UNIVERSITY OF CALIFORNIA
SANTA CRUZ

CASE, AGREEMENT, AND SENTENCE PROCESSING IN GEORGIAN

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

LINGUISTICS

by

Steven Foley

June 2020

The Dissertation of Steven Foley
is approved:

_____
Professor Matt Wagers

_____
Professor Sandy Chung

_____
Professor Amanda Rysling

_____
Professor Maziar Toosarvandani

_____
Quentin Williams
Acting Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

Case, agreement, and sentence processing in Georgian

by

Steven Foley

This dissertation examines a few dimensions of morphosyntactic complexity in Georgian. Central are the language's split-ergative case system, whereby clausal arguments are assigned different case morphology across different tense–aspect–mood categories, and its verbal agreement paradigm, in which φ-agreement morphemes interact in complex but systematic ways. Three pairs of self-paced reading experiments probe the ramifications of Georgian split ergativity for online sentence processing, in ordinary transitive root clauses and also relative clauses. The Georgian comprehender is often faced with arguments whose case morphology does not unambiguously indicate their syntactic roles. Results show that comprehenders navigate these temporary ambiguities by harmonically aligning animacy and syntactic role — all else being equal, human arguments are parsed as transitive subjects, and inanimates as direct objects. Interpreted as garden path effects, the specific distribution of reading-time disruptions add nuance to this parsing heuristic, suggesting that comprehenders have fine-grained by not unlimited access to their abstract grammatical knowledge.

Moving to relative clauses, I find that priorities are subtly readjusted during the comprehension of filler–gap dependencies. With both Accusative- and Ergative-aligned relative clauses that can either precede or follow their head nouns, Georgian is uniquely well equipped to disentangle theories of filler–gap processing. And in relative clauses of all

stripes, cues that eliminate the possibility of a subject-gap parse regularly lead to processing difficulty. This observation lends support to theories in which the structural distance between fillers and gaps — rather than the linear distance, or the informativity of ambient morphological cues — is the primary predictor of relative-clause processing difficulty.

The thesis is rounded off with a detailed formal investigation of argument–verb agreement in Georgian and its sibling languages. I identify a few generalizations that reveal key systematicities within superficially complex paradigms. These motivate an analysis deriving agreement patterns from the interaction of narrow-syntactic and post-syntactic mechanisms. Specifically, a syntactic principle independently motivated by non-agreement phenomena in other languages permits syntactic locality constraints to be loosened in certain circumstances; resulting derivational indeterminacy is obscured by morphological constraints that filter out all but the most expressive and economical combinations of agreement morphemes.

To Jim Foley

Acknowledgments

Without the help, friendship, and boundless generosity of Marina Rcheulishvili, Ana Chutkerashvili, Elene Kadagishvili, Mariam Matiashvili, Mariam Rukhadze, Ana Kolkhidashvili, Salome Kobalia, Natia Botkoveli, and especially Irma Miminoshvili & Mariam Navadze, I could not have conducted my research on Georgian. *Gmadlobt, da gagimarǰot.*

I will always be grateful to the people of Santiago Laxopa, who welcomed me into their community, shared their beautiful language, and ensured that I didn't die of appendicitis in the mountains of northern Oaxaca. They include Fe Esperanza Silva Robles, Raquel Díaz, Vianey López, Olivia Maldonado, Amelia Vasquez Reyes, Sara Reyes, Venustiano López, Evaristo López Velázquez, Nereida Robles Vásquez, Raúl Díaz, and the presidents of Laxopa in the summers of 2018 and 2019, Celestino Robles Ramírez and Óscar Benites Celis. *Duxklenle.*

Finally, I wish to acknowledge that land where I lived and worked during graduate school is the unceded territory of the Awaswas-speaking Uypi Tribe. The Amah Mutsun Tribal Band, comprising the descendants of Indigenous people taken to missions Santa Cruz and San Juan Bautista during Spanish colonization of the Central Coast, is today working hard to restore traditional stewardship practices on these lands and heal from historical trauma.

Chapter 1

Introduction

1.1 Overview

This thesis investigates a set of interconnected questions about the mental representation of morphological and syntactic grammars. The meat of that investigation comprises a set of self-paced reading experiments and an in-depth formal analysis. Engaging primarily with one particular language, Georgian, these studies focus on three phenomena prominent in its syntax and morphosyntax — Split-Ergative case alignment, relativization, and argument–verb φ-agreement. The experiments interface most readily with theories of syntactic processing, and the formal analysis with theories of abstract grammar. Together, I hope they provide some insight into the psycholinguistic and formal consequences morphosyntactic complexity, both in a single language and more generally. The primary goals and findings of these projects are summarized in the rest of this subsection.

Chapter 2 investigates the role animacy plays in the online comprehension of

temporary syntactic ambiguities, using as a lens Georgian's Split Ergativity. Consider a verb-final language — or more precisely, a language where verb-final clauses predominate, preverbal argument order is flexible, and argument omission is frequent. An active comprehender of that language is faced with the nontrivial task of associating syntactic positions to clausal arguments (ones they have already countered, or which they may predict to encounter shortly) in order to license and interpret those arguments, all without access to the verb's specific lexical information.

But the comprehender has at their disposal a smorgasbord of features and cues — prosodic, morphological, syntactic, semantic, and pragmatic. Many of these features can be arranged into into scales oriented along a shared dimension of abstract grammatical–cognitive prominence (Silverstein 1976, Hopper and Thompson 1980, Van Valin 1990, and many others). In naturalistic discourse, features from the high-prominence ends of these scales reliably cluster together; likewise features from low-prominence ends. Features that cluster in this way can be said to be *harmonic* or to be drawn from *harmonically aligned* scales. Features with clashing prominence values, on the other hand, can be called *dissonant* or *harmonically misaligned*. Psycholinguistic evidence suggests that structures involving harmonic prominence scales are in general easier to process than ones involving dissonant scales (Bornkessel-Schlesewsky and Schlesewsky 2009). So theories which incorporate prominence calculations into the parsing mechanism can successfully predict many kinds of incremental processing difficulties.

However, our current understanding of sentence processing is rather modest when it comes to just how parsing mechanisms interface with crosslinguistic grammatical vari-

ation. There are myriad conceivable shapes a language's lexicon could take, and myriad intricacies that could gild its grammatical systems. Key to advancing our understanding is the careful collation and comparison of psycholinguistic data representing a wide distribution of points across this space of variation. As I sketch in Section 1.2 and describe in greater depth in Chapters 2–4, Georgian has many morphological and syntactic properties which make it a particularly useful target of inquiry. One is its Split-Ergative case system, which gives rise to many temporary ambiguities about arguments' syntactic roles.

Data from a pair of self-paced reading studies show that comprehenders of Georgian parse clauses with two role-ambiguous human arguments very differently from clauses with two similarly ambiguous inanimate arguments. And the particular asymmetries in the distribution of reading difficulty conform strongly to the predictions of a theory of sentence processing like the extended Argument Dependency Model (eADM; Bornkessel-Schlesewsky and Schlesewsky 2009), in which harmonic scale alignment plays a prominent role. Two types of evidence support this claim. First are reading-time disruptions interpretable as garden-path effects (Frazier 1978, 1987). We reliably observe processing difficulty at or just after words bearing grammatical cues that eliminate from consideration (relatively) harmonic parses compatible with prejacent clausal material. Second are slow reading times associated with dissonant combinations of features encountered out of the blue. Specifically, inanimate nouns in the ergative case are challenging to process, even in environments where prejacent material could not have led the comprehender down a garden path.

In light of a thorough understanding of the Georgian case system, the precise array

of reading-time effects suggests that comprehenders have detailed but restricted access to their abstract grammatical knowledge when computing prominence relations online. In particular, case morphology compatible with indirect objecthood often seems to give rise to a ditransitive garden path. If this is the correct interpretation, it must be that more clause types than just monotransitives are predicted, or at least considered, during sentence comprehension. On the other hand, what's curiously absent is evidence for the garden-path effects one would expect if certain intransitive parses were similarly accessible. More precisely, an ergative noun seems to invariably induce the expectation for an agentive transitive continuation, even though Georgian has a sizeable class of non-agentive intransitive verbs which can take ergative subjects (Holisky 1981). This suggests that the agentive transitive subject interpretation of an ergative noun phrase eclipses all its other grammatically-consistent interpretations, lending credence to notion of canonical transitivity and its relevance to sentence processing (Hopper and Thompson 1980). By allowing us to reason about which argument structures and verb classes are under consideration during online syntactic processing, these reading-time data give us a window into the non-isomorphic relationship between the abstract grammar and its avatar subliminally accessible to the comprehender during the fractions of a second between button presses of a self-paced reading experiment.

Chapter 3 turns to relative-clause processing. Relative clauses and similar filler–gap constructions involving $\overline{\text{A}}$-dependencies are a prominent object of study in psycholinguistics. One reason why is that filler–gap dependencies can create unboundedly nonlocal relationships between arguments and their licensors. As humans' short-term memory capacity is finite, such grammatical relationships strain sentence-comprehension mecha-

nisms. Testament to this is the wealth of psycholinguistic evidence showing that filler–gap constructions with gaps in subject position are easiest to process (Gibson 1998, Kwon et al. 2010) — at least when the dependency in question places the filler before the gap and belongs to a language with neutral or Nominative–Accusative case alignment. Since this typological qualifier is necessary, it is hotly debated whether the 'Subject Gap Advantage' is attributable specifically to the greater abstract structural distance between the filler and gap in an object-gap $\bar{A}$-dependency compared to a subject-gap one, or due to some independent factor — perhaps the case morphology associated with (or prejacent to) that gap (Carreiras et al. 2010, Polinsky et al. 2012), or the linear–temporal distance between filler and gap (Gibson 1998, Lewis and Vasishth 2005).

These factors are confounds when studying languages with both (and only) filler-before-gap $\bar{A}$-dependencies and Nominative–Accusative case alignment, but in principle they can be manipulated independently from gap position in a language with either flexible relative-clause position or Split-Ergative case alignment. Georgian has both. In one pair of self-paced reading experiments, I compare how $\bar{A}$-traces in transitive subject and direct object position are processed in filler-before-gap relative clauses and gap-before-filler correlative clauses; in another pair, I compare the processing profiles of postnominal subject and object relative clauses representing each of Georgian's three primary transitive case-alignment patterns. Across all four experiments, reading-time disruptions suggest comprehenders consistently wander down subject-gap garden paths. Little evidence suggests that filler–gap processing difficulty in Georgian primarily tracks either the linear order of the dependency, or the case morphology associated with a gap. Taken at face value, this means

that the abstract structural distance between a filler and a gap is the key factor predicting reading-time difficulty across an $\bar{A}$-construction that comprises them.

And while a promising interpretation for these data relies heavily on the notion of structural distance, there are important (if hesitant) caveats and qualifiers for such a theory related to animacy. It should be emphasized that the animacy of relative-clause arguments played a role in stimuli design only when counterbalancing the lexicalization of itemsets; it was not an factor manipulated systematically across the experimental conditions. However, an exploratory post-hoc analysis suggests that the magnitude of the Subject Gap Advantage is at the very least dampened as RC-argument animacy decreases (cf. Mak et al. 2002, Gennari and MacDonald 2008, Wagers and Pendleton 2016). For the most part, these observations trend in the same direction one would expect if a filler trigger comprehenders to anticipate a gap to appear in the syntactic position which most harmonically aligns with that filler's animacy: the strongest object-gap costs are found in all-human relative clauses. All-inanimate relative clauses, though, are not obviously associated with a mirror-image subject-gap cost. Tentative though this conclusion must be, I take away from this the possibility that filler–gap dependencies are processed by a set of sentence-processing mechanisms which behave in subtly but measurably different ways than the set recruited (perhaps by default) in root clauses that lack fillers or gaps.

Shifting gears, **Chapter 4** offers a formal analysis which unites certain argument–verb φ-agreement phenomena in the South Caucasian languages. I argue that some aspects of the paradigm's organization are motivated by pressures to be expressive (easier to comprehend) and efficient (easier to produce) (Horn 1984, Kiparsky 2005). Across South Cau-

casian, nearly all noun-phrase clausal arguments are capable in principle of controlling a finite verb's φ-agreement morphology. And in some cells of the paradigm, there is a neat one-to-one mapping between agreement morphemes and agreement-controlling arguments. Here and there, though, we find conspicuous cases of morpheme blocking (a well-known property of agreement in the family; Harris 1981:31, Béjar and Rezac 2009, Nevins 2011, Blix 2020). Blocking can be said to obtain when argument A is observed to control verbal agreement morpheme α, and argument B independently controls morpheme β, yet a clause with both A and B licenses a verb bearing α but not β. At least for certain species of morpheme blocking in South Caucasian, I argue that it is no coincidence just which argument fails to trigger agreement in these cases: due to independent morphological factors, a verb lacking α would be suboptimally expressive in this context, while a verb bearing both α and β would be suboptimally efficient (Foley 2017). I suggest that at some level of representation, the morphological component of the grammar has access to and filters out these suboptimal alternative forms on grounds of economy.

Assuming morphology is fed by abstract syntax (as in Distributed Morphology; Halle and Marantz 1993), some consideration must be given to what constellation of syntactic assumptions will ensure that the morphological component have access to features of arguments not ultimately exponed by grammatical surface forms. One possibility is to adopt a theory of the syntactic operation Agree (Chomsky 2000, 2001) which allows a single probe to collect features of multiple goals — as proposed by Hiraiwa (2001), Béjar and Rezac (2009), and Deal (2015), among others. Adopting any such theory could permit the morphological computation to compare locally morphemes α and β. But such an approach,

I argue, fails to capture an important global generalization about South Caucasian verbal agreement. In particular, one locus of agreement morphology is rigid, only ever tracking the highest non-dative argument; the other two loci are promiscuous, sometimes tracking arguments which are inaccessible to the rigid locus.

This very state of affairs, I argue, is predicted by the Principle of Minimal Compliance (PMC; Richards 1997, 1998). The PMC lifts syntactic constraints once they have been obeyed once at a given point in the derivation. Drawing a parallel between more familiar PMC effects like those observed in Bulgarian multiple *wh*-questions, I propose that the rigid agreement locus in South Caucasian expones a φ-probe subject to standard locality constraints which forces it to Agree with the highest non-dative argument — i.e., either a nominative or ergative subject, or a nominative object just in case the subject is dative. The first φ-probe having minimally complied with these locality constraints, the next two are free to ignore them (I stipulate that all three probes are in the same syntactic position as the first; they are 'subprobes' of T^0 in the sense that Bulgarian C^0 has multiple *wh*-subprobes). This means that South Caucasian object agreement is in a sense parasitic on subject agreement (or, in dative-subject constructions, vice versa). A consequence of the PMC is that is occasionally results in derivational indeterminacy. Here specifically, transitive clauses will have multiple well-formed syntactic outputs: if the second and third φ-probes are insensitive to locality considerations, they can presumably target any argument freely. I suggest that outputs representing all logically possible probe–goal relations are delivered to the morphological component, and exponed according to standard principles. At the very end of the morphological computation, the expressiveness and efficiency filters apply, ensuring only

a single grammatical verb form for a given argument combination.

The rest of this chapter sketches briefly relevant aspects of Georgian grammar (Section 1.2), thereby setting the stage for questions raised in subsequent chapters, and makes explicit a few practices concerning terminology, notational conventions, and data hygeine (Section 1.3).

1.2 A snapshot of Georgian morphology & syntax

Georgian belongs to the small South Caucasian language family.[1] A thorough description of the language can be found in Aronson (1990), while Harris (1981) and Tuite (1998) give excellent overviews of various morphosyntactic phenomena. It is the official language of the Republic of Georgia, where there live a few million native speakers who read and write the Georgian script fluently. Georgian has a rich literary tradition; the language has been written continuously for some fifteen-hundred years (Boeder 2005:7). Many Georgians are bilingual or multilingual to some degree. Those who grew up during the Soviet Union often speak Russian, and the members of the younger generation often speak English. Some Georgian speakers also speak a one of a handful of minority languages, like two of the smaller South Caucasian languages, Mingrelian and Svan.

It is a loosely head-final language. There are postpositions, prenominal possessors & adjectives, and object–verb word order is very common. But verb–object order is also common; auxiliary verbs can or must precede lexical verbs; complementizers are left-

[1] South Caucasian is also known as 'Kartvelian', a term derived from the Georgian ethnic endonym *kartveli* 'Georgian [person]'. Thus, since speakers of Laz, Mingrelian, & Svan might not necessarily identify as ethnically Georgian, I find the term 'South Caucasian' to be more neutral.

peripheral (or at least preverbal); finite clausal complements follow verbs; and finite relative clauses follow the nouns they modify. There is an immediately preverbal position where foci, negative indefinites, and certain *wh*-elements appear (Borise and Polinsky 2018). Null pronouns are very common, as is scrambling of overt verbal arguments, which generally expresses information-structural properties (Skopeteas and Fanselow 2010). Georgian has productive passive, applicative, and causative constructions, which are mostly accompanied by special verbal morphology. There are no definite articles.

Morphology in Georgian is fairly complex. Verbs take forms which express features including tense–aspect–mood (TAM), argument structure, and φ-agreement with subjects & objects. Verbs are agglutinative insofar as it is often possible to segment many morphemes. But verbal inflection is highly morphomic (in the sense of Aronoff 1994), so it is often difficult to associate an individual morpheme with specific morphosyntactic features. For instance, we find two major classes of agreement affixes; one includes the prefix *v*– '1.DIR', while the other includes the prefix *m*– '1SG.OBL'. Typically former flags a first-person subject and the latter a first-person objects (1a). But in dative-subject constructions, the roles of *v*– '1.DIR' and *m*– '1SG.OBL' invert (1b).

(1) a. *me* *v*-*naxe* *bu,* *da* *bu-m* *m*-*naxa* *me*.
1SG 1-see:AOR owl.NOM and owl-ERG 1-see:AOR 1SG
'I saw the owl, and the owl saw me.'

b. *me* *unda* *m*-*enaxa* *bu,* *da* *bu-s* *unda* *v*-*enaxe* *me*.

1SG AUX 1-see:PLU owl.NOM and owl-DAT AUX 1-see:PLU 1SG

'I should have seen the owl, and the owl should have seen me.'

(Constructed examples)

Nominal morphology is a great deal simpler than verbal morphology, at least measured in terms of paradigmatic and allomorphic complexity. In the core alignment system for clausal arguments, only three case categories are used. Yet, computing just which case a given argument should bear is a more complicated task in Georgian than in the average case language. TAM, argument structure, and the verb's lexical class all play a role in case assignment. These factors yield an Accusative-like case alignment in some contexts, and an Ergative-like alignment in others; thus Georgian can be classified as a Split-Ergative language.[2] The following examples illustrate. Notice in particular that the 'nominative' case appears on subjects in some clause types (2a), and objects in others (2b, c); likewise we find 'dative' on both subjects (2c) and objects (2d).

(2) a. *q̇oveldġe* *čemi* *mezobl-eb-i* *gazet-s* *ḳitxuloben.*

every_day my neighbor-PL-NOM newspaper-DAT read:IMP.3SG

'Every day my neighbors read the newspaper.'

[2]More precisely, case alignment is sometimes Nominative–Accusative and sometimes Split-S (Harris 1990). Thus it would be more precise to call Georgian Split Split-S. And since there are in fact two distinct Split-S patterns, preciser still might be the term Split Split Split-S.

b. *gušin, Amiran-ma xuti sṭaṭia çaiḳitxa.*

yesterday Amiran-ERG five article.NOM read:AOR.3SG

'Yesterday, Amiran read five articles.'

c. *Barbare-s ḳi arc ert-i daumtavrebia.*

Barbare-DAT CONTR not_a_single-NOM finish:PERF.3SG

'Barbare, on the other hand, didn't even finish one.' (Constructed examples)

Despite the complex case system, there is little evidence of syntactic Ergativity in Georgian. Especially pertinent is a lack of extraction asymmetries (of the kind found in many Mayan languages, for example; Aissen 2017). $\bar{A}$-traces can appear in all verbal argument positions, no matter which case category is licensed there. More specifically, ergative-subject gaps are found in all of the major relative clause types. These include *wh*-relatives (3a), in which a relative pronoun occupies a left-peripheral position; *rom*-relatives (3b), which involve null-relerator movement and the middle-field complementizer *rom* 'COMP'; and *rom*-correlatives (3c), which belong to a more exotic genus of relativization structures involving left-peripheral adjunct clauses (Lipták 1996).

(3) a. *Wh*-relative

momeçona is msaxiob-i, [$_{RC}$ *romel-ma-c ṗoliṭiḳos-i*

like:AOR.1SG DIST actor-NOM which-ERG-REL politician-NOM

mibaja.]

imitate:AOR.3SG

'I liked the actor [$_{RC}$ who ___ imitated the politician.]'

b. *Rom*-relative

momeçona is msaxiob-i, [$_{RC}$ *ṗoliṭiḳos-i rom mibaja.*]

like:AOR.1SG DIST actor-NOM politician-NOM COMP imitate:AOR.3SG

'I liked the actor [$_{RC}$ that ___ imitated the politician.]'

c. *Rom*-correlative

[$_{CORR}$ *ṗoliṭiḳos-i rom mibaja,*] *momeçona is msaxiob-i.*

politician-NOM COMP imitate:AOR.3SG like:AOR.1SG DIST actor-NOM

≈'[$_{CORR}$ The one that ___ imitated the politician,] I liked that actor.'

(Constructed examples)

1.3 Clarifications and caveats

1.3.1 Terminological notes

In this thesis it will be necessary to discuss phenomena relating to argument structure and morphosyntax from multiple perspectives. To avoid ambiguity, I adopt some non-standard terms and symbols which deserve definitions.

First, an inventory of syntactic positions is necessary. I take for granted that the notion of *subject* is relevant for Georgian, defining a subject as the most structurally prominent argument of a clause. And while noun phrase's structural prominence cannot be defined in terms of its morphology (given the language's Split-Ergative case system) or its linear position (since scrambling and null pronouns are very common), there are a number of other convergent syntactic diagnostics. For example, whatever the case alignment or word order, only the subject can bind a reflexive or reciprocal anaphor in a two-argument clause. (Section 4.3.4 discusses subjecthood in Georgian in greater depth.)

(4) a. *msaxiob-eb-i ertmanet-s mibajaven.*
actor-PL-NOM each_other-DAT imitate:FUT.3PL
'The actors will imitate each other.'

b. **msaxiob-eb-s ertmanet-i mibajaven.*
actor-PL-DAT each_other-NOM imitate:FUT.3PL
Attempted: 'Each other will imitate the actors.'

c. *msaxiob-eb-s ertmanet-i ar miubajavt.*
actor-PL-DAT each_other-NOM NEG imitate:PERF.3PL
'The actors didn't imitate each other.'

d. **msaxiob-eb-i ertmanet-s ar miubajavt.*
actor-PL-NOM each_other-DAT NEG imitate:PERF.3PL
Attempted: 'Each other didn't imitate the actors.' (Constructed examples)

I use the abbreviation S to refer to any subject, whatever the argument structure of

the clause. S_{TR} and S_{IN} are used when it is necessary to distinguish transitive and intransitive subjects, respectively. (This departs from Dixon's 1994 familiar conventions, where 'S' indicates intransitive subjects specifically, and 'A' transitive subjects.) Based on their morphosyntactic properties in certain TAMs, intransitive subjects can be further divided into two classes, labeled S_{INA} and S_{INP}. Section 2.2 discusses in depth the nature of this split across intransitive subjects. I use the term *object* and its abbreviation O to refer to any noun-phrase argument of a verb other than the subject. Object can be classified more specifically as direct (O_{DO}) or indirect (O_{IO}). For Georgian, the O_{DO}–O_{IO} distinction can be defined without too much trouble in purely morphological terms. For instance, the O_{IO} of an aorist clause will always appear in the dative case; a O_{DO} in this TAM will be nominative (5). And by this metric, Georgian is fairly rich in two-argument verbs which take a subject and an indirect object, but not a direct object (5b–c).

(5) a. *mçeral-ma* ***mxaṭvar-s*** *roman-i* *miujġvna.*
writer-ERG **painter-DAT** novel-NOM dedicate:AOR.3SG≫3IO
'The writer dedicated the novel [O_{DO}] to the painter [O_{IO}].'

b. *mçeral-ma* ***mxaṭvar-s*** *aḳoca.*
writer-ERG **painter-DAT** kiss:AOR.3SG(≫IO)
'The writer kissed the painter [O_{IO}].'

c. *mçeral-i* ***mxaṭvar-s*** *daexmara.*
writer-NOM **painter-DAT** help:AOR.3SG≫3IO
'The writer helped the painter [O_{IO}].' (Constructed examples)

It should be emphasized that I will use these terms (subject, object, etc.) to refer exclusively to syntactic roles — i.e., sets of syntactic and morphosyntactic properties. Here and there, though, it will be also useful to discuss arguments' semantic roles — i.e., sets of presuppositions imposed on arguments by the lexical semantics of a licensing verb. A precise inventory of semantic roles is notoriously difficult to define (Newmeyer 2010), but for concreteness I adopt the following terminology.

(6) a. *Agent:* an event participant that acts volitionally and/or intentionally

b. *Patient:* an event participant that undergoes a change of state or which is directly affected by an agent

c. *Experiencer:* an event participant that experiences a psychological state

d. *Affectee:* an entity that an event positively or negatively affects; affectees include possessors another event participant

e. *Instrument:* an event participant used or manipulated by an agent in order to perform an action

f. *Theme:* any other event participant

Finally, I will use capitalization to distinguish two different uses of case terminology. Capitalized case names (Nominative, Accusative, Absolutive, Ergative, etc.) are used to talk about abstract morphosyntactic alignment: that is to say, the distribution of a particular morphological or syntactic phenomenon across different syntactic roles. In contrast, all-lowercase terms will refer to specific morphological case categories. In many languages, and in Georgian especially, the term used for a given case morpheme may not transparently

convey its syntactic distribution. For example, the nominative case affix *–i* 'NOM' in Georgian has a Nominative function in some contexts (marking S_{IN} or S_{TR}), and an Absolutive one in others (marking S_{IN} or O_{DO}). Georgian's ergative affix *–ma* 'ERG', on the other hand, cannot strictly speaking be said to have Ergative function at all, since there are ergative-marked intransitive subjects.

1.3.2 Glossing and transliteration conventions

Interlinears in this thesis largely follow the Leipzig glossing recommendations, but I adopt a few idiosyncratic practices. Hierarchical argument relationships are expressed with the much-greater-than symbol; for example, 1≫2 indicates a structure with a first-person subject and a second-person object, or a morpheme triggered by such a structure. Many glosses will make empty categories explicit: '___' indicates a gap (i.e., $\bar{A}$-trace); '◌' is a null pronoun; '∅' is a paradigmatically-justified null morpheme. For languages like Georgian with relatively free word order, I put these empty categories in linear positions that accord with dominant word order facts. For clarity, many agreement dependencies will be color-coded (especially in Chapter 4); subjects and the agreement morphemes they control will be shaded gray : objects and object-agreement morphemes will be in a white box . Finally, I use singular *they* (always flagged [SG]) to gloss third-singular arguments in contexts where their referents' gender or animacy is ambiguous.

Words in the South Caucasian languages can be very morphologically complex, but unless necessary I refrain from decomposing them fully. The convention 'μ:XYZ' signals that a morphological object μ can be decomposed into subparts which in concert express fea-

tures XYZ. Finite verbs, for example, will usually be glossed 'lemma:TAM.AGR'. So glosses like (7a) are shorthands for more detailed ones like (7b)

(7) a. *im maġal msaxiobebs,* [$_{RC}$ *romlebic dagvelaṗaraḳebodnen…*]

DEM tall actors:DAT which:NOM speak_to:COND.3PL≫1PL

'those tall actors who would speak to us…'

b. *im maġal-∅ msaxiob-eb-s,* [$_{RC}$ *roml-eb-i-c* ——

DIST.DAT tall-DAT actor-PL-DAT which-PL-NOM-REL GAP.NOM

◌ *da-gv-e-laṗaraḳ-eb-od- nen …*]

1PL.DAT PVB-1PL.OBL-APPL.INTR-speak-TH-IMP.INTR- PAST.INTR.3PL

'those tall actors who would speak to us…' (Constructed example)

Table 1.1 shows the transliteration system I use to present data from Georgian, Mingrelian, and Svan. (For Laz data, I retain the Roman orthographies used in primary sources; these are based on the Turkish alphabet.)

1.3.3 Data transparency

Transparency in data and acceptability judgement collection — whether conducted in an experimental or fieldwork setting — is of upmost importance. Thus, I provide sources for every non-English data point in this thesis. Often this means copying an interlinear gloss wholesale from other published sources. But whenever I need to adapt other authors' glossing styles to my own, I indicate that I have done so with the annotation 'glosses adapted from `[source]`'.

Grapheme	Transliteration	IPA value	Grapheme	Transliteration	IPA value
ა	*a*	/a/	ს	*s*	/s/
ა̄	*ā*	/aː/	ტ	*ṭ*	/t’/
ა̈	*ä*	/æ/	უ	*u*	/u/
ა̈̄	*ā̈*	/æː/	უ̄	*ū*	/uː/
ბ	*b*	/b/	უ̂	*w*	/w/
გ	*g*	/ɡ/	ქ	*ə*	/ə/
დ	*d*	/d/	ქ̄	*ə̄*	/əː/
ე	*e*	/e/	ფ	*p*	/pʰ/
ე̄	*ē*	/eː/	ქ	*k*	/kʰ/
ვ	*v*	/v/	ღ	*ġ*	/ʁ/
ზ	*z*	/z/	ყ	*q̇*	/q’/
თ	*t*	/tʰ/	ჱ	*ʔ*	/ʔ/
ი	*i*	/i/	შ	*š*	/ʃ/
კ	*ḳ*	/k’/	ჩ	*č*	/tʃʰ/
ლ	*l*	/l/	ც	*c*	/tsʰ/
მ	*m*	/m/	ძ	*j*	/dz/
ნ	*n*	/n/	წ	*ç*	/ts’/
ო	*o*	/o/	ჭ	*č̣*	/tʃ’/
ო̄	*ō*	/oː/	ხ	*x*	/χ/
პ	*ṗ*	/p’/	ჴ	*q*	/qʰ/
ჟ	*ž*	/ʒ/	ჯ	*ǰ*	/dʒ/
რ	*r*	/ɾ/	ჰ	*h*	/h/

Table 1.1: Transliteration scheme for Georgian, Mingrelian, and Svan

Some examples have been constructed by piecing together different parts of a grammatical description. I indicate this with the annotation 'after `[source]`'. For example, the following Svan data point was constructed by consulting (i) an agreement paradigm that lists only affixes (Topuria 1967:23); and (ii) a description of verbal TAM inflection that provides only a few forms for a few verbs, including the lexeme I give (Topuria 1967:73). I cannot be certain that this is indeed the grammatical form for this particular verb (unpredictable allomorphy or paradigm gaps are always live possibilities), but it reflects my best understanding of Topuria's (1967) grammatical description.

(8) *ǰ-amār-äs.*

2.OBL-prepare-IMP.1/2

'I was preparing you' (Upper Bal Svan, after Topuria 1967:23, 73)

Georgian data which are not directly taken or triangulated from a primary source (usually Aronson 1990) come from two places. Some were gathered from fieldwork I conducted in Tbilisi, or via remote video conferencing. A 'fieldnotes' tag will include the initials of the consultant I elicited the data point from, and the date of elicitation. Here and there I modify the original collected data point slightly, either for brevity or lexical clarity, but I indicate when I have done so. Other Georgian data points are annotated as 'constructed examples'. This signals a data point that I have constructed myself, without having a native speaker to verify its acceptability. While Georgian has been the primary locus of my research for about a decade, and I studied it in a formal classroom setting for several years, I do not consider myself even a terribly proficient speaker of the language. So I stress that

all of these data points should be approached with a healthy degree of scepticism — and in general, the more complex a constructed example, the more grains of salt it deserves. However, I have tried to resort to constructed examples only when I need to illustrate a very basic morphological or syntactic phenomenon, especially when minimal pairs are desired. Where space permits it, I cite dictionary entries from Rayfield (2006) for the lexical items I use in constructed examples.

As for the stimuli used in my self-paced reading experiments, all were constructed in consultation with at least one, and usually more, native Georgian speaker(s). My methods for doing so are the following. After devising a stimulus template (i.e., a sentence-shape which distributes pertinent morphosyntactic cues in a carefully designed way), I compile a list of lexemes with desired properties by searching Rayfield (2006). Then, Mad-Libs style, I construct a list of candidate itemsets. To the best of my ability, I try to choose combinations of words which are anodyne, plausible, and do not bias comprehenders towards particular syntactic-role assignments. For example, I avoid clauses meaning things like *the teacher scolded the student*, favoring ones like *the dentist scolded the sculptor*. This is important because all of my experiments rely on argument reversibility, and often arguments' roles are temporarily ambiguous in Georgian — so if one condition means *the teacher scolded the student*, another will mean *the student scolded the teacher*. Participants' real-world-knowledge will likely bias them away from initially parsing 'student' as a subject when reading the latter clause, and this would be an undesirable confound for an experiment probing comprehenders' general parsing strategies.

Having prepared candidate itemsets, I elicit acceptability judgements from my

consultants, emphasizing that the sentences must be grammatical, and they should be natural & idiomatic to the greatest degree possible. Logistically speaking, consultants usually only ever see one experimental condition of each itemset — not an optimal practice, but one which reduces tedium — though which condition is chosen for each itemset is randomized, and when more than one consultant is involved in stimulus vetting, each usually sees a different condition. The primary purpose of this exercise is to avoid words or combinations of words which are unusual (implausible, rare, dialectal, archaic, stilted, poetic, obscene, comical, etc.). Of course, certain psycholinguistic questions — no matter the language being studied — can only be addressed by looking at sentence shapes which are to some degree peculiar or artificial. But to the extent that I am aware of such unavoidable awkwardnesses, I point them out throughout the thesis.

This stimulus-creation process, I have found, is an illuminating method of field-work in its own right. Eliciting acceptability judgements for dozens of lexicalizations of the same sentence shape often reveals patterns which may not arise when conducting fieldwork with purposes other than designing an experiment. And indeed, my first-draft stimulus template inevitably turns out to need several rounds of revision. In the future, I would like to hand over more of the stimulus-creation process to native Georgian speakers — and, ideally, also run formal acceptability studies to norm stimuli. In the mean time, though, I can only maximize transparency. So to that end, I note that all of my stimuli — along with the experimental data and `R`-scripts I used to process and analyze them — are freely accessible on my professional website.

Chapter 2

Harmonizing animacy & Split Ergativity

2.1 Case alignment and incremental processing

It's been long observed that case morphology must serve an important role in incremental sentence comprehension, helping signal a noun phrase's syntactic and/or semantic role (Sapir 1917, Kibrik 1985, Comrie 1989, Dixon 1994, and many others). This is especially true in verb-final languages, where comprehenders typically encounter arguments before knowing the argument structure and lexical-semantic properties of the verb — a fact reflected in Greenberg's Universal 41, for instance, which strongly correlates verb-finality and case (Greenberg 1963:75).

Best for facilitating comprehension in a verb-final language would be a case alignment system with a one-to-one mapping between syntactic roles and morphological case categories. Assume for the purposes of illustration that the only syntactically distinct roles are intransitive subjects (symbolized 'S_{IN}'), transitive subjects (S_{TR}), and direct objects (O_{DO}).

A canonical Tripartite system (9), then, is the most optimal case-alignment strategy, at least from the comprehender's perspective. Encountering a noun phrase in any of the three case categories in this language would unambiguously signal both the clause's argument structure, and also where the noun phrase fits into that argument structure.[1]

(9) Tripartite case alignment

$$\underbrace{S_{TR}}_{ERG}\ \underbrace{S_{IN}}_{INTR}\ \underbrace{O_{DO}}_{ACC}$$

Slightly suboptimal are Nominative–Accusative systems (10). An Accusative-marked argument can be immediately and unambiguously mapped to a unique syntactic role and argument structure. In contrast, a Nominative argument must be a subject, but it cannot independently diagnose a clause's transitivity.

(10) Nominative–Accusative case alignment

$$\underbrace{S_{TR}\ S_{IN}}_{NOM}\ \underbrace{O_{DO}}_{ACC}$$

Less optimal still is Ergative–Absolutive case alignment (11). Like Accusative case, Ergative case can be linked with certainty to a specific role (S_{TR}). But Absolutive case is sometimes found on subjects (intransitive ones, specifically) and sometimes on direct objects. So if a clause's linearly first argument is Absolutive, the comprehender can be certain neither of that argument's broad syntactic role nor the clause's transitivity.

[1] This takes for granted, of course, that Absolutive subjects and objects do not necessarily form a syntactic natural class; i.e., that morphological Ergativity does not entail syntactic Ergativity.

(11) Ergative–Absolutive case alignment

$$\underbrace{S_{TR}}_{ERG} \quad \underbrace{S_{IN} \quad O_{DO}}_{ABS}$$

Of course, natural languages are replete with temporary (and even global) role-ambiguities akin to those arising from Ergative case alignment. Yet in general they pose little challenge to the parser. After all, many properties other than case morphology are strongly correlated with a noun phrase's syntactic role — including animacy, linear order, definiteness, information structural properties, person, and referential type (Hopper and Thompson 1980, Bornkessel-Schlesewsky and Schlesewsky 2009). A noun phrase is more likely to be an S_{TR} than an O_{DO} if it is highly animate, linearly prominent[2], definite, topical, first- or second-person, and/or pronominal. Conversely, a low-animacy noun, or one which is linearly non-prominent, indefinite, non-topical, third-person, and/or non-pronominal is more likely to be an O_{DO} than an S_{TR}. These correlations seem to be very robust across languages, and they motivate *prominence scales* like the following.

(12) Prominence Scales (where $x \succ y$ means 'x is more prominent than y')

a. $S_{TR} \succ O_{IO} \succ O_{DO}$ *Syntactic Role*

b. Human $\succ$ Animal $\succ$ Inanimate *Animacy*

c. Linearly prominent $\succ$ Linearly non-prominent *Linear position*

d. Definite $\succ$ Specific indefinite $\succ$ Nonspecific indefinite *Definiteness*

e. Topical $\succ$ Non-topical *Information structure*

[2]One might imagine that the linear prominence dimension covaries with dominant word order. In SOV languages, perhaps the clause's earliest argument is its most prominent; in VOS languages, perhaps the latest argument is.

f. 1st person ≻ 2nd person ≻ 3rd person *Person*

g. Pronoun ≻ Proper name ≻ Common noun *Referential type*

h. Agent ≻ Experiencer ≻ Patient *Thematic role*

Prominence scales are particularly common in formal linguistics, where they can help capture many generalizations about morphosyntax and the lexicon (e.g., Silverstein 1976, Dowty 1991, Aissen 2001, 2003). That's because languages often to employ special morphology or syntactic structures to mark (or downright avoid) clauses with misaligned scales — i.e., clauses whose arguments are highly prominent on one scale but non-prominent on another. Conversely, structures with harmonically aligned scales are often unmarked or in some sense default (Hopper and Thompson 1980). Differential object marking is a classic illustration of this observation. In many varieties of Spanish, for instance, O_{DO}s that are both animate and specific — i.e., O_{DO}s that are highly prominent on two scales — are flagged with the differential object marker *a* 'DOM' (13a). All O_{DO}s less prominent than that are unmarked (13b).

(13) Differential object marking in Spanish

a. *No veo *(a) María.*

NEG see:PRES.1SG *(DOM) María

'I don't see María.'

b. *No veo {el, *al} problema.*

NEG see:PRES.1SG {DEF, *DEF.DOM} problem

'I don't see the problem.' (glosses adapted from Zagona 2002:140)

But prominence scales are not only relevant to the grammar. Much experimental evidence suggests that harmonically-aligned structures — or at least ones where syntactic roles and animacy harmonize — are easier to process in real time and more likely to be predicted than misaligned ones. For example, electrophysiological studies often find that inanimate S_{TR}s induce an N400, an ERP associated with lexical–semantic processing difficulty (Weckerly and Kutas 1999, Roehm et al. 2004, Philipp et al. 2008), and filled-gap effects in reading experiments show that animate relative-clause head nouns give rise to much stronger subject-gap expectations than inanimate ones do (Mak et al. 2002, Traxler et al. 2005, Gennari and MacDonald 2008, Wagers and Pendleton 2016).

An especially articulated theory of sentence processing that features prominence scales is the extended Argument Dependency Model (eADM; Bornkessel and Schlesewsky 2006, Born-kessel-Schlesewsky and Schlesewsky 2009), whose architecture is schematized in Figure 2.1.

Especially relevant are the *Compute Prominence* and *Compute Linking* stages. In the former, noun phrases are assigned semantic proto-roles (in the sense of Dowty 1991 and Primus 1999) based on their prominence features. In the latter, arguments are linked to the thematic roles in the lexical entry of the governing verb via harmonic alignment. The easiest transitive structures to process are those where arguments' prominence features and the predicate's thematic roles are harmonically aligned and maximally distinct (i.e., at the extremes of each scale). The more a structure deviates from this ideal — the more dissonantly aligned the prominence scales are — the larger the processing cost.

Bornkessel-Schlesewsky and Schlesewsky (2008a, 2008b, 2009) propose that the

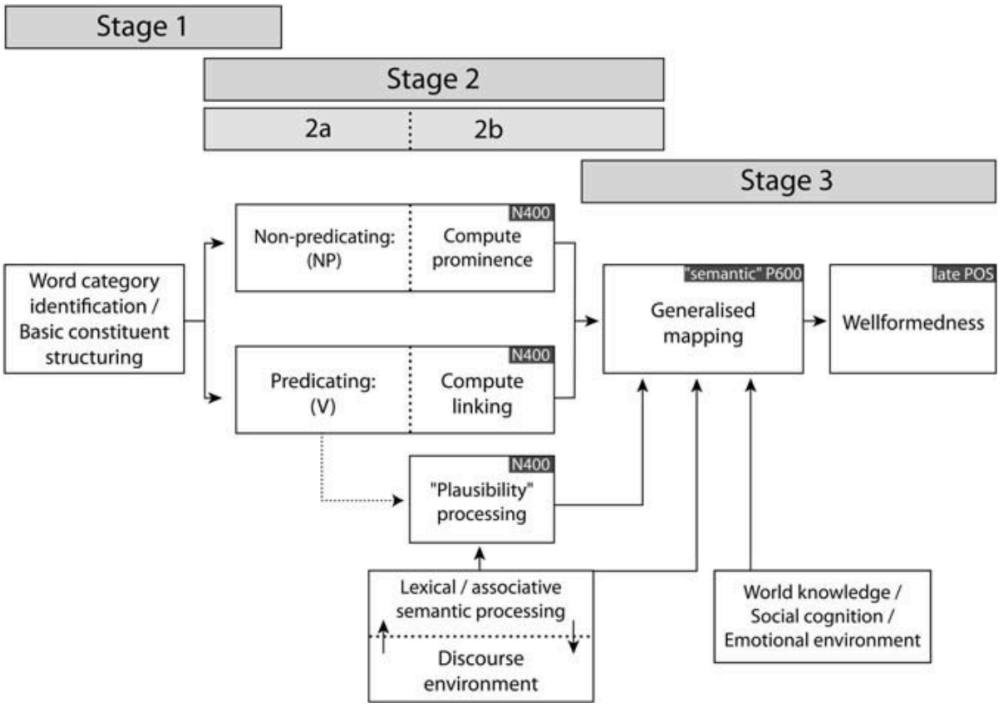


Figure 2.1: Architecture of eADM (Bornkessel-Schlesewsky and Schlesewsky 2009:42)

harmonic alignment of transitive arguments helps satisfy a *Distinctness* principle. Maximally-prominent transitive subjects and minimally-prominent direct objects are as distinct from one another as two arguments can be, since they have as few overlapping syntactic–semantic properties as possible. The less arguments overlap along prominence dimensions, the less likely it is that wires will be crossed during comprehension, with (e.g.) the true actor misparsed into object position or the patient into subject position. An important corollary of Distinctness, one which I return to in Section 2.5, is that intransitive subjects are vacuously distinct. Thus the best strategy for avoiding misparses is to prioritize intransitive parses.

