Differential case-marking in Ecuadorian Siona

Case, J.T.

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]
Máíhīkì [*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:

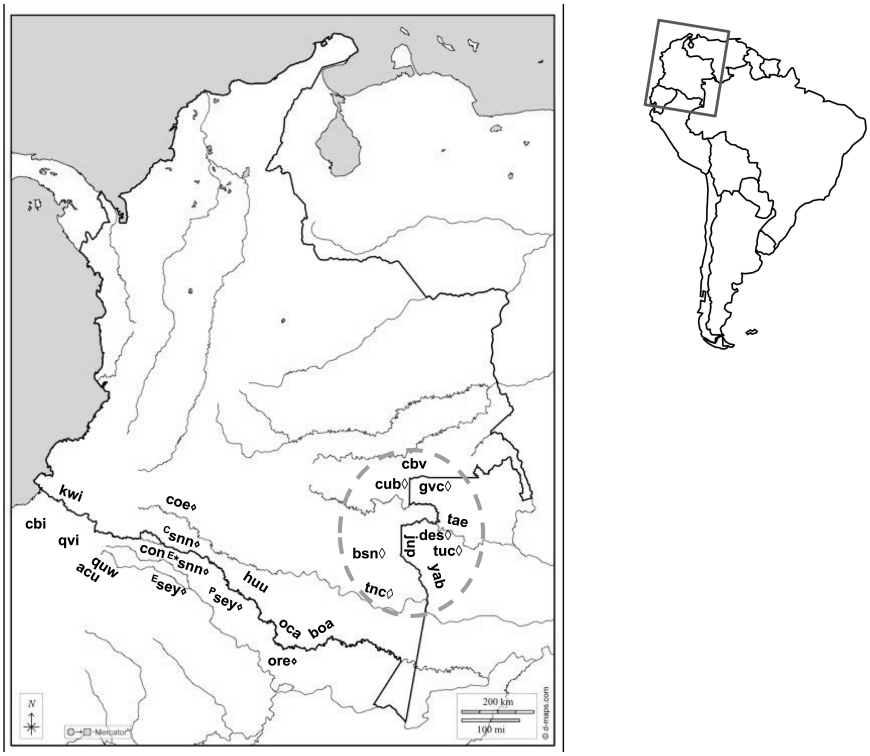

Figure A.1: Geographical distribution of languages . Map template downloaded from *https://d-maps.com* (Colombia borders, including rivers).

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* THERE (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 with his wife.'

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 with them.'

(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 only toward Pedro's house.'

b. ***wa'í maú-si'-kɨ*** ***nakóni*** *koká ká-wɨ*
fish catch-COMP-CL:M COM word say-N3S.PST.ASS
'(I) conversed with the one who had caught the fish.'

(16) Kakua^{V} (*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 up right to the middle of the front yard (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. ***ɟɨ'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* *ɟɨ'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*** ***ɟɨko=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=wɨ̃-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 ɨ̃bɨ̃ tota-ta-RExa-Abɛ̃ **õpõ-jɨ̃-ke***
other-CL:M man hit-ITER-RMT.PST-3M.SG explode-CL:FUNNEL-INST
'ɨ̃-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
'xi-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 bʉto asi bahi-ro-dɛ̃ hʉ̃a-ko-aka-hʉ*
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. ***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?'

(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. ***m'ăc-ãt** pɨ́d hɨd bib'-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. ***'ãh='íp-ít** 'ãh ni-'e'-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 with my brother.'

(40) Shiwiar (*acu* — CHICHAM) [K20:263-264, exx. (7.220),(7.223)]

a. *nukútsəí-rú* *wíi* ***tsuwír=hʲãĩ***
grandmother-1SG.P+VOC 1SG hot-COM
i-mʲa-i-hmɨ̃
CAUS-bather-PFV-1SG.S>2SG.O+DECL
'Grandmother, I just bathed you with the hot (water).'

b. *wíi* *win-i-tʲát-ha-i* ***huríhuri=hʲãĩ***
1SG come-PFV-FUT-1SG.S-DECL cannibal-COM
'I will come with the cannibals.'

