ni hã **beo-wẽ'ña**
say-S.M.PST.DEP-DS listen-SS DISC.PART NEG.EXIS-LOC.DERIV
ba'ĩ-sih-kɨ-tã'ã
be-COMP-CL:M-CNT.EXP
'. . . after (she) said (that), (her son) listened but (his father) was nowhere.' [NAT*: 20151112orapi001.022]

b. [**Context:** the shaman protagonist knows to stay still inside as the demon *cuehuati* eats another man at the house.]
. . . bëani jarëre huecaquëña caina io huë'ena ti'tani anijeoña ire
bɨa-ni **hã-rɨ-re** wee
stay-SS hammock-CL:MAIZE-N.SBJ be_lying
+ kãh-kɨ-ña kãi-ɨ-na ĩ-o
+ sleep-N2/3S.M.PST.N.ASS-REP sleep-S.M.PRS.DEP-DS 3PRO-CL:F
wɨ'e-na tĩ'tã-ni ãni + hẽõ-o-ña
house-GOAL arrive-SS eat + leave_behind-2/3S.F.PST.N.ASS-REP
ĩ-ɨ-re
3PRO-CL:M-N.SBJ
'. . . (he) stayed there, lying and sleeping in a hammock when she arrived at the house to eat him up.' [NAT*: 20151112oespa001.020]

B.2.4 Miscellaneous examples for discussion

This section lays out a selection of examples to supplement the discussion in Chapter 5, which do not concern the S-marking, the P-marking, or the L-marking sub-samples

analyzed above. Below instances of coded R-tokens and illustrative spatiotemporal tokens are provided. Two tables are found at the close of this Appendix: Table B.9 reports sentential word order patterns in the sample, and Table B.10 reports the distribution of nominal discourse markers on the tokens in the dataset.

Illustrative R-tokens for discussion

There are 70 R-tokens coded in the Siona narrative dataset, and this argumental DCM pattern displays a global PROMINENT marking-rate of 24.29%. This may be straightforwardly ascertained given that there are no inanimate R-tokens attested in the sample. Impressionistically, patterns of PROMINENT *-ni* marking align with the description of R-oriented DCM patterns in Section 3.2.3 of the *preliminary description*. The lifted examples in (47) demonstrate an instance of PLAIN *-re* marking and PROMINENT *-ni* marking respectively:

(47) a. *... cua'coni ñatasi' mo'se neahuẽ'ñareba ịọre ọcuaoña neco*
kwa'ko-ni ñahta-si + mo'se nea-wẽ'ña-reba
cook-SS dawn-COMP + day black-LOC.DERIV-INTENS
ĩ-o-re õhkwa-o-ña nehko
DEM.PROX-CL:F-N.SBJ give_to_drink-2/3S.F.PST.N.ASS-REP neco_drink
'... after cooking (it), the next day (she) gave her *neco*[17] to drink, right at dawn.' [NAT: 20110328slicr001.023]

b. [Context: A family is arguing over who gets to eat the eggs found on the path.]
si'a tsiago'o ịsijẽ'ẽ yë'ni caoña
si'a tsia + go'o ĩhsi-hĩ'ĩ **jɨ'ɨ-ni** kaa-o-ña
all egg + bundle give-IMPER 1SG-N.SBJ2 say-2/3S.F.PST.N.ASS-REP
' "Give the whole bundle of eggs to ME !", (she) said (so they say).'
[NAT*: 20150811sfryi001.121]

Impressionistically, on the basis of examples such as (47b), the prevalence for PROMINENT-marking on *contrastive* R-tokens aligns with the findings laid out for each tested sub-sample spelled out above.

Illustrative spatiotemporal nominal tokens for discussion

Outside the primary L-marking sub-sample, the discussion in Chapter 5 largely sidesteps an analysis of spatiotemporal tokens in the sample. Only a brief commentary regarding the distribution of spatiotemporal noun is given in Section 5.3.1 — the lifted examples below supplement that discussion, as presented in the same order as in Chapter 5.

Firstly, spatial oblique tokens are strictly found with their obligatory case-markers. This is shown with the SOURCE token exemplified in (48a), and with the PATH token shown in (48b):

(48) a. *... tsiaya sa'nihuëbi dëmëña*
tsia-ja **+ sa'niwɨ-bi** dimɨ-ɨ-ña
river-CL:RIV + inside-SRC emerge_from_water-2/3S.M.PST.N.ASS-REP
'(He) emerged from inside the river.' [NAT*: 20151023orocr001.040]

[17] *Neco* refers to an anti-anemic agent brewed from locally cultivated plants.

b. *. . . sai airo saquëña huao huë'equë'roja'a*
sai-ɨ ai-ro sah-ki-ña **wa-o**
go-S.M.PST.DEP big-CL:LOC go-2/3S.M.PST.N.ASS-REP parent_in_law-CL:F
+ wɨ'e-kɨ'ro-hã'ã
+ house-LOC.DERIV-PATH
'. . . going to the jungle (to hunt), (he) went past his mother-in-law's house.' [NAT*: 20151001oolpi001.009]

The next mentioned category concerns Point-of-Contact tokens (N=28). These are described as *pseudo*-spatial argument in Section 3.2.4 in the *preliminary description* in this dissertation. The bulk of these tokens bear *-na* marking, which is generally the obligatory marking pattern — as exemplified in (49). Two outliers tokens, which deviate from this pattern are lifted in (50) and (51): The former arises with *-hã'ã* (LIMIT) marking, which has a compatible spatial meaning; and the latter example is an instance of exceptional zero-marking, as noted with the spatial pronoun, *to*:

(49) *cacani ire sa'nahuë ahuëna to'tehueaña ba'ë ire yi'yebi*
kahka-ni ĩ-ɨ-re + sa'nawɨ **a-wɨ-na**
enter-SS DEM.PROX-CL:M-N.SBJ + inside heart-CL:CONT-GOAL
to'te + wẽã-a ba-ɨ'ɨ ĩ-ɨ-re ji'je-bi
sting + kill-NEG be-N3S.PST.ASS DEM.PROX-CL:M-N.SBJ wasp(spec)-SBJ
'(He) entered inside him, and the wasp stung him on the heart.'
[NAT*: 20150811sfryi001.541]

(50) *anijeoni yureta'a tsiubëja'a titotëani . . .*
ãni + hẽõ-ni jure-tã'ã **tsĩũ-bɨ-hã'ã** tihto + tɨa-ni
eat + leave_behind-SS now-CNT.EXP head-CL:RND-PATH cut + set_aside-SS
'(She) ate (him) and cut (him) up to (his) head, setting (it) aside . . .'
[NAT*: 20151112oespa001.035]

(51) *guebe irecato cuëa baquë tsiubëna ñacuana to të'caëña*
gwẽbe ĩ-ɨ-re=kato kwɨɨ-a bah-kɨ
be_drunk 3PRO-CL:M-N.SBJ=TOP swim-NEG be-2/3S.M.PST.N.ASS
tsĩũ-bɨ-na ñahkwa-na **to** tɨ'ka-ɨ-ña
head-CL:RND-GOAL eye-GOAL ANA.LOC hit-2/3S.M.PST.N.ASS-REP
'(He) swam up to him like a madman and hit him [on the head, there, by the eye].' [NAT*: 20150811sfryi001.568]

Section 5.3.1 briefly considers the case-marking properties found with temporal noun tokens in the sample. The dominant pattern is one where temporal nominals are zero-marked, and this discussion also discussed temporal Start-Point and End-Point tokens. A few tokens are found where the temporal token is marked with *-na*, as in the lifted example in (52). At present it is not clear what the function of *-na* is in such instances:

(52) *huaredohuë caijë'ë esa caoña gosa ga'yoyë ñocua oyasi'i cacona neatona ioje bëani uiño yo'co caoña*
ware-dowɨ kãĩ-hĩ'ĩ ehsa kaa-o-ña gõhsa
child-PL.AN sleep-IMPER fast say-2/3S.F.PST.N.ASS-REP ungurahua
gã'jo-jɨ ñohkwa oja-si'-i kaa-ko-na
play-N3S.PRS.ASS chambira_palm press-FUT-N3S say-S.F.PRS.DEP-DS

nea-to-na ĩ-o-hẽ bɨa-ni ũĩ-no
be_dark-CL:LOC-GOAL 3PRO-CL:F-ADD stay-SS be_lying-S.F.PST.DEP
jo'-ko kaa-o-ña
do-S.F.PRS.DEP say-2/3S.F.PST.N.ASS-REP
‘ “Go to sleep, children! Quickly!” (she) said. They are playing with the ungurahua leaves, (she) wanted to press chambira very early the next day (∼at dawn), she said (this) as she remained lying there.’
[NAT*: 20151001oolpi001.053-054]

The next grammatical relation concerns in Section 5.3.1 concerns GOAL (G) tokens — this spatial grammatical relation displays a bona DCM pattern, described in Section 3.3.2. The pattern found on canonical, non-promoted, G-tokens is exemplified by the derived spatial nouns in (53): PLAIN zero-marking is shown in (53a), and PROMINENT *-na* marking in (53b):