A language's grammar can influence how prominence scales (12) are weighted

in prominence calculations. For example, linear prominence will be highly weighted in a language like English, where word order is fairly rigid; person prominence will be especially important in a language like Lummi (Salish), where a person hierarchy constrains what argument combinations are possible in active transitive clauses (Jelinek and Demers 1983). Moreover, Compute Prominence might operate differently across constructions within a single language.

To date, the predictions of eADM can only be tested with data from a handful of languages — not terribly many more than those surveyed by Bornkessel-Schlesewsky and Schlesewsky 2009: English, Dutch, German, Chinese, Tamil, and Hindi. But to better understand the sentence processing mechanisms which rely on harmonic scale alignment, and to what extent arbitrary grammatical variation influences them, it is of course necessary to cast a wider net. Especially informative will be data from languages which have verb-final word order and case-alignment systems which do not optimally facilitate comprehension. Nearly every verb, of course, comes with presuppositions and implicatures that often tightly constrain what kinds of nouns it can plausibly take as arguments. A verb-final clause forestalls comprehenders' access to this lexical-semantic content; thus they must then either hold off on assigning grammatical roles to noun phrases until encountering the verb — a strategy which would be hard to reconcile with the wealth of psycholinguistic evidence indicating rapid, active, incremental parsing — or they must take a leap of faith guided more general parsing heuristics. As for case alignment, the more a language's morphosyntactic alignment deviates from the ideal one-to-one case–role mapping described above (9), the less reliable its case cues will be for associating arguments with specific syntactic roles. Thus case cues

will be less instrumental during the comprehension of an Ergative language, for example.

Georgian checks both boxes. It is mostly verb-final language with null pronouns and very flexible argument order. And, as I describe in the next section (Section 2.2), it has a Split-Ergative case system which is (from the perspective of the comprehender) apparently maladaptive, giving rise to many incremental ambiguities. This chapter presents data from two self-paced reading experiments which capitalize on these properties. They test how animacy and case cues interaction when parsing a Georgian sentence by manipulating clausal case alignment and word order. Experiment 1a tests parsing strategies triggered by human (i.e., high-animacy) arguments; Experiment 1b contrasts minimally, investigating inanimate (low-animacy) arguments. (These are discussed in 2.3 and 2.4, respectively.) If Georgian comprehenders rely on animacy during real-time parsing, then the two experiments should have markedly different results. Human nouns should be preferentially mapped to the S_{TR} role, and inanimate nouns to the O_{DO} role; cues that are incompatible with those associations should result in processing difficulty.

And this is indeed what we find. Reading times are disrupted at regions which force a dissonant animacy–role alignment is impossible, or which necessitate revising a previous harmonic parse — just as predicted by eADM. Indeed, the results are compatible with a theory in which comprehenders bend over backwards to avoid misaligned scales, going so far as to posit ditransitive structures so as to avoid human O_{DO}s and inanimate S_{TR}s. This is a somewhat surprising result, but one which follows from a very simple scale-alignment algorithm, detailed in Section 2.5. That concluding section also identifies a few ways in which the data are inconsistent with eADM, and discusses the inconsistencies' theoretical

consequences.

2.2 Case alignment in Georgian

Georgian is often described as a split ergative language, where Split Ergativity is a morphosyntactic alignment system in which certain structures conform to a Nom–Acc pattern and others to an Erg–Abs one. And indeed in Georgian we do find different case alignment patterns across different structures — specifically, across finite clauses with different tense–aspect–mood features. In TAMs including the future, we observe Nom–Acc alignment (14); in the aorist (perfective past), we see Erg–Abs alignment (15). (The only wrinkle so far is how the case morphemes are glossed: the 'nominative' suffix *–i* 'NOM' subsumes Nominative and Absolutive roles; the 'dative' suffix *–s* 'DAT' has an Accusative function.)

(14) a. *santel-i* *ainteba.*
candle-NOM light.INTR:FUT.3SG
'The candle will be lit.'

b. *mġvdel-i* *santel-s* *aantebs.*
priest-NOM candle-DAT light.TR:FUT.3SG
'The priest will light the candle.' (Constructed examples)

(15) a. *santel-i* *ainto.*

candle-NOM light.INTR:AOR.3SG

'The candle was lit.'

b. *mġvdel-ma* *santel-i* *aanto.*

priest-ERG candle-NOM light.TR:AOR.3SG

'The priest lit the candle.' (Constructed examples)

Deeper investigation, though, soon reveals that both the 'split' and 'ergativity' components of Georgian Split Ergativity are more complicated than these data points suggest. First, the Ergative–Absolutive alignment we see in the aorist (15) is more precisely 'Split-S' or 'Active' alignment (Harris 1990), since nominative is not the only case we find on S_{IN}s. A large class of intransitive subjects are assigned nominative in the future, but ergative in the aorist (16).

(16) a. *mġvdel-i* *daisvenebs.*

priest-NOM rest:FUT.3SG

'The priest will rest.'

b. *mġvdel-ma* *daisvena.*

priest-ERG rest:AOR.3SG

'The priest rested.' (Constructed examples)

Let's use the term 'S_{INP}' for a S_{IN} which appears in the nominative case in all finite clauses (like *santel-i* 'candle-NOM' in example 15a), and 'S_{INA}' for one which does not (like

mġvdel-ma 'priest-ERG' in 16b). It may be tempting to label S_{INP}s and S_{INA}s unaccusative and unergative subjects, respectively. And indeed many S_{INP}s in Georgian do indeed seem to be unaccusative subjects (i.e., internal arguments; complements of the verb), and many S_{INA}s seem to be unergative subjects (i.e., external arguments; specifiers of the verb). But I hesitate to equate the case properties of an intransitive subjects with argument structure, as I am not aware of many trustworthy diagnostics for unergativity (i.e., external-argumenthood) in Georgian that are independent from case assignment.[3]

An intransitive verb's Aktionsart and/or telicity, on the other hand, do seem to be good predictors of the case properties of its subject (Holisky 1981). Intransitive achievements and accomplishments, especially ones assigning the semantic role patient, belong to the S_{INP} class. (Many intransitive stative predicates and verbs of motion or posture do too.) The S_{INA} class comprises mostly activities. Many these are events and behaviors canonically associated with humans or animals, but there is also a remarkably sizable class of non-agentive atelic S_{INA} verbs, often describing meteorological phenomena or the emission of light or sound. A few examples follow.

(17) a. *Levan-ma itamada supra-ze.*
Levan-ERG act_as_tamada:AOR.3SG feast-on
'Levan was the tamada [i.e., performed toastmaster duties] at the feast.'

[3] A few unaccusativity (internal-argumenthood) diagnostics have been identified, though: Harris (1982).

b. *arçiv-ma iq̇ašq̇aša.*

eagle-ERG screech:AOR.3SG

'The eagle screeched.'

c. *šen-ma tval-eb-ma ibrčq̇viala.*

your-ERG eye-PL-ERG sparkle:AOR.3SG

'Your eyes sparkled.'

d. *ca-m daigriala.*

sky-ERG thunder:AOR.3SG

'The sky thundered down; rattled with thunder.' (Constructed examples)

An additional source of complexity comes from the factors that condition case-alignment splits in Georgian. TAM has already been mentioned. Most verbs in Georgian have distinct morphological forms for about ten distinct TAM categories. Each is associated with one of three case alignment patterns. We've seen NOM–ACC alignment in the future, but this pattern also obtains in the present, imperfective past, conditional, and two irrealis TAMS. As for the ERG–ABS pattern, it is found in the aorist and also the optative, another irrealis category. In the remaining TAMS — the perfect (often used as a non-witnessed past evidential) and the pluperfect (often used in counterfactual irrealis clauses) — we see a third pattern. S_{TR}s and S_{INA}s are assigned dative; O_{DO}s and S_{INP}s, nominative (18). I'll refer to this alignment mnemonically as the Dative–Absolutive pattern, though technically speaking it is simply a second Split-S system.

(18) a. *mġvdel-s dausvenebia.*

priest-DAT rest:PERF.3SG

'The priest [apparently] rested.'

b. *mġvdel-s santel-i auntia.*

priest-DAT candle-NOM light.TR:PERF.3SG

'The priest [apparently] lit the candle.'

c. *santel-i antebula.*

candle-NOM light.INTR:PERF.3SG

'The candle [apparently] was lit.' (Constructed examples)

It's worth noting here that the indirect objects are assigned dative case in nearly all finite environments. The only exception is that O_{IO}s appear as PP chômeurs in clauses whose subject must be independently assigned dative (19c).

(19) a. *ekim-i bavšv-s ṣign-s miscems.*

doctor-NOM child-DAT book-DAT give:FUT.3SG≫3

'The doctor will give the child a book.'

b. *ekim-ma bavšv-s ṣign-i misca.*

doctor-ERG child-DAT book-NOM give:FUT.3SG≫3

'The doctor gave the child a book.'

c. *ekim-s bavšv-is=tvis ċign-i miucia.*

doctor-DAT child-GEN=for book-NOM give:PERF.3SG≫3

'The doctor [apparently] gave the child a book.' (Constructed examples)

There are two other verb classes in Georgian with notable case assignment properties: 'applied unaccusatives' and experiencer verbs, illustrated below. I set detailed discussion of their behavior aside until Section 4.3.4, but it suffices to say that they take one dative argument and usually also a nominative one; TAM does not change this case assignment pattern.

(20) a. *q̇vav-s baxala mouḳvda.*

crow-DAT chick-NOM die.APPL:AOR.3SG≫3IO

'The chick died on the crow; the crow's chick died on it.'

(Constructed example, after Rayfield 2006:907, 978)

b. *momġeral-s moceḳvave šeuq̇varda.*

singer-DAT dancer-NOM love.INCH:AOR.3SG≫3IO

'The singer fell in love with the dancer.'

(Constructed example, after Rayfield 2006:961, 993, 1400)

c. *sṭomaṭolog-s eceḳveboda.*

dentist-DAT dance.DESID:IMP.3SG

'The dentist felt like dancing.'

(Constructed example, after Rayfield 2006:661, 1172)

Table 2.1 summarizes the core of Georgian's case alignment system. Recall that

	S_{TR}	S_{INA}	S_{INP}	O_{DO}	O_{IO}	Case Alignment
Future…	NOM			DAT		NOM–ACC
Aorist…	ERG		NOM		DAT	Split-S ('ERG–ABS')
Perfect…	DAT		NOM		PP_{for}	Split-S ('DAT–ABS')

Table 2.1: Georgian case alignment across TAM categories

a subject (symbolized 'S') are found in transitive and intransitive clauses (call these 'S_{TR}' and 'S_{IN}', respectively), and that there are two subtypes of intransitive subjects ('S_{INA}' and 'S_{INP}'). An object ('O') might be direct ('O_{DO}') or indirect ('O_{IO}').

What consequences does this specific case alignment system have for incremental sentence processing? Georgian is a mostly SOV language, so comprehenders very frequently encounter arguments before their licensing verbs. Verb-final structures definitionally delay access to the verb's lexical semantics; without that — and in the absence of a sufficiently rich context — an active comprehender can only make predictions about arguments' syntactic relationships as the sentence unfolds. For an active comprehender of Georgian, this task has some unusual twists. Its sui generis case system in particular must guide the comprehender in unique ways. Some case morphology will be very helpful to encounter. An ergative noun phrase must occur in an aorist or optative clause, it must be a subject, and there's a good chance it's a transitive subject more specifically. Moreover, any noun phrase that follows an ergative argument would also have an unambiguous syntactic role: ergative arguments only cooccur with nominative arguments that are O_{DO}s, and

dative ones that are O_{IO}s. But prejacent to an ergative argument, and absent any other reliable TAM cues, nominative- and dative-marked noun phrases are highly ambiguous. Dative especially: in Georgian we can find dative-marked transitive subjects, intransitive subjects, and indirect objects in all TAM categories, and dative-marked direct objects in six out of ten. And given that Georgian permits scrambling fairly freely, and argument drop very liberally, the linear position of nominative or dative argument does not seem especially useful for dispelling this ambiguity (though cf. Skopeteas et al. 2012).

One may be tempted to call this case system a maladaptive one from the perspective of sentence comprehension. It has, though, very much withstood the test of time: Svan, Georgian's most distantly related sibling language, has a nearly identical case system, so it likely existed in Proto-South Caucasian too (Harris 1985). Clearly, then, Georgian Split Ergativity does not pose so insurmountable a hurdle to acquirers and comprehenders so as to drive substantive grammatical change. But it's far from obvious how the parser navigates a case cues with such unevenly distributed informativity, or how they the leverage complex grammatical dependencies linking case morphology, argument structure, and TAM features.

This chapter presents evidence from two self-paced reading experiments which suggest that comprehenders of Georgian process the temporary role ambiguities which occur very frequently in in simple root clauses in a simple and logical way which conform very closely to the predictions of eADM. Experiment 1a manipulates the word order and case alignment of transitive clauses with two human arguments; Experiment 1b does the same for transitive clauses with two inanimate arguments. Patterns of incremental reading difficulty (viz., slowed reading times) are compatible with a theory in which the comprehen-

der prioritizes parses that incrementally maximize harmonic alignment of syntactic role and animacy. A human noun phrase is assigned the most prominent unclaimed role compatible with its case morphology: best would be S_{TR}, but O_{IO} is preferable to O_{DO} if a subject parse is unavailable. An inanimate noun phrase, on the other hand, is assigned the least prominent possible role: preferably O_{DO}, potentially O_{IO}, and S_{TR} only as a last resort.

For reasons already belabored, this strategy is fairly logical and rather trustworthy given the sorts of events humans typically talk about. Occasionally, though, it leads the comprehender astray. And most of the reading-time disruptions reported below can be plausibly cast in just this way, as garden-path effects (Frazier 1978, 1987): processing costs associated with revising a previous parse in light of grammatically incompatible cues. Some evidence, though, is best explained by an independent cost associated with integrating certain dissonant role and animacy features. Specifically, an inanimate noun which cannot be parsed as anything other than an S_{TR} will be reliably difficult to process. (Unambiguous human O_{DO}s do not seem similarly vexing, though.)

These result conform remarkably well to many of eADM's predictions. Less clear is how compatible the findings are with the theory's Distinctiveness principle, which predicts that parses with fewer arguments are prioritized over ones with more arguments. Ergative-marked intransitive subjects — even inanimate ones (17c–d) — are not uncommon in the language, yet ergative morphology in these experiments seems to reliably condition a *transitive*-subject interpretation. If intransitive parses are always optimal, and if both S_{TR}- and S_{INA}-parses are equally accessible, then this should not be the case.

2.3 Experiment 1a: Humans in Transitive Clauses

The goal of Experiment 1a was to test how Georgian comprehenders parse highly-animate arguments in a variety of cases and linear positions. If they assign human noun phrase to the S_{TR} role whenever possible — as a theory employing incremental harmonic alignment predicts — then we expect to cues which eliminate an SOV parse to cause processing difficulty. In this study, such cues include non-initial ergative morphology, and future-tense verbal morphology in $O_{DAT}S_{NOM}V$ sentences. Both cues indeed condition reading-time disruptions. But verbs in $S_{NOM}O_{DAT}V$ clauses do too — suggesting that parsers entertain not just monotransitive but also ditransitive continuations when evaluating parses along a prominence harmony metric.

2.3.1 Method

Participants

58 native Georgian speakers living across Georgia (37 women, average age = 30) were recruited for this experiment. They were paid 15 GEL (approximately 5.25 USD) for participating. 19 participants were ultimately excluded analysis because they answered less than 70% of filler comprehension questions correctly.

Materials

28 item sets were constructed in a 2×2 design, crossing TAM / S_{TR}–O_{DO} case frame (FUT / NOM–DAT or AOR / ERG–NOM) and word order (SOV or OSV). These sentences conformed to the template in (21). All verbal arguments were nouns referring to human entities.

(21) Stimulus template (Experiment 1a: Humans)

Adv	NP_1	Adj	NP_2	V	$Spill_1$	$Spill_2$.
w_1	w_2	w_3	w_4	w_5	w_6	w_7

Critical regions were the arguments (NP_1 at w_2, and NP_2 at w_4) and the verb (w_5). Recall that arguments in the nominative or dative cases are generally role-ambiguous before encountering the verb's TAM morphology. Arguments marked ergative, on the other hand, are unambiguously S_{TR}s, so encountering one entails that a nominative coargument is the clause's O_{DO}. (Ergative would also makes an goal parse obligatory for a dative coargument; this experiment, however, did not feature ergatives and datives in the same clause.)

A sentence-initial adverbial adjunct phrase (either temporal or locative) ensured that the first critical region was not the first one that the participants read. In seven of the itemsets, the FUT and AOR conditions featured different adverbs (e.g., *xval* 'tomorrow' vs. *gušin* 'yesterday'). It is of course not ideal to vary non-critical lexical material across different conditions of the same itemset, but the alternating pairs were chosen to be semantically parallel and closely matched in orthographic length. The adjective at w_3 served as a spillover region for NP_1. Only adjectives with vowel-final stems were chosen, as these do not to participate in case concord with the noun they modify (unlike adjectives with consonant-final stems). The lack of case concord there meant that the second argument's case cues were localized to a single SPR region, just as the first's were.

An example itemset follows.

(22) a. NOM–DAT, SOV

ianvar-ši pexburtel-i ġru dedopal-s gaicnobs iṭaliur oṗera-ši.

January-in footballer-NOM deaf queen-DAT meet:FUT.3SG Italian opera-in

'In January, the soccer player will meet the deaf queen at the Italian opera.'

b. NOM–DAT, OSV

ianvar-ši pexburtel-s ġru dedopal-i gaicnobs iṭaliur oṗera-ši.

January-in footballer-DAT deaf queen-NOM meet:FUT.3SG Italian opera-in

'In January, the deaf queen will meet the soccer player at the Italian opera.'

c. ERG–NOM, SOV

ianvar-ši pexburtel-ma ġru dedopal-i gaicno iṭaliur oṗera-ši.

January-in footballer-ERG deaf queen-NOM meet:AOR.3SG Italian opera-in

'In January, the soccer player met the deaf queen at the Italian opera.'

d. ERG–NOM, OSV

ianvar-ši pexburtel-i ġru dedopal-ma gaicno iṭaliur oṗera-ši.

January-in footballer-NOM deaf queen-ERG meet:AOR.3SG Italian opera-in

'In January, the deaf queen met the soccer player at the Italian opera.'

These experimental items were presented among 80 filler sentences, which comprised 28 experimental items for the Inanimate Experiment (see Section 2.4), and 52 more sentences with varied syntactic properties. Each of the 108 sentences was followed by a yes–no comprehension question. All of the stimuli were constructed in consultation with two native speakers.

Procedure

Subjects participated online via Ibex Farm (Drummond 2016). Upon accessing the experiment, participants read a brief introduction describing the general purpose of the task, filled in demographic information, and consented to participation. To familiarize them with the self-paced reading task and experimental procedure, participants were presented with three practice items consisting of a sentence and a comprehension question. After this, the experiment proper began. Experimental items were distributed to participants using a Latin Square, and randomized along with the fillers. Feedback was provided after each comprehension question. After finishing all 108 sentence–question pairs, an optional debriefing question appeared.

Analysis

Reading times and comprehension question response latency were analyzed using linear mixed-effects regression; question accuracy was analyzed using logistic mixed-effects regression. The case frame conditions were coded by two coefficients using centered sum contrasts: NOM–DAT (−½) and ERG–NOM (+½). Likewise for the word order conditions: SOV (+½), OSV (−½). Unless otherwise stated, maximal random effects structure was included (Barr et al. 2013). Models were estimated using the `lme4` package in `R` (Bates et al. 2015); *t*-tests were calculated using Satterthwaite's method via the `lmerTest` package in `R` (Kuznetsova et al. 2017).

Only reading times from trials with correctly-answered comprehension questions were analyzed. Six filler questions were answered correctly less than half of the time (obvi-

ous typos were found in two of these); data from all these trials were excluded from analysis. Of the resulting subset of the data, the slowest 1% (those above 3451.04 ms) and fastest 1% (those below 26.14 ms) of reading times were also excluded from analysis.

2.3.2 Results

Figure 2.2 reports mean RTs region by region. Visual inspection reveals few stark effects on reading at times preverbal regions, but noticeably longer reading times at all verbs except for those in $S_{ERG}O_{NOM}V$ clauses.

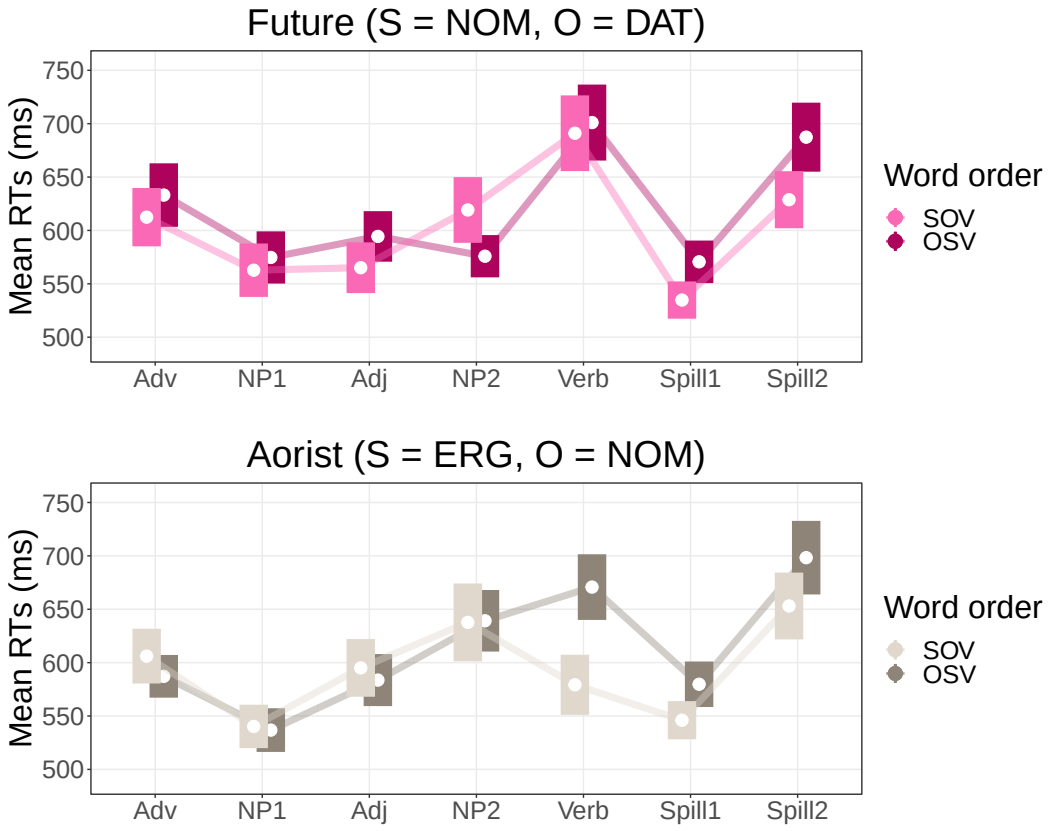


Figure 2.2: Mean readings times by region for Experiment 1a (Humans), separated by case frame

Linear mixed-effects modeling finds no significant effects of experimental manipulations on reading times at the adverb (w_1), NP_1 (w_2), or the second spillover region (w_7). A model without by-item random intercepts or slopes for word order finds no effects at the adjective (w_3); more complex models fail to converge here.

A few marginal effects were found. At NP_2, a full model fails to converge, but one omitting random by-item intercepts or slopes for case frame finds a marginal main effect of case frame at NP_2 ($\beta = 49$, $t = 1.9$, $p = 0.060$; see Table 2.2), hinting at a cost associated with the Erg–Abs alignment pattern here. At the verb, a full model finds a marginal main effect of word order ($\beta = -54$, $t = -1.9$, $p = 0.056$; see Table 2.3); on average, OSV verbs trended slower than SOV verbs. And at the first spillover region a model omitting by-item intercepts finds another trending cost for the OSV word order ($\beta = -30$, $t = -1.8$, $p = 0.071$; see Table 2.4).

	β	SE	df	*t*	*p*
(Intercept)	600	47	39	12	< 0.001
Case Frame	49	26	56	1.9	0.060
Word Order	24	29	36	0.82	< 1
Case:Order	−48	51	77	−0.93	< 1

Table 2.2: Results of linear mixed-effects modeling of reading times (in ms) at the NP_2 region of Experiment 1a (Humans), omitting random by-item intercepts or slopes for case frame

In a vacuum, none of these results is especially compelling. But there's good reason to think something meaningful really is going on at the verb region. First, feeding

	β	SE	df	t	p
(Intercept)	640	52	39	12	< 0.001
Case Frame	-55	33	36	-1.6	< 1
Word Order	-54	27	53	-1.9	0.056
Case:Order	-78	64	38	-1.2	< 1

Table 2.3: Results of linear mixed effects modeling of reading times (in ms) at the Verb region of Experiment 1a (Humans)

	β	SE	df	t	p
(Intercept)	640	52	39	12	< 0.001
Case Frame	20	19	26	1.0	< 1
Word Order	-30	16	60	-1.8	0.071
Case:Order	-5.1	34	58	-0.14	< 1

Table 2.4: Results of linear mixed effects modeling of reading times (in ms) at the first postverbal spillover region of Experiment 1a (Humans). Because of convergence issues, by-item intercepts were removed.

the model a dataset trimmed using a slightly more stringent criterion for culling reading time extrema (viz. clipping the slowest 2% and fastest 2%: those lower than 28 ms or greater than 2278.24 ms) yields a significant main effect of word order ($\beta = -62$, $t = -2.6$, $p < 0.05$), and a significant word order – case frame interaction ($\beta = -140$, $t = -2.8$, $p < 0.01$; see Table 2.5). Perhaps more than anything this exercise highlights this experiment's noisy data and lack of statistical power, but it at least lends some credence to the impression that aorist verbs are read on average slightly faster than future verbs, but mostly because aorist $S_{ERG}O_{NOM}V$

verbs are read especially fast.

	β	SE	df	*t*	*p*
(Intercept)	610	41	36	14	< 0.001
Case Frame	-34	22	43	-1.5	< 1
Word Order	-62	23	25	-2.6	< 0.05
Case:Order	-140	48	36	-2.8	< 0.01

Table 2.5: Results of linear mixed effects modeling of reading times (in ms) at the verb region of Experiment 1a (Humans), after discarding the fastest and slowest 2% of reading times

However, there is a alternative analysis whose results justify us to take seriously the reading times differences at the verb region. I reran the model replacing the simple Helmert contrast coding described in the methods section above with a more theoretically-informed coding scheme that employed three contrasts. The first contrast (COND1) contrasted the $S_{ERG}O_{NOM}V$ condition ($+¾$) with all others (each $-¼$). That encodes the analytical intuition that $S_{ERG}O_{NOM}V$ are the easiest of all to process, since they involve no incremental ambiguities. The second contrast (COND2) contrasted the $O_{NOM}S_{ERG}V$ condition ($+⅔$) against the future conditions (each $-⅓$) — encoding the intuition that $O_{NOM}S_{ERG}V$ condition is easier to process than either condition with no ergatives, since so many clause types are compatible with seeing one nominative and one dative argument. The final contrast (COND3) contrasted $S_{NOM}O_{DAT}V$ ($+½$) and $O_{DAT}S_{NOM}V$ clauses ($-½$), assuming (not innocently) that highly ambiguous strings in the canonical word are easier than ones with scrambling. A model calculating interactions between word order and these three conditions (as opposed

to interactions) finds a significant main effect of COND1 ($\beta = -99$, $t = -2.7$, $p < 0.01$; see Table 2.6). This shows that $S_{ERG}O_{NOM}V$ verbs were indeed read significantly faster than verbs in any other condition.

	β	SE	df	*t*	*p*
(Intercept)	640	52	36	12	< 0.001
Cond1	−99	36	45	−2.7	< 0.01
Cond2	−8.1	37	28	−0.21	< 1
Cond3	−15	44	33	−0.33	< 1

Table 2.6: Results of linear mixed effects modeling of reading times (in ms) at the verb region of Experiment 1a (Humans), using the special contrast coding scheme described in the main text

Descriptive statistics of comprehension question response accuracy and response time are reported in Figure 2.7. No significant effects due to experimental manipulations were found on comprehension question accuracy or question response latency, but for accuracy only a model omitting any by-participant effect structure can show this without convergence issues.

2.3.3 Discussion

The most important observation from this experiment is that verbs in $S_{ERG}O_{NOM}V$ clauses with two human arguments are read considerably faster than those with any shape. It is not surprising that OSV verbs were difficult to process. Recall that a sentence-initial O_{DO} can always be parsed as an S_{TR} in principle: all case categories are at least temporarily

	Accuracy	Latency
$\mathbf{S}_{\text{NOM}}\mathbf{O}_{\text{DAT}}\mathbf{V}_{\text{FUT}}$	83% (2.3%)	3122 (280)
$\mathbf{O}_{\text{DAT}}\mathbf{S}_{\text{NOM}}\mathbf{V}_{\text{FUT}}$	76% (2.5%)	2888 (110)
$\mathbf{S}_{\text{ERG}}\mathbf{O}_{\text{NOM}}\mathbf{V}_{\text{AOR}}$	80% (2.4%)	2747 (88)
$\mathbf{O}_{\text{NOM}}\mathbf{S}_{\text{ERG}}\mathbf{V}_{\text{AOR}}$	77% (2.5%)	2843 (100)

Table 2.7: Comprehension question response accuracy and latency (in ms) for Experiment 1a (Humans); standard errors are in parentheses

compatible with transitive-subjecthood, and none is uniquely linked to direct-objecthood. In the case of an initial dative O_{DO}, future-TAM morphology on the verb will betray its less prominent syntactic role. Thus increased reading times at the verb for $O_{DAT}S_{NOM}V$ conditions are expected; before the verb, the string is compatible with an $S_{DAT}O_{NOM}V$ parse. As for the conditions with initial nominative O_{DO}s (which were only found in ERG–NOM aorist clauses, not DAT–NOM perfect ones), we might expect reading difficulty to arise as soon as its ergative coargument. That's because ergative case is of course incompatible with a nominative S_{TR}. However, this ergative cost does not manifest until the following region, the verb. In SPR experiments, it's not at all unusual for effects to appear downstream from the cues that trigger them, but nearly every other effect in these Georgian experiments (in this section, in Section 2.4, and in Chapter 3) is observed extremely locally, at the SPR region containing the triggering cue. So it remains to be explained why in this experiment processing difficulty associated with an ergative noun phrases only surfaces at the immediately following word.

Much more surprising is the fact that verbs in $S_{NOM}O_{DAT}V$ clauses were read so

much slower than ones in $S_{ERG}O_{NOM}V$ clauses. Having encountered a nominative noun phrase and then a dative one, it would be entirely reasonable for the comprehender to predict a future-tense transitive verb. Why, then, does encountering that very cue engender just as much processing difficulty as the verb of an $O_{DAT}S_{NOM}V$ clause? It seems highly unlikely that a NOM–DAT string would be associated with an $O_{NOM}S_{DAT}V$ parse. That would require the parser either to initially map the nominative argument to direct-objecthood, or to revise an initial nominative-O_{DO} parse to an $O_{NOM}S_{DAT}V$ one upon encountering the dative noun phrase. The former possibility is incompatible with our interpretation of the ergative effect found in $O_{NOM}S_{ERG}V$ clauses; the latter is far-fetched because it requires positing a scrambled clause in a relatively rare TAM (the perfect) when more frequent structures have not been ruled out.

Perhaps the $S_{NOM}O_{DAT}V$ verb effect is not due to some disambiguating property of the future tense. Instead, it could indicated that the parser correctly links a human NP_2 to O_{DO} position, but doing so comes at a cost. That cost of course would only manifest one region later in the reading times, but we've already seen a very plausible case of a spillover effect caused by ergative NP_2s.

This an explanation makes a lot of sense from the perspective of harmonic alignment. After all, human O_{DO}s involve dissonant animacy and role scales, so perhaps merely positing a dissonant structure is taxing for the parser. This interpretation would be crucially distinct from the garden path analysis offered for the slow reading times for the verbs in OSV stimuli. Instead of a cue necessitating the comprehender's abandonment of a previous harmonic parse, here a cue leads the parser to posit a dissonant one from the get go. Such

effects are predicted by eADM; its Compute Prominence module can lead to a P400 if no harmonic parse of a noun phrase is available.

However, there is an important reason to be skeptical of this interpretation: no parallel effect tied to human O_{DO}s is observed in the $S_{ERG}O_{NOM}V$ condition. Indeed, if anything ERG–NOM–AOR seems to be the easiest transitive structure of all to process in Georgian. Why would it be that a human-O_{DO} parse causes processing difficulty in NOM–ACC clauses, but not ERG–ABS ones? This is an especially difficult question to answer because a nominative human noun which follows an ergative one (as in the $S_{ERG}O_{NOM}V$ stimuli) can *only* be an O_{DO}. Two arguments into the $S_{NOM}O_{DAT}V$ stimuli, a human O_{DO} parse is entirely avoidable; a dative human noun following a nominative one could in principle be an S_{TR}, goal, or O_{DO}.

So I suggest that the difficulty associated with $S_{NOM}O_{DAT}V$ verbs is rooted neither in a foiled expectation for a dative S_{TR}, nor in inherent difficulty in processing a human O_{DO}. Instead, I suggest that the $S_{NOM}O_{DAT}V$ verb effect is due to a different garden path. Specifically, the NOM–DAT sequence of the $S_{NOM}O_{DAT}V$ condition leads to the expectation for an $S_{NOM}IO_{DAT}DO_{DAT}V$ ditransitive clause. In other words, the human dative argument is interpreted as the clause's *goal*, rather than its S_{TR} or O_{DO}. Recall that Georgian goals are always in the dative case (23a–b), except when the S_{TR} is assigned dative (23c).

(23) a. *ekim-i* ***bavšv-s*** *ɕign-s* *miscems.*
doctor-NOM **child-DAT** book-DAT give:FUT.3SG≫3
'The doctor will give the child a book.' FUT: NOM–DAT–DAT

b. *ekim-ma* ***bavšv-s*** *çign-i* *misca.*

doctor-ERG **child-DAT** book-NOM give:FUT.3SG≫3

'The doctor gave the child a book.' AOR: ERG–DAT–NOM

c. *ekim-s* ***bavšv-is=tvis*** *çign-i* *miucia.*

doctor-DAT **child-GEN=for** book-NOM give:PERF.3SG≫3

'The doctor [apparently] gave the child a book.' PERF: DAT–PP$_{for}$–NOM

(Constructed examples)

But just why would a NOM–DAT sequence of two human arguments cause the parser to posit a ditransitive structure? If the parser prioritizes incremental harmonic alignment over all other considerations, and if ditransitive parses are not a last resort option, then this is actually the most optimal interpretation of that string. At NP$_1$, the comprehender reads a human noun in the nominative case. Nominative is compatible with S$_{TR}$ and O$_{DO}$ positions; the more prominent role harmonizes with NP$_1$'s likewise prominence animacy, so the comprehender parses NP$_1$ as an S$_{TR}$. Later, they encounter NP$_2$: again human, but now in the dative — morphology found on S$_{TR}$s, goals, and O$_{DO}$s. The most harmonic parse of NP$_2$ would be as a dative S$_{TR}$, but since the comprehender is already committed to a nominative S$_{TR}$, that parse is off the table. Therefore the most harmonic grammatical interpretation of the string involves assigning to NP$_2$ the goal role of a S$_{NOM}$IO$_{DAT}$DO$_{DAT}$V ditransitive clause.[4] However, the immediately following monotransitive verb does not a license a goal, and the parser must reparse to a S$_{NOM}$O$_{DAT}$V structure. The surprisingly high

[4] A ditransitive parse is not strictly necessary, since Georgian has many verbs that license goals (i.e., indirect objects) but not O$_{DO}$s (direct objects). The following examples illustrate; the aorist version demonstrates that the verb's object is indeed a goal, since a O$_{DO}$ would be assigned nominative in that TAM.

reading times of the $S_{NOM}O_{DAT}V$ verb region reflects this revision.

Of course, since the $S_{NOM}O_{DAT}V$ verbs in this experiment were all in the future TAM, there one final obvious alternative explanation. Perhaps $S_{NOM}O_{DAT}V$ verbs are difficult to process not because they foil a ditransitive parse, but rather because processing future-tense semantics is difficult in an experimental setting with no context that would accommodate them. This confound could be easily avoided with experimental materials using another of Georgian's NOM–ACC TAMS, perhaps the present. Short of that, though, results from Experiment 1b will give us some confidence that the $S_{NOM}O_{DAT}V$ verb cost is not merely attributable to challenges associated with processing certain TAMS.

2.4 Experiment 1b: Inanimates in Transitive Clauses

The goal of Experiment 1b was to test how Georgian comprehenders parse low-animacy arguments in a variety of cases and linear positions. If they assign inanimate noun phrase to the O_{DO} role whenever possible — as a theory employing incremental harmonic alignment predicts — then we expect to cues which eliminate a object-initial parses to cause processing difficulty. In this study, such cues include initial ergative morphology, and future-tense verbal morphology in $S_{NOM}O_{DAT}V$ sentences. Long reading times at ergative NP_1s shows that certain prominence-scale misalignments are difficult to process qua mis-

(24) a. *ekim-i* *masçavlebel-s* *aḳocebs*
doctor-NOM teacher-DAT kiss:FUT.3SG
'The doctor will kiss the teacher.'
b. *ekim-ma* *masçavlebel-s* *aḳoca*
doctor-ERG teacher-DAT kiss:AOR.3SG
'The doctor kissed the teacher.' (Constructed examples)

alignments. Scale harmony, in other words, isn't merely a metric by which to rank possible continuations. The very dissonance of dissonant cues can cause processing difficulty — not just because they might reveal garden paths.

2.4.1 Method

Participants, procedure, and analysis

Experiments 1a and 1b were run as each other's fillers in the same session, so they share the same participants and procedure. And because stimuli for the two experiments were also designed identically, reading times, comprehension question accuracy, and question response latency were analyzed in the same way. As above, the slowest and fastest 1% (those not between 26 ms and 3162.25 ms) reading times were discarded. See Section 2.3.1 for full details.

Materials

28 item sets were constructed in a 2×2 design, crossing TAM / S_{TR}–O_{DO} case frame (FUT / NOM–DAT or AOR / ERG–NOM) and word order (SOV or OSV). These sentences conformed to the template used for Experiment 1a, repeated here as (25). But unlike in that experiment, here all verbal arguments were inanimate nouns. An example itemset is given in (26).

(25) Stimulus template (Experiment 1b, Inanimates)

Adv	NP_1	Adj	NP_2	V	$Spill_1$	$Spill_2$.
w_1	w_2	w_3	w_4	w_5	w_6	w_7

(26) a. NOM–DAT, SOV

agaraḳ-ze *sarḳe* *parto* *panǰara-s* *gaṭexs* *miçisjvr-is*

dacha-on mirror.NOM wide window-DAT break:FUT.3SG earthquare

dro-s.

time-DAT

'At the summer home, the mirror will break the wide window during the earthquake.'

b. NOM–DAT, OSV

agaraḳ-ze *sarḳe-s* *parto* *panǰara* *gaṭexs* *miçisjvr-is*

dacha-on mirror-DAT wide window.NOM break:FUT.3SG earthquare

dro-s.

time-DAT

'At the summer home, the wide window will break the mirror during the earthquake.'

c. ERG–NOM, SOV

agaraḳ-ze *sarḳe-m* *parto* *panǰara* *gaṭexa* *miçisjvr-is*

dacha-on mirror-ERG wide window.NOM break:AOR.3SG earthquare

dro-s.

time-DAT

'At the summer home, the mirror broke the wide window during the earthquake.'

d. ERG–NOM, OSV

agaraḳ-ze *sarḳe* *parto* *panǰara-m* *gaṭexa* *miçisjvr-is*

dacha-on mirror.NOM wide window-ERG break:AOR.3SG earthquare

dro-s.

time-DAT

'At the summer home, the wide window broke the mirror during the earthquake.'

These experimental items were presented among 80 filler sentences, which comprised 28 experimental items for Experiment 1a (see Section 2.3), and 52 more sentences with varied syntactic properties. Each of the 108 sentences was followed by a yes–no comprehension question. All of the stimuli were constructed in consultation with two native speakers.

2.4.2 Results

The following plot reports mean RTs region by region (Figure 2.3). Impressionistically, ergative NP_1 in the $S_{ERG}O_{NOM}V$ condition are read more slowly than their nominative and dative counterparts. There also seems to be a crossover interaction at the verb: SOV is the harder word order in future clauses, but OSV is harder in the aorist.

Linear modeling reveals no significant effects of experimental manipulations at the Adverb (w_1), the adjective (w_3), NP_2 (w_4), or either spillover region (w_6, (w_7)). It does find a significant word order – case frame interaction at the verb ($\beta = -161$, $t = 2.0$, $p < 0.05$; see Table 2.8), confirming statistically what can be gleaned through visual inspection of the

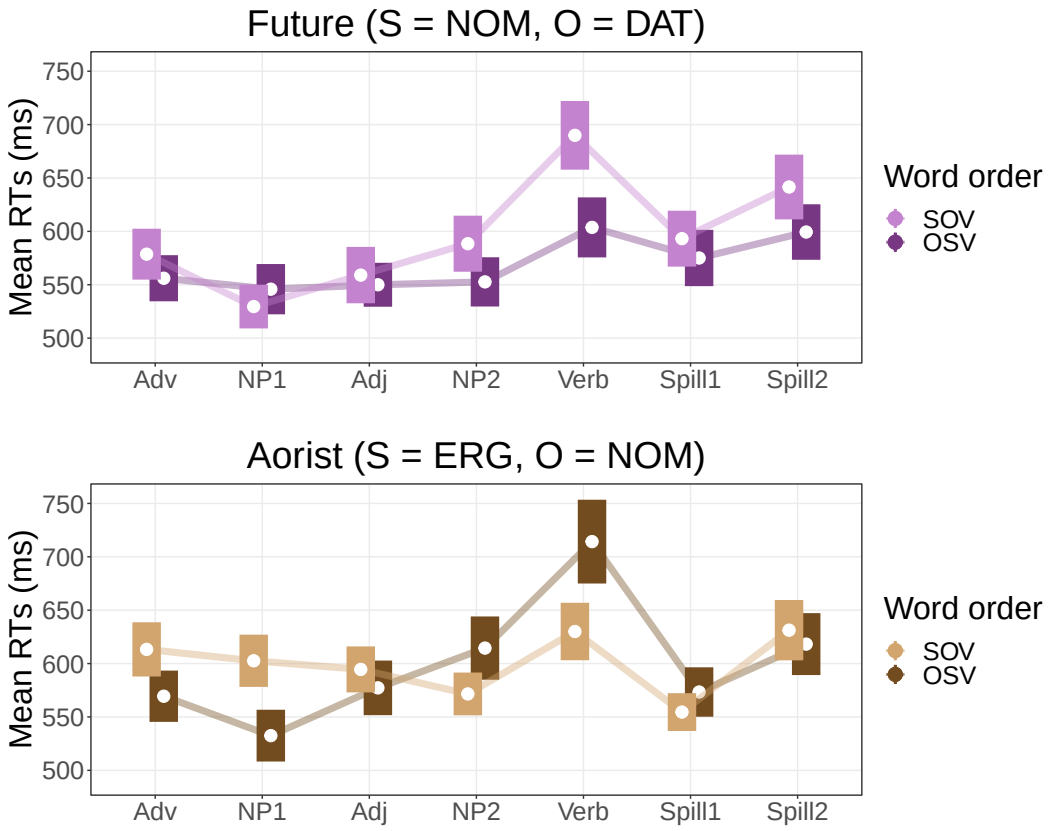


Figure 2.3: Mean readings times by region for Experiment 1b (Inanimates), separated by case frame

data.