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) 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.'

(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 with cassava.'

b. *i'-kó* *wã-rũ-kó* ***Juan-'ka***
3PRO-CL:F study-FUT-F J-COM
'She is studying with Juan.'

(43) Yuhup∨ (*yab* — NADAHUPAN)[12] [O02:158,161, exx. (30b), (36b)]

a. ***pŏh-ót*** *jâp* *wédⁿí*
hand-PRED 3PRO eat
'He eats with his hands.'

[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*** *jɨ̀də̌h* *wɨ̀dnɛ̀d^{m}ɨ́*
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*$^{♦}$ *con huu boa quw qvi acu cbv*^{V} *tae*^{V} *yab*^{V}	*e-snn*$^{♦}$ *c-snn*$^{♦}$ *coe*$^{♦}$ *cub*$^{◇}$ *tnc*$^{◇}$ *tuc*$^{◇}$ *des*$^{◇}$ *bsn*$^{◇}$ *gvc*$^{◇}$ *kwi jup*^{V}

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$^{♦}$ (*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^{V} (*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***, *-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*, *-v̋t*, *-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*** *kãĩ-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'ĩ-hi*
canoe-N.SBJ be_seated-3S.M.PRS.ASS
'(He) is seated in the canoe.'

c. ***wɨ'é-de*** *ba'ĩ-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=bʉ̌=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=bʉ̌*** *fĩ=beh-ep=hĩ*
M=LOC 1PL=go-PST=RMT.PST
'We left to Mitú.'

c. ***tʃena=bʉ̌*** *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$^{\Diamond}$ [*cub*]	***-re***, *-i*, *-rã*	***-re***, *-i*, *-rã*	***-re***, *-i*, *-rã*
Tukano$^{\Diamond}$ [*tuc*]	***-re***, *-i*, *-pʉ*	***-re***, *-i*, *-pʉ*	***-re***, *-i*, *-pʉ*
Desano$^{\Diamond}$ [*des*]	***-re***, *-ge*	***-re***, *-ge*	***-re***, *-ge*
Barasano$^{\Diamond}$ [*bsn*]	***-re***, *-hʉ*	***-re***, *-hʉ*	***-re***, *-hʉ*
Wanano$^{\Diamond}$ [*gvc*]	***+re***, *+i*, *-pʉ*	***+re***, *+i*, *-pʉ*	***+re***, *+i*, *-pʉ*
Kakua^{V} [*cbv*]	*=bʉ̌*	*=bʉ̌*	*=bʉ̌*
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*]	*-Ít*, *-an*, *có*	*-Ít*, *-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◇ (*-i*)	**e-snn*♦
tnc◇ (*-ri'ká*)	*c-snn*♦
tuc◇ (*+yaá*)	*coe*♦
des◇ (*=ya*)	*ore*♦
bsn◇ (*-ya*)	*gvc*◇
quw (*-pa*)	*con*
qvi (*-paj*)	*huu*
acu (*=n/=n^{y}a*)	*boa*
kwi (*=pa*)	ᵛ*tae*
*cbv*ᵛ (*-i'*)	
*jup*ᵛ (*=nɨ́h*)	
*yab*ᵛ (*-˜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◇ (*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◇ (*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) Hupᵛ (*jup* — NADAHUPAN) [E08:225-226, exx. (34a)-(34b)]

a. ***pedú=nɨ́h*** *cug'ăt*
P=POSS book
‘Pedro's book.’

b. *cug'ăt* ***pedú=nɨ́h***
book P=POSS
‘Pedro's book.’