(53) a. [**Context:** The children of Baina are ready to drink yage and ascend to be with their father. One of the sisters does not want to drink any.]
me basiconi saija'cua'ia ? ja'quëquë'ro ti'aja'cua'ia'ë caëña
me baa-sih-ko-ni sai-hã-ko-wa'i-a'
how have-COMP-CL:F-N.SBJ2 go-PRP-CL:F-PL.AN-COP.N.ASS
ha'-kɨ-kɨ'ro tĩ'a-hã'-ko-wa'i-a'-ɨ
parent-CL:M-DERIV.LOC arrive-PRP-CL:F-PL.AN-COP-N3S
kaa-ɨ-ña
say-2/3S.M.PST.N.ASS-REP
‘ “How will we go because of HER (lit. the one who was (there))? We will arrive where father is”, (he) said.’ [NAT*: 20150811sfryi002.279]

b. [**Context:** The great tree has come down and now the creature-people can bathe in the newly formed river. God tells the hummingbird to bring his family there.]
. . . sanëcajai ba'quëña mimi − yëiñëquë'rona sanëcajaquëña
saa + nɨhka-hai-ɨ ba-∼'-kɨ-ña
bring + stand-VBLZ-CL:M be-RMT.PST-2/3S.M.PST.N.ASS-REP
mimi **jɨi-ñɨ-kɨ'ro-na** saa
hummingbird cotton-CL:TREE-DERIV.LOC-GOAL bring
+ nɨhka-hai-kɨ-ña
+ stand-VBLZ-2/3S.M.PRS.N.ASS-REP
‘. . . the hummingbird brought (his family). (He) brought (them) to the kapok tree and stands (there).’
[NAT*: 20151023orocr001.236-237]

Word order in the narrative sample

Table B.9 indicates the order of main clausal arguments relative to the main verbal word — i.e., the verbal element in the main clause which bears inflectional morphology. Following conventions established elsewhere in this dissertation: S refers to the clausal SUBJECT, P to the DIRECT OBJECT, R to the INDIRECT OBJECT, whereas X refers to any remaining grammatical relations recorded in the narrative sample, including spatial arguments — consult Table 5.5 in Chapter 5 for an enumeration of these categories. For S, P, and R; the large capital refers to the proportion of overt

tokens (e.g., SV — refers to an overt, pre-verbal S-argument), whereas the small capital refers to covert arguments of that category (e.g., sV indicates that there is no overt S-argument in this clause):

	N	%	%SV	%VS	%sV	%PV	%VP	%pV	%RV	%VR	%rV	%XV	%VX
Intransitive	746	(41.51)	35.79	11.53	52.68							73.29	26.71
Transitive	999	(55.59)	17.82	8.51	73.67	54.15	14.61	31.23				77.14	22.86
Ditransitive	52	(2.89)	9.62	11.54	78.85	40.38	3.85	55.77	44.23	42.31	13.46		
Total	**1797**		**25.04**	**9.85**	**65.11**	**53.47**	**13.89**	**32.45**				**74.46**	**25.54**

Table B.9: Descriptive overview: Word order and argument dropping patterns in corpus sample (main clauses)

Nominal discourse markers in narrative sample

Table B.10 reports the distribution of the discourse markers as found in the narrative sample, which include the following four nominal morphemes, discussed in detail in Section 6.4.2: the exclusive marker *-se'e* (only), the additive marker *-hẽ* (also), the counter-expectative marker *-tã'ã* (even), and the contrastive topic clitic, *=kato* (as for...). As indicated in Section 5.3.2.1, *=kato* marking is a criterion for excluding tokens in the S-marking sub-sample — (54) lifts a few tokens with *=kato*:

CASE-MARKER/ROLE	*-se'e* (EXCL)	*-hẽ* (ADD)	*-tã'ã* (CNT.EXP)	*=kato* (TOP)
ZERO-MARKING	43	44	48	76
S	23 (53.49%)	32 (72.73%)	5 (10.42%)	52 (68.42%)
P	17 (39.53%)	8 (18.18%)	0	3 (3.95%)
L	2 (4.65%)	2 (4.55%)	0	5 (6.58%)
G	1 (2.33%)	2 (4.55%)	0	2 (2.63%)
TIME	0	0	43 (89.58%)	14 (18.42%)
re-MARKING (P)	5	2	3	5
bi-MARKING (S)	3	0	0	0
TOTAL ($N = 229$)	**51**	**46**	**51**	**81**

Table B.10: Distribution of discourse markers in Siona narrative dataset

Table B.10 demonstrates that there are a total of 229 nominal tokens bearing a discourse marker — i.e., 11.96% of the total 1914 tokens in the sample. These nominal markers are less frequent than case-marking, which arises on 44.1% of nominal tokens in the Siona narrative dataset, as shown in Table 5.5.

(54) a. *i̲o neato ñami ti̲'a̲ni a̲isicobi a̲ni chaoni i̲recato a̲coña ...*
ĩ-o nea-to ñahmi tĩ'ã-ni ãĩ-sih-ko-bi ãni + chao-ni
3PRO-CL:F dark-COND night arrive-SS eat-COMP-CL:F-SBJ eat + finish-SS
ĩ-i-re=kato ãh-ko-ña
3PRO-CL:M-N.SBJ=TOP eat-2/3S.F.PST.N.ASS-REP
'...she came later in the night, the one who was eating (him) came back to finish eat HIM (i.e., not the shaman).'
[NAT*: 20151112oespa001.025]

b. *ja̲ susi jecabi sëyëja̲'coa'ë jecarebacato ba̲ñë yë'ë sëyëye caoña*
hã suhsi + hehka-bi sɨji-hã'-ko-a'-ɨ
DEM.MED carachama_palm + firewood-INST dry_out-PRP-CL:F-COP-N3S
hehka-reba=kato bãã-ñɨ jɨ'ɨ sɨji-je
firewood-INTENS=TOP NEG.AUX-N3S.PRS.ASS 1SG dry_out-CL:GEN
kaa-o-ña
say-2/3S.F.PST.N.ASS-REP
'..."(I) must dry (it) out with the carachama firewood. I cannot dry out the WOOD (we have now)", (she) said.' [NAT*: 20150811sfryi001.623]

c. [**Context:** The demon *cu̲e̲huati* left the half-eaten corpse of a man when she terrorized the village. There is another man, the shaman protagonist, who is in the village hiding.]
i̲o̲ neato ñami ti̲'a̲ni a̲i̲sicobi a̲nichaoni i̲recato a̲conña
ĩ-o neato ñami tĩ'ã-ni ãĩ-sih-ko-bi ãni + chao-ni
PRO3-CL:F later night arrive-SS eat-COMP-CL:F-SBJ eat + finish-SS
ĩ-i̵-re=kato ãh-ko-ña
3PRO-CL:M-N.SBJ=TOP eat-2/3S.F.PST.N.ASS-REP
'She arrived later in the night, and she (lit. the one who had eaten) finished eating him.' [NAT*: 20151001oolpi001.092-093]

d. [**Context:** The water from the flood finally subsides and the surviving man, the protagonist, begins to explore the world.]
ñadojai ñaëña jero cani ñatoje̲ − tocato ja̲re so̲quëñëa̲ beoëña
ñaa + dohai-i̵ ñaa-i̵-ña he-ro
see + wander-S.M.PST.DEP see-2/3S.M.PST.N.ASS-REP INTER-CL:LOC
kaa-ni ñaa-to-hẽ **to=kato** hãre sõhki̵-ñi̵-ã
say-SS see-COND-ADD ANA.LOC=TOP like_that wood-CL:TREE-PL.INAN
beo-i̵-ña
NEG.AUX-2/3S.M.PST.N.ASS-REP
'He took a look around, seeing where (he) was speaking of − there were no trees there.' [NAT*: 20151023orocr001.108]

Bibliography

Aannestad, Aidan A. 2021. A typology of morphological argument focus marking. Doctoral Dissertation, University of North Dakota.

Adelaar, Willem F.H., and Pieter C. Muysken. 2004. *The languages of the Andes*. Cambridge, New York: Cambridge University Press.

Aihenvald, Alexandra. 2018. *The Oxord handbook of evidentiality*. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. 1999. Areal diffusion and language contact in the Içana-Vaupés basin, north-west Amazonia. In *The amazonian languages*, ed. Robert M.W. Dixon and Alexandra Y. Aikhenvald, 384–416. Cambridge University Press.

Aikhenvald, Alexandra Y. 2001. Areal diffusion, genetic inheritance, and problems of subgrouping: A north Arawak case study. In *Areal diffusion and genetic inheritance: Problems in comparative linguistics*, ed. Alexandra Y. Aikhenvald and Robert M.W. Dixon, 167–194. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. 2003. *A grammar of Tariana, from northwest Amazonia*. Cambridge, New York: Cambridge University Press.

Aikhenvald, Alexandra Y. 2007. Semantics and pragmatics of grammatical relations in the Vaupés linguistic area. In *Grammars in contact: A cross-linguistic typology*, ed. Alexandra Y. Aikhenvald and R.M.W. Dixon, 237–266. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. 2015. Differential case in Yalaku. *Oceanic Linguistics* 54:240–269.

Aissen, Judith. 1999. Markedness and subject choice in optimality theory. *Natural language & linguistic theory* 17:673–711.

Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural language & linguistic theory* 21:435–483.

Bibliography

Andrews, Avery. 2007. The major functions of the NP. In *Language typology and syntactic description*, ed. T. Shopen, 132–223. Cambridge: Cambridge University Press.

Arias, Leonardo, Nicholas Q. Emlen, Sietze Norder, Nora Julmi, Magdalena Lemus Serrano, Thiago Chacon, Jurriaan Wiegertjes, Austin Howard, Matheus C.B.C. Azevedo, Allison Caine, Saskia Dunn, Mark Stoneking, and Rik van Gijn. 2022. Interpreting mismatches between linguistic and genetic patterns among speakers of Tanimuka (Eastern Tukanoan) and Yukuna (Arawakan). *Interface focus* 13.

Ariel, Mira. 2014. *Accessing noun-phrase antecedents*. London, New York: Routledge.

Aristar, Anthony R. 1997. Marking and hierarchy types and the grammaticalization of case-markers. *Studies in language* 21:313–368.

Arkadiev, Peter. 2008. Two-term case systems: Typology and theoretical implications.

Arnold, Jennifer E. 2010. How speakers refer: The role of accessibility. *Language and linguistics compass* 4:187–203.

Arregi, Karlos. 2016. Focus projection theories. In *The oxford handbook of information structure*, ed. Caroline Féry and Shinichiro Ishihara, 185–202. Oxford: Oxford University Press.

Asher, Ronald, and T. Kumari. 1997. *Mayalayam*. London: Routledge.

Assmann, Muriel, Daniel Büring, Izabela Jordanska, and Max Prüller. 2023. Towards a theory of morphosyntactic focus marking. *Natural language & linguistic theory* 41:1349–1396.

Baker, Mark C. 2015. *Case: Its principles and its parameters*. Cambridge: Cambridge University Press.

Baker, Mark C, and Ruth Kramer. 2014. Rethinking Amharic prepositions as case markers inserted at PF. *Lingua* 145:141–172.

Baker, Mark C., and Nadya Vinokurova. 2010. Two modalities of case assignment: Case in Sakha. *Natural language & linguistic theory* 28:593–642.

Barnes, Janet. 1999. Tucano. In *The amazonian languages*, ed. Robert M.W. Dixon and Alexandra Y. Aikhenvald, 207–226. Cambridge, New York: Cambridge University Press.

Barnes, Janet. 2006. Tucanoan languages. In *Encyclopedia of language and linguistics*, volume 2, 130–142. Oxford, Boston: Elsevier.

Becker, Laura, and Gertrud Schneider-Blum. 2020. Morphological marking of contrast in Tima. *Glossa* 5:1–35.

Biber, Douglas, Susan Conrad, and Randi Reppen. 1998. *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.

Bickel, Balthasar. 2007. Typology in the 21st century: Major current developments. *Linguistic typology* 11(1):239–251.

Bibliography

Bickel, Balthasar, and Johanna Nichols. 2009. Case marking and alignment. In *The oxford handbook of case*, ed. Andrej Malchukov and Andrew Spencer. Oxford: Oxford University Press.

Bielec, Dana. 2015. *Basic Polish: a grammar and workbook*, volume 2. London: Routledge.

Birchall, Joshua. 2014. Argument marking patterns in South American languages. Doctoral Dissertation, Radboud University, Nijmegen.

Blake, Barry. 2001. *Case (2nd edition)*. Cambridge, New York: Cambridge University Press.

Bochnak, M Ryan, and Lisa Matthewson. 2015. *Methodologies in semantic fieldwork*. New York: Oxford University Press.

Bolaños, Katherine. 2016. A grammar of Kakua. Doctoral Dissertation, Leiden University.

Borgman, Denis M. 1990. Sanuma. In *Handbook of Amazonian languages*, ed. Derbyshire D.C. and G.K. Pullum, volume 2, 215–248. Berlin: Mouton de Gruyter.

Bossong, Georg. 1984. Animacy and markedness in universal grammar. *Glossologia* 2-3:7–20.

Bossong, Georg. 1985. *Empirische Universalienforschung: differentielle Objektmarkierung in den neuiranischen Sprachen*. Number 14 in Ars linguistica. Tübingen: Narr.

Bowe, Heather. 1990. *Categories, constituents, and constituent order in Pitjantjatjara, an Aboriginal language of Australia*. London, New York: Routledge.

Brugè, Laura, and Gerhard Brugger. 1996. On the accusative a in Spanish. *Probus* 8:1–52.

Bruil, Martine. 2012. *Documentation of ecuadorian siona*. Endangered Languages Archive. URL http://hdl.handle.net/2196/00-0000-0000-000D-EA53-3.

Bruil, Martine. 2014. Clause-typing and evidentiality in Ecuadorian Siona. Doctoral Dissertation, Leiden University.

Bruil, Martine. 2015a. Possible Quechua influence on Ecuadorian Siona. *Paper presented at REELA Leiden University, September 7th, 2015* .

Bruil, Martine. 2015b. When evidentials are not evidentials: The case of the Ecuadorian Siona reportative. *Linguistic typology* 19:385–423.

Bruil, Martine. 2018. The development of the portmanteau verbal morphology in Ecuadorian Siona: A story of the formal merger of linguistic categories. *Journal of Historical Linguistics* 8:128–167.

Bruil, Martine. 2019. The rise of the nominalizations: The case of the grammaticalization of clause types in Ecuadorian Siona. In *Nominalization in languages of the americas*, ed. R Zariquiey, M Shibatani, and D W Fleck, 391–417. Amsterdam: John Benjamins.

Bibliography

Bruil, Martine, and Justin Case. 2023. The subject-like dependent verb in Ecuadorian Siona. In *Congreso internacional IX, Amazónicas*.

Bruil, Martine, and Jesse Stewart. 2022. Phonetics and phonology of nasality in Ecuadorian Siona. *Phonological Data and Analysis* 4:1–34.

Bárány, András. 2018. DOM and dative case. *Glossa* 3:1–40.

Büring, Daniel. 2003. On D-trees, beans, and B-accents. *Linguistics & philosophy* 26:511–545.

Büring, Daniel. 2009. Towards a typology of focus realization. In *Information structure*, ed. Malte Zimmermann and Caroline Féry, 177–205. Oxford: Oxford University Press.

Büring, Daniel. 2016. (contrastive) topic. In *The oxford handbook of information structure*, ed. Caroline Féry and Shinichiro Ishihara, 64–85. Oxford: Oxford University Press.

Cann, Ronnie, and Merilin Miljan. 2012. Differential case-marking: Syntactic descriptions and pragmatic explanations. *The linguistic review* 29:585–605.

Carnie, Andres. 2006. Some remarks on markedness hierarchies: A reply to Aissen 1999 and 2003. In *Coyote papers: Working papers in linguistics*, ed. Dan Siddiqi and B.V. Tucker. Unviersty of Arizona.

de Carvalho, Francismar Alex Lopes. 2022. *Missionizing on the edge: Religion and power in the Jesuit missions of Spanish Amazonia*, volume 37. Leiden, Netherlands: Brill.

Case, Justin, and Paloma Jeretič. 2021. The interaction of number and information structure in Siona DOM. In *Proceedings of Semantics of Under-studied Languages of the Americas 11*, ed. Seung Suk Lee and Yixiao Song, 43–52. UNAM, Mexico, virtual.

Chacon, Thiago. 2012. The phonology and morphology of Kubeo: The documentation, theory, and description of an Amazonian language. Doctoral Dissertation, University of Hawai'i, Manoa.

Chacon, Thiago. 2013. On Proto-Languages and Archaeological Cultures: pre-history and material culture in the Tukanoan Family. *Revista Brasileira de Linguística Antropológica* 5:217–245.

Chacon, Thiago. 2014. A revised proposal of proto-tukanoan consonants and tukanoan family classification. *International journal of American linguistics* 80:275–322.

Chacon, Thiago. 2016. The reconstruction of laryngealization in proto-tukanoan. In *The phonetics and phonology of laryngeal features in native american languages*, 258–284. Brill.

Chacon, Thiago Costa, and Carol Genetti. 2019. Grammatical relations in Kubeo. In *Argument selectors: A new perspective on grammatical relations*, 1–34. Amsterdam: John Benjamins.

Bibliography

Chacon, Thiago Costa, and Johann-Mattis List. 2015. Improved computational models of sound change shed light on the history of the Tukanoan languages. *Journal of Language Relationship* 13:177–204.

Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In *Subject and topic*, ed. Charles N. Li, 25–56. New York: Academic Press Inc.

Chafe, Wallace L. 1987. Cognitive constraints on information flow. In *Coherence and grounding in discourse*, ed. Russell Tomlin, 21–51. Amsterdam: John Benjamins.

Chafe, Wallace L. 1994. *Discourse, consciousness, and time*. Chicago: Chicago University Press.

Chappell, Hilary, and Jean-Christophe Verstraete. 2019. Optional and alternating case marking: Typology and diachrony. *Language & linguistic compass* 13:1–42.

Chiriacescu, Sofiana, and Klaus von Heusinger. 2010. Discourse prominence and pê-marking in Romanian. *International review of pragmatics* 2:298–332.