	β	SE	df	*t*	*p*
(Intercept)	640	52	38	12	< 0.001
Case Frame	21	31	34	0.67	< 1
Word Order	-10	28	30	-0.37	< 1
Case:Order	-160	78	35	-2.0	< 0.05

Table 2.8: Results of linear mixed effects modeling of reading times (in ms) at the verb region of Experiment 1b (Inanimates)

The situation is murkier at NP_1. The normal model finds no significant effects (smallest $p > 0.1$). Employing the more theoretically sophisticating Helmert coding contrasts described in Section 2.3.2, results in only a trending effect of COND1 ($\beta = 47$, $t = 1.9$, $p = 0.053$; see Table 2.9); recall that contrast compared $S_{ERG}O_{NOM}V$ conditions to all others, so this suggests that there's something legitimate to conclude from the conspicuously slow reading times associated with ergative NP_1s.

However, when we exclude slightly more data — culling the slowest 2% (> 2242 ms) and fastest 2% (< 27 ms) — a standardly-coded model returns significant effects of word order ($\beta = -63$, $t = 2.7$, $p < 0.05$) and the interaction between word order and case frame ($\beta = -140$, $t = 2.9$, $p < 0.01$; see Table 2.10).

	β	**SE**	**df**	***t***	***p***
(Intercept)	530	35	40	14	< 0.001
Cond1	47	23	62	1.9	0.053
Cond2	−18	30	39	−0.63	< 1
Cond3	−2.47	30	26	−0.080	< 1

Table 2.9: Results of linear mixed effects modeling of reading times (in ms) at the NP_1 region of Experiment 1b (Inanimates), using the special contrast coding scheme described in the Section 2.3.2

Descriptive statistics for comprehension question response accuracy and latency are given in Table 2.11. No significant effects of experimental manipulations were found on these measures, though only a response-time model with by-item random slopes and intercepts for word order converges.

	β	SE	df	*t*	*p*
(Intercept)	610	42	37	14	< 0.001
Case Frame	-35	22	40	-1.5	< 1
Word Order	-63	23	26	-2.7	< 0.05
Case:Order	-140	47	37	-2.9	< 0.01

Table 2.10: Results of linear mixed effects modeling of reading times (in ms) at the verb region of Experiment 1b (Inanimates), after discarding the fastest slowest 2% of reading times

	Accuracy	**Latency**
$\mathbf{S}_{\text{NOM}}\mathbf{O}_{\text{DAT}}\mathbf{V}_{\text{FUT}}$	81% (2.3%)	2793 (96)
$\mathbf{O}_{\text{DAT}}\mathbf{S}_{\text{NOM}}\mathbf{V}_{\text{FUT}}$	82% (2.2%)	2875 (87)
$\mathbf{S}_{\text{ERG}}\mathbf{O}_{\text{NOM}}\mathbf{V}_{\text{AOR}}$	82% (2.2%)	2921 (180)
$\mathbf{O}_{\text{NOM}}\mathbf{S}_{\text{ERG}}\mathbf{V}_{\text{AOR}}$	83% (2.2%)	2791 (99)

Table 2.11: Comprehension question response accuracy and latency (in ms) for Experiment 1b (Inanimates); standard errors are in parentheses

2.4.3 Discussion

There are two main takeaways from Experiment 1b. First, when the first noun phrase of a clause is ergative, it is read much slower than when it is nominative. (In should be noted, however, that the most conservative statistical practices do not support this conclusion.) Insofar as this effect is legitimate, it's clear how it's related to harmonic alignment. Ergative morphology is found only on the most prominent syntactic role (S_{TR}), but inanimate is the least prominent animacy category. In contrast to all the effects seen in the human

experiment, which all plausibly signal garden-path reanalysis, this inanimate ergative cost must be rooted squarely in scale disharmony. In other words, this is somewhat equivocal evidence for something like the eADM's Compute Prominence step, which predicts that the mere integration of dissonant cues can lead to processing difficulty.

It's striking that unambiguous inanimate S_{TR}s seem difficult to process when unambiguous human O_{DO}s are not. Recall from Experiment 1a that nominative human nouns following ergative ones are relatively effortless to parse — even though the comprehender has no choice but to link that nominative human to direct-objecthood. This asymmetry has been observed before in other languages, and is hypothesized to be related to more stringent conditions on agenthood than on patienthood (Bornkessel-Schlesewsky and Schlesewsky 2009, Fauconnier 2012).

The second takeaway has to do with reading times at the verb region. Focusing on the NOM–DAT conditions, we see that this region is read more slowly in $S_{NOM}O_{DAT}V$ clauses than $O_{DAT}S_{NOM}V$ clauses. This difference can be rationalized in a harmonic alignment theory if we permit the parser to freely entertain ditransitive structures. Whether an inanimate NP_1 is nominative (in the $S_{NOM}O_{DAT}V$ condition) or dative ($O_{DAT}S_{NOM}V$), the most harmonic way to parse it is as a O_{DO}. However, this does not mean that the comprehender will process the rest of the sentence identically across these conditions. After all, nominative O_{DO}s and dative ones are found in very different case alignment systems. Having parsed a nominative O_{DO}, the comprehender has some degree of flexibility in how to treat the dative noun phrase they encounter at NP_2. In clauses with nominative O_{DO}s, a dative argument can either be an S_{TR} (e.g., in the DAT–ABS-aligned perfect; see example 23c) or a goal (in the ERG–ABS

aorist; 23b). Neither of these is an especially good role for an inanimate argument, but a O_{IO} parse is ultimately more harmonic ones because it minimizes scale dissonance.[5] But in this experiment, parsing a post-nominative inanimate noun in the dative case as a O_{IO} leads in this experiment will always lead the comprehender astray: the monotransitive verb will require a radical reparse, changing the argument structure and reassigning the nominative argument to the S_{TR} position and the dative to O_{DO}.

On the other hand, if the comprehender first encounters a dative inanimate noun phrase and parses it as O_{DO}, their subsequent choices are much more constrained. Dative O_{DO}s are only found in NOM–ACC TAMs like the future, where nominative-marked arguments can only be S_{TR}s. Therefore, an IN.DAT–IN.NOM string is not likely to garden path the comprehender in the way a IN.NOM–IN.DAT string will. This observation cuts both ways, though — it explains why there would be a garden path effect only in the SOV condition in this case frame, but it doesn't explain why a countenancing a nominative-marked inanimate S_{TR} does not seem to come with the same processing cost that positing an ergative-marked inanimate

[5] If inanimate S_{TR}s are intrinsically difficult to process because the semantic roles typically associated with transitive subjects (especially agent) are difficult to square with the semantic properties associated with inanimacy (non-sentience, immobility, lack of individuation), one might expect inanimate O_{IO}s to be likewise inherently difficult. After all, crosslinguistically we typically find O_{IO}s with recipient, affectee, or experiencer semantic roles — roles tightly associated with high animacy.

And indeed, many O_{IO}s in Georgian are high-animacy recipients, affectees, or experiencers. However, there is a considerable class of ditransitive verbs whose O_{DO} argument is semantically an instrument and whose O_{IO} is a patient (27). Perhaps these ditransitives in particular are the most easily accommodated structure to entertain when alignment concerns dictate positing two inanimate objects.

(27) a. *monadire-m mizan-s ṭqvia esrola.*
hunter-ERG target-DAT bullet.NOM shoot:AOR:3SG≫3
'The hunter shot a bullet at the target.'

b. *cxen-ma xe-s çixl-i hḳra.*
horse-ERG tree-DAT kick-NOM strike:AOR:3SG≫3
'The horse kicked the tree.' (literally 'struck the tree with a kick') (Constructed examples)

S_{TR} does. There are two conceivable reasons that we do not observe a marked alignment cost in the $O_{DAT}S_{NOM}V$ condition that parallels what we observed at NP_1 in the $S_{ERG}O_{NOM}V$ condition. First, an inanimate S_{TR} parse is not truly necessary for IN.DAT–IN.NOM parse; that sequence is also compatible with an ERG–ABS-aligned ditransitive clause whose ergative argument has not yet been encountered. Perhaps the parser indeed assigns a inanimate nominative NP_2 to S_{TR} position, but the marked alignment cost in this situation is attenuated by the glimmer of hope that a higher-animacy S_{TR} might still appear.

Alternatively, word order might be playing a role. Since SOV and SVO are the most canonical shapes for a transitive clause in Georgian, it's conceivable that linear precedence plays a key role in prominence calculus. An inanimate NP_1 in the ergative case has two high-prominence properties (transitive-subjecthood and clause-initiality) that clashes with its one low-prominence property (inanimacy). In contrast, a post-dative inanimate NP_2 in the nominative case has just one high-prominence property (transitive-subjecthood). Perhaps the prominence scales in Georgian are weighted in such a way that reading-time disruptions only occur when more than two scales are dissonantly aligned. However, this second interpretation makes it difficult to understand the long reading times at $O_{NOM}S_{ERG}V$ verbs in this experiment. Unlike the parallel effect observed at the same environment in Experiment 1a, the slowdown here can't be cast as a garden path effect — assuming, at least, that the initial inanimate nominative is correctly predicted to be an O_{DO}, as harmonic alignment would dictate.

2.5 General discussion

Georgian's split ergative case system is a double-edged sword for the incremental comprehender. How the language's three core case categories (nominative, ergative, and dative) are assigned to different syntactic roles in different case frames means that case cues vary in their informativity, both intrinsically and in relation to previously cues. Ergative is only found on subjects (typically S_{TR}s) of aorist or optative clauses; nominative and dative are found in all case frames, and on a wider variety of syntactic roles, so preverbal arguments in these cases are often compatible with multiple parses.

Results from the two reading time experiments shows that comprehenders rely on animacy cues to supplement case cues' inconsistent informativity. When reading a human noun phrase (as in Experiment 1a), it seems the parser bends over backwards to give it the most prominent available syntactic role. This behavior must be tied specifically to animacy, and to not a more general desire to identify the clause's subject. That's because clauses with two inanimate arguments are read in a clearly different way; reading times in Experiment 1b suggest inanimates are assigned the least prominent syntactic role possible. Evidence in favor of this interpretation comes from reading time disruptions at or just after clause-internal ergative morphology and clause-final verbal TAM morphology. These cues often serve as disambiguators, potentially foiling a previous parse that had maximized incremental harmonic alignment. Besides these garden path effects, though, there is modest evidence that there is inherent difficulty associated with processing inanimate S_{TR}s: arguments which violate harmonic alignment in a particularly egregious way.

Over all, these results conform rather closely to eADM's predictions. Parsers seek to maximize incremental animacy–role harmonic alignment, and are vexed by certain misalignments. However, we do not obviously observe effects attributable to eADM's Distinctiveness principle. Distinctiveness says that comprehenders should always prefer intransitive parses to monotransitive ones, and monotransitive parses to ditransitive ones. That's because the fewer arguments there are in a clause, the easier it is to keep each isolated in working memory. S_{TR}s and O_{DO}s are definitionally less distinct than S_{IN}s, simply because they compete for processing resources with another argument.

In Georgian, we find nominative-, ergative-, and dative-assigned S_{IN}s. And intransitive verbs' lexical semantics seem heterogeneous enough that the animacy of a potential S_{IN} alone should not exclude any particular intransitive parses (recall, for instance, the many non-agentive atelic intransitives which readily take inanimate ergative subjects; 17c–d). So, insofar as all intransitive structures are equally easy for the parser to access — not an innocent assumption, since nominative S_{IN}s are at least impressionistically much more common than ergative or dative ones — it seems like all clauses should initially be assigned an intransitive structure. In other words, whenever a comprehender prioritizing Distinctiveness encounters more than one noun phrase within a clause, they will have to revise their parse. But across both experiments, NP_2 does not seem to reliably and uniformly cause processing difficulty; if that argument really does force the parser away from an intransitive interpretation, then merely adding another argument position does not seem to be very cognitively effortful. Moreover, the trending cost associated with ergative inanimates in NP_1 position is difficult to understand if these noun phrases are parsed as S_{INA} arguments. Any theory of

harmonic alignment will classify inanimate S_{TR}s as dissonant, but it's less clear why the such a theory would classify the inanimate S_{IN} of an atelic verb the same way. Of course, it may just be that ergative-S_{INA} parses are given less weight during incremental prediction than ergative-S_{TR} parses, perhaps because they are less frequent. If this is the case, perhaps the parser would prefer an intransitive parse all things being equal, but independent constraints on lexical access prevent that in all cases.

Chapter 3

Processing relative & correlative clauses

3.1 Introduction

A wealth of evidence indicates that a relative clause with a gap in subject position (28a) (a Subject Relative Clause; SRC) is generally easier to process than one with a gap in object position (28b) (an Object Relative Clauses; ORC).

(28) a. the writer [$_{RC}$ who ___ inspired the painter] *SRC*

b. the writer [$_{RC}$ who the painter inspired ___] *ORC*

This Subject-Gap Advantage (SGA) manifests across both a wide range of measurements — including acceptability judgements, reading times, disambiguation preferences, eye movements, and ERPs — and also a wide range of languages — from Dutch (Frazier 1987) to Turkish (Kahraman et al. 2010) to Zapotec (Foley et al. 2019). For detailed reviews of the literature on this processing asymmetry, see Gibson (1998) and Kwon et al. (2010).

However, the underlying source of the SGA remains unclear. A number of factors have been proposed to explain the processing asymmetry, but here we focus on just three: syntactic structure, the informativity of morphological cues, and linear/temporal distance. Within the relative clause processing literature, languages with Nominative–Accusative case alignment and postnominal relative clauses are overrepresented — but these are the very morphosyntactic properties which happen to be least informative for teasing apart the effects of structure, morphological informativity, and distance. And even studies on languages with underrepresented features (in particular ergative–absolutive alignment or prenominal relatives) have their limitations, since, internal to most languages, it is not possible to independently manipulate all these factors.

This chapter presents reading-time data from Georgian, whose properties are practically tailor-made to study relative clause processing. It is a language with a Split-Ergative case system, flexible word order, a wealth of relativization strategies, and speakers that read fluently and are generally computer literate. Together, these factors make it possible to compare reading times of pre- and post-nominal relative clauses, or ones with either Nominative–Accusative or Ergative–Absolutive case alignment. The four self-paced reading experiments detailed here do just that. Together, they lend support to theories which derive the SGA from structural and cue-based principles.

A Structural Theory of the SGA posits that certain syntactic positions are inherently more accessible for filler–gap dependencies than others. Keenan and Comrie's (1977) Accessibility Hierarchy is one implementation of this idea (29).

(29) The Accessibility Hierarchy (Keenan and Comrie 1977:66)

Subject > Direct Object > Indirect Object > Oblique > Possessor > Object of Comparison

This scale of grammatical relations is theorized to have cognitive ramifications: the farther to the left a position is, the easier it will be to process a dependency with a gap there.[1] Evidence for the Accessibility Hierarchy comes from robust typological implications: if a given relativization strategy in a language permits relativization from one point on the hierarchy, it will generally also permit relativization from all positions above it on the hierarchy. If all languages' grammatical relations are organized according to this hierarchy, then the Structural Theory predicts the SGA to be universal.

A second factor hypothesized to contribute to the SGA is the informativity of morphology (especially case morphology) in and around the relative clause. This Case Informativity Theory posits that the cognitive effort necessary to integrate an argument into a syntactic structure is proportional to the amount of syntactic structure it entails (Polinsky et al. 2012; cf. Hale 2003, 2006). Thus, arguments in unmarked cases (nominative or absolutive) will be easier to process than ones in dependent cases (accusative or ergative). To illustrate, consider the following Russian relative clauses. The SRC (30a) is introduced by a relative pronoun in the nominative case. This morphology is compatible with a wide range of continuations: the gap must be in subject position, but otherwise the relative clause might be transitive or intransitive, past or non-past, active or passive. In contrast, the ORC

[1] Most filler–gap processing research has focused only on constructions with gaps in subject or direct object position; little work has investigated gaps lower on the Accessibility Hierarchy (though see Lin 2008 for a recent exception).

(30b) begins with an accusative-case relative pronoun. Compared to nominative, accusative is more informative, as it entails the presence of upcoming transitive structure. The Case Informativity Theory therefore predicts the accusative relative pronoun of the ORC to be more difficult to process than the nominative one in the SRC — in other words, it predicts an SGA for a Nominative–Accusative language like Russian.

(30) a. *sobak-a,* [$_{RC}$ ***kotor-aja*** ___ *košk-u* *dogonjaet*]

dog-F.NOM **which-F.NOM** cat-F.ACC chase:PRES.3SG

'the dog [$_{RC}$ which ___ is chasing the cat]'

b. *sobak-a,* [$_{RC}$ ***kotor-uju*** ___ *košk-a* *dogonjaet*]

dog-F.NOM **which-F.ACC** cat-F.NOM chase:PRES.3SG

'the dog [$_{RC}$ which the cat is chasing ___]'

(Russian; glosses adapted from Polinsky 2011)

In contrast, consider the processing profile of relative clauses with Ergative–Absolutive alignment, like the following from Hindi (31). Here, the morphology of the SRC's relative pronoun (ergative) is more informative than that of the ORC's (absolutive), since absolutive case appears in more syntactic environments than ergative case does in Hindi (cf. Dillon et al. 2012). By the metric of case informativity, then, we should observe an object-gap advantage (OGA) across this pair of sentences, manifesting at or just after the relative pronouns.

(31) a. *maĩ-ne ek ciṛiyā khīñcā* [$_{RC}$ ***jis-ne*** ___ *ek cūhā khā*

1SG-ERG one bird.NOM draw.PFV **REL-ERG** one rat.NOM eat

liyā.]

take.PFV

'I drew a bird [$_{RC}$ which ___ ate a rat.]'

b. *maĩ-ne ek ciṛiyā khīñcā* [$_{RC}$ ***jo*** *ek cūhe-ne* ___ *khā*

1SG-ERG one bird.NOM draw.PFV **REL.NOM** one rat-ERG eat

liyā.]

take.PFV

'I drew a bird [$_{RC}$ which a rat ate ___.]'

(Hindi; Pranav Anand, p.c.)

Crucially, the Informativity Theory predicts that the difficulty associated with accusative or ergative should not be limited to environments involving filler–gap dependencies. Even in root clauses, dependent case morphology will license predictions about upcoming structure that are relatively taxing to integrate.

Finally, the SGA may be due to the length of a relative–clause dependency. This length, between the filler and the gap, might be measured in some number of linguistic units (Gibson 1998) or in temporal distance (Lewis and Vasishth 2005). Such a theory assumes that fewer processing resources are available to other parsing operations while a filler is held in active memory, waiting to be linked to a gap. Returning to English, it is clear how the Distance Theory predicts an SGA. In an SRC (32a), the head noun filler can be associated

with its gap as soon as the parser crosses the complementizer, observing that the subject position is empty. In an ORC (32b), the subject and verb additionally intervene, making for a dependency which is longer, and therefore more difficult.

(32) a. the writer [$_{RC}$ who ___ inspired the painter]

b. the writer [$_{RC}$ who the painter inspired ___]

Considering a language with prenominal relative clauses, like Korean (33), the situation is different. Assuming gaps to be posited in canonical argument positions (Korean is an SOV language), then an SRC dependency will be longer than an ORC dependency. All things being equal, then, the Distance Theory predicts an OGA for a language with prenominal relatives. It is worth noting, however, that the picture is likely more complicated for Korean, since it has null pronouns, argument scrambling, and its relative clauses are not unambiguously marked until their right edge (by the verbal suffix *–n* 'ADNOMINAL'); together these factors mean that it is not a trivial task for the parser to determine where a relative clause dependency even starts. See Yun et al. (2015) for an account that takes such uncertainties into consideration for Korean and other typologically similar languages.

(33) a. [$_{RC}$ ___ *uywon-ul* *kongkyekha-n*] *enlonin-i*

senator-ACC attack-ADN journalist-NOM

'the journalist [$_{RC}$ that ___ attacked the senator]'

b. [$_{RC}$ *uywon-i* ___ *kongkyekha-n*] *enlonin-i*

senator-NOM attack-ADN journalist-NOM

'the journalist [$_{RC}$ that the senator attacked ___]'

(Korean; Kwon et al. 2010)

Table 3.1 summarizes the predictions made by the three theories discussed above for languages of various typological profiles. Two parameters are manipulated: case alignment pattern, and relative clause position. It is clear that a language represented by the first column (namely, a Nominative–Accusative language with postnominal relatives, like English), is the least informative kind when testing these theories. But the other parameter settings also involve some degree of analytical uncertainty. If a language represents just one of these columns, then, it may not be able to unambiguously adjudicate among the three theories.

	Accusative alignment		**Ergative alignment**	
	N [**RC** …]	[**RC** …] **N**	**N** [**RC** …]	[**RC** …] **N**
Structure	SGA	SGA	SGA	SGA
Informativity			OGA	OGA
Distance		OGA	SGA	

Table 3.1: Processing asymmetries predicted by three theories of the SGA, across relative clauses varying in gap site, linear position, and internal case alignment.

Enter Georgian. While it has only been the focus of a handful of previous psycholinguistic studies (Skopeteas et al. 2012; Foley and Wagers 2017b,a; Lau et al. 2018,

Submitted), it is a language especially well-suited to compare these theories of the SGA. With Split-Ergative case alignment with both pre- and post-nominal relative clauses, Georgian instantiates all four columns in Table 3.1. Taking advantage of this fact, we manipulated the position of a relative clause and its internal case alignment in four self-paced reading experiments in order to pinpoint loci of real-time processing difficulty. Across these experiments, the most consistent effect is a slowdown associated with relative-clause coarguments in the ergative case (an effect also found in ERPs and reading times by Lau et al. Submitted). This is compatible with both the Structural and Informativity Theories. Ergative coarguments eliminate the possibility of an SRC parse — thereby forcing a parser that privileges subject gaps into a reparse — but ergative morphology is also highly informative in Georgian — so it triggers taxing predictions about the upcoming structure. Neither theory alone can account for the full array of other effects, though. For example, ergative coarguments that precede cues for a filler–gap dependency (i.e., ergatives that could be confused for root-clause arguments) also evince processing difficulty: a result only predicted by the Informativity Theory. On the other hand, ergative relative pronouns are read no slower than uninformative nominative ones: a result only predicted by the Structural Theory. So, it seems that both factors are at play in Georgian — as seems to be the case in Avar (Polinsky et al. 2012) and Niuean (Longenbaugh and Polinsky 2016). As for the Distance Theory, we find scant evidence that supports it. Regions at the right edge of prenominal relative clauses and just after exhibit a numerical trend in favor of ORCs, but this does not reach significance.

The chapter is structured as follows. We first delve into some previous research

on cross-linguistic RC processing (Section 3.2) and describe morphosyntactic properties of Georgian relevant to the experiments (Section 3.3). From there we dive into the experiments themselves, first discussing the pair manipulating RC position (Sections 3.4 and 3.5), then the pair manipulating RC-internal case alignment (Sections 3.6 and 3.7). Finally, we discuss the results of these experiments in a broader context and their implication for the cross-linguistic picture of relative clause processing (Section 3.8).

3.2 Previous research

This section reviews cross-linguistic findings in relative clause processing, reviewing studies on languages with prenominal relative clauses, and ones on ergative languages. The processing of prenominal relatives has been most thoroughly investigated in Chinese (Hsiao and Gibson 2003, Lin and Bever 2010, Vasishth et al. 2013, and others), Japanese (Ueno and Garnsey 2008), and Korean (Kwon et al. 2010, 2013). These languages have generally exhibited an SGA, though traces of an OGA are occasionally reported. Experimental evidence from Chamorro Wagers et al. (2018), where pre- and post-nominal relatives can be compared directly, shows that the SGA is weaker in prenominal RCs than postnominal RCs, but this asymmetry may be due to a confluence of morphosyntactic properties, rather than just distance.

Relative clause processing has been studied in a growing number of ergative languages, including Basque Carreiras et al. (2010), Avar (Polinsky et al. 2012, Polinsky 2016), Ch'ol and Q'anjob'al (Clemens et al. 2015), and Niuean (Longenbaugh and Polinsky 2016,

Tollan et al. 2019). A mixed picture has emerged: some studies find an SGA, others an absolutive-gap advantage (i.e., an OGA in transitive clauses), and yet others find evidence for both.

3.2.1 Prenominal relative clauses and the subject gap advantage

Among the existing studies on languages with prenominal relative clauses, most have focused on Chinese, Japanese, and Korean. In all three languages, the left edge of a relative clause is not directly signaled by a complementizer or relative pronouns (34). Since all of these languages have null pronouns and can scramble arguments, this means some relative clauses (in particular, ones that modify clause-initial noun phrases) will be temporarily ambiguous and may be initially parsed as a root clause.

(34) a. [$_{RC}$ ___ *Gōngjī yìyuán-de*] *jìzhě* *chéngrèn-le cuòwù.*

attack senator-ADN reporter admit-PFV error

'The reporter [$_{RC}$ that ___ attacked the senator] admitted the error.' **SRC**

(Chinese; Kwon et al. 2013:539)

b. [$_{RC}$ *Giin-ga* ___ *hanashita*] *kisha-ga* *ayamari-o mitometa.*

senator-NOM attack reporter-NOM error-ACC admitted

'The reporter [$_{RC}$ that the senator attacked ___] admitted the error.' **O_{ACC}RC**

(Japanese; Kwon et al. 2013:539)

c. [$_{RC}$ ___ *Uywon-ul konghkyekha-n*] *kica-ka silswu-lul siinhayssta.*

senator-ACC attack-ADN reporter-NOM error-ACC admitted

'The reporter [$_{RC}$ that ___ attacked the senator] admitted the error.' **S_{NOM}RC**

(Korean; Kwon et al. 2013:539)

Processing studies on Japanese and Korean have routinely found an SGA, despite the linear order of relative clause and head noun (Miyamoto and Nakamura 2003, Ishizuka et al. 2006, Ueno and Garnsey 2008; Kwon 2008, Kwon et al. 2010, 2013). For example, in two eye-tracking experiments on Korean, Kwon et al. (2010) found that ORCs were read significantly more slowly than SRCs. This finding was observed across several measures: reading times of the whole sentence, regression-path duration at head nouns, and rereading times at regions across the sentence. The authors interpret their results as strong evidence in favor of a Structural Theory of the SGA; the OGA predicted by Distance Theories was not borne out.

Ueno and Garnsey's (2008) self-paced reading and ERP experiments on Japanese yielded similar results. Reading times at the head noun were significantly higher in the ORC condition than the SRC condition. As for ERPs, object gaps elicited a greater bilateral anterior negativity than subject gaps during the relative clause and a greater centro-posterior positivity after the relative clause. Both effects the authors interpret as reflecting an SGA. However, Ishizuka et al. (2006) suggest that findings such as these in Japanese are not due to the structural position of the gap, but rather the greater temporary ambiguity of the language's ORCs compared to its SRCs. They attempt to eliminate this confound by providing

contexts which lessen the ambiguity, and after this adjustment they indeed find an OGA. But, as Kwon et al. (2010) discuss, there are several issues with Ishizuka et al.'s context sentences that cast doubt on the findings. Furthermore, Ishizuka et al. did not replicate the OGA result in subsequent experiments (Kwon et al. 2010:563, fn. 12), and one of Kwon et al.'s 2010 experiments, which utilized similar disambiguating contexts for Korean, also did not find an OGA. Overall, then, data from both Korean and Japanese support a Structural Theory for the processing of prenominal relative clauses.

Processing work on Chinese relative clauses has yielded mixed results (see discussion in Kwon et al. 2010, Vasishth et al. 2013, Wagers et al. 2018), some studies finding an SGA (Lin 2006, 2008, Lin and Bever 2010, 2011; Vasishth et al. 2013, experiments 1–2), others an OGA Hsiao and Gibson 2003, Hsu and Chen 2007, Lin and Garnsey 2010, Gibson and Wu 2013; Vasishth et al. 2013, experiment 3). For instance, both Hsiao and Gibson (2003) and Lin and Bever (2010) compare reading times of SRCs and ORCs, using stimuli very similar to the example in (34a). Hsiao and Gibson (2003) find a significant difference in RTs only at the second word of the relative, with SRCs being slower (they also test doubly-embedded relatives, which showed comparable RT patterns). They interpret this as an OGA, in line with the predictions of the Distance Theory. But Lin and Bever (2010) did not replicate these results; instead, they found a consistent SGA starting at the right edge of the relative clause, a result more in line with the Structural Theory. In light of the muddy empirical landscape, (Vasishth et al. 2013) conduct a meta-analysis of fifteen RC processing studies on Chinese, including three novel experiments. They conclude that the evidence favors an SGA for the language, and that apparent OGA findings are likely due to

the local structural ambiguities inherent to SRCs that modify matrix subjects.

The languages discussed so far have rigid relative clause placement: there is no optionality as to the linear order of RC and head noun. There are languages, though, which admit both prenominal and postnominal relatives. Chamorro, a verb-initial Austronesian language of the Mariana Islands, is an example. Besides flexible RC position, Chamorro also boasts *wh*-agreement, a special type of morphology that registers the syntactic role of an extracted element (Chung 1994, 1998). *Wh*-agreement is optional on relative clause verbs, but when it appears it reliably indicates the position of the gap. Two Chamorro examples, with and without *wh*-agreement, are given in (35).

(35) a. *i* [$_{RC}$ *k‹um›uentutusi* *yu'* ___ *nigap*] *na palåo'an*

DET ‹WH.SUBJ›speak:PROG 1SG yesterday LK woman

'the woman [$_{RC}$ who ___ was speaking to me yesterday]' **SRC, *wh*-agr**

b. *i* *lalåhi* [$_{RC}$ *ni* *ma* *kakassi* ___ *i* *ma'estra*]

DET men COMP AGR tease:PROG DET teacher

'the men [$_{RC}$ that ___ were teasing the teacher]' **SRC, no *wh*-agr.**

(Chamorro; glosses adapted from Wagers et al. 2018:210)

Wagers et al. (2018) conducted a picture-matching experiment that took advantage of these morphosyntactic properties. They compared prenominal, postnominal, and headless relative clauses, which were either ambiguous (one condition: transitive without *wh*-agreement), or unambiguous (two conditions: passive, or transitive with *wh*-agreement). Stimuli were presented auditorily and participants used tablet computers to select a picture

most appropriate for the item. Ambiguous RCs were assigned subject-gap parses more frequently than object-gap parses, but this preference was much stronger for postnominal RCs than prenominal RCs. In the unambiguous conditions, the error rate mirrored trends in disambiguation. Among RCs disambiguated by object *wh*-agreement, the most errors occurred in postnominal relatives, which were most frequently associated with an SRC parse in the ambiguous conditions; prenominal relatives, which were more likely to be parsed with object gaps when ambiguous, had the most errors among RCs disambiguated by subject *wh*-agreement. However, in terms of latency subject-gap interpretations corresponded to the earliest responses, even in conditions where object-gap responses were more common.

According to Wagers et al., results from this experiment indicate that a constellation of parsing principles is at work in Chamorro, working together to maximize incremental well-formedness. Within a relative clause, the parser will have several dependencies to satisfy: the movement relationship between the filler and its gap, the φ-agreement relationship between the subject and the verb, and potentially also a *wh*-agreement relationship linking a verb and a gap. In a postnominal RC, these dependencies unfold in such a way to strongly favor a subject-gap parse: very early in the relative, at the verb region, the parser can satisfy two dependencies (filler–gap and subject–verb agreement) by projecting a gap in subject position and associating it with the head noun they just encountered. In a prenominal relative, though, this is not possible. The complementizer–verb sequence signals the very same dependencies, but having not yet encountered the head noun, the parser cannot satisfy the outstanding filler–gap and agreement dependencies. Instead, parsing continues, dependencies unresolved, until the RC coargument is encountered. If this coargument has φ-features

matching verb agreement, choosing to link it to subject position will be the most economical parsing decision, since that satisfies the agreement dependency. Later on in the string, the relative clause ends and the parser encounters the head noun. Only then can the filler–gap dependency be satisfied: and since the coargument has already taken the subject position, the next best option (assuming the Accessibility Hierarchy) is to put a gap in object position.

So, while Wagers et al. do find an OGA in prenominal relative clauses (or at least an attenuated SGA), they do not interpret this as evidence in favor of the Distance Theory. Instead, a confluence of factors — some language-specific (like idiosyncratic morphological cues) and others apparently language-general (like the preference for subject gaps) — guide relative clause processing. The challenge of integrating multiple dependencies at staggered time points is one we return to in the discussion section for Experiment 2a (Section 3.4.3).

3.2.2 The Subject-Gap Advantage in Ergative languages

Ergative languages have increasingly been the subject of psycholinguistic investigation (Longenbaugh and Polinsky 2016). Because they associate informative morphology with S_{TR}s rather than O_{DO}s, Ergative languages can tease apart the Structural and Case Informativity Theories of the SGA. However, the empirical picture that has emerged for the processing of transitive relative clauses in ergative languages is mixed: some studies report an SGA (Clemens et al. 2015), others an OGA (Carreiras et al. 2010, Tollan et al. 2019), and still others find evidence for both effects (Polinsky et al. 2012, Longenbaugh and Polinsky 2016).

The earliest studies on the interaction of relative clause processing and ergativity

are on Basque, where Carreiras et al. (2010) claim to find evidence for an absolutive-gap advantage (i.e., an OGA in transitive RCs). In two SPR studies and one ERP experiment, they compare prenominal relatives like those in (36). Their stimuli are designed to take advantage of a particular quirk of Basque morphology: a noun's definite absolutive plural form (suffixed with *–ak* 'DEF.ABS.PL') is homophonous with its definite ergative singular form (*–a-k* 'DEF-ERG'). Notice that their SRC and ORC conditions are string-identical until the very last word of the sentence. There, the agreement on an auxiliary (*have* or *be*) disambiguates the number and case features of the matrix subject — and since the matrix subject is the RC head noun, this auxiliary also disambiguates the features and position of the relative clause gap, several words back.

(36) a. [$_{RC}$ ___ *Irakasle-ak* *aipatu* *ditu-en*] *ikasle-a-k*

teacher-ABS.PL mention have:3SG≫3PL-REL student-DEF-ERG

lagun-ak *ditu.*

friend-ABS.PL have:3SG≫3PL

'The student [$_{RC}$ that ___ mentioned the teachers] has friends.' **S_{ERG}RC**

b. [$_{RC}$ *Irakasle-a-k* ___ *aipatu* *ditu-en*] *ikasle-ak*

teacher-DEF-ERG mention have:3SG≫3PL-REL student-ABS.PL

lagun-ak *dira.*

friend-ABS.PL be:3PL

'The students [$_{RC}$ that the teacher mentioned ___] are friends.' **O_{ABS}RC**

(Basque; glosses adapted from Carreiras et al. 2010:82)

The authors found significantly higher reading times in the SRC condition than in the ORC condition at the disambiguating sentence-final auxiliary. The EEG study corroborated these findings: at the same region, the SRC condition had a significantly larger P600, an ERP linked to syntactic processing difficulty. The authors interpret these results as an OGA, and suggest that something like a 'Case Accessibility Hierarchy' (cf. Moravcsik 1974) guides Basque parsers: gaps associated with absolutive case (whether subjects or objects) are more accessible than ones associated with ergative case.

However, as Clemens et al. (2015) discuss, there are at least two reasons to be wary of this conclusion. First, it could be that the processing difficulty observed at the SRC auxiliary isn't due to the gap site of relative clause, but rather the argument structure of the matrix clause. In the reported materials, the ORC conditions always contain an intransitive, copular matrix auxiliary (*be*), while the SRCs have a transitive auxiliary (*have*). It is plausible that the transitive argument structure of have is more difficult to process than the intransitive copular structure of be, and that this difference explains the asymmetry in the SRC condition. Second, the apparent ORC preference could stem from a morphological disambiguation preference, assuming Basque speakers, confronted a noun ambiguous between ERG.SG and ABS.PL, are more inclined to parse it as ERG.SG. This seems like a reasonable hypothesis — corpus data indicate ERG.SG nouns are indeed more common than ABS.PL ones in Basque (Austin 2007; via Clemens et al. 2015:428, fn. 8). Given such an inclination, participants will be more likely to choose an ORC parse during the first four words of the stimulus. Increased processing difficulty at the disambiguating sentence-final auxiliary could simply indicate a garden path effect, since that auxiliary forces the parser to

revise the original ORC parse. For these reasons, Carreiras et al.'s findings do not unequivocally support their claim that Basque parsers are guided by a Case Hierarchy.

Another set of processing studies has been conducted on Avar ((Polinsky et al. 2012, Polinsky 2016)), a Northeast Caucasian language with prenominal RCs. Data from these experiments support both the Structural and Informativity Theories: within RCs, informative morphology on ergative coarguments leads to processing difficulty; at the head nouns, structural and cue-based factors cancel each other out, leading to an apparent null result across transitive conditions.

(37) [$_{RC}$ ___ *ʕoloqana-y yas repetici-yal-de y-ač:-un y-ač'-ara-y*
unmarried-II girl(II).ABS rehearsal-OBL-LOC II-bring-GER II-come-PTCP-II
] *artistika bercina-y y-igo.*
actress(II).ABS beautiful-II II-AUX
'The actess [$_{RC}$ that ___ brought the young girl to the rehearsal] is pretty.' **S_{ERG}RC**
(Avar; gloss adapted from Polinsky et al. 2012:271)

Using a self-paced reading task, Polinsky et al. (2012) compared relative clauses with gaps in ergative-marked S_{TR} position, absolutive-marked O_{DO} position, and absolutive-marked S_{IN} position. They find two important effects. First, the absolutive coargument of an S_{ERG}RC was read significantly faster than either the ergative coargument of an O_{ABS}RC or the oblique coargument of an S_{ABS}RC. This conforms to the predictions of the Informativity Theory — absolutive is the least informative case value, and therefore should be easiest to integrate into the structure. Note that the slowdown here occurs even before the parser

has reason to anticipate a relative clause dependency; two words in, the sentence is still compatible with a declarative root-clause interpretation.

The second important finding occurs at the first spillover region after the head noun. Here the intransitive S_{ABS}RC condition is read significantly faster than either of the conditions involving a transitive RC, which are not read at significantly different speeds. Polinsky et al. 2016:179–178 describes a picture-matching experiment on Avar that replicates this finding: response latencies are shorter for S_{ABS}RC than either S_{ERG}RC or O_{ABS}RC conditions. To explain this difference, Polinsky et al. appeal to an interaction of structure and case informativity. On the one hand, if subject gaps are intrinsically better than object gaps, no matter their case value, we should expect the S_{ERG}RC to be read faster than the O_{ABS}RC. On the other hand, if morphological cues facilitate predictive parsing, we expect the reverse. That is because in an O_{ABS}RC, the ergative coargument back at the second word of the relative prompted the parser to anticipate a syntactic position for an absolutive object. Since this absolutive position has already been projected, it is relatively easy to link it to the head noun. As for the S_{ERG}RC, its absolutive coargument does not necessarily license a structural position for an ergative element. Therefore, the parser is burdened with projecting a transitive subject position, and also satisfying the relative clause dependency by linking the head noun to that position. In other words, the Accessibility Hierarchy favors the S_{ERG}RC (whose gap is a subject) but predictive parsing principles favor the O_{ABS}RC (whose structural position was licensed several words back). If weighted roughly equally, these factors will cancel each other out across the two transitive conditions. As for the S_{ABS}RC condition, both principles are on the side of a gap in absolutive subject position. Thus Polinsky et al.

explain this second effect by appealing to both the Structural and Case Theories.

A very similar set of results has been observed for Niuean (Austronesian), which is the subject of a picture-matching task conducted by Longenbaugh and Polinsky (2016). Participants were presented with illustrations depicting characters interacting in various ways. Auditory stimuli prompted participants to select one of the characters; these stimuli consisted of questions containing relative clauses of various types, as in (38). Answers containing intransitive relative clauses were answered significantly faster than ones containing transitive relative clauses. S_{ERG}RC and O_{ABS}RC response latencies, though, were not significantly different. This directly mirrors the results for Avar, where intransitive conditions were consistently easier to process than transitive ones.

(38) a. *Ko fe e kulī* [$_{RC}$ *ne epoepo* ___ *e puti?*]

where ABS dog NFUT lick ABS cat

'Where is the dog [$_{RC}$ that ___ is licking the cat?]' **S_{ERG}RC**

b. *Ko fe e puti* [$_{RC}$ *ne epoepo he kulī* ___*?*]

where ABS cat NFUT lick ERG dog

'Where is the cat [$_{RC}$ that the dog is licking ___?]' **O_{ABS}RC**

(Niuean; Longenbaugh and Polinsky 2016:111–112)

Tollan et al. (2019) observe a slightly different pattern for Niuean. Using the visual-world paradigm, they tracked eye movements during subject- or object-extracted *wh*-questions like those in (39). Note that these are ambiguous for extraction site up until the case-marked coargument. Each experimental trial began with an auditorily presented con-

text sentence that established which figures in the visual world acted on which others. Participants then heard a *wh*-question asking about one of those figures. A second experimental factor besides gap site was the verb's case frame / argument structure: the question contained either a transitive verb with an ergative subject and an absolutive object (an $S_{ERG} \gg O_{ABS}$ case frame), a transitive verb with quirky an $S_{ABS} \gg O_{OBL}$ case frame, or an intransitive verb with an ABS subject and an OBL adjunct.

(39) a. *Ko e pusi fē ne tutuli tūmau ___ e lapiti?*
PRED cat which PAST chase always ABS rabbit
'Which cat ___ always chases the rabbit?' **ERG-subject *wh*-gap**

b. *Ko e pusi fē ne tutuli tūmau he kulī ___?*
PRED cat which PAST chase always ERG dog
'Which cat does the dog always chase ___?' **ABS-object *wh*-gap**

(Niuean; Tollan et al. 2019:4)

During the verb+adverb sequence (when the extracted argument's structural position is still ambiguous), Tollan et al. observe the following eye-movement patterns. In conditions with $S_{ERG} \gg O_{ABS}$ transitive verbs, there were significantly fewer looks to visual-world figures compatible with a subject-question parse than there were in conditions with either transitive or intransitive $S_{ABS} \gg X_{OBL}$ verbs. The authors interpret this as an absolutive-gap advantage, since Niuean speakers are inclined to anticipate that a gap will be in a position associated with absolutive case during all three conditions. This corresponds to a subject gap for questions with $S_{ABS} \gg X_{OBL}$ case frames (hence the high proportion of subject-

compatible gazes in these conditions), but an object gap for questions with $S_{ERG} \gg O_{ABS}$ case frames (hence the lower proportion of such gazes in this condition). Tollan et al. (2019) do not attempt to reconcile their eye-tracking results in Niuean with Longenbaugh and Polinsky's (2016) picture-matching data, but the empirical discrepancy between the methodologies is an intriguing one.

Complicating the picture on RC processing in Ergative languages further is Clemens et al.'s (2015) study on Ch'ol and Q'anjob'al, two verb-initial head-marking Mayan languages. Unlike the previously discussed experiments, these authors find a clear subject-gap advantage for Ch'ol and Q'anjob'al, with no evidence for an absolutive-gap advantage.

Clemens et al. conduct two picture-matching experiments, one on each Mayan language. Auditory stimuli contained relative clauses which were either ambiguous for gap-site, RCs which were biased towards a particular gap-site interpretation by an animacy discrepancy between the head noun and RC coargument, or (for Q'anjob'al) RCs where were structurally unambiguous. Ch'ol relative clauses involving two 3SG arguments are always structurally ambiguous (40), but Q'anjob'al's relatives are typically unambiguous; this language requires special 'agent-focus' morphosyntax when a transitive subject undergoes $\bar{A}$-movement, as (41) shows.