(59) Yuhupᵛ (*yab* — NADAHUPAN) [O02:243, ex. (101e)]
dɨd^{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 ɨkɨ́-taiᴺ kintʲa=**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) Yuhupⱽ (*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 *tʰû∼tʰû-'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	
Tarianaⱽ [*tae*]	∅✓[d]	{-∅, *-n(h)e*}			✓TOP	

NO DCM ON S (-∅): Máíhɨ̃kìᵛ [*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*]; Kakuaⱽ [*cbv*]; Hupⱽ [*jup*]; Yuhupⱽ [*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=di*** *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. ***yi'i-ga*** *saí-yi*
1SG-SBJ.EMPH go-N3S.PRS.ASS
'And I, I am going.'

b. ***ča'kádi-wa'-na-bi*** *bá-gi-na* *gahe-ni*
fire_ant-ANIM-PL-SBJ PRO.HUM-CL:M-GOAL go_down-SS
si'á-wi
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]
ji'i-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. *ɸo'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*** *áɨ́-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*** *ɨ́ɨ́-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. ***tɨ̀nɨ̀h*** ***cápu-ăn=yɨ́'*** *tɨh 'éy-cud'ɨ̃hníy*
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'* ***tí*** *bûd=na=ka* mi=tʃãhap
3SG.F firewood cut=DECL=ASS 3SG.F=PROG
'She is cutting firewood.'

b. ***tí-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)]
bɯ̃-re *sĩa-gɯ-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)]
∼bɯ'ɯ ***∼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. ***jìdə̃h-~dìh*** *tíh* îhkéjí
3PL-OBJ 3SG greet
'He greets them.'

b. ***tìjí'*** *~úb^m í* ~jàb^m běh
man hit dog
'The man hits the dog.'

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 *ati-ñe-iti-o* ***asukar-na*** *ati-iti-o*
rice bring-NEG-FUT.LK-2SG sugar-N.SBJ.TOP bring-FUT.LK-2SG
'(You) won't bring rice. (You) will bring the SUGAR.'

(94) Koreguaje♦ (*coe* — WESTERN-TUKANOAN) [CC93:49, ex. (201)]
k^hura-wa'-ɨ-ni *k^hu'e-mo* *ji'i*
chickens-creature-M.SG-N.SBJ2 look.for-F.SG 1SG
'I (F) am looking for a chicken (a particular one, not just any).'

(95) Ecuadorian Siona♦ (^e*snn* — WESTERN-TUKANOAN)
ba̲huë yohuë hueroye te'e guë̲narore 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: 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 bikɨ-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).'