Chomsky, Noam. 1961. Some methodological remarks on generative grammar. *Word* 17:219–239.

Chosmky, Noam. 2004[1970]. Remarks on nominalization. In *Morphology: Primes, phenomena and processes*, ed. Francis Katamba, volume vol. 2, 130–168. New York: Routledge.

Cipolletti, María Susana. 1999. Dos escritos inéditos del jesuíta Pablo Maroni sobre el noroeste amazónica (indígenas encabellados, tucano, 1739-40). *Jahrbuch für Geschichte Lateinamerikas* 36:151–172.

Coghill, Eleanor. 2014. Differential object marking in Neo-Aramaic. *Linguistics* 52:335–364.

Cole, Peter. 1982. *Imbabura Quechua*, volume 5 of *Lingua descriptive studies*. Amsterdam: North-Holland Publishing Company.

Comrie, Bernard. 1989. *Language universals and linguistic typology: Syntax and morphology*. Oxford: Blackwell.

Cook, Dorothy. 1990. *Textos en koreguaje*. Bogotá: Editorial Alberto Lleras Camargo.

Cook, Dorothy, and Linda L. Criswell. 1993. *El idioma Koreguaje (Tucano occidental)*. Bogotá: Summer Institute of Linguistics.

Cook, Dorothy, and Stephen Levinsohn. 1985. Coreguaje: The domain of focus markers. In *From Phonology to Discourse: Studies in six Colombian languages*, ed. Ruth M. Brend, volume 9 of *Amerindian Series*, 91–116. Dallas: Summer Institute of Linguistics.

Corbett, Greville G. 1991. *Gender*. Cambridge, New York: Cambridge University Press.

Creissels, Denis. 2009. Spatial cases. In *The oxford handbook of case*, ed. Andrej Malchukov and Andrew Spencer, 609–625. Oxford: Oxford University Press.

Bibliography

Croft, William. 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Croft, William. 2003. *Typology and universals*, volume 2. Cambridge UK: Cambridge University Press.

Croft, William. 2009. Methods for finding language universals in syntax. In *Universals of language today*, ed. Sergio Scalise, Elisabetta Magni, and Antonietta Bisetto. Berlin: Springer.

Curnow, Timothy J. 1997. *A grammar of Awa Pit (Cuaiquer): An indigenous language o f south-western Colombia*. Canberra: The Australian National University.

Dabir-Moghaddam, Mohammad. 1992. On the (in)dependence of syntax and pragmatics: Evidence from the postposition -rā in Persian. In *Cooperating with written texts: The pragmatics and comprehension of written texts*, ed. Dieter Stein, 549–573. Berlin: Mouton de Gruyter.

Dalrymple, Mary, and Irina Nikolaeva. 2011. *Objects and information structure*. Cambridge, New York: Cambridge University Press.

Danon, Gabi. 2006. Caseless nominals and the projection of DP. *Natural language & linguistic theory* 24:977–1008.

Danon, Gabi. 2010. The definiteness feature at the syntax-semantics interface. In *Features: Perspectives on a key notion in linguistics*, ed. Anna Kibort and Greville G. Corbett, 143–165. Oxford, New York: Oxford University Press.

Danon, Gbi. 2011. Agreement and DP-internal feature distribution. *Syntax* 14:297–317.

Davidson, Kathryn. 2020. Is "experimental" a gradable predicate? In *Proceedings of nels 50*, ed. Mariam Astaryam, Yixiao Song, and Ayana Whitmal, 125–144. University of Massechusetts, Amherst.

Dayal, Veneeta. 2011. Hindi psuedo-incorporation. *Natural language & linguistic theory* 29:123–167.

Dench, Alan, and Nicholas Evans. 1988. Multiple case-marking in Australian languages. *Australian Journal of Linguistics* 8:1–47.

Diesing, Molly. 1992. *Indefinites*. Cambridge US: MIT Press.

Dik, Simon C. 1997. *The theory of functional grammar. Part 1: The structure of the clause*. Berlin, New York: Mouton de Gruyter.

Dik, Simon C., Maria E. Hoffman, Jan R. de Jong, Sie Ing Djiang, Harry Stroomer, and Lourens de Vries. 1981. On the typology of focus phenomena. In *Perspectives on functional grammar*, ed. Teun Hoekstra, Harry van der Hulst, and Michael Moortgat, 41–74. Berlin, Boston: Foris.

Dimmendaal, Gerrit J. 2010. Differential object marking in Nilo-Saharan. *Journal of African languages and linguistics* 31:13–46.

Bibliography

Dixon, R.M.W. 1977. *A grammar of Yidin*. Cambridge: Cambridge University Press.

Dixon, R.M.W. 2002. *Australian languages*. Cambridge: Cambridge University Press.

Dixon, R.M.W., and Alexandra Y. Aikhenvald. 1999. Introduction. In *The Amazonian languages*. Cambridge University Press.

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

Driemel, Imke. 2023. *Pseudo-noun incorporation and differential object marking*, volume 82. Oxford: Oxford University Press.

Dryer, Matthew S. 2013. Position of case affixes (v2020.3). In *The world atlas of language structures online*, ed. Matthew S. Dryer and Martin Haspelmath. Zenodo. URL https://doi.org/10.5281/zenodo.7385533.

Duarte, Fábio Bonfim, and Braulio Brandão de Oliveira Lopes. 2020. Differential object marking in Tukano. *Revista Letras* 7:18–44.

Dunn, John, and Shamil Khairov. 2009. *Modern Russian grammar: a practical guide*. New York: Routledge.

Dunn, Vivien, and Felicity Meakins. 2023. Ergativity. In *The Oxford Guide to Australian Languages*, ed. Claire Bowern, 217–225. Oxford: Oxford University Press.

Durant, Alessandro. 1979. Object clitic pronoun in Bantu and the topicality hierarchy. *Studies in African Linguistics* 10:31–45.

D'Alessandro, Roberta. 2017. When you have too many features: Auxiliaries, agreement and clitics in Italian varieties. *Glossa* 2:1–36.

Echeverri, Juan Alvaro. 1997. The people of the center of the world: A study in culture, history, and orality in the Colombian Amazon. Doctoral Dissertation, New York School for Social Research.

Epps, Patience. 2007. The Vaupés melting pot: Tukanoan influence on Hup. In *Grammars in contact: A cross-linguistic typology*, ed. Alexandra Y. Aikhenvald and Robert M.W. Dixon, 267–289. Oxford University Press.

Epps, Patience. 2008. *A grammar of Hup*. Berlin, New York: Mouton de Gruyter.

Epps, Patience, and Lev Michael. 2017. The areal linguistics of Amazonia. In *The cambridge handbook of areal linguistics*, ed. RaymondEditor Hickey, 934–963. Cambridge University Press.

Eraso, Natalia. 2015. Gramática tanimuka, una lengua de la Amazonía colombiana. Doctoral Dissertation, Université Lumière Lyon 2.

Evans, Nicholas D. 1985. *A grammar of Kayardild: With historical-comparative notes on Tangkic*, volume 15. Berlin; New York: Mouton de Gruyter.

Farmer, Stephanie Jo. 2015. Establishing reference in máíhīkì. Doctoral Dissertation, UC Berkeley.

Bibliography

Fauconnier, Stefanie. 2011. Differential agent marking and animacy. *Lingua* 121:533–547.

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

Fawcett, Alexia Zara. 2023. Wao Terero. In *Amazonian languages: Language isolates ii: Kanoé to yurakaré*, ed. Patience Epps and Lev Michael, volume 2, 1191–1242. Berlin: De Gruyter: Mouton.

Fischer, Rafael, and Kees Hengeveld. 2023. A'ingae (Cofán). In *Language isolates: Aikanã to Kandozi-Shapra*, ed. Patience Epps and Lev Michael, volume 1, 65–124. Berlin: De Gruyter.

Fábregas, Antonio. 2013. Differential object marking in Spanish: State of the art. *Borealis* 2:1–80.

Gerner, Matthias. 2008. Ambiguity-driven differential object marking in Yongren Lolo. *Lingua* 118:296–331.

Gerner, Matthias. 2016. Differential subject marking in Azhee. *Folia Linguistica* 50:137–173.

Gernsbacher, Morton Ann. 1989. Mechanisms that improve referential access. *Cognition* 32:99–156.

van Gijn, Rik. 2014. The Andean foothills and adjacent Amazonian fringe. In *The native languages of South America: Origins, development, typology*, ed. Loretta O'Connor and Pieter Muysken, 102–125. Cambridge: Cambridge University Press.

van Gijn, Rik, Justin Case, Martine Bruil, Simon A. Claassen, Karolina Grzech, and Nora Julmi. 2023. Lexically driven patterns of contact in alignment systems of languages of the northern Upper Amazon. *Open linguistics* 9:1–40.

van Gijn, Rik, Harald Hammarström, Simon van de Kerke, Olga Krasnoukhova, and Pieter Muysken. 2017. Linguistic areas, linguistic convergence and river systems in South America. In *The cambridge handbook of areal linguistics*, ed. Raymond Hickey, 964–996. Cambridge: Cambridge University Press.

van Gijn, Rik, and Pieter Muysken. 2020. Highland–lowland relations: A linguistic view. In *Rethinking the andes–amazonia divide: A cross-disciplinary exploration*, ed. A.J. Pearce, D.G. Beresford-Jones, and P. Heggarty, 178–210. London: University College Press.