(40) *Ta' jul-i jiñi x'ixik* [$_{RC}$ *ta'-bä i-tsäk'-ä jiñi wiñik.*]

SC arrive-INTR DET woman PFV-REL 3.ERG-cure-TR DET man

'The woman [$_{RC}$ that ___ cured the man] arrived.' **S_{ERG}RC**

or 'The woman [$_{RC}$ that the man cured ___] arrived.' **O_{ABS}RC**

(Ch'ol; Clemens et al. 2015:437)

(41) a. *Max jay ix ix* [$_{RC}$ *max h-el-a'* ○ ___.]

PFV arrive DET woman PFV 2.ERG-see-TR 2SG

'The woman [$_{RC}$ that you saw ___] arrived.' **O_{ABS}RC**

b. *Max jay ix ix* [$_{RC}$ *max-ach il-**on**-i* ___ ○.]

PFV arrive DET woman PFV-2.ABS see-AF-INTR 2SG

'The woman [$_{RC}$ that ___ saw you] arrived.' **S_{TR}RC: agent focus necessary**

(Q'anjob'al; Clemens et al. 2015:438)

For both Ch'ol and Q'anjob'al, accuracy and response latencies indicate an SGA: ambiguous RCs elicited SRC-congruent responses a majority of the time, responses to SRC-biased transitive RCs were more accurate than responses to ORC-biased ones, and response times were shortest after relative clauses compatible with a subject-gap parse. As for the structurally unambiguous conditions in Q'anjob'al, these items elicited SRC-compatible and ORC-compatible responses in roughly equal proportion. This is a surprising result, given that the presence of an agent-focus morpheme in a transitive relative clause should only be compatible with an SRC interpretation, and its absence with an ORC interpretation. The authors suggest that this result may stem from a preference to extract O_{DO}s as passive

subjects rather than active objects, and therefore a transitive ORC is relatively unusual. Alternatively, it may indicate that the extraction asymmetry in (41) is not as strict as has been previously reported for the language.

Setting aside the issue regarding the Q'anjob'al agent-focus construction, these experiments demonstrate a clear SGA, results most compatible with the Structural Theory of relative-clause processing. So why didn't the Mayan languages exhibit the kind of ergative penalty observed in Avar, Basque, and (possibly) Niuean? Clemens et al. propose that head-marking morphology (like the ergative-aligned verbal agreement in Mayan) is less useful for incremental processing than dependent-marking morphology (like the ergative-aligned case system of Avar, Basque, and Niuean). Indeed, if the head noun and RC coargument have identical φ-features in Ch'ol, RC-internal verbal agreement plays no disambiguating role; both DPs in the clause could potentially control either the ergative or absolutive agreement affixes. Dependent marking, on the other hand, is instrumental for assigning arguments their structural positions, given the tight correlation of argument structure and presence of particular case categories. In the absence of case cues, as in Mayan, it seems the parser must default to more general parsing strategy, such as using the Accessibility Hierarchy.

3.3 Background on Georgian

This section details two areas of Georgian grammar relevant to the present study (see Aronson 1990 for a detailed description of the language). The first is the Split-Ergative case system, which strongly dissociates case morphology and syntactic role. As described

in Section 3.3.1, a given argument may appear in different cases depending on the tense–aspect–mood (TAM), and each of the three core case categories varies in which syntactic roles, and how many, they map to. The processing profile of this complex case system is the focus of Skopeteas et al. (2012), whose results are summarized in Section 3.3.2. In Section 3.3.3 I turn to Georgian's relativization constructions. This study employs three such constructions — *rom*-relatives, *rom*-correlatives, and *wh*-relatives — which allow us to manipulate both the linear order of the relative clause and the head noun it modifies, and also the informativity and distribution of morphological cues within the relative clause itself. Finally, Section 3.3.4 lays out the predictions made by the Structural, Case Informativity, and Distance hypotheses for Georgian relative clauses of various types.

3.3.1 Case alignment

Georgian has a complex TAM-conditioned Split-Ergative case system (Harris 1985, Nash 2017b). Practically speaking, this means that different TAMs are associated with different verbal-argument case frames. For example, in the future we find that transitive subjects are assigned nominative while direct objects are assigned dative (42a); in the aorist, we see ergative subjects and nominative objects (42b); in the perfect, dative subjects and nominative objects (42c). (For the present purposes it suffices to consider only case-assignment facts in normal monotransitive clauses; for a detailed description which also discusses the behavior of intransitive subjects, experiencer subjects, and indirect objects, see Section 1.2, Section 2.2, or Table 2.1.)

(42) Monotransitive case frames in Georgian

a. *datv-i* *çav-s* *dainaxavs.*

bear-NOM otter-DAT catch_sight:FUT.3SG

'The bear will catch sight of the otter.' **NOM–DAT case frame**

b. *datv-ma* *çav-i* *dainaxa.*

bear-ERG otter-NOM catch_sight:AOR.3SG

'The bear caught sight of the otter.' **ERG–NOM case frame**

c. *datv-s* *çav-i* *daunaxavs.*

bear-DAT otter-NOM catch_sight:PERF.3SG

'The bear [apparently] caught sight of the otter.' **DAT–NOM case frame**

(Constructed examples, after Rayfield 2006:494, 503, 1562)

The following table summarizes the patterns. What's especially pertinent is the asymmetrical distribution of the three case categories. Only subjects in the aorist (or optative) will bear ergative case; nominative and dative are found on both subjects and objects, though in different TAMS. Encountering ergative morphology, then, will be especially informative for the Georgian comprehender, since it is incompatible with a large swath of possible continuations.

3.3.2 Case processing in Georgian

With the aim of investigating how Georgian's Split-Ergative case system is processed, Skopeteas et al. (2012) conduct two grammaticality judgement experiments manipulating word order and case alignment. Stimuli consisted of written three-word sentences,

	S_{TR}	O_{DO}	Alignment Pattern
Future	NOM	DAT	NOM–ACC
Aorist	ERG	NOM	ERG–ABS
Perfect	DAT		DAT–ABS

Table 3.2: Case assignment in monotransitive clauses across TAM categories in Georgian

in either SOV or OSV word order, presented incrementally on a computer display. The first two words, the noun-phrase arguments, were presented by themselves for 5,000 ms. Then, the verb appeared, and participants were asked to judge the sentence as grammatical or ungrammatical. (All experimental items were grammatical, but fillers were either grammatical or ungrammatical.)

Both experiments compared verbs with $S_{NOM} \gg O_{DAT}$ case frames to ones with $S_{DAT} \gg O_{NOM}$ frames. For the first experiment these were transitive verbs in either NOM–ACC- or DAT–ABS-aligned TAMs. A main effect of case alignment affecting response latencies obtained in this experiment, with DAT–ABS sentences leading to significantly slower responses than NOM–ACC ones. The second experiment compared $S_{DAT} \gg O_{NOM}$ experiencer-subject verbs to $S_{NOM} \gg O_{DAT}$ applied intransitives. Here, the authors found main effects of case alignment (responses to experiencer verbs being slower), word order (OSV being slower), and a significant interaction between the two factors; all three effects were driven by the dramatically slower response latencies to $O_{NOM}S_{DAT}V$ sentences with experiencer verbs.

Skopeteas et al. conclude the following. First, absent any disambiguating information, Georgian speakers are biased to parse nominative DPs as subjects and dative DPs as objects — even though the language has nominative objects and dative subjects. This explains why $S_{DAT} \gg O_{NOM}$ case frames take longer to process than $S_{NOM} \gg O_{DAT}$ ones. This bias persists even when a dative DP is linearly first, suggesting that word order is a lower-ranked cue to grammatical role than case is. Second, upon encountering a verb that does require a dative subject, revising the links between case and role is easier for stimuli containing experiencer verbs than for stimuli containing Dat–Abs transitive verbs. Moreover, the authors observe that unscrambled SdatOnomV clauses are relatively easy to process given a experiencer verb (at least compared to scrambled OnomSdatV clauses), but both word orders are difficult to process given a Dat–Abs transitive verb. The obvious difference between experiencer verbs and Dat–Abs transitive verbs is that the former license their dative subjects lexically (i.e., in all tams), whereas the latter do so syntagmatically, by virtue of the mechanism enforces the language's Split Ergativity. However, it remains an open question just why lexical dative subjects have a privileged processing status.

3.3.3 Relativization strategies

Georgian has an impressive array of relativization strategies (Foley 2013, Nash 2017a, Bhatt and Nash 2018). Three types of relative clauses are relevant for the present experiment. First is the *wh*-relative. As in English, a *wh*-relative involves the $\bar{A}$-movement of a *wh*-phrase relative pronoun. This *wh*-phrase appears at the left edge of the relative clause, hosts the relativizing enclitic *–c* 'rel', and bears the case morphology associated

with the corresponding gap site. Any verbal argument relativized using the *wh*-relative strategy, as can a wide variety of adjuncts.

(43) *vnaxe* (*is*) *mxaṭvar-i* [$_{RC}$ *romel-ma-c* *gušin* ___ *mçeral-s*
see:AOR.1SG (DEM) painter-NOM which-ERG-REL yesterday writer-DAT
momġeral-i *gaacno.*]
singer-NOM introduce:AOR.3SG≫3
'I saw that painter [$_{RC}$ who ___ introduced the singer to the writer yesterday.]'
(Constructed example based on fieldnotes; MN 170804)

Second is the *rom*-relative, which resembles a *that*-relative in English. *Rom*-relatives are postnominal, involve Ā-movement of a null operator, and contain the declarative complementizer *rom*. While *rom* appears at the left edge of complement clauses, the complementizer has a different distribution in relatives and certain other subordinate clauses. As (44) shows, *rom* may appear in any non-initial position, as long as it is before the verb and does not split up a constituent. As for the gap, it may appear in any argument position and at least some adjunct positions. Consultants often prefer a demonstrative to appear with the head noun of a *rom*-relative, but this element does not seem to be absolutely obligatory.

(44) *vnaxe* ?(*is*) *mxaṭvar-i* [$_{RC}$ {**rom*} *gušin* {*rom*} *mçeral-s*

see:AOR.1SG ?(DEM) painter-NOM {*COMP} yesterday {COMP} writer-DAT

{*rom*} *momġeral-i* {*rom*} *gaacno* {**rom*}.]

{COMP} singer-NOM {COMP} introduce:AOR.3SG≫3 {*COMP}

'I saw that/the painter [$_{RC}$ that ___ introduced the singer to the writer yesterday.]'

(Constructed example based on fieldnotes; MN 170804)

A third relativization strategy is the *rom*-correlative. Broadly speaking, a correlative is a species of adjunct clause which introduces a referent (or multiple referents) that is picked up in the matrix clause by anaphoric proform (Bittner 2001:39; see also Srivastav 1991, Dayal 1996, Lipták 1996). If the proform is a demonstrative, appearing within a matrix-clause DP, then the correlative has a very similar function to a relative clause. One kind of correlative in Georgian is illustrated in (45).[2] Like a *rom*-relative, it contains the non-initial complementizer *rom*, and a gap in argument position formed by $\bar{A}$-movement of a null operator. But unlike a *rom*-relative, it is separated from the head noun, appearing at the left periphery of the matrix clause, and a demonstrative is obligatory on the head noun. The correlative can be translated roughly into English with a left-dislocated hanging topic containing a relative clause.

[2]Argument-modifying correlatives with other shapes also exist, for instance head-internal ones. See Bhatt and Nash (2018) for more details.

(45) [CORREL {*rom} gušin {rom} mçeral-s {rom} momġeral-i {rom}

{*COMP} yesterday {COMP} writer-DAT {COMP} singer-NOM {COMP}

gaacno {*rom},] vnaxe *(is) mxaṭvar-i

introduce:AOR.3SG≫3 {*COMP} see:AOR.1SG (*DEM) painter-NOM

≈ '[CORREL The one that ___ in the singer to the writer yesterday,] I saw that painter.'

(Constructed example)

It seems that both the position of a correlative-internal gap and the position of the matrix-clause correlate noun phrase are constrained by subjacency; neither can appear within an island like a relative clause, for example. This is evidence for movement both of a null operator within the correlative, and also movement of the correlative itself from a position more local to the correlate (cf. Bhatt and Nash 2018 for a more nuanced view taking into account more types of correlatives from Georgian).

(46) a. [CORR Dato rom akebda,] Nino-m is msaxiob-i supra-ze

Dato.NOM COMP praise Nino-ERG DEM actor-NOM feast-on

gaicno.

met

'[CORR The one that Dato was praising ___,] Nino met **that actor** at the feast.'

b. *[$_{\text{CORR}}$ *Dato-s rom uqvars kal-i* [$_{\text{RC}}$ *romelic akebda,*]]

Dato-DAT COMP loves woman-NOM which:NOM praised

Nino-m is msaxiob-i supra-ze gaicno.

Nino-ERG DEM actor-NOM feast-on met

Attempted: '[$_{\text{CORR}}$ The one that Dato loves the woman [$_{\text{RC}}$ who was praising ___,]] Nino met **that actor** at the feast.'

c. *[$_{\text{CORR}}$ *Dato rom akebda,*] *Levan-s uqvars kal-i,* [$_{\text{RC}}$

Dato.NOM COMP praised Levan-DAT loves woman-NOM

romelmac is msaxiob-i supra-ze gaicno.]

which:ERG DEM actor-NOM feast-on met

Attempted: '[$_{\text{CORR}}$ The one that Dato loves ___,] Levan loves the woman [$_{\text{RC}}$ who met **that actor** at the feast.]'

(Fieldnotes, MN 180313)

This particular correlative construction will be useful in the present study, as it is maximally parallel to a *rom*-relative: at least in terms of strings of words, the two differ only in the position of the subordinate clause. Thus, throughout this chapter we will refer to both correlatives and relative clauses proper as 'relatives', 'SRCs', or 'ORCs' for the sake of terminological simplicity. It is important to keep in mind, though, that correlatives are not simply prenominal relatives; indeed, the syntax and semantics of correlativization and relativization are importantly distinct (Dayal 1996, Bhatt 2003, Bhatt and Nash 2018). It may be, then, that their processing profiles are also distinct too. But insofar as both types

of constructions involve Ā-movement, and require the parser to link a filler to a gap — whatever linear order the filler and gap appear in — I will assume that they are sufficiently similar to warrant direct comparison in Experiments 2a and 2b.

One final caveat: Georgian noun-modifying correlatives are superficially similar to a range of adjunct clauses, including one strategy for forming when- and because-clauses (47a), and also counterfactual conditionals (47b). These clauses are also left-peripheral and contain a non-initial rom. And while they do not contain gaps in argument position, they may contain null pronouns. Consequently, these adjunct clauses may be locally or even globally ambiguous with a noun-modifying correlative (48).

(47) a. [$_{CP}$ *laṭaria rom moigo,*] *mcxobel-ma šeçq̇viṭa mušaoba.*
lottery.NOM COMP win:AOR.3SG baker-ERG quit:AOR.3SG work.NOM
'[$_{CP}$ When they won the lottery,] the baker quit their job.'
or '[$_{CP}$ Because they won the lottery,] the baker quit their job.'

b. [$_{CP}$ *laṭaria rom moego,*] *mcxobel-i šeçq̇veṭda mušaoba-s.*
lottery.NOM COMP win:PLU.3SG baker-NOM quit:COND.3SG work-DAT
'[$_{CP}$ Had they won the lottery,] the baker would have quit their job.'

(After fieldnotes; MN 180228)

(48) a. [$_{CP}$ ◌ *laṭaria rom moigo,*] *im mcxobel-ma*

3SG.ERG lottery.NOM COMP win:AOR.3SG DEM baker-ERG

šeçq̇viṭa mušaoba.

quit:AOR.3SG work.NOM

'[$_{CP}$ When/because they won the lottery,] that baker quit their job.'

Adjunct clause with a null pronoun

b. [$_{CP}$ ___ *laṭaria rom moigo,*] *im mcxobel-ma šeçq̇viṭa*

lottery.NOM COMP win:AOR.3SG DEM baker-ERG quit:AOR.3SG

mušaoba.

work.NOM

'[$_{CP}$ The one that won the lottery,] that baker quit their job.'

Noun-modifying correlative with a gap

(After fieldnotes; MN 180228)

It is an open question whether the similarity between these adjunct clauses and noun-modifying correlatives is skin deep, or if it indicates a deeper syntactic parallel. One possibility is that the adjunct clause are essentially correlatives with gaps in adjunct position. The connection between correlatives and conditionals is a well-studied one (e.g., Izvorski 1996, Bhatt and Pancheva 2006), making this an attractive analytical possibility, but future research will be necessary to confirm it.

3.3.4 Predictions

To contextualize the design and results of the relative clause experiments, it will be useful to spend some time unpacking the predictions made by the Structural, Informativity, and Distance Theories for SRCs and ORCs in Georgian. Consider first a *wh*-relative. One possible *wh*-relative is schematized in (49): first comes a head noun, then a *wh*-phrase relative pronoun, the coargument, and the verb. Each of these DPs will bear a case suffix (κ_1–κ_3), and the verb will bear TAM morphology.

(49) HdN-κ_1, [$_{RC}$ *wh*P-κ_2 CoArg-κ_3 V-TAM]

Only the Informativity Theory predicts processing differences at the head noun. The more informative its case value (κ_1), the harder it will be to process, since informative cues will license the parser to anticipate a more specific structure for the matrix clause. Ergative head nouns, then, should be read more slowly than either nominative or dative ones. The same prediction applies at the relative pronoun and coargument, since their case morphology gives cues to the structure of the relative clause. There should be no processing differential at the relative clause verb. The relative clause will be unambiguously transitive by the time the parser encounters the verb, as two DPs (the relative pronoun and the coargument) will have already been incorporated into the structure.

The Structural Theory, on the other hand, predicts that the parser defaults to a subject-gap parse, and will only abandon that parse in the face of unambiguous evidence that the gap cannot be in subject position. As subjects can be nominative, ergative, or dative in Georgian, this means (all things being equal) that the value of κ_2 should not affect

processing; a relative pronoun in any case could in principle be linked to a subject gap. Of course, Skopeteas et al.'s (2012) findings complicate this picture slightly. If the Georgian parser is biased to treat any dative DP as an object, including a relative pronoun, then a dative relative pronoun would be associated with an object gap. If this is the case, the Structural Theory predicts dative relative pronouns to be read more slowly than nominative or ergative ones.

Moving on to the coargument, the only possible value of κ_3 which would necessitate a revision of an initial SRC parse as ergative. This is because a nominative *wh*P could indicate an SRC in the NOM–ACC alignment; ergative case, which can only appear on subjects, foils this hypothesis.

If dative relative pronouns are treated by the parser as compatible with a subject gap, then neither $wh\text{P}_{\text{NOM}}$–$\text{CoArg}_{\text{DAT}}$ nor $wh\text{P}_{\text{DAT}}$–$\text{CoArg}_{\text{NOM}}$ sequences will give the parser any strong reason to suspect the gap to be in a specific position; after all, there are both $\text{S}_{\text{DAT}} \gg \text{O}_{\text{NOM}}$ and $\text{S}_{\text{NOM}} \gg \text{O}_{\text{DAT}}$ clauses. So, it will not be until the verb that the string is disambiguated. Assuming an initial subject-gap parse, reading times will increase at a perfect verb that follows a $wh\text{P}_{\text{NOM}}$–$\text{CoArg}_{\text{DAT}}$ string, and at a future verb that follows a $wh\text{P}_{\text{DAT}}$–$\text{CoArg}_{\text{NOM}}$ string. Those verbs disambiguate to the ORC parse, dashing the parser's hopes. Of course, if $wh\text{P}_{\text{DAT}}$ initially incurs an object-gap penalty, it will actually be sequences of the shape $wh\text{P}_{\text{DAT}}$–$\text{CoArg}_{\text{NOM}}$–$\text{V}_{\text{PERF}}$ that require a reparse; $wh\text{P}_{\text{DAT}}$–$\text{CoArg}_{\text{NOM}}$–$\text{V}_{\text{FUT}}$ strings would conform to the initial (pessimistic) ORC prediction.

The tree diagram in Figure 3.1 schematizes these predictions (setting aside the possibility that dative is linked with objecthood). The Structural Theory predicts processing

difficulty at every ORC cue.

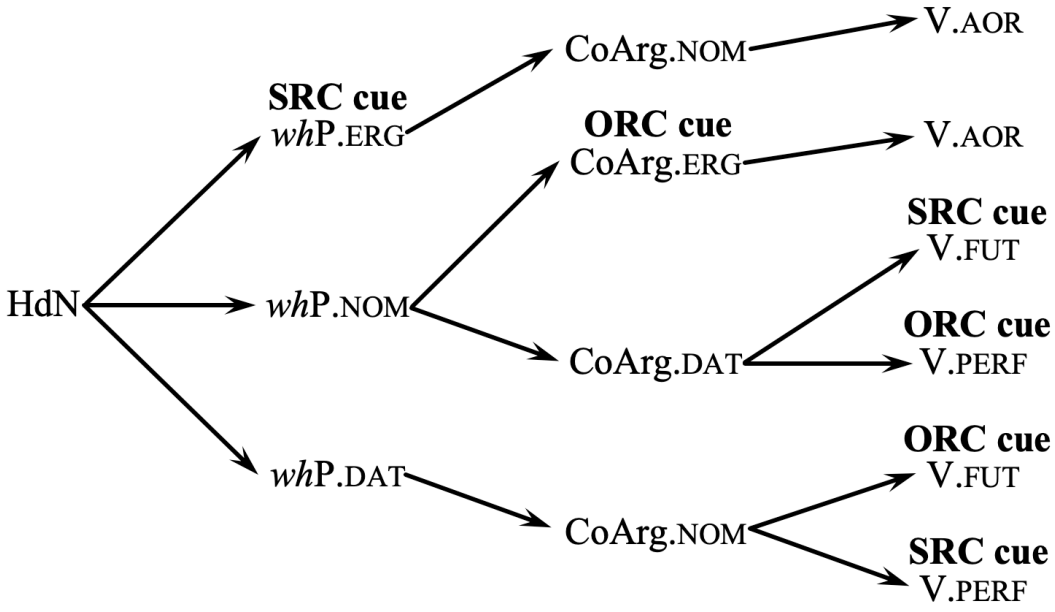


Figure 3.1: Incremental gap-site disambiguation points for a hypothetical *wh*-relative (49)

What about linearity? Assuming the distance between fillers and gaps is computed relative to the default SOV word order, the Distance Theory makes identical predictions to the Structural Theory for postnominal *wh*-relatives. Subject gaps will be preferred because they make for shorter dependencies than object gaps.

We now turn to a *rom*-relative, like the one schematized in (50). Our three theories make very similar predictions for this kind of relative as they did for the *wh*-relative. The only difference is that the gap's case and grammatical role must be triangulated from the coargument's case and the verb's TAM morphology. But this does not affect the predicted incremental processing differentials: points at which SRCs are eliminated (ergative coarguments and the verbs in FUT and PERF ORCs) will still be hurdles for parsers using

structural- or distance-based heuristics; head nouns and coarguments in the ergative will still be difficult if informativity leads to processing trouble.

(50) HdN-κ_1, [$_{\text{RC}}$ CoArg-κ_2 *rom* V-TAM]

Finally, let us consider the predictions for *rom*-correlatives. There are two types that should be considered, schematized below. These crucially differ in whether the correlative coargument is unambiguously in a subordinate structure or not. If it follows *rom* (51a), the coargument must belong to an embedded clause. If it precedes *rom* (51b), it is possible to initially parse that DP as a part of a root clause.

(51) a. [$_{\text{RC}}$ XP *rom* CoArg-κ_1 V-TAM] … DEM HdN-κ_2

b. [$_{\text{RC}}$ CoArg-κ_1 *rom* XP V-TAM] … DEM HdN-κ_2

In the case of the temporarily ambiguous coargument (51b), only the Informativity Theory predicts a difference in processing caused by the value of κ_1. The other theories only expect RT differences in structures that unambiguously involve filler–gap dependencies.

But at what point does a correlative unambiguously involve a filler–gap dependency, exactly? As shown above, Georgian correlatives can be string-identical to certain adjunct clauses — which do not involve gaps, at least not in argument position. One possibility is that the parser waits for unambiguous bottom-up evidence to posit a filler–gap dependency. This seems to be the case in English, as Staub et al. (2018) show. In a series of eye-tracking experiments, they compare the processing profile of DPs like those in (52).

(52) a. the information [$_{RC}$ that the health department provided ___]

b. the information [$_{CP}$ that the health department provided a cure]

Here the embedded clauses are temporarily ambiguous between a relative clause (i.e., a structure that involves a filler–gap dependency) and a complement clause (one that does not). Staub et al. observed consistent processing difficulty at regions where a string that had been temporarily compatible with a complement-clause parse was disambiguated as a relative clause. This effect obtained no matter how biased a particular lexical item was to occur with a complement or relative clause (as calculated by corpus frequencies). They interpret their results as support for the hypothesis that the parser avoids the maintenance of long-distance filler–gap dependencies whenever possible.

With that in mind, it may be the case the Georgian speaker assumes that a sentence-initial *rom*-clause is an adjunct clause — i.e., that it is not a correlative, which would involve an argument gap that needs to be matched to a filler. A garden-path effect will obtain at the point in the string when it becomes clear that the *rom*-clause was a correlative all along. Since Georgian allows null pronouns in all argument positions, the mere fact that an argument is missing from the *rom*-clause is not a sufficient signal. Instead, the demonstrative-modified head noun in the matrix clause is likely the cue that disambiguates the string towards a correlative parse. (Though even this cue might not be foolproof, as examples in 48 show.) If this hypothesis about how *rom*-clauses are interpreted is correct, the head noun will trigger a garden-path effect. And, since the head noun requires the empty category within the *rom*-clause to be reinterpreted as the gap of a filler–gap dependency, any costs

associated with the structural position of that gap, or the distance between it and the filler, will compound the difficulty of this garden path.

In other words, if the Structural Theory is on the right track, we expect the head nouns of object-gap correlatives to be read more slowly than those of subject-gap correlatives. The Distance Theory predicts the opposite, assuming that an object gap in a prenominal correlative structure counts as being closer to the head noun than a subject gap does.

3.4 Experiment 2a: Processing relatives & correlatives with *rom*

The goal of Experiment 2a was to compare the predictions of the Structural and Distance theories through contrasting postnominal *rom*-relatives and prenominal *rom*-correlatives. Changing the order of the head noun and relative clause affects the length of ORC and SRC dependencies. If dependency length is a driving factor in Georgian, then postnominal relatives should exhibit an SGA, while correlatives should exhibit an OGA. If structure is most important, though, an SGA will emerge in both orders.

3.4.1 Method

Participants

57 native Georgian speakers living in Tbilisi, Georgia (45 women, average age = 23) were recruited via social media. They were paid 40 GEL (approx. 13 USD) for their participation. One participant was excluded from subsequent analysis because they answered less than 70% of comprehension questions for the fillers correctly, and because their median RT was much slower than the rest (more than two standard deviations beyond the mean of

participant medians).

Materials

24 item sets were constructed in a 2×2 design, crossing Relative Clause Position (postnominal relative vs. prenominal correlative) and Gap Site (subject gap vs. object gap). These sentences conformed to the template in (53).

(53) a. Postnominal relative template (Experiment 2a)

D+HdN	[RC	CoArg+C^0	XP$_1$	XP$_2$	V	]		Spill$_1$	Spill$_2$	Spill$_3$	Spill$_4$
w$_1$		w$_2$	w$_3$	w$_4$	w$_5$			w$_7$	w$_8$	w$_9$	w$_{10}$

b. Prenominal correlative template (Experiment 2a)

[RC	CoArg+C^0	XP$_1$	XP$_2$	V	]	D+HdN	Spill$_1$	Spill$_2$	Spill$_3$	Spill$_4$
	w$_2$	w$_3$	w$_4$	w$_5$		w$_6$	w$_7$	w$_8$	w$_9$	w$_{10}$

The relative clause itself (w$_2$–w$_5$) consisted of a coargument and the complementizer rom (presented in a single SPR window, w$_2$), a two-word adjunct phrase, and a clause-final verb. The relative clause verb was always in the aorist, a TAM which triggers the ERG–ABS case alignment. Consequently, the coargument of the SRCs appeared in the NOM case, and the coarguments of the ORCs appeared in the ergative case. Matrix clause material included the head noun (appearing either at w$_1$ or w$_6$) and a four-word continuation (w$_7$–w$_{10}$) to capture potential spillover effects. The head noun was always the subject of the matrix clause, but its case was counterbalanced between nominative, ergative, and dative. Syntax and argument structure of the matrix clause varied across item sets. The animacy of the head noun and relative clause coargument were equal, and were counterbalanced across itemsets: half had human arguments and half animal. Examples in (54) give the RC-modified DP from

a representative item set.

(54) a. Postnominal SRC

is ***gogo,*** [$_{CP}$ ***biǯ-i*** *rom bnel ṭq̇e-ši naxa,*] …

DEM **girl.NOM** **boy-NOM** COMP dark forest-in saw

'the girl [$_{RC}$ that ___ saw the boy in the dark forest] …'

b. Postnominal ORC

is ***gogo,*** [$_{CP}$ ***biǯ-ma*** *rom bnel ṭq̇e-ši naxa,*] …

DEM **girl.NOM** **boy-ERG** COMP dark forest-in saw

'the girl [$_{RC}$ that the boy saw ___ in the dark forest] …'

c. Prenominal SRC

[$_{CP}$ ***biǯ-i*** *rom bnel ṭq̇e-ši naxa,*] ***is*** ***gogo*** …

boy-NOM COMP dark forest-in saw **DEM** **girl.NOM**

'[$_{CORR}$ the one that ___ saw the boy in the dark forest,] that girl…'

d. Prenominal ORC

[$_{CP}$ ***biǯ-ma*** *rom bnel ṭq̇e-ši naxa,*] ***is*** ***gogo*** …

boy-ERG COMP dark forest-in saw **DEM** **girl.NOM**

'[$_{CORR}$ the one that the boy saw ___ in the dark forest,] that girl…'

These experimental items were presented among 76 filler sentences, which comprised 36 experimental items for Experiment 2b (see Section 3.5), and 40 more sentences which did not contain relative clauses. Each of the 100 sentences was followed by a yes–no

comprehension question. All of the stimuli, in this experiment and the others, were constructed by the first author in consultation with three native speakers.

Procedure

Subjects participated online via Ibex Farm (Drummond 2016). Upon accessing the experiment, participants read a brief introduction describing the general purpose of the task, filled in demographic information, and consented to participation. To familiarize them with the self-paced reading task and experimental procedure, participants were presented with three practice items consisting of a sentence and a comprehension question. After this, the experiment proper began. The experimental items were distributed in a Latin Square, and randomized along with the fillers. Feedback was provided after each comprehension question. After finishing all 100 sentence–question pairs, an optional debriefing question appeared.

Analysis

Reading times and comprehension question response latency were analyzed using linear mixed-effects regression; question accuracy was analyzed using logistic mixed-effects regression. The Gap conditions were coded by two coefficients using centered sum contrasts: SRC (−½) and ORC (+½). Likewise for the Relative Clause Position conditions: postnominal relative (−½), prenominal correlative (+½). Unless otherwise stated, maximal random effects structure was included (Barr et al. 2013). Models were estimated using the `lme4` package in `R` (Bates et al. 2015); *t*-tests were calculated using Satterthwaite's method

via the `lmerTest` package in `R` (Kuznetsova et al. 2017).

3.4.2 Results

Figure 3.2 reports mean reading times for each SPR window, partitioned by relative clause position. The most striking effect is that ergative coarguments are read significantly more slowly than nominative coarguments for both prenominal correlatives and postnominal relatives. Results from linear mixed-effects model are given in Table 3.3 (random by-participant intercepts were removed because of convergence issues, but slopes were retained). Another significant effect emerges at the verb region; these results are shown in Table 3.4 (for convergence issues, the by-participant slope was removed). Here we see a main effect of relative clause type, with correlatives being faster. Finally, correlative head nouns in the subject-gap condition appear to be read slower than those in the object-gap condition, but this effect was not significant ($\beta = 55 \pm 44$, $t(29) = -1.2$, $p = 0.23$).

Turning to performance on comprehension questions, we found no effects of Gap or RC position on either accuracy or response time. The average accuracy was 88% and the average latency was 3587 ms. For experimental conditions, mean accuracy spanned a narrow range from 86% to 89%, and it was comparable for the fillers (88%).

3.4.3 Discussion

We observed a large RT slowdown when participants read the ergative coargument in both postnominal and prenominal RCs. This ergative coargument cost is consistent with the predictions of both the Case Informativity and Structural Theories. Ergative case is

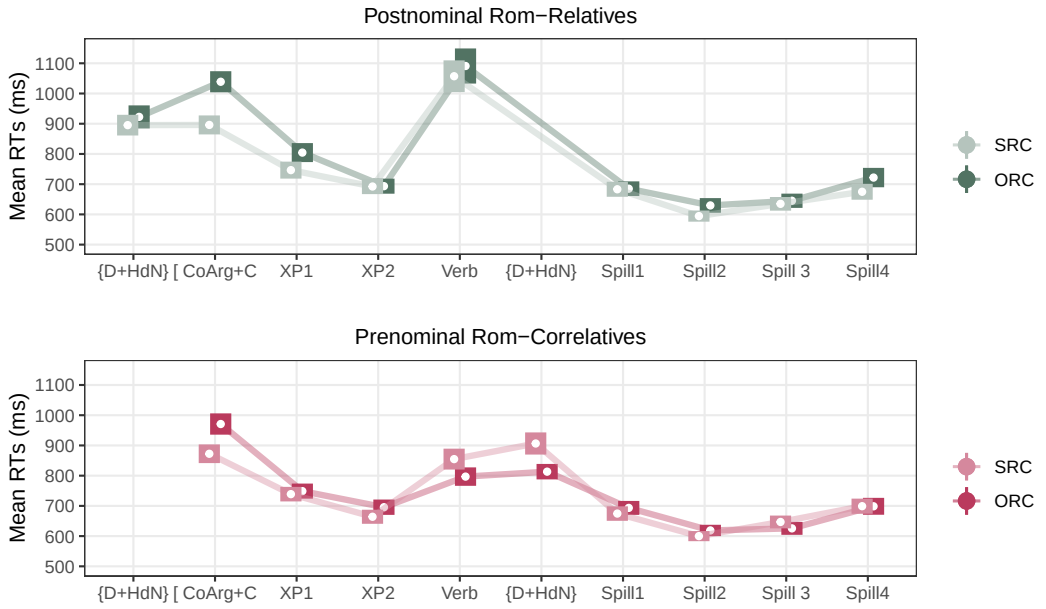


Figure 3.2: Mean reading times for Experiment 2a (*Rom*-relatives)

	β	SE	df	*t*	*p*
(Intercept)	962	51	67	19	< 0.001
Gap	114	31	104	3.7	< 0.001
RC Position	−40	31	22	−1.3	< 1
Gap:Position	−27	59	30	−0.45	< 1

Table 3.3: Experiment 2a (*Rom*-relatives) linear mixed-effects model for reading times at the relative-clause coargument (random by-participant intercepts removed).

highly informative, since it entails a specific argument structure and set of TAMS, and it also serves as an effective gap-site disambiguator, since an overt subject within a relative eliminates the possibility of an SRC parse.

It is notable that the ergative cost manifests even in the prenominal correlatives.

	β	SE	df	t	p
(Intercept)	970	64	56	15	< 0.001
Gap	-2.3	50	26	-0.047	< 1
RC Position	-250	65	45	-3.8	< 0.001
Gap:Position	-86	91	27	-0.94	< 1

Table 3.4: Experiment 2a (*Rom*-relatives) linear mixed-effects model for reading times at the relative-clause verb (by-participant slopes removed).

From the perspective of the Case Informativity Theory, this is entirely expected; ergative case is more informative than nominative in all syntactic environments. From the perspective of the Structural hypothesis, though, the effect is surprising, since the sequence DP+*rom* at the beginning of a sentence is not an unambiguous structural cue to a correlative. As discussed in Section 3.3.3, this string is also consistent with an adjunct clause, an environment where the preference for subject gaps over object gaps is presumably irrelevant. So, if gap position in the key factor determining processing cost here, it must be that parsers treat a DP+*rom* sequence as a correlative (a structure with an argument gap) right off the bat. However, this seems to challenge Staub et al.'s 2018 findings that parsers avoid positing filler–gap dependencies whenever possible. Therefore, the simplest explanation for the fact that an ergative coargument cost obtains in both postnominal relatives and prenominal correlatives is that the relative informativity of ergative case is taxing for the parser.

Of course, since the coargument and complementizer were presented in a single SPR window, Experiment 2a's correlatives were more like those schematized in Section 3.3 as (51a) than (51b). Decisive evidence in favor of the Informativity Theory would be if

the ergative cost lingers in a correlative whose coargument could be initially parsed as a root-clause argument. The design of Experiment 2b aims to test for just this possibility.

The second significant result of this experiment was that verbs in prenominal correlatives were read faster than verbs in postnominal relatives. To understand why this might be, let's first consider why relative-final verbs might be read slower than other regions in the first place. We suggest this could be a kind of integration cost, of the kind Wagers et al. (2018) discuss for Chamorro. Upon encountering a RC-final verb, the parser is confronted with a number of tasks: linking previously-ambiguous case morphology to the appropriate syntactic roles, integrating the arguments and adjuncts into the verb's argument structure and lexical semantics, conclusively resolving the filler–gap dependency, and shifting gears back to processing the matrix clause. The processing cost of these demands compound and cause the parser to slow down.

So why is there no similar slowdown in the correlative verbs? I hypothesize that the processing burden found at a single region of a postnominal RC is spread across two regions in a prenominal one. The verb of a correlative triggers the parser to disambiguate case morphology, integrate argument structure, and turn to a different clause, but the filler–gap dependency can only be partially resolved: the gap site may be disambiguated, but the lexical content associated with it will not be encountered until the next word, the head noun. And in fact, the head noun region of the correlatives of this experiment were read, at least impressionistically, slower than comparable regions in other conditions. This is compatible with the fact that cluster of processing tasks, which are resolved by a single word in postnominal RCs, are stretched across two words in prenominal RCs. Consequently the

processing burden is distributed across two words here.

3.5 Experiment 2b: Processing more (cor)relatives with *rom*

This being among the first studies on Georgian filler–gap processing, I sought to replicate the findings of the previous experiment in Experiment 2a. The main change to the design was that the complementizer rom was presented in its own SPR window, rather than in the same window as the relative clause coargument. This made the coargument of the prenominal relative clause temporarily compatible with a root-clause parse, thereby addressing the fact that Experiment 2a cannot adjudicate fully between the Structural and Case Informativity Theories. And with this change, Experiment 2b replicates the main findings of Experiment 2a. Postnominal RC conditions show another dramatic ergative coargument effect. In prenominal RCs, ergative coarguments also condition a slowdown, but a smaller one. This suggests that ergative qua ergative is indeed relatively difficult to parse, as predicted by the Case Informativity Theory.

3.5.1 Method

Participants

63 native Georgian speakers were recruited for Experiment 2b (44 women, average age = 23). One participant lived in Kutaisi, Georgia; the rest were from Tbilisi. They were paid 20 GEL (approx. 8.00 USD) for their participation. Seven participants were ultimately excluded from analysis, either due to comprehension scores lower than 70%, or because their median RT was much slower than the rest (more than two standard deviations

beyond the mean of participant medians).

Materials

24 item sets were constructed in a 2×2 design, crossing Relative Clause Position (postnominal relative vs. prenominal correlative) and Gap Site (subject gap vs. object gap). Stimuli conformed to the following template.

(55) a. Postnominal relative template (Experiment 2b)

Dem+HdN,	[$_{RC}$	Adj	CoArg	C^0	XP	V]	$Spill_1$	$Spill_2$	$Spill_3$.
w_1		w_2	w_3	w_4	w_5	w_6	w_8	w_9	w_{10}

b. Prenominal correlative template (Experiment 2b)

[$_{RC}$	Adj	CoArg	C^0	XP	V,]	Dem+HdN	$Spill_1$	$Spill_2$	$Spill_3$.
	w_2	w_3	w_4	w_5	w_6		w_8	w_9	w_{10}

The materials differed from Experiment 2a's in the following ways. The relative clause consisted of five words, each with its own SPR window. The first two words were an adjective and a noun (w_2 and w_3), together making up the coargument DP. The noun was either in the nominative case (for the SRC conditions) or the ergative case (for the ORC conditions); the adjective was selected from a morphological class that does not show case concord with the head noun (cf. the adjectives in Experiments 2a and 3a, which did participate in case concord). At w_4 was the complementizer *rom*. This was given its own window in order to delay the cue to embeddedness in the prenominal conditions, as in the schematized correlative above (51b). Since the initial string [Adj N] is temporarily compatible with a root-clause parse, delaying the presentation of *rom* until after the coargument DP allows us to test the role of case informativity outside of disambiguated embedded environments.

Rounding out the relative is a one-word adjunct (an adverb, locative/temporal PP, or noun in the instrumental case) and the clause-final verb (always in the aorist TAM).

The head noun was presented together with the demonstrative (w_1 or w_7). It was always the matrix clause subject, but its case was counterbalanced (nominative, ergative, or dative). The head noun and coargument were matched in animacy, and itemsets were counterbalanced for the animacy of these nouns: they were either human, animal, or inanimate. Representative relative clauses and head nouns follow.

(56) a. Postnominal SRC

is gogo, [$_{CP}$ *axalgazrda* ***biǯ-i*** *rom ṭq̇e-ši naxa,*] …

D girl.NOM young **boy-NOM** COMP forest-in saw

'that girl [$_{RC}$ that ___ saw the young boy in the forest] …'

b. Postnominal ORC

is gogo, [$_{CP}$ *axalgazrda* ***biǯ-ma*** *rom ṭq̇e-ši naxa,*] …

D girl.NOM young **boy-ERG** COMP forest-in saw

'that girl [$_{RC}$ that the young boy saw ___ in the forest] …'

c. Prenominal SRC

[$_{CP}$ *axalgazrda* ***biǯ-i*** *rom ṭq̇e-ši naxa,*] ***is gogo***…

young **boy-NOM** COMP forest-in saw **D girl.NOM**

'[$_{CORR}$ the one that ___ saw the young boy in the forest,] that girl…'

d. Prenominal ORC

[$_{CP}$ *axalgazrda* ***biǯ-ma*** *rom ṭqe-ši naxa,*] ***is gogo***…

young **boy-ERG** COMP forest-in saw **D girl.NOM**

'[$_{CORR}$ the one that the young boy saw ___ in the forest,] that girl…'

These items were presented among 76 filler sentences, including 36 experimental items for Experiment 3b (3.7) and 40 more sentences which did not contain relative clauses. Each of these 100 sentences was followed by a yes–no comprehension question.

Procedure and Analysis

The procedure and analysis were the same as in Experiment 2a (Section 3.4).

3.5.2 Results

Mean reading times are shown in Figure 3.3. Table 3.5 gives results from linear mixed-effect modeling on RTs at the coargument noun region. We observe a significant main effect of gap site at this region, with the ORC condition again being slower. This is the only significant effect. We did not replicate the main effect of relative clause type at the verb region in this experiment ($\beta = 41 \pm 48$, $t(28) = -0.86$, $p = 0.40$).

Comprehension accuracy ranged from 79% to 82% across experimental conditions; it was 80% on average. These were slightly lower than the average accuracy for filler question (85%). Response latencies were on average 3,049 ms. There was no significant effect on either comprehension measure caused by the experimental manipulations.