(98) 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-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♦ [E*snn*]	+✓[a,d]	{***-re***, *-ni*, (*-na*)}	✓ANIM		✓FOC	
Colombian Siona♦ [C*snn*]	+✓[d(a)]	{***-re***, *-ni*, *-na*}	✓(ANIM)		✓FOC	
Koreguaje♦ [*coe*]	+✓[d(a)]	{***-re***, *-ni*}	✓(ANIM)		✓CONTR	
Murui [*huu*]	∅✓[d]	{*-*∅, *-mo*}		✓PRO	✓CONTR	
Tena Kichwa [*quw*]	+✓[o]	{*-ta*, *-ma*}				✓TBD
Tariana^V [*tae*]	+✓[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^V [*cbv*] (*=di'*); Hup^V [*jup*] (*=ăn*) Yuhup^V [*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	+∅✓[d(r,o)]	{-∅, **-re**, *-ni*, *-na*}	✓(PRO, SPEC)	✓FOC	✓(PRED)
	G	∅✓[d(r)]	{-∅, *-na*}	✓(PRO, SPEC)	✓FOC	
Colombian Siona♦ [C*snn*]	L	+∅✓[d(r,o)]	{-∅, **-re**, *-na*}	✓PRO, (SPEC)	✓TOP	✓(PRED)
	G	∅✓[r]	{-∅, *-na*}	✓PRO, (SPEC)		
Koreguaje♦ [*coe*]	L	+∅✓[d,r(o)]	{-∅, **-re**, *-na*}	✓SPEC	✓TOP	✓(PRED)
	G	∅✓[r,d]	{-∅, *-na*}	✓SPEC	✓TOP	
Kubeo◇ [*cub*]	L=G+	+✓[r(d)]	{**-re**, *-i*, *-rã*}	✓SPEC	✓(FOC)	
Tukano◇ [*tuc*]	L=G+	+∅✓[r(d)]	{-∅, **-re**, *-pʉ*}	✓SPEC	✓(TOP)	
Desano◇ [*des*]	L=G+	+∅✓[r(d)]	{-∅, **-re**, *-ge*}	✓SPEC	✓(FOC)	✓(PROTO)
Barasano◇ [*bsn*]	L=G+	+∅✓[r(d)]	{-∅, *-hʉ*, (**-re**)}	✓(SPEC)	✓(FOC)	
Wanano◇ [*gvc*]	L	+✓[r(o)]	{**-re**, *-i*, *-pʉ*}	✓REF		✓(DIST)
	G+	+✓[r]	{**-re**, *-pʉ*}	✓(SPEC)		
Murui [*huu*]	L=G	+∅✓[(r,o)]	{-∅, *-mo*, *-na*}	✓(PRO)		✓(PROTO)
Tena Kichwa [*quw*]	L	+✓[(o)]	{*-pi*, *-ma*}			✓(PRED)
Imbabura Kichwa [*qvi*]	G	+✓[(o)]	{*-man*, *-ta*}			✓TBD
Shiwiar [*acu*]	L=G	+∅✓[r,o]	{-∅, *=nam*, *=i*}	✓PRO		✓PROTO
Awa Pit [*kwi*]	L	+✓[o]	{*=ta*, *=pi*, *=mal*, *=ki*}			✓DIST
	G	+✓[o]	{*=ta*, *=pi*, *=mal*}			✓DIST
Tariana^V [*tae*]	L=G	∅✓[(o)]	{-∅, *-se*}			✓(PROTO)
Hup^V [*jup*]	L	+✓[d(o)]	{*-an*, -V*t*, *-có*}		✓CONTR	✓(PRED)
Yuhup^V [*yab*]	L=G+	+✓[o]	{*-bâh*, *-'áh*, *-'áhà*}			✓DIST

No spatial case-markers: Máíhɨ̃kì♦ [*ore*] (-∅(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*] (*=bʉ̃*(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 in the house yard and made her dance, spinning.'

b. ***yo-i*** ***kĩrãmi-i*** *eda-ma*
here-LOC house-LOC arrive-3PL.ANIM
'They arrived at the house.'

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 here.'

(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 here, after which (we) went down to a farm (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 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'tĩ-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).'

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~~ʉ~~* *sai-na'-me* *ch~~ʉ~~k~~ʉ~~na*
DEM.MED leaf-CL:RIV turtle go-PRP-N3S 1PL.EXCL
'We (EXCL) go to the Caquetá River for turtles ...'

b. ***meja-huʉ-na*** *co~~ʉ~~* *mai-na'-me*
sand-CL:CONT-GOAL turtle come-PRP-N3S
'... the turtles come up on the beach ...'

c. *jain~~ʉ~~ko* ***meja-huʉ*** *co~~ʉ~~* *mai-me*
many.PL.ANIM sand-CL:CONT turtle come-N3S
'... many are the turtles which come up on the beach.'

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~~i~~-do* ***nofi~~i~~ko-mo*** *i-t~~i~~-kue*
this.CTS-CL:DAY-PL-INST La_Chorrera-LOC exist-LK-1SG
'Nowadays, I live in La Chorrera.'

ii. *kue* *a~~i~~* ***i~~y~~i*** *jaai-d-e*
1SG wife jungle_garden go-LK-3
'My wife went to the jungle garden.'

b. i. ***beno-mo*** *i-t~~i~~-ka~~i~~*
here-LOC exist-LK-1PL
'We live here (in this specific place, e.g., in this village).'

ii. ***beno-na*** *i-t~~i~~-ka~~i~~*
here-N.SBJ.TOP exist-LK-1PL
'We are here (passing through this place, e.g., in the jungle).'

iii. *jaai-ño-kai-ñe-no* *bu-e* ***beno*** *ñee∼ñee-d~~i~~-o*
go-?-RAPID-NEG-SEQ Q1-CL:GEN here do∼RED-LK-2SG
'You didn't go, what are you doing (and doing) here?'