Givón, Talmy. 1976. Topic, pronoun, and grammatical agreement. In *Subject and topic*, ed. Charles N. Li, 149–188. New York: Academic Press.

Givón, Talmy. 1983. Topic continuity in spoken English. In *Topic continuity in discourse: A quantitative cross-language study*, 343–363. Amsterdam: John Benjamins.

Glushan, Zhanna. 2010. Deriving case syncretism in differential object marking systems. *University of Connecticut (Ms.)* .

Gomez-Imbert, Elsa. 2011. La famille tukano. In *Dictionnaire des langues*, ed. Emilio Bonvini, Joëlle Busuttil, and Alain Peyraube, 1454–1460. Paris: Presses Universitaires de France.

Good, Jeff. 2011. The typology of templates. *Langauge and Linguistics Compass* 5:731–747.

Good, Jeff. 2016. *The linguistic typology of templates*. Cambridge: Cambridge University Press.

Gordon, Matthew. 2007. The intonational realization of contrastive focus in Chickasaw. In *Topic and focus: Cross-linguistic perspectives on meaning and intonation*, ed. Chungmin Lee, Matthew Gordon, and Daniel Büring, volume 82, 69–82. Dordrecht: Springer.

Grzech, Karolina Z. 2016. Discourse enclitics in Tena Kichwa: A corpus-based account of information structure and epistemic meaning. Doctoral Dissertation, University of London.

Guntsetseg, Dolgor. 2016. *Differential case marking in Mongolian*. Weisbaden: Harrassowitz Verlag.

Hacohen, A., O. Kagan, and D. Plaut. 2021. Differential object marking in Modern Hebrew: Definiteness and partitivity. *Glossa* 6:1–34.

Haegeman, Liliane. 1991. *Introduction to government and binding theory*. Oxford: Blackwell.

Halle, Morris, and Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel Jay Keyser, 111–176. Cambridge US: MIT Press.

Hartmann, Katharina, and Malte Zimmermann. 2009. Morphological focus marking Gùrùntùm (West Chadic). *Lingua* 119:1340–1365.

Haspelmath, Martin. 2005. Argument marking in ditransitive alignment types. *Linguistic discovery* 3:1–21.

Haspelmath, Martin. 2007. Prominence scales in differential object marking: A critique of Aissen (2003). *MPI-EVA, Leipzig (ms)* .

Haspelmath, Martin. 2008. Feature 105: Ditransitive construction, the verb 'give'. In *The world atlas of lanugage structures online*. Max Planck Digital Library (Munich).

Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86:663–687.

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

Haspelmath, Martin. 2015. Descriptive scales versus comparative scales. In *Scales and hierarchies: A cross-disciplinary perspective*, ed. Ina Bornkessel-Schlesewsky, Andrej L. Malchukov, and Marc D. Richards, 45–59. Berlin: De Gruyter.

Bibliography

Haspelmath, Martin. 2019. Differential place marking and differential object marking. *STUF - Language Typology and Universals* 72:313–334.

Haspelmath, Martin. 2021. Role-reference associations and the explanation of argument coding splits. *Linguistics* 59:123–174.

Hastings, Rachel. 2003. The semantics of discontinuous noun phrases in Quechua. In *Proceedings of conference on the Semantics of Under-represented Languages in the Americas*, ed. Jan Anderssen, Paula Menéndez-Benito, and Adam Werle, volume 2, 35–55. Vancouver, Canada.

von Heusinger, Karlos. 2008. Verbal semantics and the diachronic development of differential object marking in Spanish. *Probus* 20:1–33.

von Heusinger, Karlos, and Georg Kaiser. 2003. Animacy, specificity, and definiteness in Spanish. In *Proceedings of the workshop "semantic and semantic aspects of specificity in romance languages"*, 41–65. Universität Konstanz: Arbeitspapier 113.

von Heusinger, Klaus, and Georg A. Kaiser. 2011. Affectedness and differential object marking in Spanish. *Morphology* 21:593–617.

von Heusinger, Klaus, and Jaklin Kornfilt. 2005. The case of the direct object in Turkish: Semantics, syntax and morphology. *Turkic Languages* 9:3–44.

von Heusinger, Klaus, and Petra B. Schumacher. 2019. Discourse prominence: Definition and application. *Journal of Pragmatics* 154:117–127.

de Hoop, Helen. 1996. *Case configuration and noun phrase interpretation*. New York: Garland.

de Hoop, Helen, and Monique Lamers. 2006. Incremental distinguishability of subject and object. *Case, valency and transitivity* 269–287.

de Hoop, Helen, and Andrej L. Malchukov. 2007. On fuild differential case marking: A bidirectional OT approach. *Lingua* 117:1636–1656.

de Hoop, Helen, and Andrej L Malchukov. 2008. Case-marking strategies. *Linguistic inquiry* 39:565–587.

de Hoop, Helen, and Peter de Swart. 2008. Cross-linguistic variation in differential subject marking. In *Differential subject marking*, 1–16. Dordrecht: Springer.

Iemmolo, Giorgio. 2010. Topicality and differential object marking: Evidence from Romance and beyond. *Studies in language* 34:239–272.

Iemmolo, Giorgio. 2013. Symmetric and asymmetric alternations in direct object encoding. *STUF-Language Typology and Universals* 66:378–403.

Iggesen, Oliver A. 2013. Number of cases (v2020.3). In *The world atlas of language structures online*, ed. Matthew S. Dryer and Martin Haspelmath. Zenodo. URL https://doi.org/10.5281/zenodo.7385533.

Irimia, Monic Alexandrina, and Alexandru Mardale. 2023. *Differential object marking in Romance: Towards microvariation*. Amsterdam: John Benjamins.

Irimia, Monica Alexandrina. 2020. Variation in differential object marking: On some differences between Spanish and Romanian. *Open linguistics* 6:424–462.

Irimia, Monica Alexandrina. 2024. Asymmetric coordination in Romanian: A diagnostic for DOM position? *Linguistic inquiry* 1–22.

Jake, Janice L. 1983. Grammatical relations in Imbabura Quechua. Doctoral Dissertation, University of Illinois, Urbana-Champaign.

Janic, Katarzyna, and Martin Haspelmath. 2023. Questionnaire on reflexive constructions in the world's languages. In *Reflexive constructions in the world's languages*, ed. Katarzyna Janic, Nicoletta Puddu, and Martin Haspelmath, 847–853. Berlin: Language Science Press.

Jeretič, Paloma. 2021. Neg-raising modals and scaleless implicatures. Doctoral Dissertation, New York University.

Jerro, Kyle. 2017. The causative-instrumental syncretism. *Journal of linguistics* 53:751–788.

Johnson, Orville E, and Stephen H Levinsohn. 1990. *Gramática secoya*. Quito: Summer Institute of Linguistics.

Jolkesky, Marcelo Pinho de Valhery. 2016. Estudo arqueo-ecolinguístico das terras tropicais sul-americanas. Doctoral Dissertation, Universidade de Brasília (Brazil).

Jones, Christian, and Daniel Waller. 2015. *Corpus linguistics for grammar: A guide for research*. London, New York: Routledge.

Jones, Wendell, and Paula Jones. 1991. *Barasano syntax*, volume 101. Arlington US: Summer Institute of Linguistics.

Kalin, Laura. 2014. Aspect and argument licensing in Neo-Aramaic. Doctoral Dissertation, UC Los Angeles.

Kalin, Laura. 2017. Nominal licensing is driven by valued (phi-)features. *Nordlyd* 43:15–29.

Kalin, Laura. 2018. Licensing and differential object marking: The view from Neo-Aramaic. *Syntax* 21:112–159.

Kalin, Laura, and Philipp Weisser. 2018. Asymmetric DOM in coordination: A problem for movement-based approaches. *Linguistic inquiry* 50:662–676.

Karlsson, Fred. 2015. *Finnish: An essential grammar (third edition)*. London: Routledge.

Kasimir, Elke. 2005. Question-answer test and givenness: Some question marks. *Working Papers of the SFB632, Interdisciplinary Studies on Information Structure* 3:1–52.

Keine, Stefan. 2010. *Case and agreement from fringe to core: A minimalist approach*. Berlin: De Gruyter.

Bibliography

Keine, Stefan, and Gereon Müller. 2015. Differential argument encoding by impoverishment. In *Scales and hierarchies*, ed. Ina Bornkessel-Schlesewsky, Andrej L. Malchukov, and Marc D. Richards, 75–130. Berlin: De Gruyter Mouton.

Kiss, Katalin É. 1998. Identificational focus versus information focus. *Languages* 74:245–273.

Kittilä, Seppo. 2005. Optional marking of arguments. *Language sciences* 27:483–514.

Kittilä, Seppo. 2008. Animacy effects on differential goal marking. *Linguistic typology* 12:245–268.

Kizilkaya, Semra, Zarina Levy-Forsythe, and Klaus von Heusinger. 2022. Affectedness and differential object marking in Turkish and Uzbek. *Linguistics* 60:1907–1941.

Klein, Udo, and Peter de Swart. 2011. Case and referential properties. *Lingua* 121:3–19.