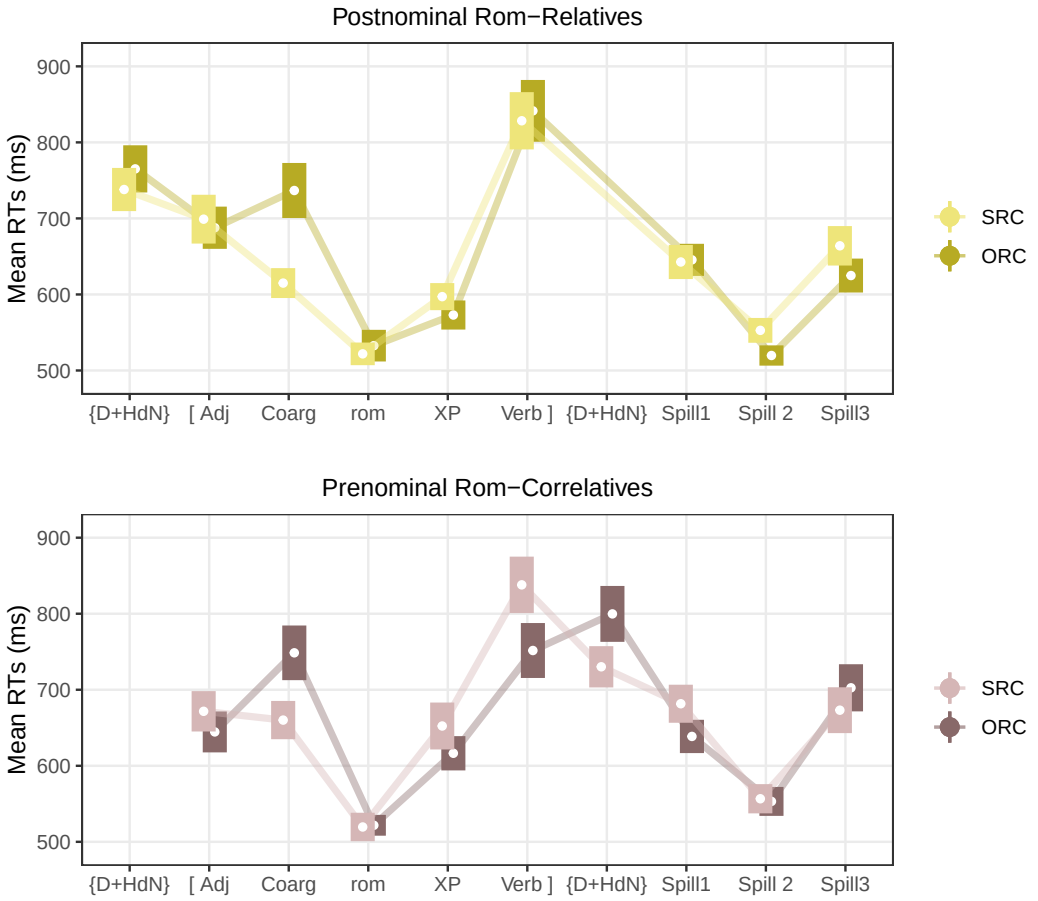


Figure 3.3: Mean reading times by region for Experiment 2b (More *rom*-relatives), separated relative clause position

3.5.3 Discussion

As in Experiment 2a, relative-clause coarguments in the ergative case were read more slowly than coarguments in the nominative case. The most straightforward interpretation of this effect stems from informativity: ergatives are taxing not (just) because they eliminate SRC parses, but because they supply the parser with more information about the ambient clause than nominatives do. This is bolstered by the fact that the correlative coar-

	β	SE	df	*t*	*p*
(Intercept)	690	35	59	20	< 0.001
Gap	100	39	31	2.6	< 0.05
RC Position	30	37	24	0.82	< 1
Gap:Position	−19	62	38	−0.31	< 1

Table 3.5: Mean reading times by region for Experiment 2b (More *rom*-relatives), separated relative clause position

gument was presented before a cue to embeddedness (the complementizer *rom*).

3.6 Experiment 3a: Processing *wh*-relatives

Experiment 3a investigates the processing profile of *wh*-relatives. Unlike a *rom*-relative, a *wh*-relative provides immediate information about the gap site: the *wh*-phrase relative pronoun that appears at its left edge bears the case associated with the syntactic position of the gap. Of course, Georgian's Split-Ergative case system means that this is often an ambiguous cue; the gap site might not be fully disambiguated until the coargument or verb is encountered.

3.6.1 Method

Participants and Materials

The same 57 participants from Experiment 2a participated in this experiment, since Experiment 2a and Experiment 3a were conducted in the same session.

36 itemsets were constructed in a 2×3 design, crossing Gap Site (SRC vs. ORC) and relative-clause internal TAM / Case Alignment (Future/NOM–ACC vs. Aorist/ERG–ABS vs. Perfect/DAT–ABS). The items conformed to the following template.

(57) Stimulus template (Experiment 3a)

Adv	HdN,	[$_{RC}$	*wh*P	XP_1	XP_2	Adj	CoArg	V,]	$Spill_1$	$Spill_2$	$Spill_3$	$Spill_4$.
w_1	w_2		w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}	w_{11}	w_{12}

The relative clause consisted of a *wh*-phrase (w_3), which bears the case associated with the gap site; a two-word adjunct phrase (w_4 and w_5); a two-word coargument DP, consisting of an adjective which shows case concord (w_6) and a noun (w_7); and finally the verb (w_8). Matrix clause material included a sentence-initial adverb (w_1), the head noun (w_2), and a four-word continuation (w_9–w_{12}). The head noun always served as the subject of the matrix clause, but across itemsets its case was counterbalanced, rotating between nominative, ergative, and dative. The head noun and coargument were of equal animacy, either both being human or both animal nouns.

The head noun DP of a sample itemset is given below. Note that the gap position is manipulated simply by swapping the case morphology of the *wh*-phrase and the coargument.

(58) a. Future (NOM–DAT case frame) SRC

gogo, [$_{RC}$ ***romelic*** *bnel ṭqe-ši* ***maġal biǯs*** *naxavs*] ...

girl.NOM **which:NOM** dark forest-in **tall boy:DAT** see:FUT

'the girl [$_{RC}$ who ___ will see the tall boy in the dark forest] …'

b. Future (NOM–DAT case frame) ORC

gogo, [$_{RC}$ ***romelsac*** *bnel ṭqe-ši* ***maġali biǯi*** *naxavs*] ...

girl.NOM **which:DAT** dark forest-in **tall boy:NOM** see:FUT

'the girl [$_{RC}$ who the tall boy will see ___ in the dark forest] …'

c. Aorist (ERG–NOM case frame) SRC

gogo, [$_{RC}$ ***romelmac*** *bnel ṭqe-ši* ***maġali biǯi*** *naxa*] ...

girl.NOM **which:ERG** dark forest-in **tall boy:NOM** see:AOR

'the girl [$_{RC}$ who ___ saw the tall boy in the dark forest] …'

d. Aorist (ERG–NOM case frame) ORC

gogo, [$_{RC}$ ***romelic*** *bnel ṭqe-ši* ***maġalma biǯma*** *naxa*] ...

girl.NOM **which:NOM** dark forest-in **tall boy:ERG** see:AOR

'the girl [$_{RC}$ who the tall boy saw ___ in the dark forest] …'

e. Perfect (DAT–NOM case frame) SRC

gogo, [$_{RC}$ ***romelsac*** *bnel ṭqe-ši* ***maġali biǯi*** *unaxavs*] ...

girl.NOM **which:DAT** dark forest-in **tall boy:NOM** see:PERF

'the girl [$_{RC}$ who ___ saw the tall boy in the dark forest] …'

f. Perfect (DAT–NOM case frame) ORC

gogo, [$_{RC}$ ***romelic*** *bnel ṭqe-ši* ***maġal bičs*** *unaxavs*] ...

girl.NOM **which:NOM** dark forest-in **tall boy:DAT** see:PERF

'the girl [$_{RC}$ who the tall boy saw ___ in the dark forest] …'

These experimental items were embedded among 64 filler sentences, comprising 24 experimental items for Experiment 2a (see Section 3.4.1) and 40 more sentences which did not contain relative clauses. Each of the 100 sentences was followed by a yes–no comprehension question.

Procedure and Analysis

The procedure was identical to Experiments 2a and 2b (see Section 3.4.1).

Reading times and comprehension question response latency were analyzed using linear mixed-effects regression; question accuracy was analyzed using logistic mixed-effects regression. The gap conditions were coded using centered sum contrasts: SRC (−½) and ORC (+½). TAM conditions were coded by Helmert contrasts: the first coefficient (TAM1) compared the aorist condition (+⅔) with the mean of the future (−⅓) and perfect conditions (−⅓); the second coefficient (TAM2) compared the future condition (+½) with the perfect (−½). Unless otherwise stated, maximal random effects structure was included (Barr et al. 2013). Models were estimated using the `lme4` package in `R` (Bates et al. 2015); *t*-tests were calculated using Satterthwaite's method via the `lmerTest` package in `R` (Kuznetsova et al. 2017).

3.6.2 Results

The following plot reports mean RTs region by region (Figure 3.4). Significant effects are found at the first region of the coargument (the adjective which bears case concord, w_6) and at the RC-final verb (w_8). Results from linear mixed-effects models are given in Tables 3.6 and 3.7, respectively (the latter model removes the by-participant slope and by-group intercept for convergence issues). The significant interaction of gap and TAM2 indicates that the coargument was read more slowly in the SRC condition, but only in the aorist condition (i.e., when the coargument was ergative). The significant main effects of TAM1 and TAM2 show that perfect verbs were read the most slowly, future verbs were of intermediate speed, and aorist verbs were read the most quickly. This scale corresponds to both the length and morphological complexity of these three TAM categories. Crucially, in the aorist conditions there was no effect at the relative pronoun ($\beta = 26 \pm 29$, $t(41) = -0.90$, $p = 0.38$) or the subsequent spillover region ($\beta = 22 \pm 35$, $t(24) = 0.62$, $p = 0.54$).

	β	SE	df	t	p
(Intercept)	690	31	71	22	< 0.001
Gap	23	29	45	0.77	< 1
Case1	-11	25	37	-0.45	< 1
Case2	40	24	39	1.7	< 1
Gap:Case1	-64	50	38	-1.3	< 1
Gap:Case2	160	45	69	3.6	< 0.001

Table 3.6: Experiment 3a (*Wh*-relatives) linear mixed-effects modeling for reading times at the adjective region of the relative clause coargument

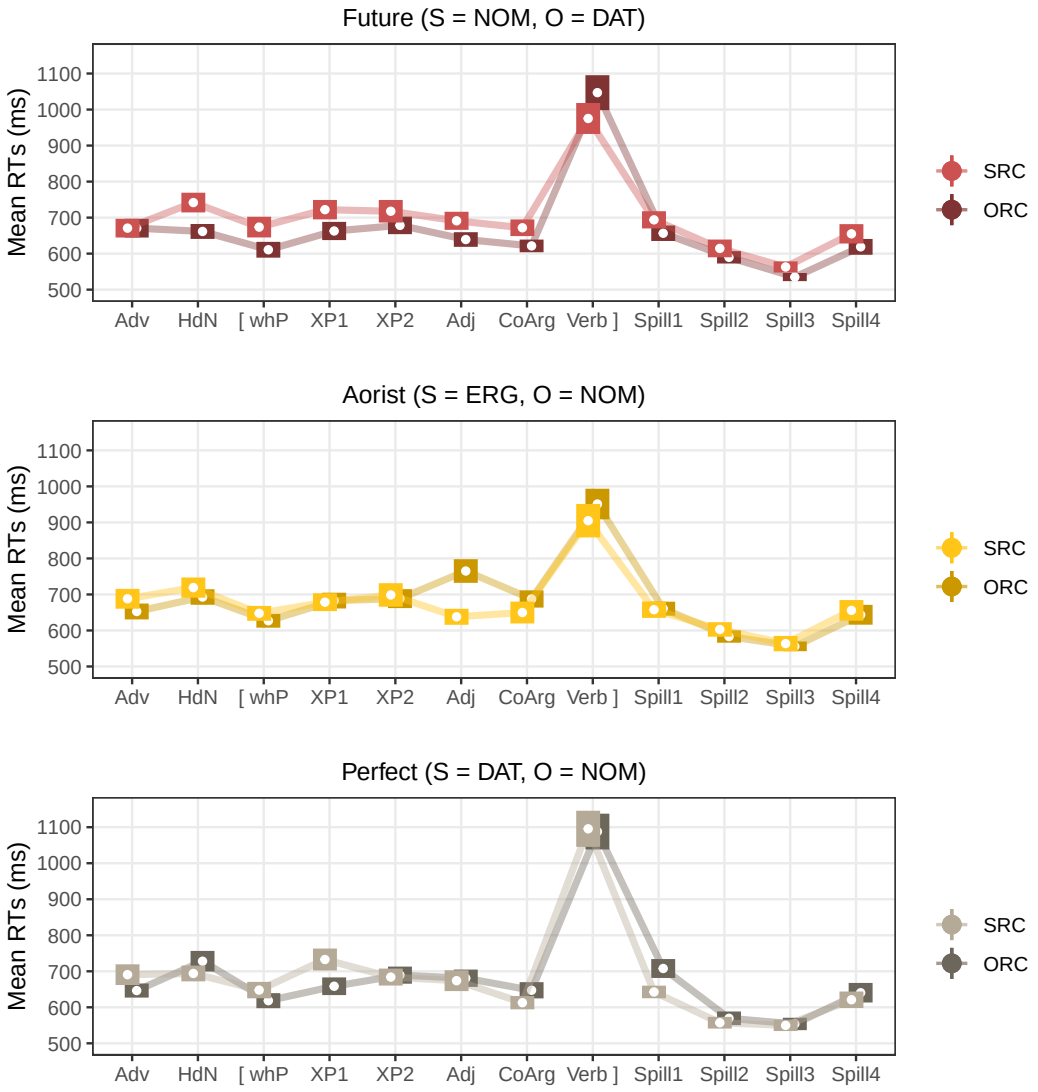


Figure 3.4: Mean readings times by region for Experiment 3a (*Wh*-relatives), separated by relative clause case frame

In the future conditions, note that a baseline difference between the SRC and ORC conditions starts at the head noun and lingers through the relative clause. This must be a spurious effect, because at the head noun window there is no evidence that a filler–gap dependency of any kind will follow. This baseline error is likely related to a technical issue

	β	SE	df	t	p
(Intercept)	1050	72	57	15	< 0.001
Gap	53	42	44	1.3	< 1
Case1	-100	39	60	-2.6	< 0.05
Case2	-130	33	150	-3.8	< 0.001
Gap:Case1	69	73	120	0.94	< 1
Gap:Case2	-15	66	120	-0.23	< 1

Table 3.7: Experiment 3a (*Wh*-relatives) linear mixed-effects modeling for reading times at the verb region of the relative clause coargument (model removes the by-participant intercept and by-item interaction slope)

in this experiment which distributed the lists unequally among participants.

Turning to performance on comprehension questions, we found no effects of Gap or RC position on either accuracy or response time. The average accuracy was 83% and the average latency was 2878 ms. For experimental conditions, mean accuracy spanned a narrow range from 80% to 87%, slightly lower than accuracy for the fillers (88%).

3.6.3 Discussion

The effects observed in Experiment 3a are most compatible with the Structural Theory of RC processing. Recall that this theory predicts that cues which eliminate the possibility of an SRC parse will cause RTs to slow: these include the ergative coargument of aorist ORCs, and the verbs of future and perfect ORCs (Figure 3.1). And indeed, a cost associated with ergative coarguments is found in this experiment. The ergative coargument

cost is of course also compatible with the Case Informativity Theory, but the lack of an effect at the *wh*-phrase region is unexpected from an informativity perspective. Whatever syntactic position an ergative-marked DP appears in, it will drastically narrow down the clause's possible argument structures and TAMs. Ergative *wh*-phrases, then, should be just as slow as ergative coarguments. Equal reading times for nominative and ergative *wh*-phrases is predicted by the Structural Theory, though, as both are at least temporarily compatible with subject-gap parses.

3.7 Experiment 3b: Processing more *wh*-relatives

Experiment 3b aims to replicate the main findings of Experiment 3a, while making a few minor changes to the design. The most notable design change is how the itemsets were counterbalanced for animacy: as in all previous experiment, head nouns and coarguments were matched in animacy, but itemsets drew from pairs of nouns that were either human, animal, or inanimate. As we will see, post-hoc analysis reveals that animacy affects processing in revealing ways.

3.7.1 Method

Participants and Materials

The same 63 participants from Experiment 2b participated in this experiment, as Experiments 2b and 3b were run in the same session (Section 3.5.1).

36 itemsets were constructed in a 2×3 design, crossing Gap Site (SRC vs. ORC) and relative-clause internal TAM / case frame (FUT/NOM–DAT VS. AOR/ERG–NOM VS. PERF/DAT–

NOM). The items followed the following template.

(59) Stimulus template (Experiment 3b)

HdN,	[$_{RC}$	*wh*P	Adj	CoArg	XP$_1$	XP$_2$	V,]	Spill$_1$	Spill$_2$	Spill$_3$.
w$_1$		w$_2$	w$_3$	w$_4$	w$_5$	w$_6$	w$_7$	w$_8$	w$_9$	w$_{10}$

The materials differ from those in Experiment 3a in a few ways. First, the order of the coargument DP and the adjunct XP were swapped; this was to ensure space between the coargument and the verb to disentangle effects that might emerge at both locations.

Second, the adjectives used in this experiment all belonged to a morphological class which does not show case concord. Syncretisms across the adjectival concord system mean that not all agreeing adjectives will indicate the case of their containing DP unambiguously. This experiment gets around this complication by ensuring that all case the morphology present in the coargument DP is unambiguous and appears on the noun.

Third, the animacy of head nouns and coarguments was counterbalanced across itemsets with three categories (human, animal, and inanimate) rather than just two. It has been observed that the animacy of a relative clause head noun can modulate the strength of the SRC advantage, with inanimate head nouns potentially neutralizing the advantage altogether (Mak et al. 2002, Traxler et al. 2005, Gennari and MacDonald 2008, Wagers and Pendleton 2016, and others). Thus, including head nouns from across the animacy spectrum means the data represent a wider array of parsing strategies.

The RC-modified DP from a representative itemset follows.

(60) a. Future (NOM–DAT case frame) SRC

gogo, [$_{RC}$ ***romelic*** *axalgazrda* ***biǯ-s*** *bnel ṭqe-ši naxavs*] ...

girl.NOM **which:NOM** young **boy-DAT** dark forest-in see:FUT

'the girl [$_{RC}$ who ___ will see the young boy in the dark forest] …'

b. Future (NOM–DAT case frame) ORC

gogo, [$_{RC}$ ***romelsac*** *axalgazrda* ***biǯ-i*** *bnel ṭqe-ši naxavs*] ...

girl.NOM **which:DAT** young **boy-NOM** dark forest-in see:FUT

'the girl [$_{RC}$ who the young boy will see ___ in the dark forest] …'

c. Aorist (ERG–NOM case frame) SRC

gogo, [$_{RC}$ ***romelmac*** *axalgazrda* ***biǯ-i*** *bnel ṭqe-ši naxa*] ...

girl.NOM **which:ERG** young **boy-NOM** dark forest-in saw

'the girl [$_{RC}$ who ___ saw the young boy in the dark forest] …'

d. Aorist (ERG–NOM case frame) ORC

gogo, [$_{RC}$ ***romelic*** *axalgazrda* ***biǯ-ma*** *bnel ṭqe-ši naxa*] ...

girl.NOM **which:NOM** young **boy-ERG** dark forest-in saw

'the girl [$_{RC}$ who the young boy saw ___ in the dark forest] …'

e. Perfect (DAT–NOM case frame) SRC

gogo, [$_{RC}$ ***romelsac*** *axalgazrda* ***biǯ-i*** *bnel ṭqe-ši* *unaxavs*] ...

girl **which:DAT** young **boy-NOM** dark forest-in see:PERF

'the girl [$_{RC}$ who ___ saw the young boy in the dark forest] …'

f. Perfect (DAT–NOM case frame) ORC

gogo, [$_{RC}$ ***romelic*** *axalgazrda* ***biǯ-s*** *bnel ṭqe-ši* *unaxavs*] ...

girl **which:NOM** young **boy-DAT** dark forest-in see:PERF

'the girl [$_{RC}$ who the young boy saw ___ in the dark forest] …'

These experimental items were embedded among 64 filler sentences, comprising 24 experimental items for Experiment 2b (see 3.5.1) and 40 more sentences which did not contain relative clauses. Each of the 100 sentences was followed by a yes–no comprehension question.

Procedure and Analysis

The procedure was identical to all other experiments (see Section 3.4.1). The analysis was identical to Experiment 3a (Section 3.6.1), which had a similar design.

3.7.2 Results

Reading times are shown below (Figure 3.5). A linear mixed-effects model reveals no significant effects at the coargument region (largest $\beta = 111 \pm 71$, $t(51) = -1.6$, $p = 0.12$), at the relative pronoun (largest $\beta = 35 \pm 18$, $t(29) = 1.9$, $p = 0.06$), or at the adjective

spillover region (largest $\beta = 36 \pm 22$, $t(510) = 1.6$, $p = 0.11$). At the verb region, though, there were main effects of TAM1 and TAM2 (Table 3.8), just as in Experiment 3a. Again, RTs correlate with verbs' length/morphological complexity.

As in the previous experiments, there were no significant effects of experimental conditions on comprehension measures. Responses to experimental conditions were between answered correctly between 80% and 87% of the time, with a mean accuracy of 83% — close to the average accuracy in the filler questions (85%). The average response latency was 2,878 ms.

	β	SE	df	*t*	*p*
(Intercept)	910	54	65	17	< 0.001
Gap	28	37	31	0.76	< 1
Case1	−92	42	45	−2.2	< 0.05
Case2	−120	41	44	−2.9	< 0.01
Gap:Case1	84	82	57	1.0	< 1
Gap:Case2	7.1	68	67	0.11	< 1

Table 3.8: Experiment 3b (More *wh*-relatives) linear mixed-effects model for reading times at the relative clause verb

3.7.3 Discussion

Given the robust ergative coargument effects in Experiments 2a, 2b, and 3a, the null result at this region for this experiment is notable. However, I believe this lack of result is a consequence of the animacy counterbalancing described in Section 3.7.1. It has

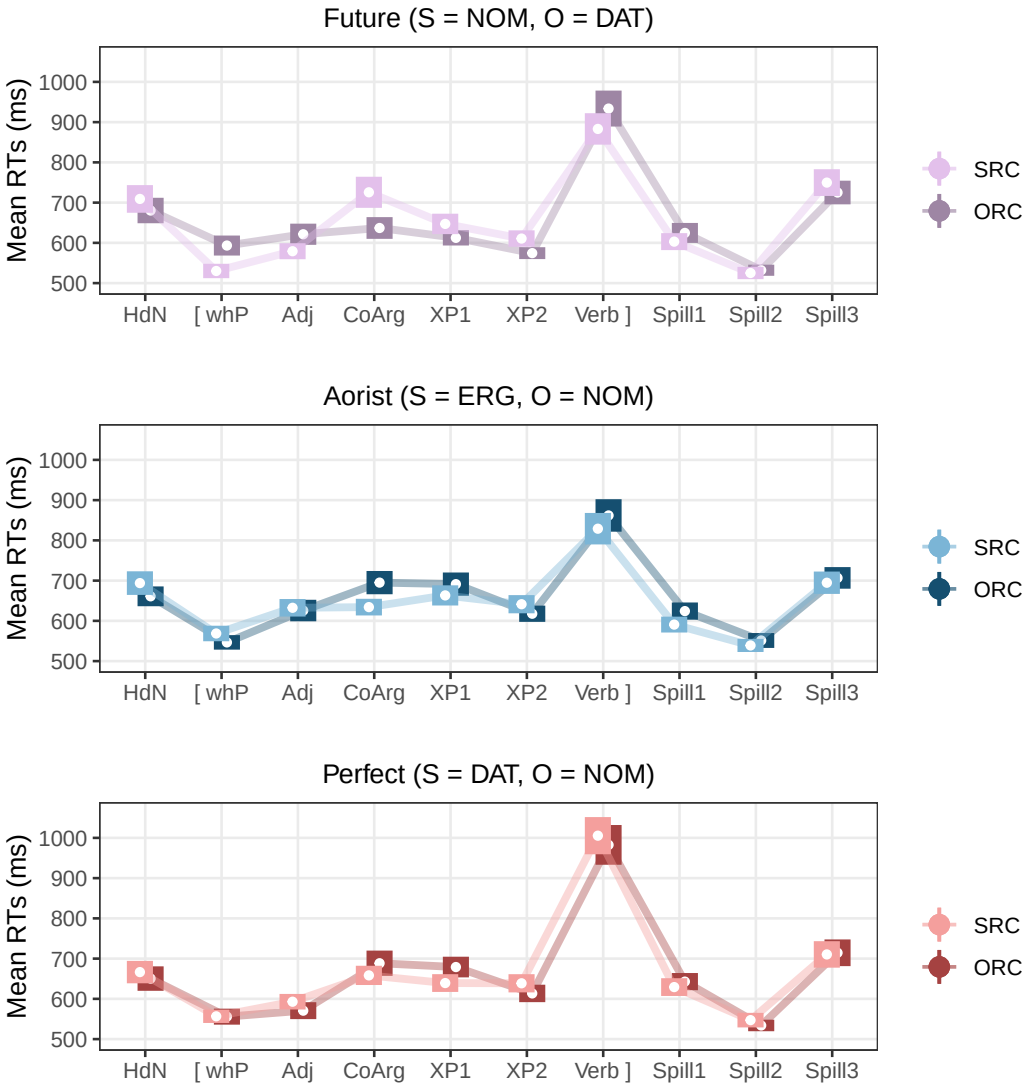


Figure 3.5: Mean readings times by region for Experiment 3b (More *wh*-relatives), separated by relative-clause TAM

been observed that expectations regarding the structure of a relative clause can be modulated by adjusting the animacy of the head noun: nouns high on the animacy scale lead to strong subject-gap expectations; ones low on the animacy scale lead to weak subject-gap expectations, or even object-gap expectations (Gennari and MacDonald 2008, Wagers and

Pendleton 2016). Recall that a third of the itemsets in Experiment 3b had inanimate head nouns. If such a head noun leads to the parser to expect an object gap, then in these trials an ergative RC coargument — an unambiguous subject — will come as no surprise. This ORC expectation, I believe, dampens the ergative coargument cost that arises in trials with human and animal head nouns, which are more likely to condition SRC expectations, and therefore lead to a garden-path effect at ergative coarguments.

Suggestive evidence in favor of this interpretation comes from exploratory analyses of the animacy counterbalancing. Figure 3.6 shows how RTs are modulated by gap site, TAM, and animacy. Especially revealing are the patterns at the RC-final verbs in the aorist and future conditions. As reported in Table 3.9, ORC verbs are markedly slower than SRC verbs given human arguments, but this trend evens out for animals, and reverses for inanimates. This pattern is in line with our thoughts above: human head nouns lead to a strong subject-gap expectation, and inanimate head nouns lead to a moderate object-gap expectation.

	Verb.AOR (ERG–NOM case frame)			**Verb.FUT (NOM–DAT case frame)**		
	SRC	**ORC**	Δ_{O-S}	**SRC**	**ORC**	Δ_{O-S}
Human	723 (42)	868 (66)	115 (78)	840 (56)	974 (87)	134 (103)
Animal	868 (82)	873 (74)	5 (110)	940 (74)	1009 (80)	69 (109)
Inanimate	896 (72)	836 (69)	−60 (100)	836 (56)	823 (64)	−44 (85)

Table 3.9: Results at the verb region for the aorist and future conditions of Experiment 3b (More *wh*-relatives), separated by argument animacy. Reported are mean RTs in ms (with standard errors), and mean differences ($RT_{ORC} - RT_{SRC}$)

Also, recall that the head noun was further counterbalanced across itemsets for

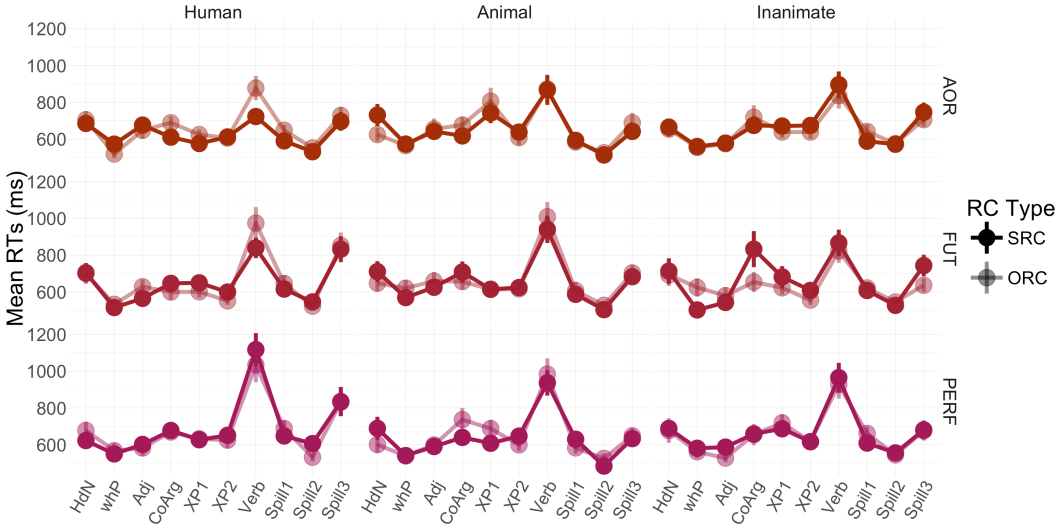


Figure 3.6: Mean RTs (ms) for Experiment 3b (More *wh*-relatives), partitioned by TAM and argument animacy

case. Table 3.10 gives mean RTs for the head noun region partitioned by animacy and case. These data complicate the picture for the Informativity Theory. Ergative head nouns are indeed read faster than nominative or dative ones overall, but breaking down the averages by animacy reveals that the ergative penalty only applies to non-human nouns. For humans, it is dative case that causes a slowdown.

	NOM	ERG	DAT	
Human	648 (32)	670 (31)	**750 (37)**	683 (19)
Animal	642 (31)	**725 (42)**	633 (35)	667 (21)
Inanimate	630 (35)	**754 (48)**	663 (34)	681 (23)
	640 (19)	714 (23)	677 (20)	

Table 3.10: Mean RTs in ms (with standard errors) at the head noun region of Experiment 3b (More *wh*-relatives), separated by case and animacy. Numbers outside the table represent by-column and by-row means.

This pattern cannot be accounted for by the strongest version of the Informativity Theory. The animacy of an argument does not entail anything about its grammatical role, or about upcoming syntactic structure (at least not in Georgian, where there are no grammaticized animacy hierarchy effects). In other words, just based on the syntactic distribution of case categories, which is in principle orthogonal to animacy, all ergatives should be equally informative and costly. The picture that is emerging, however, suggests the parser is aware of canonical relationships between grammatical role and animacy, and also between grammatical role and case. Non-humans are canonical objects and non-canonical subjects. Therefore, seeing them in a case which entails subjecthood (ergative) is more surprising than seeing them in a case with is compatible with objecthood (nominative or dative). As for humans, what is unexpected is that there is any RT difference at all — human nouns are canonical subjects, and nominative, ergative, and dative are all possible subject cases. But the fact that dative human nouns are read slowest recalls Skopeteas et al.'s 2012 findings: they discovered that dative case in Georgian, for one reason or another, is linked to objecthood. Humans being non-canonical objects, and dative case apparently being a canonical object case, the combination of dative and human will be relatively surprising.

3.8 General discussion

In this study, four self-paced reading experiments on Georgian manipulated the position of a relative clause and its internal case alignment. We observe RT patterns which are most amenable to a combination of the Structural and Informativity Theories of the SGA.

The Structural Theory posits that subject gaps are universally preferred over object gaps during filler–gap processing Keenan and Comrie (1977). Thus, it predicts that cues which entail an object gap — or at least eliminate the possibility of a subject gap — should incur a penalty. One such cue in Georgian is ergative morphology on a relative-clause internal DP. And across three experiments (Experiments 2a, 2b, and 3a), we observe a robust cost due to ergative coarguments (an effect replicated in Lau et al.'s Submitted ERP and reading-time studies), bearing out this prediction. Also harmonious with the Structural Theory is the fact the case of a relative pronoun does not affect how quickly it is read (Experiments 3a and 3b). Since nominative, ergative, and dative relative pronouns are all at least temporarily compatible with a subject-gap parse, they all satisfy the parser's desire for a subject gap equally well.

However, one observed effect cannot be explained by the Structural Theory alone. In Experiment 2b, correlative coarguments were presented before the complementizer *rom*, which signals an embedded structure. The ergative cost also emerged here, even though there is no reason to posit a filler–gap dependency of any kind before encountering *rom*. The Case Informativity Theory, on the other hand, accounts for this effect straightforwardly. It proposes that an element's processing cost is proportional to how specific a prediction the parser is licensed to make by that element's morphology, especially case morphology. Ergative in Georgian is found in a very restricted set of environments, so encountering this morphology, whether inside or outside an unambiguous filler–gap context, will be taxing, explaining the ergative coargument effect in Experiment 2b. However, we observe one type of ergative DP which did not cause a slowdown: the relative pronouns of the *wh*-relatives

in Experiments 3a and 3b. So informativity alone cannot explain the full range of effects, either.

How do our results compare to those from other ergative languages? Unlike Basque (Carreiras et al. 2010) or Niuean (Tollan et al. 2019), Georgian does not seem to simply privilege gaps associated with the unmarked case (nominative). If it did, then nominative relative pronouns would have been read faster than either ergative or dative ones. Indeed, Georgian's postnominal relative clauses seem to pattern with Ch'ol and Q'anjob'al's (Clemens et al. 2015), insofar as subject gaps enjoy a processing advantage. As for the prenominal relatives in Georgian, here the results most closely resemble those from Polinsky et al.'s 2012 study on Avar. In that language ergative coarguments also conditioned a slowdown, and RTs at the right edge of transitive prenominal relative clauses exhibited no significant relationship to gap position. One direction for future research would be to test how intransitive subject gaps are processed in Georgian. If things behave like Avar, they should be easier to process than either transitive-subject or direct-object gaps. But a potential source of variation is Georgian's Split-S alignment, which means some intransitive subject gaps will be associated with nominative case, and others with ergative (Table 2.1).

Another avenue for follow-up work is animacy. Exploratory analyses in Experiment 3b led to some interesting preliminary observations surrounding the interactions of animacy and case, and how they influence the parser's expectations. A very simple design would be to manipulate the case (nominative, ergative, or dative) and animacy (human or nonhuman) of a root-clause argument and track differences in reading times. Especially intriguing is the fact that dative seems to be linked to object position, even though dative

subjects are plentiful in Georgian (a finding that echoes Skopeteas et al.'s 2012 results). So diving deeper into the processing profile of dative case in particular may shed light on how grammatical knowledge and real-world knowledge influence predictive parsing.

Georgian is a language with typologically unusual properties, but one with literate, computer-savvy speakers. This makes the language especially well-suited to psycholinguistic research. The present study — along with Skopeteas et al. (2012), Foley and Wagers (2017b), Lau et al. (Submitted) — has given proof to this concept. Our experiments have capitalized on a unique constellation of grammatical properties in Georgian, which make the language an ideal testing ground for various theories of relative clause processing. Results suggest that multiple factors guide relative clause processing: syntactic structure, morphological cues, and also potentially arguments' animacy.

Chapter 4

Balancing subject and object agreement

4.1 Introduction

Agreement can be defined as the covariation between the morphosyntactic features of a *controller* (typically, a core argument) and the form of a morphological *locus* (e.g., a templatic slot within a verbal complex). Often, agreement relations are simple enough to describe. For example, in Sambaa (Bantu, Tanzania), the outermost agreement locus on the verb is always controlled by the subject, while inner loci are controlled by objects (61). (In glosses throughout, agreement loci and their controllers are color-coded.) Call this kind of agreement relation, in which the controller of a given locus does not vary, *rigid agreement* (62).

(61) a. *Mbegha* *a* *-za-* *mw* *-ona* *ng'wanae*.

1.Mbegha 1 -PERF- 1 -see 1.child:POSS.3SG

'Mbegha saw his child.' (Sambaa, glosses adapted from Riedel 2009:3)

b. ◌ *n* -*za*- *mw* -*ona* *ng'wana*.

1SG 1SG -PERF- 1 -see 1.child

'I saw the child.' (Sambaa, glosses adapted from Riedel 2009:46)

c. ◌ *n* -*za*- *ji* -*ona* *kui*.

1SG 1SG -PERF- 5 -see 5.dog

'I saw the dog.' (Sambaa, glosses adapted from Riedel 2009:46)

(62) *Rigid agreement*

Across a paradigm that only manipulates the φ-feature of core arguments (i.e., other features like tense–aspect–mood, definiteness, information structure, etc. are held constant), if a given locus of agreement is always controlled by a particular argument (e.g., the subject, the object, the absolutive-marked DP), then that locus expresses rigid agreement.

In contemporary syntactic theory, rigid agreement is straightforward to account for. Indeed, there is generally little analytical flexibility given standard assumptions about locality conditions on Agree (including the Minimal Link Condition and the Phase Impenetrability Condition; Chomsky 1995, 2000, 2001, et seq.) and the operation's directionality (namely, that it only operates upwards — e.g., Preminger 2013; cf. Baker 2008, Zeijlstra 2012, Bjorkman and Zeijlstra 2019). If a locus displays rigid subject agreement, then it is taken to express an Agree relation launched by a functional head higher than the external argument, say T^0. If it is dedicated to object agreement, it must expone features collected

by v^0, or another head lower than the external argument.

Of course, not all agreement is rigid. Many languages exhibit loci whose controller varies across cells of a paradigm. After Béjar (2003), I call this *promiscuous agreement* (63).

(63) *Promiscuous agreement*

Across a paradigm that only manipulates the φ-feature of core arguments (i.e., other features like tense–aspect–mood, definiteness, information structure, etc. are held constant), if a given locus of agreement has different controllers in different cells of that paradigm, then that locus expresses promiscuous agreement.

Promiscuity takes many forms. It may be as simple as an agreement affix that tracks whichever argument has a particular feature (so-called *omnivorous agreement*; Nevins 2011). Take Ayutla Mixe (Mixe–Zoquean, Mexico). In this language, either the subject (64a) or the object (64b) may control the agreement suffix *–të* 'PL'.

(64) a. ◌ *m⟨y⟩äjts-të-p* ◌.

3PL ⟨3⟩grab-PL-INDEP.TR 3SG

'They [PL] grab it.'

b. ◌ *m⟨y⟩äjts-të-p* ◌.

3SG ⟨3⟩grab-PL-INDEP.TR 3PL

'They [SG] grab them [PL].'

(Ayutla Mixe, glosses adapted from Romero-Méndez 2008:317)

Another common type of promiscuity is agreement that references a feature hier-

archy. Ayutla Mixe also illustrates this phenomenon. In addition to its omnivorous plural suffix, the language displays a prefixal agreement locus that tracks either the subject or the object, whichever outranks the other on the hierarchy 1ST > 2ND > 3RD (Romero-Méndez 2008:461–467).

(65) a. *japom* *ëjts* *mejts* *n* *-'ex-a'am-py.*

tomorrow 1SG 2SG 1 -see-DESID-INDEP.TR

'Tomorrow I will see you.' (Ayutla Mixe, Romero-Méndez 2008:465)

b. *meets* *të* *ëëtst* *x* *-tsaan-* *t* .

2PL PST 1PL 1 -hug- DEP.PL

'You [PL] hugged us.' (Ayutla Mixe, Romero-Méndez 2008:465)

c. *mejts* *Pedro* *m* *-nas-tej-ë-p.*

2SG Pedro 2 -HORZ-knead-INCH-INDEP.TR

'You are pushing Pedro.' (Ayutla Mixe, Romero-Méndez 2008:359)

d. *yë'ë* *japom* *m* *-päät-äjn-ë-p* ◌ .

DEM tomorrow 2 -meet-DESID-INV-INDEP 2SG

'They [SG] will meet you tomorrow.' (Ayutla Mixe, Romero-Méndez 2008:466)

e. *yë' kiixy* *yë' yuk* *y* *-'ixy-py.*

DEM girl DEM dog 3 -see-INDEP.TR

'The girl sees the dog.' (Ayutla Mixe, Romero-Méndez 2008:466)

Compared to rigid agreement, promiscuous agreement is less straightforward to

derive, but several contemporary syntactic theories have been developed to account for it. A common tactic is to modify Agree in such a way that a single probe can interact with more than one goal. This might be accomplished through Multiple Agree (Hiraiwa 2001, Nevins 2011), cyclic domain expansion (Béjar and Rezac 2009), or insatiable probing (Deal 2015). Under this approach, the probe collects features of both the subject and the object. But since a promiscuous agreement locus only ever expresses features of a single argument in the surface morphology, postsyntactic mechanisms will be necessary to sort out just which argument's features are exponed.

The present chapter offers a different account of certain promiscuous agreement phenomena, one which requires no novel Agreement technology. Instead, I recruit an independently motivated mechanism, the Principle of Minimal Compliance (PMC; Richards 1997, 1998). Put simply, the PMC permits a syntactic constraint to be ignored at a particular syntactic position after it has been obeyed once there. While conceived primarily with movement, binding, and ellipsis phenomena in mind, the PMC makes a prediction about agreement: should a syntactic head bear multiple φ-probes, standard locality constraints on Agree should apply to the first probe, but not subsequent ones.

I argue this prediction is borne out in the South Caucasian languages. Verbs in this family display complex agreement morphology spread across several templatic slots (see Foley to appear for descriptions). Crucially, one of these agreement loci is rigid, always tracking the highest non-dative argument, while the rest are promiscuous. This pattern follows if the loci each expone a φ-probe borne by a relatively high functional head, like T^0. The first probe, constrained by locality conditions, has no choice but to Agree with the

subject. Should the subject be in the dative case, and thus by hypothesis be encased in phasal functional structure, the probe must instead bypass it and Agree with the nominative object (Rezac 2008). But now that T^0 has minimally complied with locality constraints, its subsequent probes are free to ignore them, potentially interacting with erstwhile inaccessible arguments like objects and dative subjects.

Of course, once locality constraints are lifted, there might not be any other syntactic principles which favor probing one argument over the other. This results in derivational indeterminacy. While indeterminacy might be welcome for other minimal compliance phenomena, where free variation is observed, there is no such free variation in South Caucasian agreement. To account for this, I suggest that outputs of the derivation compete for realization on postsyntactic grounds. Expressiveness and economy filters (Kiparsky 2005) ensure that only the output whose verb expresses as many arguments with as few morphemes as possible surfaces (cf. Foley 2017).

I do not claim this to be a general theory of promiscuous agreement. Indeed, promiscuity is likely not a homogeneous phenomenon, and it may have many derivational sources. However, the proposed theory is especially well suited to the constellation of facts observed in South Caucasian. First, a PMC-based approach explains why there is exactly one rigid agreement locus and multiple promiscuous ones; stipulating that different probes have different interaction properties is not necessary. Second, this account derives the agreement patterns of both normal constructions and those with dative subjects; the latter prove to be a sticking point for alternative theories. Third, the promiscuous agreement observed in South Caucasian cannot be stated in either featural or syntactic terms. Unlike Ayutla Mixe,

for instance, where a feature hierarchy determines whether the subject or object is targeted for agreement, the core generalization about South Caucasian promiscuity is a morphological one. Thus a postsyntactic filtration mechanism is especially appropriate.