(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-ãn* *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)	*hũĩ-* (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, *bẽhẽtub**i**-**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-tub~~**i**~~**-na**$^{<\text{ Non-Specific }>}$ m~~i~~ni + tui-k~~i~~
be_fallen-CL:STICK-GOAL go_up + sit_on_something-CL:M
ba-∼'-k~~i~~-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-tub~~**i**~~**-na**$^{<\text{ Specific }>}$ ja'o hẽna-k~~i~~
be_fallen-CL:STICK-GOAL mud stick-CL:M
ba-∼'-k~~i~~-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 ja̲maca i̲omaca 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 – 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. *…ja̲re 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, *pihpɨri* (vulture), is treated as non-contrastive; whereas the other two token, *ĩõmahka* (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 *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
hɨje-kɨ̈ nee-ni ha'o-bi pɨ̈∼pɨ̈-ni
break.INTRS-S.M.PRS.DEP make-SS leaf-INST pile∼REDUP-SS
ha'ru-ɨ̈-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-kɨ̈-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-kɨ̈-bi hãre tres to'to sani + ha'ru-ni ɨ̃-ɨ̈
go-CL:M-SBJ like_that three watering_hole go + sit_down-SS 3PRO-CL:M
beo + hẽ̃õ-ɨ̈-ña **gõ'no** ũhku + chao-kɨ̈
NEG.EXIS + leave_behind-3S.M.PST.N.ASS-REP chicha drink + finish-CL:M
ba-∼'-kɨ̈-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-hɨ̈-na jɨ'ɨ̈ **bia** **+ da'ka-re** ũhku-si'-ɨ̈
see-PL.PRS.DEP-DS 1S aji_pepper + liquid-N.SBJ drink-FUT-N3S
kaa-ɨ̈-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 Input: 0.393			* p = < 0.001 ! 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) *jaro utibëbi tomeji huare ja 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 cuju'i 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 i hueqüeje
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 cujü gachoquëje jü'iñe baü'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 tsimacabi goeo goeo goeo 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-hẽ̃ 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 tsĩĩ 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 dë̤j̲o̲ bacoña
go'i-ni **[ĩ-o tsɨ̤'bo]** baa-ɨ̤-ña
return-SS 3PRO-CL:F ax have-2/3S.M.PST.N.ASS-REP
baa-kɨ̤-na ĩ-ɨ̤ 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 b̲a̲quëña j̲a̲mubi*
dɨ̤mɨ̤-kɨ̤ ba-∼-kɨ̤-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-kɨ̤-ña **a'ri + tsia-ja**
kill + go-2/3S.M.PST.N.ASS-REP small + river-CL:RIV
bah-kɨ̤-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 Input: 0.606			* p = < 0.001 ! 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 **tsĩĩ-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. *jo̱jẹ̱ yo'co a̱co'ë tsi̱dohuëre yureto̱'a̱ quëajo̱'cua'ia'ë ja'quëre careña*
hãhẽ jo'-ko ãh-ko'ɨ **tsĩĩ-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 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 ju̱quëna yureta̱'a̱ mamajëre cayabare baëña ë̱më
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.]
. . . a̱nichaojëna tsiare baoña io̱
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-ɨ̶ hãmu ti ha'o daa-sih-kɨ̶**] hãre sa'niwɨ̶-na
3PRO-CL:M armadillo ANA mud bring-COMP-CL:M like_that inside-GOAL
bɨa-ɨ̶
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-ɨ̶-ña
armadillo=TOP like_that not_be-2/3S.M.PST.N.ASS-REP
neewesɨ̶-ɨ̶-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 i̱o̱ni 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 **wahɨ̶-ko-ni** hɨa-si'-i kaa-ni yo'-kɨ̶-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 **sõhkɨ̶tene + wa'i-re**
lie_in_hammock-S.F.PST.DEP 1S=TOP catfish(spec) + meat-N.SBJ
ãĩ-ko-a'-ɨ̶ kaa-ko hã-ɨ̶-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 jɨ'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 hɨjo-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 tsɪ̃ɪ̃-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 ëmë
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 ɪ̃mi̵
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-i-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 taitotaʼa tsiaya baʼijaʼcoa
↑**ih-ko-re** kwẽẽ-ni hẽõ-ñuʼũ kwẽẽ-ni
DEM.PROX-CL:F-N.SBJ take_down-SS leave_behind-HORT take_down-SS
hẽõ-i-na ↑hã-o-bi tãi-to-tãʼã
leave_behind-S.M.PST.DEP-DS DEM.MED-CL:F-SBJ fall-COND-CNT.EXP
tsia-ja baʼi-hãʼ-ko-aʼ
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 daijẽʼẽ cani sañuʼu mësaruni caquë daajẽʼẽ siʼa tsëcabëre
kaa-ni dai-hɨ̃ʼɨ̃ kaa-ni saa-ñuʼũ **mɨhsaru-ni** kaa-kɨ
say-SS come-IMPER say-SS bring-HORT 2PL-N.SBJ2 say-S.M.PRS.DEP
daa-hɨ̃ʼɨ̃ siʼa tsɨhka-bi-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.]
sahuëʼhuë mëareña yaibaicato iohuaʼire cucujoñuʼu cani . . .
sawɨʼwɨ + mɨa-re-ña jai + bãĩ=kato
fly + go_up-N2/3S.PST.N.ASS-REP jaguar + person.COL=TOP
ĩ-o-waʼi-re kũh∼kũ + hõã-ñuʼũ 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 aideʼhuani i dëjore. . .*
[**ĩ-ɨ** **kori-jo-re**] duhta-ni baa-kɨ-bi aideʼwa-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 **ĩhĩ-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 huareje̲ 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-ɨ-na [ti tu-bɨ-hẽ] jure-tã'ã
sit_on_something-S.M.PRS.DEP-DS ANA trunk-CL:RND-ADD now-CNT.EXP
hɨa + tu-bɨ de'o-ɨ-ña
hard + trunk-CL:RND become-2/3S.M.PST.N.ASS-REP
'(He) sat where there was a river. 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'niwɨ hai kwa'ko-ro tsĩũ-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) grew inside the big covered pot.' [NAT*: 20150811sfryi001.147]