Kohlberger, Martin. 2020. A grammatical description of Shiwiar. Doctoral Dissertation, Leiden University.

Krasnoukhova, Olga. 2012. The noun phrase in the languages of South America. Doctoral Dissertation, Radboud University Nijmegen.

Kratzer, Angelika, and Edlisabeth Selkirk. 2020. Deconstructing information structure. *Glossa* 5:1–53.

Krifka, M. 2001. For a structured meaning account of questions and answers. *Audiatur Vox Sapientia. A Festschrift for Arnim von Stechow* 287–319.

Krifka, Manfred. 2011. Questions and question-acts. Presented at the 18th Amsterdam Colloquium, December 2011, available via `amor.cms.hu-berlin.de/~h2816i3x/talks`.

Krifka, Manfred, and Renate Musan. 2012. Information structure: Overview and linguistic issues. In *The expression of information structure*, ed. Manfred Krifka and Renate Musan, 1–43. De Gruyter.

Kurumada, Chigusa, and T. Florian Jaeger. 2015. Communicative efficiency in language production: Optional case-marking in Japanese. *Journal of Memory and Language* 83:152–178.

König, Elisabeth. 1991. *The meaning of focus particles: A comparative perspective*. London: Routledge.

Labov, William. 1966. *The social stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.

Labov, William. 1970. The study of language in its social context. *Studium generale* 23:30–87.

Lambrecht, Knud. 1994. *Information structure and sentence form*. Cambridge: Cambridge University Press.

Bibliography

LaPolla, Randy J. 1995. 'Ergative' marking in Tibeto-Burman. *Senri ethnological studies* 41:189–228.

Lefebvre, Carl, and Pieter Muysken. 1988. *Mixed categories: Nominalization in Quechua*. Dordrecht: Kluwer.

Leonetti, Manuel. 2004. Specificity and differential object marking in Spanish. *Catalan journal of linguistics* 3:75–114.

Leonetti, Manuel. 2008. Specificity in clitic doubling and in differential object marking. *Probus* 20:33–66.

Levin, Theodore. 2019. On the nature of differential object marking: Insights from Palauan. *Natural language & linguistic theory* 37:167–213.

Lidz, Jeffrey. 2006. The grammar of accusative case in Kannada. *Language* 82:10–32.

Lyutikoval, E., and A. Pereltsvaig. 2023. Towards a formal model of differential object marking in Tatar. *Rhema* 4:9–26.

López, Luis. 2012. *Indefinite objects: Scrambling, choice function, and differential marking*. Cambridge US: MIT Press.

Malchukov, Andrej L. 2008. Animacy and asymmetries in differential case marking. *Lingua* 118:203–221.

Malchukov, Andrej L., and Helen de Hoop. 2011. Tense, aspect, and mood based differential case marking. *Lingua* 121:35–47.

Maldonado, Violeta Vázquez-Rojas. 2011. Case marking and semantic incorporation in Tarascan. In *Semantics of under-represented languages in the americas (sula) 5*. UMOP (UMass Occasional Papers in Linguistics).

Malone, Terry. 1987. Proto-Tucanoan and Tucanoan genetic relationships. Doctoral Dissertation, Bogotá: Summer Institue of Linguistics.

Mason, Jason Alden. 1950. The langauges of the South American indians. In *handbook of south american indians: Physical anthropology, linguistics and cultural geography of south american indians*, ed. Julian H. Steward, volume 6, 157–317. Washington, D.C.: Smithsonian Institution, Bureau of American Ethnology.

Massam, Diane. 2001. Pseudo noun incorporation in Niuean. *Natural language & linguistic theory* 19:153–197.

Matić, Dejan, and Daniel Wedgwood. 2013. The meanings of focus: The significance of an interpretation-based category in a cross-linguistic perspective. *Journal of linguistics* 49:1–37.

Matthewson, Lisa. 2004. On the methodology of semantic fieldwork. *International journal of American linguistics* 70:369–415.

Matthewson, Lisa. 2022. Semantic fieldwork: How experimental should we be? *Semantic field methods* 4:1–24.

Bibliography

McGregor, William B. 2009. Typology of ergativity. *Langauge and linguistics compass* 3:480–508.

McGregor, William B. 2010. Optional ergative cae marking in a typological-semiotic perspective. *Lingua* 120:1610–1636.

Mejeant, Lucia. 2001. Culturas y lenguas indígenas del Ecuador. *Revista Yachaikuna* 1:1–11.

Miestamo, Matti, Dik Bakker, and Antti Arppe. 2016. Sampling for variety. *Linguistic typology* 20:233–296.

Miller, Marion. 1999. *Desano grammar*, volume 6. Arlington US: Summer Institute of Linguistics.

Miyagawa, Shigeru. 2009. *Why agree? why move?: Unifying agreement-based and discourse-configurational languages*. Cambridge, US: MIT Press.

Monahan, T. 1990. Arguments in Hindi. Doctoral Dissertation, Stanford University.

Moravcsik, Edith A. 1978. On the case marking of objects. In *Universals of human languages, vol. 4: Syntax*, ed. Joseph H. Greenberg, 249–289. Stanford US: Stanford University Press.

Morse, Nancy L., and Michael B. Maxwell. 1999. *Gramática del cubeo*. Bogotá: Editorial A. Lleras Camargo.

Murphy, A., and S. Meyase. 2022. Licensing and anaphora in Tenyidie. *Glossa* 7.

Mursell, Johannes. 2018. Object marking in swahili is topic agreement. *Jezikoslovlje* 19:427–455.

Muysken, Pieter. 1989. Predication chains: Case and argument status in Quechua and Turkish. *Linguistic inquiry* 20:627–645.

Muysken, Pieter, Harald Hammarstöm, Joshua Birchall, Rik van Gijn, Olga Karsnoukhova, and Neele Müller. 2015. Linguistic areas, bottom-up or top-down? The case of the Guaporé-Mamoré. In *Language contact and documentation*, ed. Bernard Comrie and Lucía Golluscio, 205–238. Berlin; Munich: De Gruyter Mouton.

Muysken, Pieter C. 2012. Contacts between indigenous languages in South America. In *The indigenous languages of south america: A comprehensive guide*, ed. Lyle Campbell and Verónica Grondona, 235–258. Berlin: De Gruyter.

Myhill, John, and Zhiqun Xing. 1996. Towards an operational definition of discourse contrast. *Studies in language* 20:303–360.

Næss, Åshild. 2004. What markedness marks: The markedness problem with direct objects. *Lingua* 114:1186–1212.

Nedjalkov, Igor. 1997. *Evenki*. London: Routledge.

Neely, Kelsey. 2012. Interrogatives and question words in Máíhìkì. In *Máíjìkì Project: 2012 modules and text sampler (ms.)*. Máíjìkì Project, UC Berkeley.

Neveu, Grace. 2012a. Radial directionals. In *Máíjìkì Project: 2012 modules and text sampler (ms.)*. Máíjìkì Project, UC Berkeley.

Neveu, Grace. 2012b. Spatial relations in Máíhìkì. In *Máíjìkì Project: 2012 modules and text sampler (ms.)*. Máíjìkì Project, UC Berkeley.

Newmeyer, Frederick J. 2002. Optimality and functionality: A critique of functionally-based optimality theory. *Natural language & linguistic theory* 20:43–80.

Nichols, Johanna. 1986. Head-marking and dependent-marking grammar. *Language* 62:56–119.

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

Nichols, Johanna, and Balthasar Bickel. 2013. Locus of marking in the clause (v2020.4). In *The world atlas of language structures online*, ed. Matthew S. Dryer and Martin Haspelmath. Zenodo. URL https://doi.org/10.5281/zenodo.13950591.

Nordlinger, Rachel. 1997. Morphology building syntax: Constructive case in Australian languages. In *Proceedings of the LFG97 conference*, ed. Myriam Butt and Tracy Holloway King, 2–16.

Oomen, Antoinette. 1978. Focus in the Rendille clause. *Studies in Afroasiatic languages* 9:35–65.

Ormazabal, Javier, and Juan Romero. 2013a. Differential object marking, case and agreement. *Borealis* 1:221–239.

Ormazabal, Juan, and Javier Romero. 2013b. Non accusative objects. *Catalan Journal of Linguistics* 12:155–173.

Ospina, Ana María, ed. 2013. *Expresión de nociones espaciales en lenguas amazónicas*. 2. Bogotá: Instituto Caro y Cuervo.

Ospina Bozzi, Ana María. 2002. Les structures élémentaires du yuhup makú, langue de l'Amazonie colombienne: Morphologie et syntaxe. Doctoral Dissertation, Université de Paris 7.

Palancar, Enrique L., Andrew Spencer, and Andrej L. Malchukov. 2012. Varieties of ergative. In *The Oxford handbook of ergativity*, 562–571. Oxford, New York: Oxford University Press.

Palillo, John C. 2002. *Analyzing linguistic variation: Statistical models and methods*, volume 114. Stanford US: CSLI Publications.

Piispanen, Peter Sauli. 2021. *Inside Ecuador: Quechua and Spanish loanword etymologies for Cofán, Siona, Cayapa, Colorado and Waorani*. (unpublished).