The rest of the chapter is structured as follows. Section 4.2 delves into details of South Caucasian agreement, identifying several key generalizations and meta-generalizations. Section 4.3 gives a preliminary outline for a theory which can account for the generalizations. Special attention will be paid to the syntactic properties of dative-subject constructions, which turn out to be crucial in adjudicating between different approaches. Section 4.4 outlines the proposed theory, motivating the syntactic ingredients (including the PMC) and the postsyntactic ones (including derivational filtration). Section 4.5 concludes.

4.2 South Caucasian agreement

Verbs in the South Caucasian languages (Georgian, Laz, Mingrelian, and Svan; also known as the Kartvelian languages) register their arguments' φ-features across three major morphological loci. The templatic structure of the verb is shown in Figure 4.1.[1]

preverb	AGR prefix	verb stem	TAM+AGR	PL AGR

Figure 4.1: The South Caucasian verbal template

In the leftmost slot appear preverbs: prefixes with aspectual, directional, and/or lexical function; they do not exhibit agreement. Between preverbs and the verb stem is a

[1] Discussion in this chapter necessarily elides certain details of agreement microvariation across the family. More detailed descriptions can be found in Harris (1991), Tuite (1998), Boeder (2005), and Foley (to appear).

slot for agreement prefixes.[2] Only one morpheme may appear here, and it might register the subject or the object, depending on the argument combination. Immediately suffixal to the stem are morphemes that express tense–aspect–mood (TAM) features along with agreement with a single argument, always the highest non-dative one (i.e., the object in dative-subject constructions, and the subject elsewhere). The rightmost slot is dedicated to agreement with plural arguments, either subjects or objects.

This brief summary already alludes to the fact that the TAM suffixes instantiate rigid agreement, while the agreement prefixes and plural suffixes are promiscuous. The following three paradigms illustrate, exemplifying each of the three major case frames for transitive clauses.[3] The first two show that, in clauses whose subject is either nominative (66) or ergative (67), the verb's TAM suffix will always register the subject.

[2]I remain neutral as to the precise syntactic nature of these prefixes. In the literature on Georgian, some authors have assumed that they are instances of 'agreement proper' (i.e., simply the exponents of copied φ-features; Béjar and Rezac 2009), while others have assumed they are products of clitic doubling (i.e., they are pronominal in nature; Halle and Marantz 1993). Either assumption is ultimately compatible with my analysis.

[3]The South Caucasian languages have complex case systems characterized by TAM-based Split Ergativity. For most transitive verbs, arguments receive different cases in different TAM categories. The precise details vary slightly across the languages, but generally the present tense triggers a NOM subject and an ACC object (ACC always being syncretic with DAT in the family); the aorist (perfective past) triggers an ERG subject and a NOM object; and the perfect (past evidential) triggers a DAT subject and a NOM object. The situation is different for verbs with experiencer subjects, which always exhibit a DAT–NOM case frame, no matter the TAM. See Harris (1985) for a full description of the South Caucasian case systems and a diachronic analysis.

(66) *Agreement in a nominative-subject construction*

(Mingrelian, Chikobava 1936:161)

‘NOM measures.PRES ACC’

	1SG	2SG	3SG
1SG	—	*rzimun-* *k*	*bzimun-* *k*
2SG	*bzimun-* *k*	—	*zimun-* *k*
3SG	*bzimun-* *s*	*rzimun-* *s*	*zimun-* *s*

(67) *Agreement in an ergative-subject construction*

(Mingrelian, Chikobava 1936:163)

‘ERG measured.AOR NOM’

	1SG	2SG	3SG
1SG	—	*rzim-* *i*	*bzim-* *i*
2SG	*bzim-* *i*	—	*zim-* *i*
3SG	*bzim-* *u*	*rzim-* *u*	*zim-* *u*

Compare the distribution of TAM suffixes in a dative-subject clause (68). Here the suffixes take the same shape as we saw in the aorist ergative-subject paradigm (*–i* ‘PST.1/2’ and *–u* ‘PST.3SG’) but now they are controlled by the object. Generalization (69) summarizes the behavior of these morphemes.

(68) *Agreement in a dative-subject construction* (Mingrelian, Kipshidze 1914:85)

‘DAT wanted.IMP NOM’

	1SG	2SG	3SG
1SG	—	*moḳord-i*	*moḳord-u*
2SG	*goḳord-i*	—	*goḳord-u*
3SG	*voḳord-i*	*oḳord-i*	*oḳord-u*

(69) *Tense–aspect–mood suffix generalization*

TAM suffixes exhibit rigid agreement, always tracking the highest non-dative argument.

As for the number agreement suffixes, they exhibit a pattern of omnivory very similar to what we saw in Ayutla Mixe (64). A single suffix, shaped *–t* ‘PL’ in most of the South Caucasian languages, can agree omnivorously with either a plural subject or object. Paradigm (70) shows this for Laz; (71) summarizes.[4]

(70) *Omnivorous plural agreement* (Laz, Chikobava 1936:163)

‘ERG measured.AOR NOM’

	2SG	2PL
1SG	*gzumi*	*gzumi-t*
1PL	*gzumi-t*	*gzumi-t*[5]

[4]The details of number agreement in South Caucasian are more complicated than this paradigm suggests; see Tuite (1998) for a very thorough description.

[5]It’s ambiguous which argument controls the plural suffix in the 1PL≫2PL argument combination. I represent

(71) Plural agreement generalization

Plural suffixes exhibit promiscuous agreement (specifically omnivory), tracking a plural subject or object.

Finally, the agreement prefixes: they come in two flavors, which I'll refer to as *direct* and *oblique*. The following data from Svan illustrate their distribution.[6] In nominative- and ergative-subject constructions, the direct prefixes (which in Svan include *xw*– '1SG.DIR' and *x*– '2.DIR') register the subject, while oblique prefixes (including *m*– '1SG.OBL' and *ǰ*– '2.OBL') register the object.

(72) Prefixal agreement in a nondative-subject construction

(Upper Bal Svan, after Topuria 1967:23, 73)

'NOM was preparing.IMP ACC'

	1SG	2SG	3SG
1SG	—	*ǰ*-*amāräs*	*xw*-*amāräs*
2SG	*m*-*amāräs*	—	*x*-*amāräs*
3SG	*m*-*amāra*	*ǰ*-*amāra*	*amāra*

In the 1SG≫2SG and 2SG≫1SG cells, note that we only observe object prefixes — even though the language has the morphological resources to register the subject, either on

this ambiguity by combining gray shading (which would indicate a subject controller) with an outlined box (which would indicate an object controller).

[6] I alternate between data from the Upper Bal and Lent'ekhi dialects of Svan, since the former displays TAM-suffix allomorphy patterns which most clearly illustrate the points I wish to convey about nominative-subject constructions, and the latter displays patterns which most clearly illustrate the points I wish to convey about dative-subject constructions. As far as I am aware, the only relevant differences between the agreement systems of these dialects have to do with allomorphy.

instead of the object or in addition to it. For example, the 1SG≫2SG verb form is *ǰ-amāräs* 'I was preparing you', rather than * *xw-amāräs* or * *xw-ǰ-amāräs*. This slot competition has been called *agreement displacement* (Béjar 2003, Rezac 2003, Béjar and Rezac 2009) and *person complementarity* (Nevins 2011) in South Caucasian, but for descriptive concision I'll refer to it as *prefix blocking*.

The nature of prefix blocking is a central question in South Caucasian morphosyntax. The family is well-known for having liberal phonotactics that tolerate complex onset clusters (e.g., Buts-khrikidze 2002), so a simple phonological deletion account is not tenable. And clearly the generalization about which prefixes block which others is not a featural one, since (for instance) first-person prefixes do not always block second-person prefixes (compare Ayulta Mixe; 65a–b). Just looking at the data in (72), it is tempting to make a syntactic generalization instead: given the choice between an overt subject prefix and an overt object prefix, choose only the object prefix. But the agreement patterns in dative-subject constructions show that this generalization is falls apart. The following paradigm (73) shows that the functions of the direct and oblique prefixes are 'inverted' if the subject receives dative cause: the direct series (*xw–* '1SG.DIR' and *x–* '2.DIR') registers objects, while the oblique series (*m–* '1SG.OBL' and *ǰ–* '2.OBL') registers subjects. Concentrating on the 1SG≫2SG and 2SG≫1SG cells here, we now see a preference for *subject* agreement.

(73) Prefixal agreement in a dative-subject construction

(Lent'ekhi Svan, Topuria 1967:21)

'DAT loves.PRES [NOM]'

	[1SG]	[2SG]	[3SG]
1SG	—	*m -alät̰xi*	*m -alät̰*
2SG	*ǰ -alät̰xwi*	—	*ǰ -alät̰*
3SG	[*xw*] *-alät̰xwi*	[*x*] *-alät̰xi*	*x -alät̰*[7]

Across nondative- and dative-subject constructions, what unites the prefix blocking patterns is the fact that the oblique prefix is always favored over the direct prefix. Consider the 1SG≫2SG cell of the nominative-subject paradigm (72; [*ǰ*] *-amāräs* 'I was preparing you') and the 2SG≫1SG cell of the dative-subject paradigm (73; *ǰ -alät̰xwi* 'you love me'). In both, it's *ǰ–* '2.OBL' that beats *xw–* '1SG.DIR', even though that winning morpheme registers an object in the former verb form and a subject in the latter. This leads us to the following descriptive generalization (74).

(74) Agreement Prefix generalization

The agreement prefix slot exhibits promiscuity: sometimes the subject is agreed with, sometimes the object. Could both the subject and the object control an overt agreement prefix, though, we observe *prefix blocking*: whichever argument would control an oblique prefix is the one which is uniquely expressed.

[7] Svan has an overt prefix, *x–* '3.DAT', that registers third-person datives (as does Georgian, where the morpheme is shaped *h/s/∅–* '3.DAT').

Why might the system be organized in this way? I argue that the preference for oblique prefix has morphological roots, rather than syntactic ones. Recall that the controller of the TAM suffix also 'inverts' in dative-subject constructions. This fact is key to understanding prefix blocking. In normal-subject constructions, the TAM suffix tracks the subject, and object prefixes block subject prefixes. In dative-subject constructions, on the other hand, the TAM suffix tracks the object, and subject prefixes block object prefixes (75–76). Consequently, in these argument combinations, the prefix and the suffix each ends up tracking a different argument — something conveyed graphically by the different shades of highlighting on the agreement morphemes in the grammatical examples.

(75) a. *mi* *ǰ*-*amār*- *äs* *si*.

1SG.NOM 2.OBL-prepare- IMP.1/2 2SG.ACC

'I was preparing you.'

b. * *mi* *xw*-*amār*- *äs* *si*.

1SG.NOM 1SG.DIR-prepare- IMP.1/2 2SG.ACC

Attempted: 'I was preparing you.'

c. * *mi* *xw*-*ǰ*-*amār*- *äs* *si*.

1SG.NOM 1SG.DIR-2.OBL-prepare- IMP.1/2 2SG.ACC

Attempted: 'I was preparing you.'

(Upper Bal Svan, after Topuria 1967:23, 73)

(76) a. *si* *ǰ*-*alät̩*-*xwi* *mi*.

2SG.DAT 2.OBL-love-PRES.1 1SG.NOM

'You love me.'

b. * *si* *xw*-*alät̩*-*xwi* *mi*.

2SG.DAT 1SG.DIR-love-PRES.1 1SG.NOM

Attempted: 'You love me.'

c. * *si* *ǰ*-*xw*-*alät̩*-*xwi* *mi*.

2SG.DAT 2.OBL-1SG.DIR-love-PRES.1 1SG.NOM

Attempted: 'You love me.' (Lent'ekhi Svan, Topuria 1967:21)

No matter the clause's case frame, the preference for the oblique prefix ensures a verb with maximally expressive and economical morphology (a generalization first articulated in Foley 2017). The attested forms give just enough morphological information to reliably calculate the person-features of the subject and the object. The verb *ǰ*-*amār*-*äs* 'I was preparing you', for instance, has a TAM suffix (*–äs* 'IMP.1/2') that indicates a first- or second-person subject; the prefix (*ǰ–* '2.OBL') unambiguously signals a second-person object. In concert, the two morphemes eliminate the possibility of a second-person subject, since a 2SG≫2SG argument combination would violate Condition B. (A reflexive interpretation is ruled out independently, since anaphors trigger the same agreement as third-person arguments in South Caucasian.) Imagine instead that the 1SG≫2SG argument combination triggered the form * *xw*-*amār*-*äs*, with the direct prefix (*xw–* '1.DIR') blocking the oblique prefix, rather than vice versa. This would unambiguously signal a first-person subject,

but it would not express any features of the object. It would therefore not be as expressive as the attested form. Contrariwise, an alternative with both agreement prefixes (either * *xw* - *ǰ* -*amār*- *äs* or * *ǰ* - *xw* -*amār*- *äs*) would be redundant and therefore uneconomical, since the features of both arguments can be calculated reliably just with the oblique prefix. The same logic holds for dative-subject verbs.

4.3 Accounting for the generalizations

With descriptive generalizations in tow, this section discusses analytical approaches to the three agreement loci of South Caucasian verbs. First we examine the TAM suffixes, and observe many parallels with agreement systems of more familiar languages. The patterns observed are predicted by a bog-standard φ-probe — one to which the φ-features of dative subjects are invisible, by hypothesis because of the structure of a dative argument. The characteristic highest-nondative pattern, we'll see, follows from Relativized Minimality (Rizzi 1990). The same principle can also account for the behavior of the plural suffixes, though an additional assumption is necessary: namely, that φ-probes may be relativized to search for a specific feature (here, [+PL]). Arguments lacking that feature will simply not figure into the probe's calculus.

Finally we turn to the agreement prefixes. There exists in the literature an elegant syntactic analysis of the facts in Georgian (Béjar 2003, Béjar and Rezac 2009, Lomashvili and Harley 2011) and Laz (Atlamaz 2013). In this Cyclic Agree analysis, a low φ-probe is relativized to [PART], and the highest argument in its c-command domain will satisfy

that probe just in case that argument is first or second person. A third-person argument in that position, on the other hand, triggers a second cycle of Agree that targets a farther argument. With a few important syntactic assumptions — including, crucially, one about the structure of dative-subject constructions — this Cyclic Agree theory can account for South Caucasian prefix blocking in both normal and dative-subject contexts. However, investigating the syntactic properties of dative-subject constructions more carefully reveals that the crucial assumption does not hold. Thus a Cyclic Agree analysis of South Caucasian prefix blocking unravels. Some other mechanism, which will be articulated in Section 4.4, is necessary to allow the relevant φ-probe to Agree promiscuously, sometimes with the subject and sometimes the object.

4.3.1 Accounting for tense–aspect–mood suffixes

South Caucasian TAM suffixes reflect the φ-features of the highest non-dative argument. This means that for clauses with nominative or ergative subjects, the subject will be the controller (77a); for clauses with dative subjects, the object will be (77b).

(77) a. *šen* *nax-av-* ∅ [*mas*].
2SG.NOM see.PFV-THM- NPST.1/2 [3SG.ACC]
'I will see them [SG].' (Georgian, after Aronson 1990:171)

b. *šen* *nax-* *e* [*is*].
2SG.ERG see.PFV- PST.1/2 [3SG.NOM]
'You saw them [SG].' (Georgian, Aronson 1990:172)

c. *mas* *e-nax-* *e* *šen*.

3SG.DAT PLU-see.PFV- PST.1/2 2SG.NOM

'They [SG] had seen you.' (Georgian, Aronson 1990:273)

This is a familiar pattern — dative subjects disrupt normal agreement in many languages. Russian (78) and especially Nepali (79) are a close parallels.

(78) a. ◌ *xot-* *im* *mira*.

1PL.NOM want- NPST.1PL peace.GEN

'We want peace.' (Russian, Wade 2011:117)

b. *ja* *znaju,* *emu* *nrav-* *im* *-sja* *my*.

1SG.NOM know:NPST.1SG 3SG.M.DAT like- NPST.1PL -REFL 1PL.NOM

'I know that he likes us.'

(Russian, glosses adapted from Sigurðsson 2002:720)

(79) a. *ma* *yas* *pasal-mā* *patrikā* *kin-ch-* *u*.

1SG.NOM DEM.OBL store-LOC newspaper.NOM buy-NPST- 1SG

'I buy the newspaper in this store.' (Nepali, Bobaljik 2008:310)

b. *maile* *yas* *pasal-mā* *patrikā* *kin-* *e*.

1SG.NOM DEM.OBL store-LOC newspaper.NOM buy- PST.1SG

'I bought the newspaper in this store.' (Nepali, Bobaljik 2008:310)

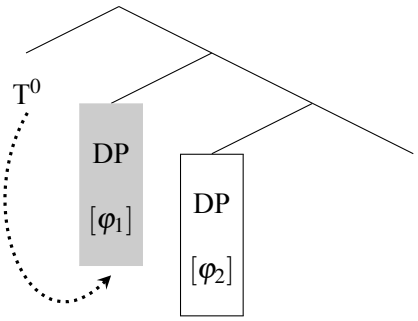
c. *malāī* *timī* *man* *par-ch-au*.

1SG.DAT 2.M.HON.NOM liking occur-NPST-2.M.HON

'I like you.' (Nepali, Bobaljik 2008:311)

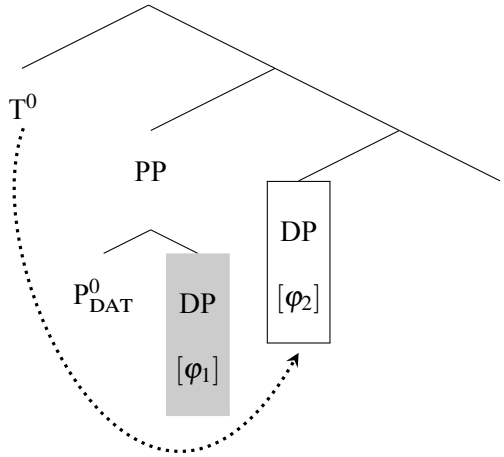
To account for this pattern, I assume that South Caucasian nominative and ergative subjects are simple DPs, while dative subjects are encased in extra functional structure, perhaps a PP (Rezac 2008) or a KP (Bittner and Hale 1996). Crucially, this extra layer is phasal (see Abels 2003 on the phasehood of PPs), and its head bears neither φ-features of its own nor a φ-probe to copy those of the embedded DP.[8] By virtue of the Phase Impenetrability Condition (Chomsky 2001), then, the φ-features of the dative subject are inaccessible — and indeed invisible — to the relevant probe on T^0. So, while the probe must interact with the closest argument (the subject) in nominative- and ergative-subject constructions, it skips over the subject and instead probes the object in dative-subject constructions. Trees in (80) illustrate.

(80) a.



[8]Exceptional object agreement in dative-subject constructions is not universal — in some languages, dative subjects instead function as defective interveners. Descriptively, the agreement locus will not be able to register the features of either the dative subject or the object; default 3SG agreement typically obtains instead. Defective intervention in dative-subject constructions is a theoretically challenging phenomenon (see Preminger 2014, Ch. 8 for discussion), but one possible analysis is to assume that the phasal head encasing a dative subject bears its own set of φ-features (perhaps default 3SG ones). This would ensure that T^0 would be able to access neither the φ-features of the subject itself (buried within the dative PP) nor those of the object.

b.



This can be seen as a Relativized Minimality effect (Rizzi 1990). Just as C^0's *wh*-probe will ignore a non-*wh* subject in an object-*wh* question, so too will T^0's φ-probe ignore the subject of a dative-subject construction. In both of these cases, the subject — while structurally closer than the object to the relevant probe — either entirely lacks the features sought for by the probe, or they are inaccessible and invisible to it.

4.3.2 Accounting for plural suffixes

South Caucasian's suffixal number agreement system is a textbook case of omnivorous agreement (Nevins 2011). In Georgian, for instance (81), the suffix *–t* 'PL' can appear if there is a plural subject, a plural object, or both.

(81) *g-nax-e-**t**.*

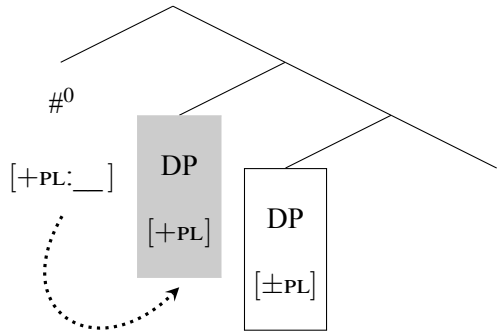
2.OBL-see.PFV-PST.1/2-**PL**

'**We** saw you [SG].' or 'I saw **you [PL]**.' or '**We** saw **you [PL]**.'

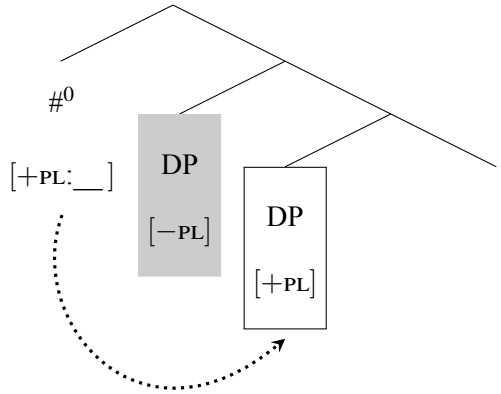
(Georgian, Aronson 1990:172)

This pattern, too, can be cast in terms of Relativized Minimality (cf. Preminger 2014:40–47). Assuming a φ-probe (call it $\#^0$) may be relativized to search for a subset of the φ-features (following Nevins 2007, 2011, Béjar and Rezac 2009), then arguments which do not bear those features will simply be ignored. This is illustrated below (82).[9]

(82) a.



b.



4.3.3 Accounting for prefixal agreement

Recall that agreement prefixes track either the subject or the object. In nondative-subject constructions, the object will control the prefix, except when the object is third person and the subject is first or second, in which case the subject will control the prefix. The reverse holds in dative-subject constructions: the subject will control the prefix, unless the subject is third person. Data from Georgian recap the generalizations.

[9] In SG≫SG configurations, the number probe will simply fail to find a goal (Preminger 2011, 2014).

(83) a. *me* *g* -*nax*- *e* *šen*.

1SG.ERG 2.OBL -see.PFV- PST.1/2 2SG.ACC

'I saw you.'

b. *me* *v* -*nax*- *e* *is*.

1SG.ERG 1.DIR -see.PFV- PST.1/2 3SG.NOM

'I saw them [SG].' (Georgian, Aronson 1990:172)

(84) a. *šen* *g* -*e*-*nax*- *e* *me*.

2SG.DAT 2.OBL -PLU-see.PFV- PST.1/2 1SG.NOM

'You had seen me.'

b. *mas* *v* -*e*-*nax*- *e* *me*.

3SG.DAT 1.DIR -PLU-see.PFV- PST.1/2 1SG.NOM

'They [SG] had seen me.' (Georgian, Aronson 1990:273)

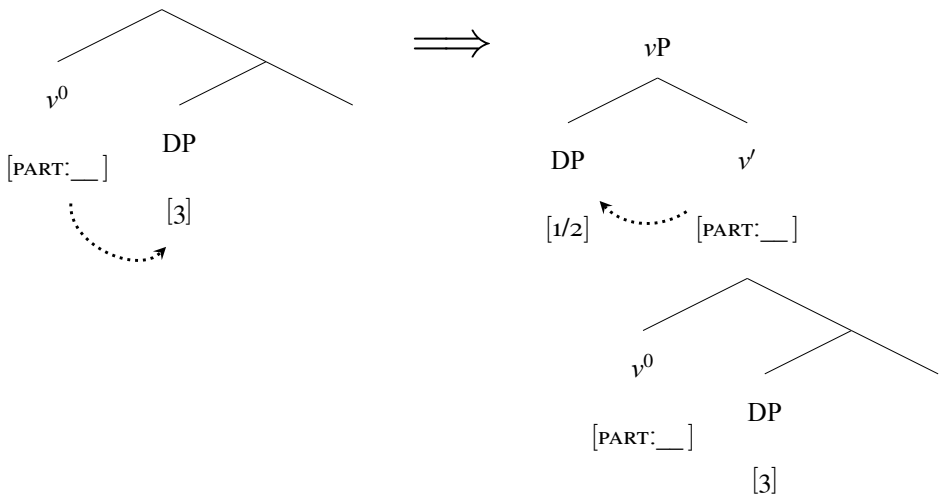
This pattern has been successfully modeled using a framework that accepts cyclic probing and cyclic domain expansion (Béjar 2003, Béjar and Rezac 2009, Lomashvili and Harley 2011, Atlamaz 2013). Such a theory assumes that the relevant φ-probe is on v^0, and that probe is relativized to Agree with [PART] (i.e., first- or second-person) arguments.

Taking nondative-subject constructions as a starting point, when v^0 merges, it will probe its c-command domain for a suitable goal, interacting with the object. If the object is first or second person, v^0's probe will be satisfied. This straightforwardly captures the general preference for object agreement in clauses with a nominative or ergative subject,

and more specifically the phenomenon of prefix blocking there: after all, in 1≫2 and 2≫1 derivations, the probe will only ever interact with the object.

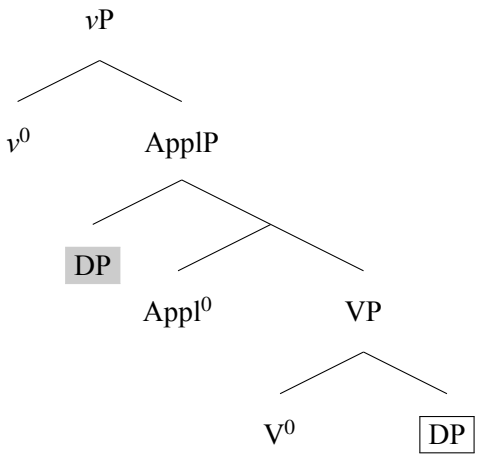
An additional assumption, though, is necessary to ensure that *subject* agreement can obtain in 1/2≫3 contexts. Invoking Bare Phrase Structure (Chomsky 1995:241–249), Béjar and Rezac (2009) assume that nonterminal projections of a head will bear all the properties of that head. Relevant for them, this means that the node traditionally labeled v' will bear a φ-probe, just like v^0 does. When the external argument is merged into the structure, then, the intermediate projection of v is the in the right position to probe it (v', of course, c-commanding the DP in Spec-vP). This interleaving of external merge and Agree — call it cyclic domain expansion — allows the probe on v to interact with the subject, just in case the object did not satisfy the probe on its first cycle of Agree. Such a derivation is schematized below. (For consistency with other diagrams in this chapter, Bare Phrase Structure labels are not adopted, though the theory is still assumed in the discussion of the Cyclic Agree analysis.)

(85)

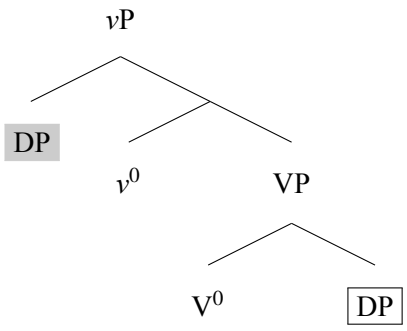


This analysis can be extended to capture the prefixal agreement patterns of dative-subject constructions. To do so, though, it is necessary to ensure that a first- or second-person dative subject can satisfy v^0 before it has a chance to probe the object. This is possible if we assume dative subjects in South Caucasian are uniformly merged below v^0, perhaps in the specifier of an applicative projection (ApplP) that v^0 takes as its complement — an analysis that echoes Hermon (1981), Marantz (1981), Belletti and Rizzi (1988), Landau (2009), and many others. This makes a dative-subject construction essentially an applied unaccusative; such a structure is contrasted with the canonical transitive structure in (86).

(86) a. *Applied Unaccusative*



b. *Canonical Transitive*



The applied unaccusative analysis is crucial to extending a Cyclic Agree theory to the agreement patterns of South Caucasian dative-subject constructions (Béjar 2003, Lomashvili and Harley 2011, Atlamaz 2013). Were dative subjects external arguments (i.e., higher than v^0), prefixal competition should behave identically in both dative- and nondative-subject constructions. That is to say, an object agreement preference should emerge everywhere. But this is not the case — subject agreement prefixes block object agreement prefixes when the subject is dative (84a).

Adopting an argument structure with low dative subjects solves the problem for a Cyclic Agree analysis. A first- or second-person dative subject will satisfy v^0's [PART] probe on the first cycle of Agree. In other words, v^0 will never have a chance to interact with the object in 1.DAT≫2 and 2.DAT≫1 contexts. In 3.DAT≫1/2 derivations, on the other hand, v^0 will launch a second cycle of Agree after the third-person dative subject fails to satisfy it. Cyclic domain expansion is not necessary here; v^0 simply continues scanning its c-command domain after interacting with the dative subject, eventually finding the lower object.

4.3.4 Another look at dative-subject constructions

The applied unaccusative analysis of South Caucasian dative-subject constructions (86a) can be traced back to Marantz (1989), and shares the spirit of Harris's (1981:119, 128) Relational Grammar analysis. And there are several compelling reasons to adopt it. The first argument is morphological. Many dative-subject verbs bear explicit unaccusative and applicative morphology. In Georgian, for example, simple unaccusatives often bear the

inchoative suffix *–d* 'INCH', and in the nonpast their 3SG TAM suffixes take the allomorph *–a* 'NPST.3SG.UNACC' (87a). Moreover, an external possessor or affectee argument may be introduced to an unaccusative with the applicative prefix *u–* 'APPL' (87b). All three of these morphemes, it turns out, can be found in certain dative-subject verbs, as in (87c).

(87) a. *baxala* *mo-ḳv-d-eb-* *a* .

chick.NOM PVB-die-INCH-THM- NPST.3SG.INTR

'The chick will die.' (Georgian, after Rayfield 2006:907, 955)

b. *baxala* *ġvav-s* *mo-u-ḳv-d-eb-* *a* .

chick.NOM crow-DAT PVB-APPL-die-INCH-THM- NPST.3SG.INTR

'The crow's chick will die.' (Georgian, after Rayfield 2006:907, 978)

c. ◌ *mezobel-i* *še-u-ġvar-d-eb-* *a* .

3SG.DAT neighbor-NOM PVB-APPL-love-INCH-THM- NPST.3SG.INTR

'They [SG] will fall in love with the neighbor.'

(Georgian, after Rayfield 2006:1400)

A second piece of evidence compatible with the applied unaccusative structure is a syntactic one. A verb cannot license both a dative subject and a dative O_{IO}.[10] These instead must be expressed as nonargumental PPs, as the following pair of examples from Georgian shows. In (88a), the O_{IO} is a noun-phrase argument in the dative case, and it triggers prefixal agreement. In (88b), the O_{IO} is expressed en chômage in a PP headed by the enclitic postposition *–tvis* 'for'; it conspicuously triggers no verbal agreement.

[10]This generalization holds at least in Georgian (Harris 1981:118, Ershova 2016). Laz seems to allow applied dative-subject constructions, though; see Öztürk (2016).

(88) a. *Rezo-m samaǰur-i* [*g*]*-ačuka* [*šen*].

Rezo-ERG bracelet-NOM [2.OBL]-give:AOR.3SG [1SG.DAT]

'Rezo gave you a bracelet.'

b. *turme Rezo-s samaǰur-i učukebia* [*šen=tvis*].

apparently Rezo-DAT bracelet-NOM give:PERF.3SG [2SG.GEN=for]

'Rezo apparently gave you a bracelet.' (Georgian, after Harris 1981:117)

If only one applicative projection is permitted per clause, it follows from the applied-unaccusative analysis that dative-subject constructions cannot host indirect or applied objects. By hypothesis, all three kinds of arguments compete for the same structural position (Spec-ApplP), so the presence of a dative subject prevents a O_{IO} from being merged into the structure.

But if all dative subjects are indeed lower than nondative subjects, one might expect to find other syntactic consequences of this structural difference. Do dative subjects have fewer subjecthood properties than nominative or ergative subjects — perhaps behaving differently with respect to binding and control phenomena? At least in Georgian, there actually seem to be few syntactic differences, if any, between dative and nondative subjects, aside from their case and agreement properties. (I am unaware of detailed work on the syntax of dative subjects in the other South Caucasian languages.) For example, dative and nondative subjects alike can bind reflexive anaphors (89), bind anaphoric possessors (90), and control PRO (91).

(89) a. *Gela irçmunebs tavis tav-s.*

Gela.NOM convince:PRES.3SG own.ACC self-ACC

'Gela$_1$ is convincing himself$_1$.'

b. *Gela turme daurçmunebia tavis-i tav-i.*

Gela-DAT apparently convince:PERF.3SG own-NOM self-NOM

'Gela$_1$ apparently convinced himself$_1$.' (Georgian, Harris 1981:125)

(90) a. *deda bans tavis švil-s.*

mother.NOM bathe:PRES.3SG own.ACC child-ACC

'The mother$_1$ is bathing her$_1$ child.' (Georgian, Harris 1981:27)

b. *deda-s turme daubania tavis-i švil-i.*

mother-NOM apparently bathe:PERF.3SG own-NOM child-NOM

'The mother$_1$ apparently bathed her$_1$ child.' (Georgian, after Harris 1981:27)

(91) a. *Rezo-m daaṗira saxl-is ašeneba.*

Rezo-ERG intend:AOR.3SG house-GEN build:NMLZ

'Rezo intended PRO to build a house.'

b. *Rezo-s turme daupirebia saxl-is ašeneba.*

Rezo-DAT apparently intend:PERF.3SG house-GEN build:NMLZ

'Rezo apparently intended PRO to build a house.'

(Georgian, constructed examples)

The lack of subjecthood differences between subjects of different cases casts some doubt on the hypothesis that dative-subject constructions have a special syntax. However, there are two more damning pieces of evidence — passivization and PCC effects — that make the strong hypothesis that all South Caucasian dative-subject constructions have applied unaccusative syntax very difficult to maintain. First, consider facts surrounding passivization. Many (but not all) transitive verbs in Georgian exhibit a synthetic passive alternation. This alternation is characterized by the promotion of the internal argument to subjecthood; the external argument is suppressed, but can be optionally expressed en chômage in a *by*-phrase. Crucially, synthetic passivization is available to these verbs both in tenses that trigger nondative transitive subjects (e.g., the aorist; 92) and in tenses that trigger dative transitive subjects (e.g., the pluperfect; 93).

(92) a. *Levan-ma šeaḳeta macivar-i.*
Levan-ERG repair:AOR.3SG refrigerator-NOM
'Levan repaired the refrigerator.'

b. *macivar-i šeḳetda Levan-is mier.*
refrigerator-NOM repair.PASS:AOR.3SG Levan-GEN by
'The refrigerator was repaired by Levan.' (Georgian, MN 190902)

(93) a. *Levan-s unda šeeḳetebina macivar-i.*
Levan-DAT MODAL repair:PLU.3SG refrigerator-NOM
'Levan should have repaired the refrigerator.'

b. *macivar-i* *unda* *šeḳetebuliqo* *Levan-is* *mier.*
refrigerator-NOM MODAL repair.PASS:PLU.3SG Levan-GEN by
'The refrigerator should have been repaired by Levan.'

(Georgian, MN 190902)

While not all dative subjects can be synthetically passivized — it seems that experiencer subjects never can be, for instance — the fact that at least some can be is a serious problem for the applied unaccusative analysis of dative-subject constructions.[11] Let's adopt the standard assumption that active clauses involve a v^0 (or perhaps Voice^0) that introduces an external argument in its specifier, while passive clauses involve a v^0 which cannot introduce an external argument (e.g., Alexiadou et al. 2018). It is difficult to imagine, then, how one could passivize an applied-unaccusative structure like (86a). Conceivably there might be a way to analogize the passive alternation here by admitting the possibility of "passive" Appl^0 — one which cannot introduce an argument in its specifier, just as a passive v^0 cannot. A simpler alternative, though, is to adopt the canonical transitive structure (86b) for at least those dative-subject constructions which are passivizable. In other words, these passivization facts strongly suggest that at least some dative subjects are bona fide external arguments.

A second challenge to the applied-unaccusative structure comes from PCC facts. The PCC, or Person–Case Constraint, is a family of restrictions that compare the person features and hierarchical position of arguments in a syntactic domain; inadmissible combi-

[11] The passivization argument is limited to synthetic passives. Georgian also has analytic passives, which consist of a past participle and an auxiliary verb. Analytic passivization is much more widely available in the language; even some dative experiencer subject verbs have analytic passives.

nations cannot be expressed with the default agreement or pronominal cliticization strategies (Bonet 1991, Anagnostopoulou 2005). In Georgian (as in many other languages), the PCC restricts combinations of objects in a ditransitive. Just which combinations Georgian's PCC rules out seems to be a matter of dialectal or idiolectal variation (given the conflicting descriptions of Harris 1981 and Wier 2011), but a combination which is ungrammatical for all Georgian speakers is a ditransitive that has a third-person subject, and two objects which are first or second person (i.e., *3.SUBJ≫1.IO≫2.DO and *3.SUBJ≫2.IO≫1.DO).

Example (94a) gives the baseline case: a 3≫1≫3 ditransitive is perfectly grammatical, and the first-person O_{IO} triggers an agreement prefix on the verb. Compare this to (94b), whose 3≫1≫2 ditransitive is ungrammatical due to the PCC, no matter which object the verb agrees with. Such an argument combination can be expressed, though, if it is repaired through 'object camouflage' (Harris 1981:48–52). This phenomenon can be thought of as an operation that converts the O_{DO} of a PCC-violating structure into a spurious reflexive anaphor; the O_{IO} is then free to control prefixal agreement (94a).

(94) a. *Vano-m me Nino še-m-a-dar-a.*

Vano-ERG 1SG.DAT Nino.NOM PVB-1SG.OBL-DITR-compare-PST.3SG

'Vano compared Nino [O_{DO}] to me [O_{IO}].' (Georgian, after Harris 1981:48)

b. * *Vano-m me šen še-m/g-a-dar-a.*

Vano-ERG 1SG.DAT 2SG.NOM PVB-1SG.OBL/2.OBL-DITR-compare-PST.3SG

Attempted: 'Vano compared you [O_{DO}] to me [O_{IO}].'

c. *Vano-m me šeni tav-i še-m-a-dar-a.*

Vano-ERG 1SG.DAT 2SG.GEN self-NOM PVB-1SG.OBL-DITR-compare-PST.3SG

Literally: 'Vano compared yourself [O_{DO}] to me [O_{IO}].'

(Georgian, glosses adapted from Harris 1981:283)

A full analysis of the Georgian PCC and its repair is beyond the scope of this chapter. It suffices to say, though, that PCC effects are generally taken to arise when a single probe must license more than one argument, and something goes awry (see, e.g., Béjar and Rezac 2003, Anagnostopoulou 2005, Nevins 2007, Foley and Toosarvandani 2019, Coon and Keine 2019). For Georgian, a reasonable assumption is that that probe is v^0, and for one reason reason or another it cannot license two internal arguments that are first or second person.

This is where dative subjects come in. If indeed dative-subject constructions are applied unaccusatives — ditransitives that lack an external argument, essentially — a natural prediction is that the PCC should rule out 1.DAT≫2 and 2.DAT≫1 dative-subject constructions, and the objects of these clauses should be repaired through object camouflage. As we have seen several times throughout this chapter, though, such structures are perfectly grammatical, even without spurious reflexive objects.

(95) a. *Gela-s vuq̇varvar me.*

Gela-DAT love:PRES.3SG≫1SG 1SG.NOM

'Gela loves me.'

b. * *Gela-s* *uq̇vars* *čemi* *tav-i.*

Gela-DAT love:PRES.3SG≫3 1SG.GEN self-NOM

Attempted: 'Gela loves myself.'

(Georgian, glosses adapted from Harris 1981:142)

We can contrast Georgian with a language like Basque, whose PCC restricts argument combinations in ditransitives and DAT≫NOM constructions in entirely parallel ways (Rezac 2008). I suggest that one possible explanation for the difference between these languages is the structure of their dative-subject constructions. Perhaps the applied-unaccusative structure is alive and well in Basque; v^0, then, will be tasked with licensing both the direct & indirect objects of a ditransitive, and the subject & object of a dative-subject construction. Whatever principle or set of principles that prevents v^0 from licensing them in PCC-violating configurations should naturally obtain in both types of clauses. In contrast, perhaps some or even all Georgian dative-subject constructions have canonical transitive structures (86b). If this is the case, there is no reason to expect PCC effects to arise in dative-subject constructions in this language.

Eschewing the applied unaccusative structure raises questions regarding the morphology of certain dative-subject verbs (87c) and the unavailability of double-dative structures (88b). The former fact is perhaps not too troublesome, since one can easily imagine a set of postsyntactic operations that would result in ersatz applied-unaccusative morphology on a verb with a genuine external argument. As for the latter issue, I suggest the Case Filter may be an explanation — dative subjects and O_{IO}s may not be competing for a single

syntactic position, but they are competing for a single case-licenser.

In sum, an applied unaccusative structure — in which dative subjects are merged low, in Spec-ApplP — is initially appealing for South Caucasian dative-subject constructions. However, there are several compelling reasons to think that dative subjects (at least for some verbs, at least in Georgian) are bona fide external arguments. And if even some dative subjects are external arguments, merged above v^0, the Cyclic Agree analysis of prefixal agreement in South Caucasian makes a fatal prediction. That's because in this system, all 1≫2 and 2≫1 verbs with external arguments should exhibit a preference prefixal object agreement, since the [PART] probe on v^0 will be satisfied by the internal argument before the external argument (whatever case it receives) is merged. This, of course, is counter to fact: dative-subject constructions exhibit for a preference prefixal subject agreement.[12] Therefore another analysis, one to be developed in the next section, is necessary.

4.4 Filtration and Minimal Compliance

With one prominent syntactic theory out of the running to derive South Caucasian agreement (Cyclic Agree; Béjar 2003, Rezac 2003, Béjar and Rezac 2009, et seq.), this section articulates an alternative that has better descriptive and explanatory adequacy. The

[12] The structure of dative-subject constructions is a major empirical problem for the Cyclic Agree theory, but it is not the only one. First, since Béjar and Rezac (2009) assume the direct–oblique prefix distinction is a second-cycle effect (i.e., features copied on the first cycle of Agree are exponed with oblique prefixes, and those copied on a second cycle are exponed with direct prefixes). Thus they predict that unaccusative subjects should control oblique agreement prefixes. But this is not the case; all (nondative) intransitive subjects control direct prefixes. Second, Bejar and Rezac predict uniform subject agreement in ditransitives with a local-person subject and two third-person objects. That's because the probe will not be satisfied until it finds a [PART] feature on the external argument. However, in Georgian, certain 2≫3≫3 ditransitives actually exhibit prefixal *indirect-object* agreement. This should not be possible if the probe interacts with the second-person subject.

theory will have both syntactic and postsyntactic components, which are in principle independent. For both components, there are several conceivable sets of assumptions that would yield the observed facts. However, I advocate for a theory which employs the Principle of Minimal Compliance (Richards 1997, 1998) in the syntactic component, and derivational filters sensitive to morphological expressiveness and economy (Kiparsky 2005) in the postsyntactic component. This analysis goes above and beyond descriptive adequacy, as it accounts for important meta-generalizations about South Caucasian agreement.

4.4.1 Motivating Postsyntactic Filtration

Recall the descriptive generalizations about prefix blocking from Section 2. Prefix blocking in dative- and nondative-subject constructions cannot be unified in syntactic terms — no one particular argument position is preferred as an agreement controller across these environments. However, the phenomenon can be unified in morphological terms: oblique prefixes (e.g., *m*– '1SG.OBL' and *ǰ*– '2.OBL' in Svan) always block direct prefixes (e.g., *xw*– '1SG.DIR' and *x*– '2.DIR' in Svan). And this fact, I argue, is ultimately rooted in a pressure to maximize morphological expressiveness. There are a few pieces to this argument: (i) direct prefixes and TAM suffixes register one set of arguments (either the subject in nondative-subject constructions or the object in dative-subject constructions), (ii) oblique prefixes register the opposite arguments (either an object in nondative-subject constructions or the subject in dative-subject constructions), and (iii) TAM suffixes are obligatory, and they agree rigidly (their controller is fixed across a paradigm). Therefore, in environments where both an overt direct prefix and and overt direct prefix could be licensed — any 1≫2

or 2≫1 verb in Svan, say (96) — choosing the oblique prefix will always result in a verb that registers two arguments instead of just one, maximizing expressiveness.