b. *. . . go'ini caquë ba̲'quëña baja'i tsiaya − ja̲o̲na ja̲ë̲hua'i* guya-yë
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**
be-3S.M.PST.ASS river-CL:RIV DEM.MED-CL:F-GOAL
hã-ɨ-wa'i guja-jɨ
DEM.MED-CL:M-PL.AN bathe-N3S.PRS.ASS
'. . . (he) came back and said, "There is a river. They are bathing there." ' [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
hɨje-kɨ nee-ni ha'o-bi pɨ∼pɨ-ni
break.INTRS-S.M.PRS.DEP make-SS leaf-INST pile∼REDUP-SS
ha'ru-ɨ-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-kɨ-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]

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 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*[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 i̱sijẽ̱'ẽ̱ yë'ni caoña
si'a tsia + go'o ĩhsi-hɨ̃'ɨ̃ **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 saija̱'cua'ia ? ja'quëquë'ro ti̱'aja̱'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 ba̱'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 siɨi-hã'-ko-a'-ɨ
DEM.MED carachama_palm + firewood-INST dry_out-PRP-CL:F-COP-N3S
hehka-reba=kato bãã-ñɨ jɨ'ɨ siɨi-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 ñatoje − tocato jare so̲quëñëa 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.)* .

Bibliography

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.

Bibliography

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.

Bibliography

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.

Bibliography

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.

Bibliography

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.

Bibliography

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.