Polinsky, Maria. 2010. Linguistic typology and formal grammar. In *The oxford handbook of typology*, ed. Jae Jung Sung, 650–665. Oxford: Oxford University Press.

Bibliography

R Core Team. 2021. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ramirez, Henri. 2019[1997]. *A fala tukano dos Ye'pâ-Masa: Tomo I (Gramática)*. Manaus: Inspetoria Salesiana Missionária da Amazônia, CEDEM.

Carrera de la Red, Micaela. 2020. Fonología del tucano occidental en un vocabulario misionero del espacio geográfico ecuato-colombiano (siglo xviii). *Lingüística* 26:51–69.

Carrera de la Red, Micaela. 2021. La dinámica del contacto lingüístico en la Amazonía ecuato-colombiana durante el siglo xviii en textos de un misionero hablante de quichua. In *Dinámicas lingüísticas de las situaciones de contacto*, ed. Azucena Palacios, María Sanchez Paraíso, and Micaela Carrera de la Red, 277–298. Berlin: De Gruyter.

Reh, Mechthild. 1985. *Die Krongo-Sprache (Nìinò-mó-dì): Beschreibung, Texte, Wörderverzeichnis*. Berlin: D. Reimer.

Repp, Sophie. 2016. Contrast: Dissecting an elusive information-structural notion and its role in grammar. In *The oxford handbook of information structure*, ed. Caroline Féry and Ishishara Shinichiro, 270–289. Oxford: Oxford University Press.

Riesberg, Sonja. 2018. Optional ergative, agentivity and discourse prominence: Evidence from Yali (Trans-New Guinea). *Linguistic typology* 22:17–50.

Roberts, Craige. 2012. Information structure: Towards an integrated formal theory of pragmatics. *Semantics and pragmatics* 5:1–69.

Rochemont, Michael S. 1998. Phonological focus and structural focus. In *The limits of syntax*, ed. Peter Culicover and Louise McNally, 337–363. Leiden: Brill.

Rochemont, Michael S. 2013. Discourse new, f-marking, and normal stress. *Lingua* 136:38–62.

Rodríguez-Mondoñedo, Miguel. 2007. The syntax of object: Agree and differential object marking. Doctoral Dissertation, University of Connecticut.

Rojas-Berscia, Luis Miguel, and Corentin Bourdeau. 2017. Optional or syntactic ergativity in Shawi? Distribution and possible origins. *Linguistic discovery* 15:50–65.

Rojas-Berscia, Luis Miguel, Corentin Bourdeau, Stefan Grondelaers, and Pieter A.M. Seuren. 2023. Antifunctionality in Shawi split ergativity, a processing analysis. *Revista Brasileira de Linguística Antropológica* 15:17–55.

Roosvall, Emilia. 2020. Goal and source in South American languages. Master's thesis, Stockholm University.

Rooth, Mats. 1985. Association with focus. Doctoral Dissertation, UM Amherst.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural language semantics* 1:75–116.

Rose, Françoise, Magdalena Lemus Serrano, Natalia Eraso, and Thiago Chacon. 2017. A new look into Arawak-Tukanoan contact: The Yukuna-Tanimuka bidirectional hypothesis. *SSILA Annual Meeting, 2017, University of Texas, Austin* .

Rufilanchas, Daniel Riaño. 2014. Differential object marking in Ancient Greek. *Linguistics* 52:513–541.

Sadler, Louisa, and Rachel Nordlinger. 2006. Case stacking in realizational morphology. *Linguistics* 44:459–487.

Samek-Lodovici, Vieri. 2016. Constraint conflict and information structure. In *The oxford handbook of information structure*, ed. Caroline Féry and Shinichiro Ishihara, 203–224. Oxford: Oxford University Press.

Sankoff, David. 1988. Sociolinguistics and syntactic variation. In *Linguistics: The cambridge survey*, ed. F.J. Newmeyer, 899–903. Cambridge: Cambridge University Press.

Sankoff, David, and William Labov. 1979. On the uses of variable rules. *Language in Society* 8:189–222.

Schmidtke-Bode, Karsten, and Natalia Levshina. 2018. Reassessing scale effects on differential case marking: Methodological, conceptual and theoretical issues in the quest for a universal. In *Diachronic typology of differential argument marking*, ed. Ilja A. Seržant and Alena Witzlack-Makarevich, volume 19 of *Studies in diversity linguistics*, 509–537. Berlin: Language Science Press.

Schwarz, Anne. 2014. 'From where or what?' – Constituent interrogatives in Secoya. In *Proceedings of conference on language documentation and linguistic theory (7-8 december 2013)*, ed. Aicha Belkadi, Kaki Chatsiou, and Kirsty Rowan, volume 4. School of Oriental and African Studies, University of London.

Schwarzschild, Roger. 1999. Givenness, avoidf and other constraints on the placement of accent. *Natural language semantics* 7:141–177.

Schütze, Carson T. 1996. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press.

Seifart, Frank. 2005. *The structure and use of shape-based noun classes in Miraña (North West Amazon)*. Radboud Universiteit Nijmegen.

Seifart, Frank. 2007. The prehistory of nominal classification in Witotoan languages. *International journal of American linguistics* 73:411–445.

Seifart, Frank. 2011. *Bora loans in Resígaro: Massive morphological and little lexical borrowing in a moribund Arawakan langauge*, volume 2. Cadernos de Etnolingüística.

Seifart, Frank, Doris Fagua, Jürg Gasché, and Juan Alvaro Echeverri. 2009. *A multimedia documentation of the languages of the people of the center*. DOBES-MPI (Nijmegen). URL https://salsa-tipiti.org/sound/sound-repository/people-of-the-center/.

Bibliography

Seifart, Frank, and Doris L Payne. 2007. Nominal classification in the North West Amazon: Issues in areal diffusion and typological characterization. *International journal of American linguistics* 73:381–387.

Selkirk, Elisabeth. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge US: MIT Press.

Selkirk, Elisabeth. 2008. Contrastive focus, givenness and the unmarked status of "discourse-new". *Acta Linguistica Hungarica* 55:331–346.

Seržant, Ilja A. 2016. Nominative case in Baltic in the typological perspective. In *Valency, argument realization and grammatical relations in baltic*, ed. Axel Holvoet and Nicole Nau, 137–198. Amsterdam: John Benjamins.

Shain, Cory, and Judith Tonhauser. 2010. The synchrony and diachrony of differential object marking in Paraguayan Guaraní. *Language variation and change* 22:321–346.

Silva, Wilson de Lima. 2012. A descriptive grammar of Desano. Doctoral Dissertation, University of Utah.

Silverstein, Michael. 1986. Hierarchy of features and ergativity. In *Features and projections*, ed. Pieter Muysken and Henk van Riemsdijk, 163–232. Berlin, Boston: De Gruyter Mouton.

Simpson, Jane. 2023. Semantic case. In *The oxford guide to australian languages*, ed. Claire Bowern, 226–242. Oxford: Oxford University Press.

Sinnemäki, Kaius. 2014. A typological perspective on differential object marking. *Linguistics* 52:281–313.

Skilton, Amalia. 2013. *A new proposal of Western Tukanoan consonants and internal classification. Senior thesis essay*. Yale University (Unpublished).

Stenzel, Kristine. 2008. Kotiria differential object marking in cross-linguistic perspective. *Amerindia* 32:154–181.

Stenzel, Kristine. 2013a. Butterflies 'leaning' on the doorframe: Expressions of location and position in Kotiria and Wa'ikhana. In *Expresión de nociones espaciales en lenguas amazónicas*, ed. Ana María Ospina Bozzi, 85–108. Bogotá: Instituto Caro y Cuervo.

Stenzel, Kristine. 2013b. Contact and innovation in Vaupés possession-marking strategies. In *Upper Rio Negro: Cultural and linguistic interaction in Northwestern Amazonia*, ed. Patience Epps and Kristine Stenzel, 353–402. Rio de Janeiro: Museo do Índio - Funai.

Stenzel, Kristine. 2013c. *A reference grammar of kotiria (wanano)*. Lincoln US: University of Nebraska Press.

Stenzel, Kristine. 2013d. Tucanoan languages. In *Oxford bibliography of languages (online)*. Oxford University Press.

Stenzel, Kristine. 2015. Considerações sobre ordem de palavras, tópico e o efeito foco em kotiria. *Revista Virtual de Estudos da Linguagem* 10:223–246.

Stolz, Thomas, Sander Lestrade, and Christel Stolz. 2014. *The crosslinguistics of zero-marking of spatial relations*. Munich: De Gruyter.

Stolz, Thomas, Cornelia Stroh, and Aina Urdze. 2013. Comitatives and instrumentals (v2020.3). In *The world atlas of language structures online*, ed. Matthew S. Dryer and Martin Haspelmath. Zenodo. URL https://doi.org/10.5281/zenodo.7385533.

de Swart, Peter. 2007. Cross-linguistic variation in object marking. Doctoral Dissertation, Radboud University, Nijmegen.

Tagliamonte, Sali A. 2006. *Analysing sociolinguistic variation*. Cambridge, New York: Cambridge University Press.

Tagliamonte, Sali A. 2012. *Variationist sociolinguistics: Change, observation, interpretation*. Oxford: Wiley-Blackwell.