(96) a. *mi* *ǰ*-*amār*-*äs* *si*.
1SG.NOM 2.OBL-prepare-IMP.1/2 2SG.ACC
'I was preparing you.' (Upper Bal Svan, after Topuria 1967:23, 73)

b. *si* *m*-*amār*-*äs* *mi*.
2SG.NOM 1SG.OBL-prepare-IMP.1/2 1SG.ACC
'You were preparing me.' (Upper Bal Svan, after Topuria 1967:23, 73)

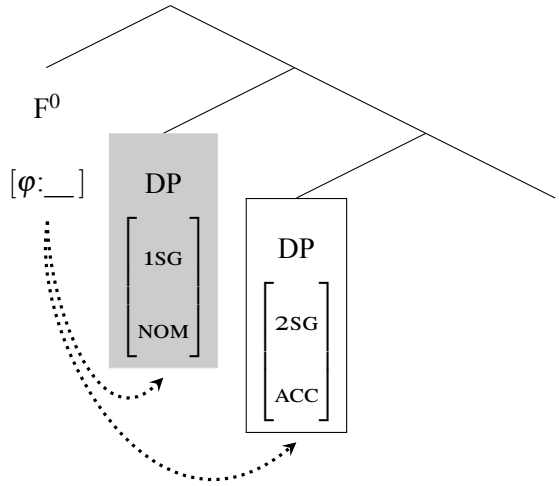
c. *mi* *m*-*aläṭ*-*xi* *si*.
1SG.DAT 1SG.OBL-love-STAT.2 2SG.NOM
'I love you.' (Lent'ekhi Svan, Topuria 1967:21)

d. *si* *ǰ*-*aläṭ*-*xwi* *mi*.
2SG.DAT 2.OBL-love-STAT.1 1SG.NOM
'You love me.' (Lent'ekhi Svan, Topuria 1967:21)

Since morphological considerations are fundamental to South-Caucasian prefixal agreement, our ultimate analysis should place some burden on the postsyntactic component of the grammar (presupposing, of course, Late Insertion, as in Distributed Morphology; Halle and Marantz 1993, et seq.). Let's assume provisionally that some probe in South Caucasian — call it F^0 — collects all the features features of all arguments in a clause. (This assumption will be interrogated and refined in Section 4.2.) F^0 might do so via Multiple Agree (Hiraiwa 2001, Nevins 2011), or by probing insatiably (Deal 2015). For purposes

of illustration, consider how this would work in a Svan clause like (96a), where there is a 1SG.NOM subject and a 2SG.ACC object.

(97)



For concreteness, I adopt the following binary features for person, number, and also case categories (98).[13] Using these, I define vocabulary entries for Svan agreement prefixes in (99).

(98) a. *Phi-features*

	±PART	±SPKR	±PL
1ST	+	+	±
2SG	+	−	±
3RD	−	−	±

b. *Case features*

	±DIR	±OBL
NOM	+	−
ACC	−	−
ERG	+	+
DAT	−	+

(99) Agreement Prefix Vocabulary Entries (Svan)

[13] On morphosemantics of binary φ-features, see Harbour (2016). Nothing hinges on these particular features, of course; this section could easily be reworked using privative φ-features arranged in a feature geometry, following Harley and Ritter (2002).

a. $\begin{bmatrix} F^0 \\ +\text{PART} \\ +\text{SPKR} \\ +\text{DIR} \end{bmatrix} \leftrightarrow$ *xw–* '1SG.DIR'

c. $\begin{bmatrix} F^0 \\ +\text{PART} \\ +\text{SPKR} \\ -\text{DIR} \end{bmatrix} \leftrightarrow$ *m–* '1SG.OBL'

b. $\begin{bmatrix} F^0 \\ +\text{PART} \\ -\text{SPKR} \\ +\text{DIR} \end{bmatrix} \leftrightarrow$ *x–* '2.DIR'

d. $\begin{bmatrix} F^0 \\ +\text{PART} \\ -\text{SPKR} \\ -\text{DIR} \end{bmatrix} \leftrightarrow$ *ǰ–* '2.OBL'

Returning to the derivation in (97), the postsyntactic component will be tasked with morphologizing the following syntactic terminal (100). Given the above vocabulary entries, two equally good morphemes could be inserted at F^0: *xw–* '1SG.DIR' exponing the subject's features, or *ǰ–* '2.OBL' exponing the object's.

(100) $\begin{bmatrix} F^0 \\ \begin{bmatrix} +\text{PART} \\ +\text{SPKR} \\ -\text{PL} \\ +\text{DIR} \\ -\text{OBL} \end{bmatrix}, \begin{bmatrix} +\text{PART} \\ -\text{SPKR} \\ -\text{PL} \\ -\text{DIR} \\ -\text{OBL} \end{bmatrix} \end{bmatrix}$

One might imagine such a tie along the Subset Principle could result in both vo-

cabulary entries being inserted (yielding a form like * *xw* -*ǰ*-*amār*- *äs* for 'I was preparing you', 96a), one or the other being inserted randomly (yielding free variation between * *xw* -*amār*- *äs* and *ǰ*-*amār*- *äs*), or even ineffability (yielding a paradigm gap).[14] None of these outcomes, of course, is the desired one. One tool in the Distributed Morphology arsenal which can solve a conundrum like this is impoverishment. This is a postsyntactic mechanism that changes features from a syntactic terminal before it is targeted for Vocabulary Insertion. The following impoverishment rule (101) would suffice for the present case, since it would change the 1SG subject's features in such a way to ensure that the vocabulary entry for *xw*– '1SG.DIR' (99a) could not be inserted. After impoverishment, only one agreement prefix could expone F^0: the desired one, *ǰ*– '2.OBL' (99d). An impoverishment analysis thereby correctly yields a verb like *ǰ*-*amār*- *äs* 'I was preparing you'.

(101) Prefixal Impoverishment (Svan)

$$\begin{bmatrix} +\text{PART} \\ +\text{DIR} \end{bmatrix} \rightarrow \begin{bmatrix} -\text{PART} \\ +\text{DIR} \end{bmatrix} \Big/ \begin{bmatrix} \underline{\quad}, & \overset{F^0}{\begin{bmatrix} +\text{PART} \\ -\text{DIR} \end{bmatrix}} \end{bmatrix}$$

This impoverishment rule is sufficiently general to prevent any direct prefix from being inserted whenever an oblique prefix could be inserted also. In other words, not it predicts the correct prefix not just for (96a), but also (96b–d). So, the analysis sketched here is descriptively adequate. However, it is not explanatory. There is no rhyme or reason to the precise formulation of (101); it could just as easily be formulated in a way to

[14] Coon and Keine (2019) assume ineffability always results when terminals bearing multiple arguments' bundles of features (like 100), not just when vocabulary entries tie in the eyes of the Subset Principle.

generate unattested forms (e.g., * *xw* *-amār-* *äs* for 'I was preparing you'). More specifically, this impoverishment rule makes no reference to the principle that I claim motivates South Caucasian prefix blocking: morphological expressiveness (I articulate a precursor to this argument in Foley 2017). Indeed, impoverishment rules are by nature arbitrary and unprincipled.[15] This isn't to say impoverishment should be abandoned — many observed morphological patterns are indeed arbitrary and unprincipled, so a powerful and relatively autonomous operation like impoverishment seems necessary. However, in this particular case, impoverishment misses out on a key generalization.

I suggest the best way to capture the generalization is to incorporate into the analysis a filter that makes direct reference to expressiveness. (102a) conveys the spirit of such a filter (cf. Kiparsky 2005); (102b) gives a more precise definition suited to the analysis developed in this section.

(102) Expressiveness (to be revised)

a. All else being equal, verb forms which register the features of more arguments with (overt) morphology are to be preferred over ones which register fewer arguments.

b. Should Vocabulary Insertion be faced with multiple vocabulary entries eligible to expone features of a terminal X^0 which tie by the metric of the Subset Principle, consider the larger morphological word M to which X^0 belongs. Insert whichever vocabulary entry at X^0 that expones features of arguments

[15] Some authors, though, have proposed that impoverishment rules can be united and/or motivated by more general constraints and principles — see, for example, Arregi and Nevins (2012), Ch. 4.

not already exponed by other vocabulary entries within M.

In functional terms, this filter disprefers morphological redundancy (cf. Foley's 2017 markedness constraint penalizing multiple exponence). And this is precisely what I wish to capture regarding prefix blocking in South Caucasian. Expressiveness is just as descriptively adequate as an impoverishment rule. In the case of the Svan Vocabulary Insertion tie discussed above, this filter compares the the whole verb forms that would result from inserting either vocabulary entry: i.e., * *xw* -*amār*- *äs* and *ǰ*-*amār*- *äs* for 'I was preparing you'. The former will be filtered out, of course, since it expresses the features of fewer arguments than the latter. But besides capturing the empirical patterns in South Caucasian, a principle like Expressiveness also captures the higher-level generalization motivating the phenomenon of prefix blocking.

4.4.2 Motivating the Principle of Minimal Compliance

The previous section outlined the general shape of the postsyntactic part of our analysis — to account for South Caucasian prefix blocking explanatorily, some principle that directly references morphological expressiveness is necessary. In this section, I refine the syntactic part of the analysis.

Let's take stock of the syntactic assumptions made so far. First, the three agreement loci in South Caucasian verb (occupied by TAM suffixes, the plural suffixes, and the agreement prefixes, respectively) each expone a distinct φ-probe along the clausal spine. The TAM-suffix probe is a fairly run of the mill one, not relativized to search for a specific φ-feature. As discussed in Section 3.1, this means that the TAM-suffix probe will always

Agree with the first argument with φ-features it can interact with. Consequently, we capture the rigidity of the TAM-suffix agreement slot — as long as we assume that dative DP subjects' φ-features are inaccessible to this probe (perhaps by virtue of a phasal PP layer encasing them), then the TAM-suffix will always Agree with the highest non-dative argument.

As for the plural-suffix probe, recall that it corresponds to a promiscuous agreement locus, which sometimes registers the features of a plural subject, and sometimes a plural object. In Section 3.2, I outlined one way to capture this pattern, supposing that the plural-suffix probe is relativized to search for [+PL]. Such a relativization allows this probe to ignore singular subjects in favor of plural objects, yielding promiscuity.

Finally, consider the prefix probe, another promiscuous locus. In Section 4.1, I assumed that this probe always Agrees with all arguments in a clause. This was necessary because there is no syntactic principle that can guide the probe to Agree uniquely with the argument whose features are ultimately exponed by the prefix. To capture the specific species of promiscuity seen in the South Caucasian agreement prefixes, then, it seemed necessary to collect all the arguments' features in the narrow syntax, and then have the postsyntactic component sort out just which argument's features to expone.

In sum, not only do we have three distinct φ-probes in South Caucasian, but each of these probes has different Agreement properties. Rigid agreement is rooted in a very broad relativization; promiscuous agreement arises either from a more narrow relativization, or a probe that Agrees more than once. There is nothing in principle wrong with such an a state of affairs. Indeed, if we believe the lexicon is the locus of all grammatical variation, and that φ-probe are ordinary lexical items, then we predict every conceivable combination

of every conceivable φ-probe.

However, we are still missing out on a notable meta-generalization about South Caucasian agreement: among all the agreement loci, exactly one is rigid, while the rest are promiscuous. This may very well be an accident of the South Caucasian lexicon. In that case, the chapter could conclude here. But it is worth asking if this meta-generalization can be derived from some independent grammatical principle.

In pursuing this goal, consider another difference between the TAM-suffix probe on the one hand, and the plural-suffix probe and the prefix probe on the other. Besides varying along the rigidity–promiscuity dimension, the two sets of probes vary in terms of how they interact with dative subjects. Consider (103). This data point shows that both the plural-suffix probe and the prefix probe can Agree with a dative subject. This fact is puzzling — indeed paradoxical — if we conceive of the TAM-suffix probe's inability to Agree with dative subjects in terms of the Phase Impenetrability Condition. If dative subjects' φ-features are sealed away behind a phase boundary, how can any probe — whatever its relativization or Agreement properties — access them?

(103) *tkven* *g*-*e*-*nax*-*a*-*t* *is*.

2PL.DAT 2.OBL-PLU-see.PFV-PST.3SG-PL 3SG.NOM

'You [PL] had seen them [SG].' (Georgian, Aronson 1990:273)

What unites these two observations — first, that one probe is rigid while the others are promiscuous; and second, that one probe (the very same one) cannot interact with dative subjects, while the other can — is locality. The TAM-suffix probe is obliged to obey standard

locality conditions (e.g., Minimal Link Condition and the Phase Impenetrability Condition), while the other probes are apparently able to ignore them in certain circumstances.

I argue that this selective application of locality constraints in South Caucasian agreement stems from the Principle of Minimal Compliance (the PMC; Richards 1997, 1998), whose original definition is given below. In a nutshell, the PMC allows a syntactic constraint to be lifted at a particular point in the derivation once it has been obeyed once there.

(104) Principle of Minimal Compliance

For any dependency D that obeys constraint C, any elements that are relevant for determining whether D obeys C can be ignored for the rest of the derivation for purposes of determining whether any other dependency D' obeys C.

An element X is relevant to determining whether a dependency D with head A and tail B obeys constraint C if:

a. X is along the path of D (that is, $X = A$, $X = B$, or A c-commands X and X c-commands B), *and*

b. X is a member of the class of elements to which C makes reference.

(Richards 1998:601)

To illustrate the PMC, consider the following data from Bulgarian. In multiple *wh*-questions in this language, all *wh*-phrases move to a left-peripheral position. A peculiar fact, though, is that while the structurally highest *wh*-phrase must be linearly first, any lower *wh*-phrases can be freely ordered after it.

(105) a. *Koj*$_1$ *kogo*$_2$ *kakvo*$_3$ *e* *pital* t_1 t_2 t_3*?*

who whom what AUX asked

'Who asked whom what?'

b. *Koj*$_1$ *kakvo*$_3$ *kogo*$_2$ *e* *pital* t_1 t_2 t_3*?*

who what whom AUX asked

'Who asked whom what?' (Bulgarian, Richards 1997:332)

Richards explains this word order variation in the following way. The first time C^0 probes for a *wh*-phrase, it is forced by locality conditions (relevant here is the Minimal Link Condition) to attract the highest such phrase (wh_1; here the subject, *koj* 'who'). In doing so, C^0 has now minimally complied with this locality constraint. So by virtue of the PMC, C^0 has a choice when probes another time. It may either attract the next-highest *wh*-phrase (wh_2; *kogo* 'whom'), continuing to obey the Minimal Link Condition, or it may skip over that one and attract the lowest *wh*-phrase (wh_3; *kakvo* 'what') instead. In other words, the PMC results in derivational fork in the road. Either wh_2 or wh_3 can be attracted to Spec-CP after wh_1 is. This indeterminacy is welcome, though, since we observe free variation. The two possible word orders (105a, b) reflect the two well-formed outputs of the single starting derivation.

This fact about Bulgarian multiple *wh*-movement parallels South Caucasian agreement in the following way. In Bulgarian, exactly one *wh*-probe is rigidly constrained by locality conditions, while the subsequent ones are not. In South Caucasian, exactly one φ-probe is rigidly constrained by locality conditions, while the others are not.

In order to use the PMC to derive the South Caucasian facts, it is necessary to adjust a few of our syntactic assumptions. First, each of the agreement loci still corresponds to a single φ-probe, but these three probes must all be in the same syntactic position in order for the PMC to lift locality constraints. I identify the position as T^0. Second, for reasons that will soon be clear, it is no longer necessary to give the three probes different relativizations or Agreement properties. All three may simply be vanilla probes that are satisfied by any set of φ-features. Such a probe I notate here [φ:__] (as I do above). Third, the three subprobes on T^0 must be ordered; subprobe n Agrees only after subprobe $n-1$ does.

Consider the predictions of these assumptions. When T^0 is merged, its first subprobe ([φ_1:__]) Agrees. Constrained by the Minimal Link Condition and the Phase Impenetrability Condition, [φ_1:__] must copy the features of the highest nondative argument. In nondative-subject constructions, the goal is therefore the subject; this step of agreement counts as minimal compliance with with the Minimal Link Condition, since [φ_1:__] indeed probes the most local goal. In dative-subject constructions, the goal will be the object; this step of agreement counts as minimal compliance with the Phase Impenetrability Condition, since [φ_1:__] indeed does not probe across a phase boundary. After this first step, one or another locality constraint has been minimally compied with. Thus the PMC permits the next subprobe ([φ_2:__]) to ignore that constraint, and if there is more than one argument in the clause, derivational indeterminacy will result. Along one forking path, [φ_2:__] obeys locality constraints again, Agreeing with the highest nondative argument. Along the other path, [φ_2:__] ignores the relevant constraint, Agreeing with the erstwhile inaccessible argument. This will either be the object in a nondative-subject construction (a possibility now

that the Minimal Link Condition is lifted), or the subject in a dative-subject construction (a possibility now that the Phase Impenetrability Condition is lifted). As for [φ_3:__], it too may Agree with any argument.

A few words on probe–morpheme mapping. It is crucial that the TAM-suffixes expone T^0's first subprobe ([φ_1:__]), as it is the one uniquely constrained by locality considerations. As for the other probes, I will assume that the agreement prefixes expone [φ_2:__] and the plural suffixes expone [φ_3:__]. This mapping is not accidental: TAM-suffixes, agreement prefixes, and plural suffixes express decreasingly many types of features. If the Subset Principle is a primary driver of morphological exponence (Halle 1997), then the morphological component should prioritize inserting morphemes which express person, number, and TAM features over ones which express merely person and number features, and those should be prioritized over morphemes which express merely number features.

A theory of South Caucasian agreement that employs the PMC enjoys an important advantage over the one previously developed: it explains the meta-generalization that there is exactly one rigid agreement locus. This fact is reduced to a PMC effect, one entirely parallel to Bulgarian multiple *wh*-question facts discussed in (105).

However, there is one glaring difference between Bulgarian multiple *wh*-movement and South Caucasian Agreement: we find free variation only in the former. The PMC can introduce derivation indeterminacy, and this was desirable for Bulgarian, since it derives the observed word order flexibility. But in South Caucasian, any given argument combination is mapped deterministically to a single set of agreement morphemes. Since there is no particular goal that T^0's second and third subprobes must target, this PMC analysis actually

predicts x^2 well-formed syntactic outputs for a derivation with x arguments.

How can we sort out this indeterminacy? Postsyntactic filtration is a suitable solution. Adjusting the definition of Expressiveness to better reflect this theory's input to the morphological module of the grammar (106), we are now prepared to tie together the syntactic and postsyntactic components of this analysis.

(106) Expressiveness (final version)

a. All else being equal, verb forms which register the features of more arguments with (overt) morphology are to be preferred over ones which register fewer arguments.

b. Should the PMC result in more than one well-formed output of a single numeration, compare how each of the outputs would be expressed morphologically. Filter out all outputs except the one(s) whose verb maximizes morphological expressiveness.

In the theory outlined in Section 4.1, whether a probe Agreed rigidly or promiscuously depended on its lexical properties: what it could Agree with, and how many times. In the present theory, promiscuity is a byproduct of the PMC. Probes — at least ones specified to Agree with any φ-bearing argument — behave rigidly by default. Under the right circumstances, though, constraints which ensure probes' rigidity may be lifted by the PMC.

4.4.3 The Analysis Illustrated

While the previous section outlined the final analysis in abstract terms, this section looks at a few specific examples to better illustrate the system. Let's first consider a 1≫2 nondative-subject construction, like (107). The derivation of such a clause is schematized in Figure 4.2.

(107) *mi* *ǰ*-*amār*- *äs* *si*.

1SG.NOM 2.OBL-prepare- IMP.1/2 2SG.ACC

'I was preparing you.' (Upper Bal Svan, after Topuria 1967:23, 73)

This being a nondative-subject construction, [φ_1:__] must probe the subject. Recall that this probe is the one that will be exponed by a TAM suffix. By Agreeing the the subject, [φ_1:__] minimally complies with the Minimal Link Condition. Next [φ_2:__], the probe which will ultimately be exponed by an agreement prefix, has a chance to Agree. With the Minimal Link Condition lifted by the PMC, the subject and the object are equally good goals for this probe, so the derivation branches. Down one path (on the left), [φ_2:__] targets the subject; down the other (on the right) it targets the object.

For reasons of space, Figure 4.2 omits [φ_3:__] entirely. An exhaustive representation of this derivation would actually branch each of these well-formed outputs two more times each, yielding four possible ways in which T^0's three subprobes could grammatically interact with the two arguments of this monotransitive clause. Recall that [φ_3:__] is exponed by the plural suffixes — morphemes that definitionally express a [+PL] feature. But since this derivation has only singular arguments, no matter which argument [φ_3:__] probes,

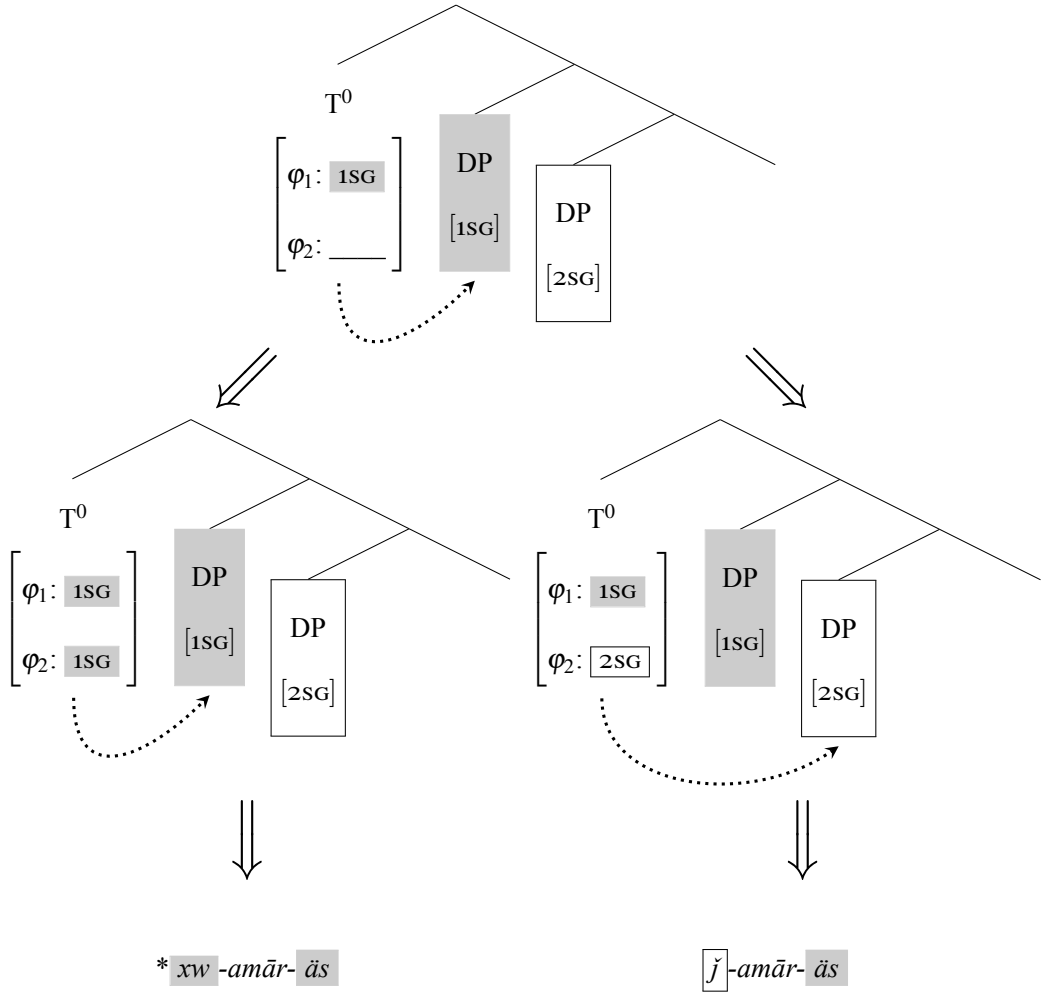


Figure 4.2: Deriving agreement in a 1≫2 nondative-subject clause (Upper Bal Svan)

there will be no vocabulary entry able to expone it. Omitting [φ_3:__] in this derivation, then, is innocuous.

Let's skip, then, directly to the postsyntactic component of the grammar. The lowest downwards-pointing arrows elide standard Distributed-Morphology style exponence operations (as in Halle and Marantz 1993, et seq.), which presumably operate cyclically. What's important is the whole verb form. After Vocabulary Insertion, the morphology will

compare how each of the well-formed syntactic outputs would be morphologized. In all the competing outputs, [φ_1:__] will be exponed by the TAM suffix *–äs* 'IMP.1/2'. In those outputs where [φ_2:__] Agreed with the subject, this probe will be exponed by *xw–* '1SG.DIR'. In those outputs where it Agreed with the object, *ǰ–* '2.OBL' will be inserted. There being no possible exponent for [φ_3:__] in a SG≫SG clause, whichever argument this third probe interacts with is immaterial for the final shape of the verb. So, across the four syntactic outputs of this derivation, the verbs in two of them will be exponed with the form * *xw* -*amār*- *äs* , and the verbs in the other two will be exponed *ǰ* -*amār*- *äs* . Expressiveness compares these competitors, and prefers the later form since it expresses the features of both the subject and the object, rather than just the subject.[16]

For a second illustration, consider a 2≫1 dative-subject verb (108, Figure 4.3).

(108) *si* *ǰ* -*aläṭ*- *xwi* *mi* .

1SG.DAT 2.OBL -love- STAT.1 1SG.NOM

'You love me.' (Lent'ekhi Svan, after Topuria 1967:21)

Besides the features of the subject and object, the only difference between Figures 4.2 and 4.3 is how locality principles guide [φ_1:__]. The Phase Impenetrability Condition prevents this first probe from targeting the subject, so this probe will have Agreed with the object in all syntactic outputs. Downstream, this syntactic difference means that prefixal subject agreement (expressed with *ǰ–* '2.OBL') will be preferred to prefixal object agreement.

[16] Technically speaking, the output where [φ_3:__] probes the subject and the one where it probes the object tie in the eyes of Expressiveness. So in a certain sense this analysis does predict free variation for this argument combination, but it is not the kind of free variation that has morphophonological consequences.

This shows how Expressiveness derives the preference for oblique prefixes, whether they register the subject or the object.

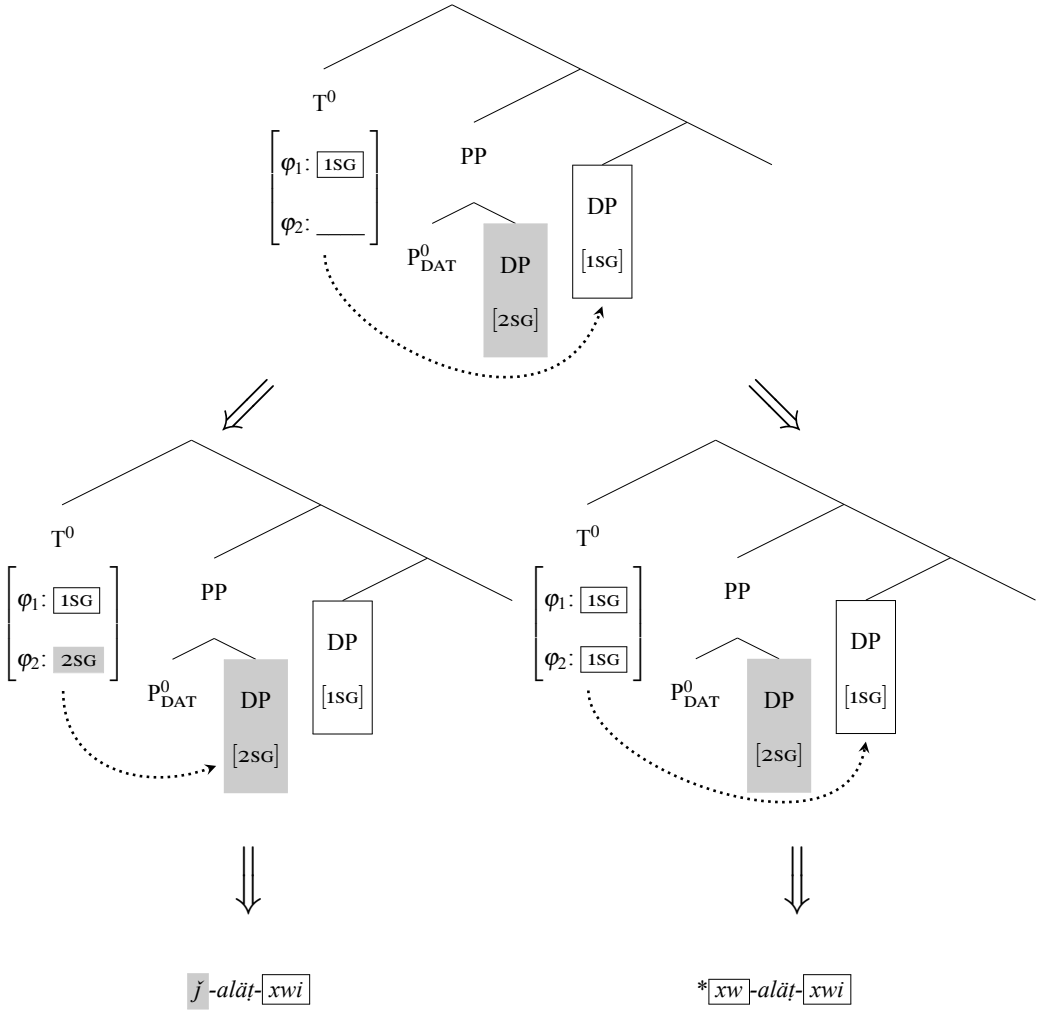


Figure 4.3: Deriving agreement in a 2≫1 dative-subject clause (Lent'ekhi Svan)

Next let's consider a derivation with a first-person nominative subject and a third-person accusative object. Recall from Section 2 that the South Caucasian languages altogether lack agreement prefixes that track third-person O_{DO}s.

(109) *mi* *xw* *-amār-* *äs* *eǰas*.

1SG.NOM 1SG.DIR -prepare- IMP.1/2 3SG.ACC

'I was preparing them [SG].' (Upper Bal Svan, after Topuria 1967:23, 73)

This morphological gap has an important ramification for the way Expressiveness will operate here. This filter is defined in such a way to prefer overt agreement morphology to null agreement morphology (or the lack of agreement morphology), all else being equal (106). We saw in Figure 4.2 that Expressiveness filters out derivations where [φ_2:__] probes a first-person subject in 1.NOM≫2.ACC contexts. Here, however, Expressiveness will ultimately favor an output where both [φ_1:__] and [φ_2:__] copy the features of the first-person subject. That's because in the alternative case, where [φ_2:__] instead probes the third-person object, the verb could not have any agreement prefix. The attested form *xw* *-amār-* *äs* is more expressive than the unattested **amār-* *äs*, since the latter does not expone the subject's [+SPKR] feature at all (the TAM suffix – *äs*, since it can be controlled by both first- and second-person subjects, expones just [+PART]).

A final example illustrates the utility of the second of Kiparsky's (2005) morphological constraints, Economy. While Expressiveness pulls much of the weight in this analysis of South Caucasian agreement, certain corners of the paradigm require us to look towards another filter — one which prefers verb forms with as few morphemes as possible. The following definition of Economy will serve our purpose.

(110) a. All else being equal, verb forms with fewer morphemes are to be preferred over ones with more morphemes.

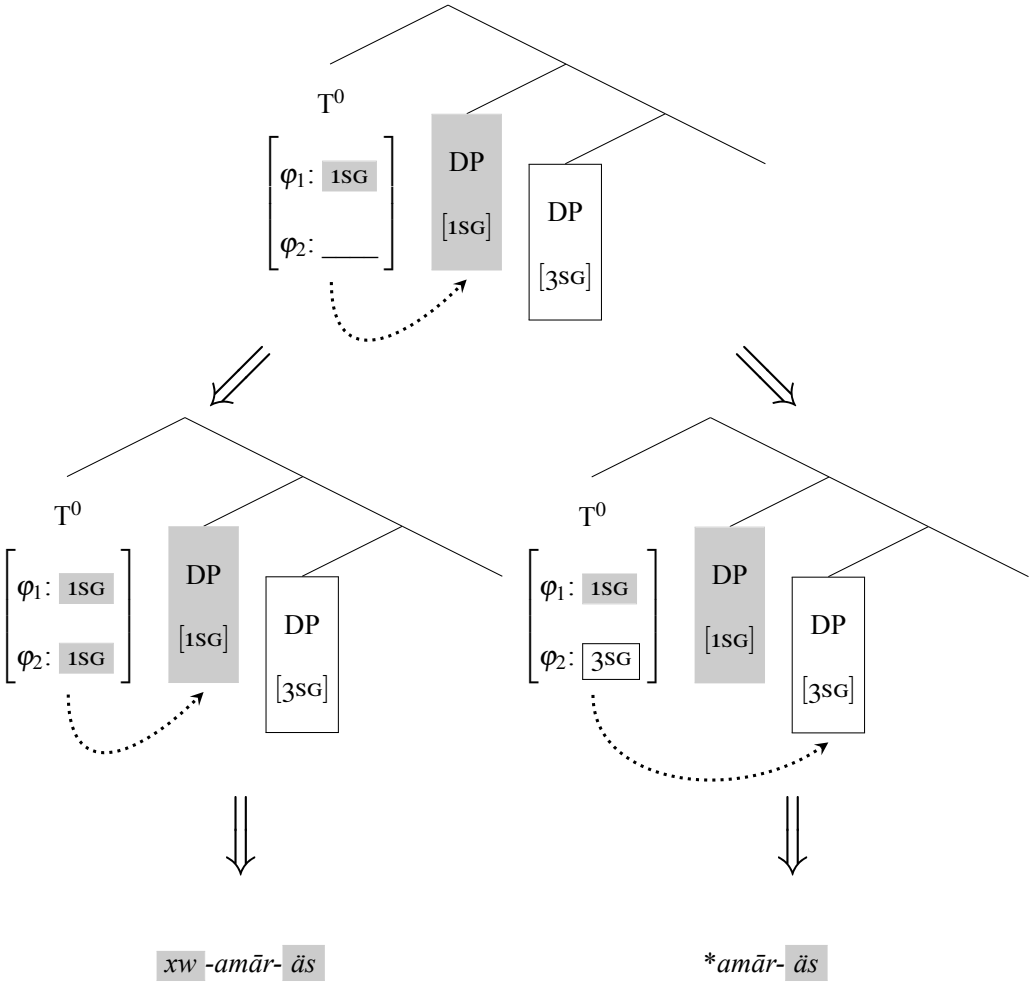


Figure 4.4: Deriving agreement in a 1≫3 dative-subject clause (Upper Bal Svan)

b. Should the PMC result in more than one well-formed output of a single numeration, compare how each of the outputs would be expressed morphologically. Filter out all outputs except the one(s) whose verb maximizes morphological economy.

Relevant forms are shown in below. Example (111a) demonstrates that the plural suffix *–x* 'PL' appears in a 3SG.NOM≫2PL.ACC verb. Compare (111b–c): while these

examples have a plural object, the plural suffix is ungrammatical. Intuitively, the important difference between these cases is that Svan has oblique agreement prefixes in its morphological inventory that distinguish number for the first person (namely *m–* '1SG.OBL', *n–* '1EXCL.OBL', and *gw–* '1INCL.OBL'), but only a single number-neutral oblique prefix for the second person (*ǰ–* '2.OBL'). Consequently, it would be morphologically redundant if a verb with a first-person plural accusative object bore both the appropriate oblique prefix and the plural suffix *–x* 'PL', as in (111c). This ungrammatical form **gw-amār-a-x* would multiply expone the object's [+PL] feature — and, as I argue in Foley (2017) for Georgian, the South Caucasian languages' agreement systems conspire against the multiple exponence of number features.

(111) a. *eǰa* *ǰ-amār-a-x* *sgäy*.
3SG.NOM 2.OBL-prepare-IMP.3SG-PL 1PL.ACC
'They [SG] were preparing you [PL].'

b. *eǰa* *gw-amār-a* *näy*.
3SG.NOM 1INCL.OBL-prepare-IMP.3SG 1PL.ACC
'They [SG] were preparing us [INCL].'

c. * *eǰa* *gw-amār-a-x* *näy*.
3SG.NOM 1INCL.OBL-prepare-IMP.3SG-PL 1PL.ACC
Attempted: 'They [SG] were preparing us [INCL].'

(Upper Bal Svan, after Topuria 1967:23, 73)

In the derivations for (111a) and (b), the behavior of [φ_3:__] is crucially impor-

tant, but to save space, I will represent the derivations in a more compact way. Consider first the 3SG.NOM≫ 2PL.ACC structure (Figure 4.5). The four syntactic outputs of this derivation would each be exponed differently. Expressiveness will filter out all but the output in which [φ_1:__] probes the 3SG subject, and both [φ_2:__] and [φ_3:__] probe the 2PL object; this output is exponed *ǰ-amār- a -x* 'they [SG] were preparing you [PL]'. Economy is not relevant for this derivation because there is no alternative morphological form which expresses the features of the verb's arguments equally well but with fewer morphemes.

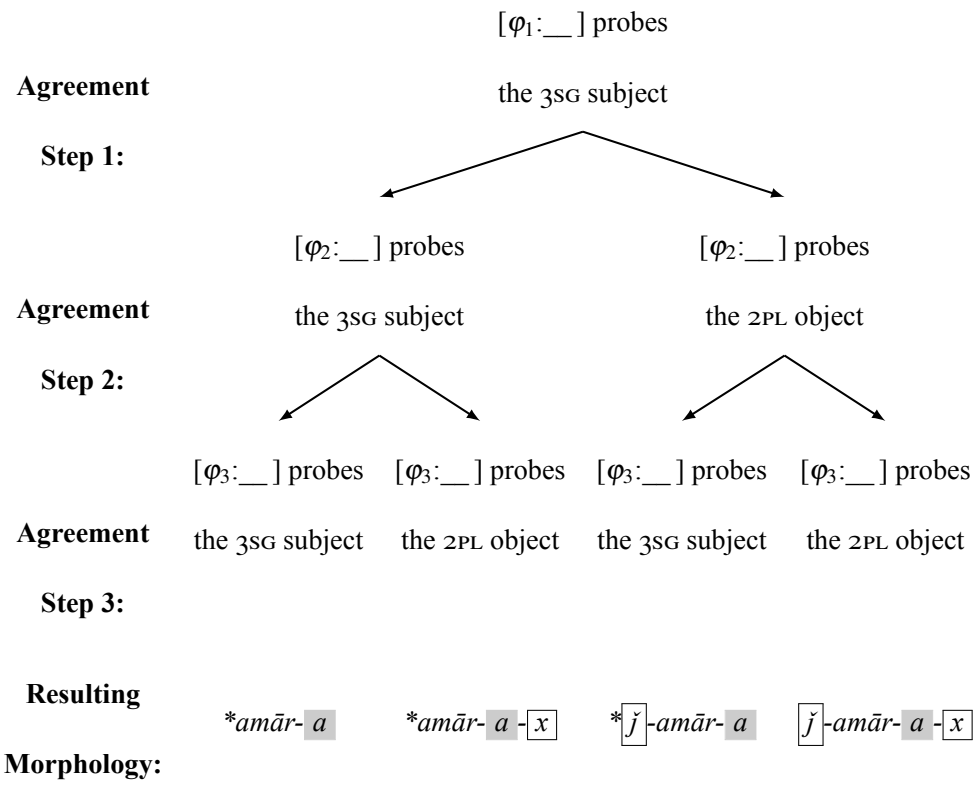


Figure 4.5: Deriving agreement in a 3SG.NOM≫2PL.ACC clause (Lent'ekhi Svan)

Now turn to the 3SG.NOM≫1INCL.ACC structure (Figure 4.6). There are two mor-

phological forms which express as many of the arguments' features as possible, including the object's [+PL] feature: *gw*-*amār*- *a* and **gw*-*amār*- *a* -*x* for 'they [SG] were preparing us [INCL]'. In other words, these two forms tie in the eyes of Expressiveness. Economy, however, prefers the first form, since it expresses those features more efficiently, using two overt agreement morphemes instead of three. Together, then, Expressiveness and Economy filter out all but the derivation in which [φ_1:__] and [φ_3:__] probe the 3SG subject, while [φ_2:__] probes the 1INCL object.

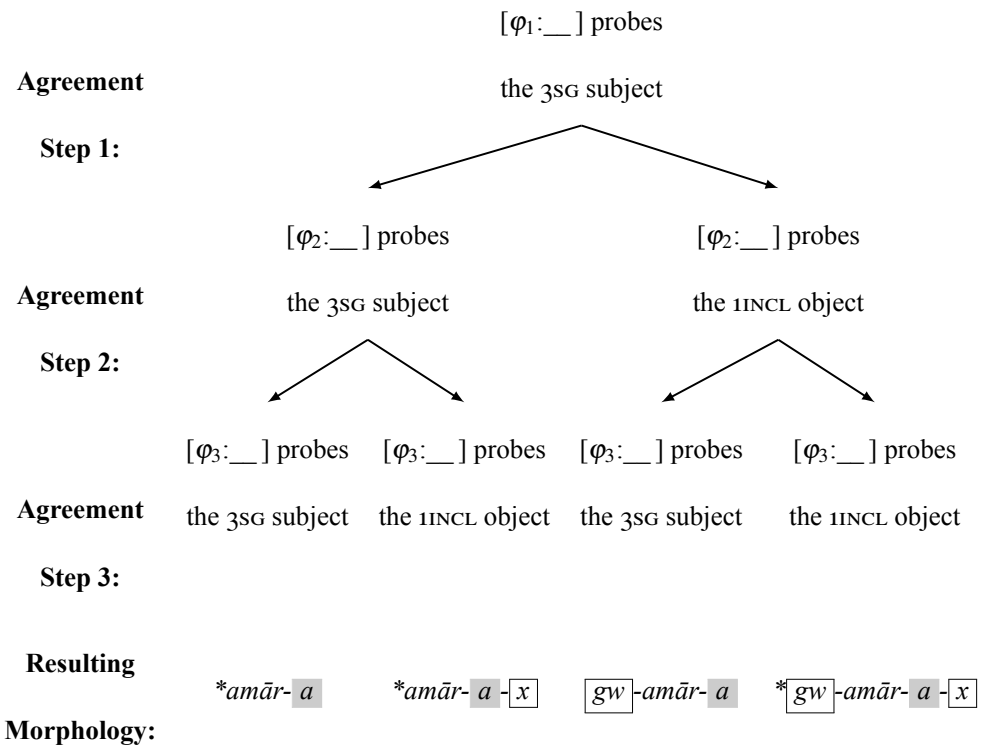


Figure 4.6: Deriving agreement in a 3SG.NOM≫1INCL.ACC clause (Upper Bal Svan)

4.5 Conclusion

This chapter has had two analytical goals. One was to articulate descriptive generalizations over the complex agreement phenomena of the South Caucasian languages, and meta-generalizations over those. Chief among these are (112) and (113). Together, these generalizations identify a remarkable amount of order and systematicity in a set of data which at first glance may seem overwhelmingly complicated.

(112) Among the three agreement loci in South Caucasian verbs, exactly one exhibits rigid agreement (agreement whose controller remains stable across a paradigm) while the other two exhibit promiscuous agreement (agreement whose controller varies across cells of a paradigm).

(113) The phenomenon of prefix blocking in these languages (whereby one agreement prefix which could conceivably appear on a verb is systematically blocked by another) is best characterized in postsyntactic terms, as a way that the agreement system maximizes morphological expressiveness.

The chapter's second analytical goal was to develop a theory of South Caucasian agreement which captures these generalizations. The proposed analysis has both a narrow-syntactic component and a postsyntactic one. The narrow-syntactic part of the analysis captures generalization (112) through a novel application of the Principle of Minimal Compliance (PMC; Richards 1997, 1998). The PMC was originally conceived primarily with movement, binding, and ellipsis phenomena in mind, but it makes predictions in the domain

of φ-agreement — predictions which I claim are clearly borne out the South Caucasian languages. This is not the first extension of the PMC in agreement-related phenomena (Preminger 2019, for instance, argues that PMC can help us better understand clitic doubling) but it strengthens the idea that the PMC is a very general grammatical principle. The PMC is very mysterious, and this chapter makes no attempt to explain its precise formulation or derive it from deeper principles. But the fact that PMC phenomena can be seen in such a menagerie of syntactic phenomena — multiple *wh*-movement, reflexive binding, weak cross over, VP ellipsis, subjacency (Richards 1997, 1998), clitic doubling (Preminger 2019), and now South Caucasian agreement — should assure us that the PMC has identified some key facet of natural language.

The analysis's postsyntactic component accounts for generalization (113) by employing derivational filters that refer explicitly to morphological Expressiveness and Economy (Kiparsky 2005). It is indeed a powerful tool to compare surface forms that correspond to distinct syntactic outputs, and to do so globally (considering whole verb-forms) rather than strictly locally. I have attempted to narrow the scope of this postsyntactic competition by invoking it only in cases where derivational indeterminacy is a consequence of the PMC. It seems to me, though, that something postsyntactic filtration it will be necessary for an explanatory account of South Caucasian agreement, given that prefix blocking so eludes a neat syntactic characterization, especially in light of the syntactic properties of dative subject constructions. Of course, it remains an open question just what level of explanation we are pursuing. It may ultimately be that Expressiveness and Economy are not part of individuals' synchronic grammars per se, but rather principles that guide acquisition of the system,

or its diachronic development.

This chapter does not claim to offer a comprehensive analysis of South Caucasian agreement. I have attempted to avoid an overly parochial perspective by focusing on patterns which are the most stable across the family, and which offer especially compelling theoretical footholds. But in doing so I have necessarily neglected corners of these languages' agreement paradigms which, at least at first glance, pose a challenge to the generalizations offered here. I refer an interested reader to the excellent and extremely thorough description of Tuite (1998), and theoretical work including Atlamaz (2013), Foley (2017), Thivierge (2019), Blix (2020) and Bondarenko & Zompì (2020), which investigates South Caucasian agreement phenomena from a variety of enlightening angles.

Likewise, this chapter does not claim to offer a general theory of promiscuous agreement. The South Caucasian facts are remarkably well suited for a derivation cast in terms of the PMC, given generalization (112). But it is plainly not the case that promiscuity only arises as a PMC effect. If this were true, promiscuous agreement should always exist alongside rigid agreement — but this can be easily refuted by, for example, Ayutla Mixe (64–65), whose verbs have two promiscuous agreement loci but no rigid loci.

Nevertheless, the analysis outlined here might still contribute to a path forward for future research on agreement. In recent years, a wealth of syntactic mechanisms have been proposed to account for complex agreement phenomena — including but not limited to upwards Agree (Baker 2008, Bjorkman and Zeijlstra 2019), Cyclic Agree & cyclic domain expansion (Béjar and Rezac 2009), Multiple Agree (Nevins 2011), conjunctive feature satisfaction (Coon and Bale 2013), interaction and satisfaction conditions on Agree (Deal

2015), and Feature Gluttony (Coon and Keine 2019). While some research has explored the predictions made by combinations of these mechanisms (e.g., Despić et al. 2019), it seems unlikely that all of them are available to the grammar. The analysis offered here needs no non-standard Agree technology, but by it does employ a syntactic principle (the PMC) which is not often associated with φ-agreement. This general tactic — adopting the minimal assumption about Agree necessary, but looking for parallels with very different domains of syntax — may prove a fruitful one in developing a theory of φ-agreement and related phenomena.

Chapter 5

Conclusion

5.1 Summary of key findings & conclusions

Harmonic alignment of syntactic role and animacy is an important strategy for parsing transitive root clauses. Experiments 1a and 1b (Chapter 2) show that Georgian comprehenders navigate clauses with two human arguments very differently than they do clauses with two inanimate arguments. I argue that the precise distribution of reading-time disruptions across these experiments follows from a straightforward parsing strategy: given a high-animacy noun phrase ambiguous for syntactic role, assign it to the most prominent available argument position; given a low-animacy noun phrase, assign it to the least prominent position. Recall that a pre-verbal nominative argument is compatible with a large set of syntactic roles (transitive subject, intransitive subject, or direct object); likewise for a pre-verbal dative argument (which might be a transitive subject, a certain kind of intransitive subject, an experiencer subject, a direct object, or an indirect object). Results from these

experiments support a theory in which parsers seek to harmonically align a scale of animacy classes (at least human $\succ$ inanimate) with a very specific scale of syntactic roles (viz., transitive subject $\succ$ indirect object $\succ$ direct object), all the while obeying the language's case-assignment rules and reparsing prejacent clausal material only as a last resort.

Processing ergative inanimates is inherently difficult. Experiment 1b (Section 2.4) shows that ergative-marked inanimate noun phrases are reliably associated with processing difficulty. This is the case even when an ergative inanimate is the first argument encountered. This ergative cost, then, cannot be characterized as a garden path effect. Rather, there must be some inherent difficulty associated with ergative-marked nouns, as has also been observed for Hindi (Choudhary et al. 2007). Even though some intransitive verbs in Georgian can take ergative subjects — a point I return to in Section 5.2 — I attribute this effect to the dissonance associated with integrating inanimacy (a low-prominence property) and transitive subjecthood (a high-prominence property). Intriguingly, we observe no mirror-image cost associated with high-animacy arguments which are unambiguously direct objects, but this asymmetry has been documented in other languages (Bornkessel-Schlesewsky and Schlesewsky 2009, Fauconnier 2012), and likely indicates that the comprehender places more stringent contraints on agents than patients.

Ergative morphology within a relative clause causes a filled-gap effect. In three out of four relative-clause processing experiments (Chapter 3), we observe reading-time disruptions at ergative-marked noun phrases within a relative clause. This must be an independent effect of the inanimate-ergative cost, since even ergative arguments referring to humans trigger slowed reading times within relative clauses. I interpret this as a

filled-gap effect (Crain and Fodor 1985, Stowe 1986). Upon recognizing a filler–gap dependency (as signaled by a relative pronoun, for instance), the comprehender immediately anticipates that the gap will be in the highest upcoming subject position. Subsequent ergative morphology within the $\overline{\text{A}}$-construction foils this prediction, since an ergative argument can only be a subject, and an overt subject within a relative clause is incompatible with a subject-gap interpretation. This evidence supports the view that Georgian comprehenders prioritize subject gaps above all others when processing $\overline{\text{A}}$-dependencies, rather than, say, gaps associated with the least informative morphological case category, as has been argued for Basque (Carreiras et al. 2010) or Avar (Polinsky et al. 2012).

Blocking effects in Caucasian agreement can be seen as maximizing morphological economy and expressiveness. In certain cells of the South Caucasian languages' verbal agreement paradigms, conspicuously absent are morphemes which one might expect to appear given independent observations. I argue it is no coincidence just which morpheme is blocked in these environments: the missing morpheme would have been redundant for comprehension, and therefore uneconomical for production (Horn 1984, Kiparsky 2005). More precisely, a comprehender can triangulate the φ-features of the subject and object just as easily from the observed verb form as a hypothetical form which includes the missing affix; including that affix would only further burden the production system. Building on this intuition, I suggest that the morphological component has access to alternative complete verb forms, which can be optimized with respect to constraints enforcing speaker- and hearer-based economy.

Asymmetrical behavior across different classes of South-Caucasian agree-

ment morphemes can be seen as a consequence of the Principle of Minimal Compliance. I argue that the distribution of agreement morphemes across verbal inflectional paradigms can be fruitfully cast as a consequence of Principle of Minimal Compliance (Richards 1997, 1998), since one locus of agreement only every tracks the highest non-dative argument, while the others are not so tightly constrained. In an abstract way, this parallels more familiar Minimal-Compliance effects, of the kind we see in Bulgarian multiple *wh*-questions, where the leftmost peripheral position must be occupied by the highest *wh*-phrase, but lower *wh*-phrases can be freely ordered after the first. Employing the Principle of Minimal Compliance in my analysis has two desirable consequences: first, it introduces the derivational indeterminacy necessary for a morphological filtration account of morpheme blocking; second, it avoids positing syntactic mechanisms apparently bespoke to φ-agreement phenomena.

5.2 Directions for future research

This dissertation has articulated a number of questions to plumb moving forward; I discuss a few briefly here. First, under what circumstances will comprehenders entertain an intransitive parse? Theoretically speaking, intransitivity seems optimal for the parser (Bornkessel-Schlesewsky and Schlesewsky 2009), as intransitive clauses require the comprehender to integrate the lexical semantics of a verb and just a single noun phrase. Without having to juggle more than one argument, there's no danger that the parser's wires might get crossed during syntactic-role assignment. And since there are nominative, ergative, and

dative intransitive subjects alike in Georgian, an intransitive parse is always available one argument into a clause. Yet, across six self-paced reading experiments there is little evidence to suggest that encountering the second argument of an ultimately transitive clause foils an initial intransitive parse. One possible explanation is that adding an argument position to a syntactic parse associated which such a small processing cost that the self-paced reading methodology cannot detect it. If this is the case, we might expect to find intransitive garden-path effects using more sensitive psycholinguistic methods, like eye-tracking while reading. A more intriguing alternative, though, is that 'canonical' transitive clauses — ones with an animate agent and an inanimate patient — for some reason enjoy a privileged status in online sentence comprehension, so much so that they eclipse intransitive and non-canonical transitive parses out of the blue. The notion of canonical transitivity is a useful one for explaining many grammatical and typological phenomena (e.g., Hopper and Thompson 1980), and it would shed important light on the nature of human cognition if such canonicity truly guided sentence comprehension in ways which are disentangleable from, say, the mere frequency distributions of different types of argument structures and animacies.

Second, where do ditransitives fit into the sentence-processing picture? My account of effects observed in Experiments 1a and 1b hinges on the assumption that NOM–DAT and DAT–NOM strings can lead comprehenders down a ditransitive garden path. This interpretation can very easily be tested in Georgian by manipulating the argument structure of the clause-final verb. (And since Georgian has productive applicative and causative constructions, monotransitive–ditransitive pairs can be readily formed from the same lexeme.) Specifically, I predict that a ditransitive verb that follows a NOM–DAT or DAT–NOM sequence

(114a) will be easier to process than a monotransitive one (114b).

(114) a. *msaxiob-i mxaṭvar-s ◌ gaacnobs.*

actor-NOM painter-DAT 3.DAT introduce:FUT.3SG

'The actor will introduce them [SG/PL] to the painter.'

(or 'The actor will introduce the painter to them [SG/PL].')

b. *msaxiob-i mxaṭvar-s gaicnobs.*

actor-NOM painter-DAT meet:FUT.3SG

'The actor will meet the painter.' (Constructed examples)

If indeed we find no garden-path effect in clauses like (114a), this further challenges Bornkessel-Schlesewsky and Schlesewsky's (2009) Distinctiveness principle. Insofar as the two arguments of a monotransitive verb are inherently less distinct than the single argument of an intransitive, the three arguments of a ditransitive will be less distinct still.

Prominence vs. accessibility: I argue that O_{IO} syntactic role is more prominent than the O_{DO} role, since certain dative noun phrases in Experiments 1a and 1b seem to lead to be parsed as indirect objects. This contrasts with Keenan and Comrie's (1977) Accessibility Hierarchy, where direct objects outrank indirect objects. Evidence for this accessibility relationship comes from languages like Beijing Chinese, Hausa, Persian, and Welsh, where certain relativization strategies permit gaps in O_{DO} position but not O_{IO} position (Keenan and Comrie 1977:76–79); there are no relativization strategies in Keenan and Comrie's sample which permit gaps in O_{IO} but not O_{DO} position. There are a few conceivable ways to square this circle.

- **Prominence and accessibility are distinct:** Keenan and Comrie's hierarchy is at least partially rooted in frequency — all clauses have subjects, fewer have direct objects, fewer still have indirect objects, etc. If parsers predict the likelihood that an upcoming gap will occur in a given syntactic position given the absolute frequency of clauses with that position, ranking direct objects over indirect objects is a logical strategy. Prominence, on the other hand, is rooted in argument prototypicality and the degree to which a multiplicity of syntactic and semantic features are harmonically aligned. Indirect objects are typically associated with the affectee semantic role, while direct objects are often patients. And since affectees are canonically animate while patients are canonically inanimate, it follows that O_{IO}s should be more prominent than O_{DO}s. This explanation, though, leaves us a with lingering mystery: why should there be a special set of parsing principles reserved for filler–gap dependencies?

- **Prominence and accessibility are the same, but typological facts about relativization may obscure their unity:** There is scant empirical evidence indicating that O_{IO} gaps are indeed harder to process than O_{DO} gaps (or vice versa). It may be that the O_{IO} position is indeed universally more accessible than the O_{DO} position during filler–gap processing. Grammatical asymmetries like those in Persian and Welsh would then need another explanation. (Consider also Ergative extraction asymmetries, which are another thorn in the side of the Accessibility Hierarchy.)

- **Indirect objects heterogeneous:** This is hard to deny — even internal to English, affectees can be expressed in multiple syntactic positions (e.g., the first noun-phrase

complement of a double-object construction or the complement of a *to*-PP in a prepositional-object ditransitive; high vs. low applicatives, Pylkkänen 2008). Perhaps, then, comparing how different non-subject gaps are processed is an inherently fraught venture, especially across languages.

Finally, I wish to speculate briefly on the nature of the Principle of Minimal Compliance. It is certainly a very peculiar grammatical mechanism; few others exempt structures from otherwise inviolable constraints on syntactic dependencies. So I hypothesize that the PMC may not be universal grammatical principle at all, but rather a epiphenomenon of limitations on syntactic processing. Perhaps calculating whether a given dependency obeys a grammatical generalization at a given position requires the parser to allocate a certain number of cognitive resources, and calculating that same dependency again at the same position overloads the parser's bandwidth. This overload results in a grammaticality illusion (Gibson and Thomas 1986), which over time becomes grammaticized.

As an illustration, consider the multiple *wh*-question paradigm from Bulgarian discussed throughout Chapter 4. Strictly speaking, the wh_1–wh_3–wh_2 worder order in (115b) constitutes a Superiority violation (Chomsky 1973). And perhaps at some previous stage in the history of Bulgarian, this sentence was ungrammatical. But comprehenders failed to recognize such Superiority violations often enough that they eventually internalized them as begin just as grammatical as Superiority-obeying questions like (115a).

(115) a. *Koj*$_1$ *kogo*$_2$ *kakvo*$_3$ *e* *pital* t_1 t_2 t_3*?*

who whom what AUX asked

'Who asked whom what?'

b. *Koj*$_1$ *kakvo*$_3$ *kogo*$_2$ *e* *pital* t_1 t_2 t_3*?*

who what whom AUX asked

'Who asked whom what?' (Bulgarian, Richards 1997:332)

This account makes a few clear predictions. First, we should not observe PMC effects in all grammars. In other words, we expect to find languages whose multiple *wh*-questions must all strictly obey Superiority, as in this hypothetical Proto-Bulgarian. Second, it may be possible to induce PMC-related grammaticality illusions in any language, given the right syntactic structure and confluence of processing factors.

Bibliography

Abels, Klaus. 2003. Successive cyclicity, anti-locality, and adposition stranding. Ph.D. Dissertation, University of Connecticut, Storrs.

Aissen, Judith. 2001. Markedness and subject choice in Optimality Theory. In *Optimality-theoretic syntax*, eds. Geraldine Legendre, Jame Grimshaw, and Sten Vikner, 61–96. Cambridge: MIT Press.

Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory* 21:435–483.

Aissen, Judith. 2017. Correlates of ergativity in mayan. In *Oxford handbook of ergativity*, eds. Jessica Coon, Diane Massam, and Lisa Travis. New York, NY: Oxford University Press.

Alexiadou, Artemis, Elena Anagnostopoulou and Florian Schäfer. 2018. Passive. In *Syntactic Structures after 60 years: The impact of the Chomskyan revolution in Linguistics*, 403–425. Mouton de Gruyter.

Anagnostopoulou, Elena. 2005. Strong and weak person restrictions: A feature checking analysis. In *Clitic and affix combinations: Theoretical perspectives*, eds. Lorie Heggie and Francisco Francisco Ordóñez, 199–235. Amsterdam: John Benjamins.

Aronoff, Mark. 1994. *Morphology by Itself: Stems and Inflectional Classes*. MIT Press.

Aronson, Howard I. 1990. *Georgian: A reading grammar*. Columbus, Ohio: Slavica.

Arregi, Karlos and Andrew Nevins. 2012. *Morphotactics: Basque auxiliaries and the structure of spellout*. Dordrecht: Springer.

Atlamaz, Ümit. 2013. Cyclic Agreement and empty slots in pazar laz. In *Proceedings of the 37th Annual Meeting of the Berkeley Linguistics Society: Special Session on Languages of the Caucasus*, eds. Shinae Kang Chundra Cathcart and Clare S. Sandy, 18–31.

Austin, Jennifer. 2007. Grammatical interference and the acquisition of ergative case in bilingual children learning Basque and Spanish. *Bilingualism: Language and Cognition* 10:315–331.

Baker, Mark C. 2008. *The Syntax of Agreement and Concord*. Cambridge: Cambridge University Press.

Barr, Dale J., Roger Levy, Christoph Scheepers and Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 63:255–278.

Bates, Douglas, Martin Mächler, Ben Bolker and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67:1–48.

Béjar, Susana. 2003. Phi-syntax: A theory of agreement. Ph.D. Dissertation, University of Toronto.

Béjar, Susana and Milan Rezac. 2003. Person licensing and the derivation of PCC effects. In *Romance linguistics: Theory and acquisition*, eds. Ana Teresa Pérez Leroux and Yves Roberge, 49–62. Amsterdam: John Benjamins.

Béjar, Susana and Milan Rezac. 2009. Cyclic Agree. *Linguistic Inquiry* 40:35–73.

Belletti, Adriana and Luigi Rizzi. 1988. Psych-verbs and θ-Theory. *Natural Language and Linguistic Theory* 6:291–352.

Bhatt, Rajesh. 2003. Locality in correlatives. *Natural Language and Linguistic Theory* 21:485–541.

Bhatt, Rajesh and Léa Nash. 2018. Deconstructing correlatives: Individuals and properties – the case of Georgian *rom*-relatives. Paper presented at The Fifth meeting of The Semantics of African, Asian and Austronesian Languages (Triple A 5). Konstanz, Germany. June 27–29.

Bhatt, Rajesh and Roumyana Pancheva. 2006. Conditionals. In *The blackwell companion to syntax*, eds. Martin Everaert and Henk van Riemsdijk, 638–687. Oxford.

Bittner, Maria. 2001. Topical referents for individuals and possibilities. In *Proceedings of the 11th Semantics and Linguistic Theory Conference (SALT xi)*, eds. Rachel Hastings, Brendan Jackson, and Zsofia Zvolenszky, 36–55.

Bittner, Maria and Ken Hale. 1996. The structural determination of case and agreement. *Linguistic Inquiry* 27:1–68.

Bjorkman, Bronwyn and Hedde Zeijlstra. 2019. Checking up on (φ-)Agree. *Linguistic Inquiry* 5:527–569.

Blix, Hagen. 2020. Spans in South Caucasian agreement. *Natural Language and Linguistic Theory* .

Bobaljik, Jonathan. 2008. Where's phi? Agreement as a post-syntactic operation. In *Phi theory: Phi-features across interfaces and modules*, eds. Daniel Harbour, David Adger,

and Susan Béjar, 294–328. Oxford: Oxford University Press.

Boeder, Winfried. 2005. The South Caucasian languages. *Lingua* 115:5–89.

Bondarenko, Tanya and Stanislao Zompì. 2020. Leftover agreement: Spelling out Kartvelian number. Ms., Paper presented at Generative Linguistics in the Old World (GLOW43), Humboldt-Universität zu Berlin.

Bonet, Eulàlia. 1991. Morphology after syntax: Pronominal clitics in Romance. Ph.D. Dissertation, Massachusetts Institute of Technology.

Borise, Lena and Maria Polinsky. 2018. Focus without movement: Syntax–prosody interface in Georgian. In *Proceedings of the Forty-Eighth Annual Meeting of the North East Linguistic Society (NELS 48)*, eds. Sherry Hucklebridge and Max Nelson.

Bornkessel, Ina and Matthias Schlesewsky. 2006. The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review* 113:787–821.

Bornkessel-Schlesewsky, Ina and Matthias Schlesewsky. 2008a. Minimality as vaccuous distinctness: Evidence from cross-linguistic sentence comprehension. *Lingua* 119:1541–1559.

Bornkessel-Schlesewsky, Ina and Matthias Schlesewsky. 2008b. Unmarked transitivity: A processing constraint on linking. In *Investigations of the syntax–semantics–pragmatics interface*, ed. Robert D. Van Valin, Jr., 413–434. Amsterdam: John Benjamins.

Bornkessel-Schlesewsky, Ina and Matthias Schlesewsky. 2009. The role of prominence information in the real-time comprehension of transitive constructions: A cross-linguistic approach. *Language and Linguistic Compass* 3:19–58.

Butskhrikidze, Marika. 2002. *The Consonant Phonotactics of Georgian*. Utrecht: LOT.

Carreiras, Manuel, Jon Andoni Duñabeitia, Marta Vergara, Irene de la Cruz-Pavía and Itziar Laka. 2010. Subject relative clauses are not universally easier to process: Evidence from basque. *Cognition* 115:70–92.

Chikobava, Arnold. 1936. *Čanuris gramaṭiḳuli analizi ṭeksṭebiturt [A grammatical analysis of Laz with texts]*. Tbilisi: SSRḲ Mecnierebata Aḳademia, Sakartvelos Piliali.

Chomsky, Noam. 1973. Conditions on transformations. In *A festschrift for Morris Halle*, eds. Stephen R. Anderson and Paul Kiparsky, 232–286. New York: Holt, Reinhart, and Winston.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, eds. Roger Martin, David Michaels, and Juan Uriagereka, 89–156. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.

Choudhary, Kamal Kumar, Matthias Schlesewsky, Dieter Roehm and Ina Bornkessel-Schlesewsky. 2007. The role of animacy in the processing of ergative constructions: Evidence from Hindi. Ms., Paper presented at 13th Annual Conference on Architectures and Mechanisms for Language Processing (AMLaP 13), Turku, Finland.

Chung, Sandra. 1994. *Wh*-agreement and "referentiality" in Chamorro. *Linguistic Inquiry* 25:1–44.

Chung, Sandra. 1998. *The design of agreement: Evidence from Chamorro*. Chicago: Uni-

versity of Chicago Press.

Clemens, Lauren Eby, Jessica Coon, Pedro Mateo Pedro, Adam Milton Morgan, Maria Polinsky, Gabrielle Tandet and Matthew Wagers. 2015. Ergativity and the complexity of extraction: a view from Mayan. *Natural Language and Linguistic Theory* 33:417–467.

Comrie, Bernard. 1989. *Language universals and linguistic typology: Syntax and morphology*. University of Chicago press.

Coon, Jessica and Alan Bale. 2013. The interaction of person and number in Mi'gmaq. *NordLyd* 41:85–101.

Coon, Jessica and Stefan Keine. 2019. Feature gluttony. Ms., McGill University and the University of Southern California.

Crain, Stephen and Janet D. Fodor. 1985. How can grammars help parsers? In *Natural language parsing: Psycholinguistic, computational, and theoretical perspectives*, eds. David Dowty, Lauri Kartunnen, and Arnold M. Zwicky, 94–128. Cambridge University Press.

Dayal, Veneeta. 1996. *Locality in Wh-Quantification: Questions and Relative Clauses in Hindi*. Kluwer: Dordrecht.

Deal, Amy Rose. 2015. Interaction and satisfaction in φ-agreement. *North East Linguistic Society (NELS)* 45:1–14.

Despić, Miloje, Michael David Hamilton and Sarah Murray. 2019. A Cyclic and Multiple Agree account: Person/number marking in Cheyenne. *Natural Language and Linguistic Theory* 37:51–89.

Dillon, Brian, Andrew Nevins, Alison C. Austin and Collin Phillips. 2012. Syntactic and

semantic predictors of tense in Hindi: An ERP investigation. *Language and Cognitive Processes* 27:313–344.

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

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67:547–619.

Drummond, Alex. 2016. Ibex Farm. http://spellout.net/ibexfarm.

Ershova, Ksenia. 2016. Dative blocking in Georgian. Poster presented at the South Caucasian Chalk Circle, UChicago Center in Paris, Sept. 22–24.

Fauconnier, Stephanie. 2012. Constructional effects of involuntary and inanimate agents: A cross-linguistic study. Ph.D. Dissertation, Katholieke Universiteit Leuven.

Foley, Steven. 2013. The syntax of Georgian relative clauses. Undergraduate honors thesis, New York University.

Foley, Steven. 2017. Morphological conspiracies in Georgian and Optimal Vocabulary Insertion. *Proceedings of the fifty-second annual meeting of the Chicago Linguistic Society (CLS 52)* 217–232.

Foley, Steven. to appear. Agreement in the languages of the Caucasus. In *The Oxford Handbook of the Languages of the Caucasus*, ed. Maria Polinsky. Oxford University Press.

Foley, Steven, Jed Pizarro-Guevarra, Kelsey Sasaki, Maziar Toosarvandani and Matt Wagers. 2019. Pronouns over gaps in parsing? The Subject Relative Clause Advantage in Santiago Laxopa Zapotec. Paper presented at the 93rd Annual Meeting of the Linguistic Society of America (LSA 93), New York City, NY. January 3–6.

Foley, Steven and Maziar Toosarvandani. 2019. Agreement, locality, and the syntax of pronouns: The Person–Case Constraint and beyond. Ms., University of California Santa

Cruz.

Foley, Steven and Matt Wagers. 2017a. The subject gap preference in a split-ergative language: Reading time evidence from georgian. Paper presented at the 48th Annual Meeting of the North East Linguistics Society (NELS 48). University of Iceland, Reykjavík, Iceland. October 27–29.

Foley, Steven and Matt Wagers. 2017b. Subject gaps are still easiest: Relative clause processing and georgian split ergativity. Poster presented at the 30th Annual CUNY Conference on Human Sentence Processing (CUNY 30). Massachusetts Institute of Technology, Cambridge, MA. March 30 – April 1.

Frazier, Lyn. 1978. On comprehending sentences: Syntactic parsing strategies. Ph.D. Dissertation, University of Connecticut.

Frazier, Lynn. 1987. Syntactic processing: Evidence from Dutch. *Natural Language and Linguistic Theory* 5:519–559.

Gennari, Silvia P. and Maryellen C. MacDonald. 2008. Semantic indeterminacy in object relative clauses. *Journal of Memory and Language* 58:161–187.

Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68:1–76.

Gibson, Edward and James Thomas. 1986. Memory limitations and structural forgetting: The perception of complex ungrammatical sentences as grammatical. *Language and Cognitive Processes* 14:225–248.

Gibson, Edward and Iris Wu. 2013. Processing Chinese relative clauses in context. *Language and Cognitive Processes* 28:125–155.

Greenberg, Joseph. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of Language*, ed. Joseph Greenberg, 73–113. Cambridge: MIT Press.

Hale, John. 2003. The information conveyed by words in sentences. *Journal of Psycholinguistc Research* 32:101–123.

Hale, John. 2006. Uncertainty about the rest of the sentence. *Cognitive Science* 30:643–672.

Halle, Morris. 1997. Distributed Morphology: Impoverishment and fission. In *MIT working papers in linguistics*, eds. Benjamin Bruening, Yoonjung Kang, and Martha McGinnis, 425–449. Cambridge, MA: MIT.

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*, eds. Kenneth Hale and Samuel Jay Keyser, 111–176. Cambridge, MA: MIT Press.

Harbour, Daniel. 2016. *Impossible persons*. Cambridge, MA: MIT Press.

Harley, Heidi and Elizabeth Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78:482–526.

Harris, Alice C. 1981. *Georgian Syntax: A study in Relational Grammar*. Cambridge: Cambridge University Press.

Harris, Alice C. 1982. Georgian and the Unaccusative Hypothesis. *Language* 58:290–306.

Harris, Alice C. 1985. *Diachronic Syntax: the Kartvelian Case*. New York: Academic Press.

Harris, Alice C. 1990. Georgian: A language with active case marking: A reply to B.G. Hewitt. *Lingua* 80:35–53.

Harris, Alice C., ed. 1991. *The Indigenous Languages of the Caucasus, Vol. 1: The Kartvelian Languages*. Delmar, N.Y.: Caravan.

Hermon, Gabriella. 1981. Non-nominative subject constructions in the Government and Binding framework. Ph.D. Dissertation, University of Illinois.

Hiraiwa, Ken. 2001. Multiple Agree and the defective intervention constraint in Japanese. *MIT Working Papers in Linguistics* 40:67–80.

Holisky, Dee Ann. 1981. *Aspect and Georgian medial verbs*. Caravan.

Hopper, Paul J. and Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 58:251–299.

Horn, Laurence. 1984. Toward a new taxonomy for pragmatic inference: Q-based and r-based implicature. *Meaning, form, and use in context: Linguistic applications* 11:42.

Hsiao, Franny and Edward Gibson. 2003. Processing relative clauses in chinese. *Cognition* 90:3–27.

Hsu, Chun-Chieh and Jenn-Yeu Chen. 2007. A new look at the subject–object asymmetry: The effects of linear distance and structural distance on the processing of head-final relative clauses in chinese. In *Interdisciplinary approaches to to relative clauses (rel07)*, eds. Rachel Hastings, Brendan Jackson, and Zsofia Zvolenszky, 36–55.

Ishizuka, Tomoko, Kentaro Nakatani and Edward Gibson. 2006. Processing Japanese relative clauses in context. Paper presented at the 19th annual CUNY Conference on Human Sentence Processing (CUNY 19). New York, NY.

Izvorski, Roumyana. 1996. The syntax and semantics of correlative proforms. In *NELS 26: proceedings of the Twenty-Sixth Meeting of the North East Linguistics Society*, ed.

Kiyomi Kusumoto, 133–147.

Jelinek, Eloise and Richard A. Demers. 1983. The agent hierachy and voice in some Coast Salish languages. *International Journal of American Linguistics* 49:167–185.

Kahraman, Barış, Atsushi Sato, Hajime Ono and Hiromu Sakai. 2010. Relative clauses processing before the head noun: Evidence for strong forward prediction in Turkish. In *Proceedings of the 6th Workshop on Altaic Formal Linguistics (WAFL6)*, eds. Hiroki Maezawa and Azusa Yokogoshi, 155–170.

Keenan, Edward L. and Bernard Comrie. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8:63–99.

Kibrik, Alexander E. 1985. Toward a typology of ergativity. *Grammar inside and outside the clause* 268–324.

Kiparsky, Paul. 2005. Blocking and periphrasis in inflectional paradigms. In *Yearbook of morphology 2004*, 113–135. Springer.

Kipshidze, Iosif A. 1914. *Grammatika mingrel'skago (iverskago) jazyka s" xrestomatīeju i slovarem" [Grammar of Mingrelian (Iverian) with Chrestomathy and Dictionary]*. St. Petersburg: Tipografīja Imperatorskoj Akademīj Nauk″.

Kuznetsova, Alexandra, Per B. Brockhoff and Rune H. B. Christensen. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 32:1–26.

Kwon, Nayoung. 2008. Processing of syntactic and anaphoric gap–filler dependencies in korean: Evidence form self-paced reading, ERP, and eye-tracking experiments. Ph.D. Dissertation, University of California San Diego.

Kwon, Nayoung, Robert Kluender, Marta Kutas and Maria Polinsky. 2013. Subject–object

processing asymmetries in Korean relative clauses: Evidence from ERP data. *Language* 89:537–585.

Kwon, Nayoung, Yoonhyoung Lee, Peter C. Gordon, Robert Kluender and Maria Polinsky. 2010. Cognitive and linguistic factors affecting subject–object asymmetry: An eye-tracking study of pre-nominal relative clauses in korean. *Language* 86:546–582.

Landau, Idan. 2009. *The Locative Syntax of Eperiencers*. Cambridge, MA: MIT Press.

Lau, Ellen, Nancy Clarke, Michaela Socolof, Rusudan Asatiani and Maria Polinsky. 2018. Early (predictive) detection of conflicting constraints? ERP evidence from georgian. Poster presented at the 31st annual CUNY Conference on Human Sentence Processing (CUNY 31). University of California Davis, Davis, CA.

Lau, Ellen, Nancy Clarke, Michaela Socolof, Rusudan Asatiani and Maria Polinsky. Submitted. A subject relative clause preference in a split-ergative langauge: ERP evidence from georgian. Ms., Manuscript, University of Maryland.

Lewis, Richard L. and Shravan Vasishth. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science* 29:375–419.

Lin, Chien-Jer Charles. 2006. Grammar and parsing: A typological investigation of relative clause processing. Ph.D. Dissertation, University of Arizona.

Lin, Chien-Jer Charles. 2008. The processing foundation of head-final relative clauses. *Language and Linguistics* 9:813–838.

Lin, Chien-Jer Charles and Thomas G. Bever. 2010. Subject preference in the processing of relative clauses in chinese. In *Proceedings of the 25th West Coast Conference on Formal Linguistics (WCCFL 25)*, eds. Donald Baumer, David Montero, and Michael Scanlon,

254–260.

Lin, Chien-Jer Charles and Thomas G. Bever. 2011. Garden path and the comprehension of head-final relative clauses. In *Processing and producing head-final structures*, eds. Hiroko Yamashita, Yuki Hirose, and Jerome L. Packard, 277–297.

Lin, Yowyu and Susan Garnsey. 2010. Animacy and the resolution of temporary ambiguity in relative clause comprehension in Mandarin. In *Processing and producing head-final structures*, eds. Hiroko Yamashita, Yuki Hirose, and Jerome L. Packard, 241–275.

Lipták, Anikó, ed. 1996. *Correlatives Cross-Linguistically*. Amsterdam: Benjamins.

Lomashvili, Leila and Heidi Harley. 2011. Phases and templates in Georgian agreement. *Studia Linguistica* 65:233–267.

Longenbaugh, Nick and Maria Polinsky. 2016. The processing of long-distance dependencies in niuean. In *The Proceedings of the 22nd Meeting of the Austronesian Formal Linguistics Association (AFLA 22)*, ed. Henrison Hsieh, 106–120.

Mak, Willem M., Wietske Vonk and Herbert Schriefers. 2002. The influence of animacy on relative clause processing. *Journal of Memory and Language* 47:50–68.

Marantz, Alec. 1981. On the nature of grammatical relations. Ph.D. Dissertation, Massachusetts Institute of Technology.

Marantz, Alec. 1989. Relations and configurations in georgian. *Ms., University of North Carolina, Chapel Hill* .

Miyamoto, Edson T. and Michiko Nakamura. 2003. Subject–object asymmetries in the processing of relative clauses in japanese. In *Proceedings of the 22nd West Coast Conference on Formal Linguistics (WCCFL 22)*, eds. Gina Garding and Mimu Tsujimura, 342–355.

Moravcsik, Edith A. 1974. Object–verb agreement. In *Working Papers on Language Universals 15*, eds. Gina Garding and Mimu Tsujimura, 25–140. Stanford, CA: Stanford University Press.

Nash, Léa. 2017a. Simple correlatives in georgian. Ms., Paper presented at Queen Mary Correlatives Workshop, London. March 13.

Nash, Léa. 2017b. The structural source of split ergativity and ergative case in Georgian. In *The oxford handbook of ergativity*, eds. Jessica Coon, Diane Massam, and Lisa deMena Travis, 175–204. Oxford University Press.

Nevins, Andrew. 2007. The representation of third person and its consequences for person-case effects. *Natural Language and Linguistic Theory* 25:273–313.

Nevins, Andrew. 2011. Multiple agree with clitics: Person complementarity vs. omnivorous number. *Natural Language and Linguistic Theory* 29:939–971.

Newmeyer, Frederick J. 2010. On comparative concepts and descriptive categories: A reply to Haspelmath. *Language* 86:688–695.

Philipp, Markus, Ina Bornkessel-Schlesewsky, Walter Bisang and Matthias Schlesewsky. 2008. The role of animacy in the real time comprehension of mandarin chinese: Evidence from auditory event-related brain potentials. *Brain and Language 105* 105:112–133.

Polinsky, Maria. 2011. Reanalysis in adult heritage language: New evidence in support of attrition. *Studies in Second Language Acquisition* 33:305–328.

Polinsky, Maria. 2016. *Deconstructing Ergativity: Two types of ergative languages and their features*. Oxford University Press.

Polinsky, Maria, Carlos Gómez Gallo, Peter Graff and Ekaterina Kravtchenko. 2012. Sub-

ject preference and ergativity. *Lingua* 122:267–277.

Preminger, Omer. 2011. Agreement as a fallible operation. Ph.D. Dissertation, Massachusetts Institute of Technology.

Preminger, Omer. 2013. That's now how you agree: A reply to Zeijlstra. *The Linguistic Review* 300:491–500.

Preminger, Omer. 2014. *Agreement and its failures*. Cambridge, MA: MIT Press.

Preminger, Omer. 2019. What the PCC tells us about "abstract" agreement, head movement, and locality. *Glossa* 4.

Primus, Beatrice. 1999. *Cases and thematic roles*. Tübingen, Germany: Niemeyer.

Pylkkänen, Liina. 2008. *Introducing arguments*. Cambridge, MA: MIT Press.

Rayfield, Donald, ed. 2006. *A Comprehensive Georgian–English Dictionary*. London: Garnett Press.

Rezac, Milan. 2003. The fine structure of Cyclic Agree. *Syntax* 6:158–181.

Rezac, Milan. 2008. Phi-agree and theta-related case. In *Phi theory: Phi-features across interfaces and modules*, eds. Daniel Harbour, David Adger, and Susan Béjar, 83–129. Oxford: Oxford University Press.

Richards, Norvin. 1997. What moves where when in which language? Ph.D. Dissertation, Massachusetts Institute of Technology.

Richards, Norvin. 1998. The principle of minimal compliance. *Linguistic Inquiry* 29:599–629.

Riedel, Kristina. 2009. The syntax of object marking in Sambaa: A comparative Bantu perspective. Ph.D. Dissertation, Leiden University.

Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, MA: MIT Press.

Roehm, Dietmar, Martin Schlesewsky, Ina Bornkessel-Schlesewsky, Stefan Frisch and Hubert Haider. 2004. Fractionating language comprehension via frequency characteristics of the human eeg. *Neuroreport* 15:409–12.

Romero-Méndez, Rodrigo. 2008. A reference grammar of Ayutla Mixe (Tukyo'm Ayuujk). Ph.D. Dissertation, University of Buffalo, the State Univeristy of New York.

Sapir, Edward. 1917. Het Passieve Karakter van het Verbum Transitivum of van het Verbum Actionis in Talen van Noord-Amerika (The Passive Character of the Transitive Verb or of the Active Verb in Languages of North America), C. C. Uhlenbeck [Review]. *International Journal of American Linguistics* 1:82–86.

Sigurðsson, Halldór. 2002. How to be an oblique subject: Russian vs. Icelandic. *Natural Language and Linguistic Theory* 20:691–724.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In *Grammatical categories in australian languages*, ed. Robert M. W. Dixon, 112–71. New Jersey: Humanities Press.

Skopeteas, Stavros and Gisbert Fanselow. 2010. Focus in Georgian and the expression of contrast. *Lingua* 120:1370–1391.

Skopeteas, Stavros, Gisbert Fanselow and Rusudan Asatiani. 2012. Case inversion in georgian: Syntactic properties and sentence processing. In *Case, Word Order, and Prominence: Interacting cues in language production and comprehension*, eds. Monique Lamers and Peter de Swart, 145–171. Springer.

Srivastav, Veneeta. 1991. The syntax and semantics of correlatives. *Natural Language and Linguistic Theory* 9:637–686.

Staub, Adrian, Francesca Foppolo, Caterina Donati and Carlo Cecchetto. 2018. Relative clause avoidance: Evidence for a structural parsing principle. *Journal of Memory and Language* 98:26–44.

Stowe, Laurie A. 1986. Evidence for on-line gap-location. *Language and Cognitive Processes* 1:227–245.

Thivierge, Sigwan. 2019. Expanding the agreement domain in Georgian. Paper presented at the 43rd Penn Linguistics Conference, Mar. 22–24.

Tollan, Rebecca, Diane Massam and Daphna Heller. 2019. Effects of case and transitivity on processing dependencies: Evidence from niuan. *Cognitive Science* 43:1–14.

Topuria, Varlam. 1967. *Svanuri ena, I.: Zmna [Svan, Vol. I: The Verb]*. Tbilisi: Mecniereba.

Traxler, Matthew, Rihana Williams, Shelley Blozis and Robin Moris. 2005. Working memory, animacy, and verb class in the processing of relative clauses. *Journal of Memory and Language* 53:204–224.

Tuite, Kevin. 1998. Kartvelian morphosyntax: Number agreement and morphosyntactic orientation in the south caucasian languages. Ph.D. Dissertation, University of Montreal.

Ueno, Mieko and Susan Garnsey. 2008. An ERP study of the processing of subject and object relative clauses in Japanese. *Language and Cognitive Processes* 23:648–688.

Van Valin, Robert D., Jr. 1990. Semantic parameters of split intransitivity. *Language* 66:221–260.

Vasishth, Shravan, Zhong Chen, Qiang Li and Gueilan Guo. 2013. Processing Chinese relative clauses: Evidence for the subject-relative advantage. *PLoS One* 8:1–15.

Wade, Terrence. 2011. *A Comprehensive Russian Grammar*. Wiley-Blackwell, 3rd edition.

Wagers, Matthew, Sandra Chung and Manuel F. Borja. 2018. Grammatical licensing and relative clause parsing in a flexible-word order langauge. *Cognition* 178:207–221.

Wagers, Matthew and Emily Pendleton. 2016. Structuring expectation: Licensing animacy in relative clause comprehension. In *Proceedings of the 33rd West Coast Conference on Formal Linguistics (WCCFL 33)*, eds. Kyeong min Kim, Pocholo Umbal, Trevor Block, Queenie Chan, Tanie Cheng, Kelli Finney, Mara Katz, Sophie Nickel-Thompson, and Lisa Shorten, 29–46. Springer.

Weckerly, Jill and Marta Kutas. 1999. An electrophysiological analysis of animacy effects in the processing of object relative sentences. *Psychophysiology* 36:559–57.

Wier, Thomas. 2011. Georgian morphosyntax and feature hierarchies in natural langauge. Ph.D. Dissertation, University of Chicago.

Yun, Jiwon, Zhong Chen, Tim Hunter, John Whitman and John Hale. 2015. Uncertainty in processing relative clauses across east asian languages. *Journal of East Asian Linguistics* 24:113–148.

Zagona, Karen. 2002. *The Syntax of Spanish*. Cambridge University Press.

Zeijlstra, Hedde. 2012. There is only one way to agree. *The Linguistic Review* 29:491–539.

Öztürk, Balkız. 2016. Applicatives in Pazar Laz. Paper presented at the South Caucasian Chalk Circle, UChicago Center in Paris, Sept. 22–24.