Thiesen, Wesley, and David Weber. 2012. *A grammar of Bora with special attention to tone*, volume 148. Dallas: Summer Institute of Linguistics.

Tonhauser, Judith, and Erika Colijn. 2010. Word order in Paraguayan Guaraní. *International journal of American linguistics* 76:255–288.

Tonhauser, Judith, and Lia Matthewson. 2016. *Empirical evidence in research on meaning*. LingBuzz (Unpublished).

Tournadre, Nicolas. 1995. Tibetan ergativity and the trajectory mode. *Senri ethnological sttudies* 41:261–275.

Tsunoda, Tasaku. 1985. Remarks on transitivity. *Journal of linguistics* 21:385–396.

Valenzuela, Pilar M. 2011. Argument coding and pragmatic marking of the transitivie subject in Shiwilu (Kawapanan). *International journal of American linguistics* 77:91–120.

Valenzuela, Pilar M. 2015. ¿qué tan amazónicas son las lenguas kawapana? Rasgos centro-andinos y evidencia para una posible sub-área lingüística. *Lexis* 39:5–56.

Vallduví, Enric. 2016. Information structure. In *The cambridge handbook of formal semantics*, ed. Maria Aloni and Paul Dekker, 728–755. Cambridge, New York: Cambridge University Press.

Valle, Daniel. 2014. Focus marking in Kakataibo (Panoan). *Oklahoma working papers in indigenous languages* 1:55–78.

Vallejos, Rosa. 2021. Nominal classification without grammatical agreement: Evidence from Secoya. *International journal of American linguistics* 87:423–455.

Vallejos, Rosa, and Hunter L Brown. 2021. Locative construals: Topology, posture, disposition, and perspective in Secoya and beyond. *Cognitive linguistics* 32:251–286.

Bibliography

Vallejos, Rosa, and Anne Schwarz. 2016. Serial verb construction types in Secoya (Tucanoan). In *Presented at Amazónicas VI (Letícia & Tabatinga, Colombia)*.

Vallejos Yopán, Rosa. 2009. The focus function(s) of =pura in Kokama-Kokamilla discourse. *International journal of American linguistics* 75:399–432.

van 't Veer, Marijn, Martine Bruil, and Oleksandra Damonte-Matveienko. 2025. Preaspiration in ecuadorian siona. In *Rarities in phonetics and phonology*, 755–794. Language Science Press. URL https://doi.org/10.5281/zenodo.15148196.

Velie, Daniel. 2008[1975]. *Orejón: Bosquejo de la fonología y gramática*, volume 10 of *Datos etno-lingüísticos*. Lima: Summer Institute of Linguistics.

Velleman, Leah, and David Beaver. 2016. Question-based models of information structure. In *The oxford handbook of information structure*, ed. Caroline Féry and Ishishara Shinichiro, 86–107. Oxford: Oxford University Press.

Velupillai, Viveka. 2012. *An introduction to linguistic typology*. Amsterdam, Philadelphia: John Benjamins.

Vickers, William T. 1981. The Jesuits and the SIL: External policies for Ecuador's Tucanoans through three centuries. In *Is God an American? an anthropological perspective on the missionary work of the Summer Institue of Linguistics*, ed. Søren Hvalkof and Peter Aaby, 51–62. London, Copenhagen: International Work Group on Indigenous Affairs and Survival International.

Vickers, William Taylor. 1976. Cultural adaptation to amazonian habitats: The siona-secoya of eastern ecuador. Doctoral Dissertation, University of Florida.

Vinokurova, Nadezhda. 2005. Lexical categories and argument structure: A study with reference to Sakha. Doctoral Dissertation, Utrecht University.

van der Wal, Jenneke. 2011. Focus excluding alternatives: Conjoint/disjoint marking in Makhuwa. *Lingua* 121:1734–1750.

van der Wal, Jenneke. 2015. What you see is (not) what you get: information structure on the interface between syntax and discourse. *MFiL,6-7 November 2015, University of Manchester* .

van der Wal, Jenneke. 2016. Diagnosing focus. *Studies in language* 40:259–301.

van der Wal, Jenneke. 2022. *A featural typology of Bantu agreement*. Oxford: Oxford University Press.

Waltz, Nathan, and Carolyn Waltz. 1997. *El agua, la roca y el humo: Estudios sobre la cultura wanana del Vaupés*. Bogotá: SIL.

Waltz, Nathan E., and Alva Wheeler. 1972. Proto Tucanoan. In *Compararive studies in Amerindian languages*, ed. Esther Matteson, Alva Wheeler, Frances L. Jackson, Nathan E. Waltz, and Diana R. Christian, 119–149. The Hague: Mouton de Gruyter.

Wasserstrom, Rober. 2014. Surviving the rubber boom: Cofán and Siona society in the Colombia-Ecuador borderlands (1875-1955). *Ethnohistory* 61:525–548.

Wheeler, Alva. 1992. Comparaciones lingüísticas en el grupo tucano occidental. In *Estudios comparativos proto tucano*, ed. Stephen H. Levinsohn, 17–53. Bogotá: Editorial Alberto Lleras Camargo.

Wheeler, Alva L. 1962. A Siona text morphologically analyzed. In *Ecuadorian Indian Languages*, ed. Benjamin Elson, volume 7, 248–268. University of Oklahoma: CNWS.

Wheeler, Alva L. 1970. Grammar of the Siona language, Colombia, South America. Doctoral Dissertation, UC Berkeley.

Wheeler, Alva L. 1987. *Gantëya bain: El pueblo siona*. Bogotá: Summer Institute of Linguistics.

Wiegertjes, Juriaan. 2020. Laryngeal oppositions of obstruents in Ecuadorian Siona: An element-based approach. *Leiden University (Unpublished)* .

Willis, Christina M. 2011. Optional case marking in Darma (Tibeto-Burman). *Linguistics of the Tibeto-Burman Area* 34:101–131.

Witzlack-Makarevich, Alena, and Ilja A Seržant. 2018. Differential argument marking: Patterns of variation. In *Diachrony of differential argument marking*, ed. Alena Witzlack-Makarevich and Ilja A Seržant, 1–49. Berlin: Language Science Press.

Wojtylak, Katarzyna I. 2021. *A grammar of Murui (Bue): a Witotoan language from Northwest Amazonia*. Leiden, Boston: Brill.

Wordick, Frank J. 1982. *The Yindjibarndi language*. Canberra: Australian National University.

Wälchli, Bernhard, and Fernando Zúñiga. 2006. Source-Goal (in)difference and the typology of motion events in the clause. *Language typology and universals* 59:284–303.

Zeller, Jochen. 2014. Three types of object marking in Bantu. *Linguistische Berichte* 239:347–367.

Zimmermann, Malte, and Edgar Onea. 2011. Focus marking and focus interpretation. *Lingua* 121:1651–1670.

Zúñiga, Fernando. 2007. The Discourse-Syntax interface in northwestern Amazonia: Differential object marking in Makú and some Tucanoan languages. In *Language endangerment and endangered languages: Linguistic and anthropological studies with special emphasis on the languages and cultures of the Andean-Amazonian border area.*, ed. W. Leo Wetzels, 209–227. CNWS, University of Leiden.

Curriculum Vitae

Justin Case was born in 1990 in Tampere, Finland; but spent his childhood in northern Ontario, Canada. After one year of music school at York University, Toronto in 2008; he completed a double-major BA degree at the University of Waterloo (Canada) in French Studies and in Spanish and Latin American Studies in 2012, including one year on exchange at the University of Navarre in Pamplona, Spain. After one year of teaching English in Kyiv, Ukraine, Justin undertook his Masters' studies in Europe.

He began his MA studies in Medieval Studies at University College London in 2014. The MA was awarded with merit in November 2015, culminating in a thesis project *Assimilation of voicing in the Alemannic Middle High German corpus of Notker III*. He spent a year working at a language school in San Francisco, USA, before returning to Europe to undertake a second MA program in Linguistics, Language and Society, at the University of Amsterdam in 2016. This second MA was awarded cum laude in June 2017, with a thesis project written in Spanish with the title *¿Alguna vez existió el adjetivo? La relevancia del adjetivo en las Artes misioneras sobre las lenguas tupí-guaraníes en la época colonial temprana (ca. 1500-1650 EC)* [English: *Were there ever any adjectives? The relevance of the adjective in the missionary grammars on Tupi-Guarani languages in the early colonial era (c. 1500-1650 CE)*].

Justin started his doctoral studies at the University of Ottawa in September 2017, and completed a full taught program and two qualifying examinations. In the spring of 2018, he participated in an on-site field course with Prof. dr. A.P. Salanova and Dr. M. Bruil at the Siona community of Sototsiaya. From January 2021 he established a cotutelle agreement with Leiden University, as an external candidate, and the doctoral dissertation project contained in these pages took shape. Since 2018, Justin has visited this Siona community for four additional field stints (2019, 2022, 2023, 2025), and has invited a speaker friend to work with him at the University of Ottawa, Canada in Fall 2024. To date he has taught a half-dozen courses on theoretical and descriptive linguistics topics at the University of Ottawa in English and in French, and co-taught two courses on language diversity at Leiden University.