Transitivity, prominence, and verb finality in sentence processing: Georgian Maze evidence

Steven Foley (University of Southern California)
foleys@usc.edu
Draft: October 18, 2023

Abstract: Bottom-up evidence is not necessary to form expectations about the thematic relations between core grammatical arguments of a sentence. Word order, case, and animacy can influence how a noun is ultimately integrated with the verb, as a subject or object. Three behavioral experiments investigate how verb-final sentences are processed in Georgian, a verb-final split-ergative language where many case–role mappings can only be disambiguated by verbal morphology. Manifesting as reaction times at verbs disambiguating a range of word orders, mappings, and argument structures, a few notable effects emerge: applied indirect objects clearly increase processing difficulty; object-initial dative-subject sentences are particularly difficult; nonactive/anticausative intransitives are of comparable difficulty to active monotransitives; there may be an meaningful cost to agreement morphology registering null first/second-person subjects. Results shed light on the roles of prominence, prototypicality, frequency, and abstract transitivity in theories of grammatical relations and sentence processing.

1. Introduction

Verb final word order is widespread in natural language: a remarkable fact, in light of the outsized contribution of a typical verb during real-time sentence comprehension. The verb root supplies a clause with its core lexical semantic relations, and just how those relations are grammatically mapped to nominal arguments can depend on verbal inflection. The verb being the locus of so many thematic and morphosyntactic dependencies, withholding that word until the end of the clause may seem maladaptive for efficient communication of who did what to whom in what kind of event (Inoue & Fodor 1995).

But natural languages tend to throw the comprehender a bone. There has been a long standing typological observation that verb finality correlates with rich case inflection (Greenberg 1963's Universal 41) and a long standing intuition that case morphology has an adaptive function for comprehension (Sapir 1917), signaling as it does grammatical role more or less obliquely. For example, take classical ergative case (Dixon 1979). Biuniquely associated with transitive subjecthood, ergative is an excellent cue to that syntactic role. In turn, transitive subjecthood is correlated with a cluster of lexical-semantic properties. After all, natural language lexicons are overwhelmingly organized so that grammatical relations are rather semantically coherent (Dowty 1991). Thus, ergative is a reliable cue not just to a narrow syntactic relation (transitive subjecthood, external argumenthood, Initial-1-hood, etc.) — it also reliably cues a cluster of entailments about the event participant's relative agency, causal relation, animacy, and so on. Encountering ergative morphology on a preverbal noun, then, allows the comprehender to reliably predict quite a bit about the morphosyntactic and thematic dependencies of the clause, well before the verb can be processed.

This paper focuses on a verb-final language where case cues are often much less reliable than classical ergative. In Georgian, the so-called nominative and dative cases have wide distributions, marking different grammatical roles in different tenses and argument structures. Consequently, sequences of preverbal nominative and dative arguments are highly ambiguous. The nominative–dative sequence in 1, for example, is temporarily compatible with a range of parses: subject- or object-initial, active or nonactive, with or without an indirect object. Generally, tense and argument-structure morphology on the verb is what disambiguates the correct parse of the clause (2).

(1) *Ambiguous preamble*[1]

ekʰim-i	msaχiob-s…
doctor-NOM	actor-DAT

(2) *Possible continuations*

a. …gaaʧʰerebs — *Monotransitive SOV parse*
stop:ACT:FUT
"The doctor will stop the actor"

b. …gauʧʰerebia — *Monotransitive OSV parse*
stop:ACT:PERF
"The doctor has (apparently) stopped the actor"

c. …gauʧʰerebs… — *Ditransitive SOV parse, anticipated O*
stop:ACT:APPL:FUT
"The doctor will stop the actor […for 3RD]"
or "The doctor will stop […3RD] for the actor"

d. …gavuʧʰere — *Ditransitive DO-IO-V parse, S-Agr*
stop:ACT:APPL:AOR:1SGS
"I stopped doctor for the actor"

e. …gauʧʰerda — *Applied nonactive S-IO-V parse*
stop:NACT:APPL:AOR
"The doctor stopped for the actor"

---

[1] Georgian illustrations are presented in broad phonological transcription using the IPA. Verbs are not morphologically decomposed, but they are glossed for a few paradigmatic dimensions: argument structure (active/nonactive; with/without applicative), tense, and any agreement with first- or second-person arguments.

Abbreviations: A 'active subject'; ACT 'active'; AOR 'aorist'; APPL 'applicative'; AV 'agent voice'; DAT 'dative'; DITR 'ditransitive'; ERG 'ergative'; Ex 'experiencer subject of psych verb'; FUT 'future'; G 'indirect object (lexical goal or applied affectee)'; GEN 'genitive'; INTR 'intransitive'; NACT 'nonactive'; NMLZ 'deverbal nominalization'; NOM 'nominative'; O '[any] object'; OBL 'oblique'; P 'direct object'; PERF 'perfect'; PLU 'pluperfect'; PV 'patient voice'; RT 'reaction time (e.g., reading time or lexical-recognition time)'; S '[any] subject'; St 'stimulus object of psych verb'; U 'nonactive subject'; XYZ *or* X–Y–Z 'linear sequence of elements'; X/Y/Z 'morphosyntactic/clausal elements, order underspecified'; X:Y:Z 'word X inflected for features Y and Z'

How do comprehenders navigate such ambiguities? To what degree do they entertain each grammatically available parse, and why? Novel behavioral data from three sentence-processing experiments address these questions, tracking reading times across out-of-the-blue sentences using an incremental lexicality-decision task (the Lexicality Maze: Freedman & Forster 1985, Boyce et al. 2020). The present study, building on Skopeteas et al. (2011), leverages Georgian's grammatical idiosyncrasies to isolate interactions between prominence scales argued to be useful in processing similar case–role ambiguities crosslinguistically (cf. the eADM theory; Bornkessel-Schlesewsky & Schlesewsky 2006, et seq). Effects emerge at disambiguating verb regions; nuanced generalizations emerge about word order, agreement, tense, and particular case–role mappings — but in sum, we find that prominence scales are not implemented symmetrically during real-time sentence processing.

Consider the scales for animacy (human > inanimate) and grammatical-role (transitive subject > indirect object > direct object). Suppose that parsers are biased in general to assign an ambiguous high-animacy argument the most prominent role allowed by the grammar. After all, they do seem inclined to give low-animacy arguments the lowest possible role, in Georgian (Foley 2020, ch. 2) and beyond (Bornkessel-Schlesewsky & Schlesewsky 2014). Results do not support a general advantage to animacy–role harmony (cf. Foley 2020). In this study, almost all preverbal nouns referred to humans. Given Georgian's tense-based split ergativity and a productive applicative alternation, any regular verb can take a dative-case subject or indirect object given the right inflection. Yet comprehenders are generally most eager to interpret high-animacy datives as direct objects: a disharmonic combination of animacy and role. In other words, harmony between animacy and role per se is not a heuristic for parsing, though animacy clearly helps identify transitive subjects (Bornkessel-Schlesewsky & Schlesewsky 2009b).

In morphosyntactic typology, differential marking is often associated with high-animacy objects (Aissen 1991). From various theoretical perspectives this is a consequence of their unexpectedness (Haspelmath 2021), their featural markedness (Kalin 2018), their non-prototypicality (Dowty 1991). Insofar as any of these concepts translate to processing difficulty, the comprehender might seek to avoid positing them at all absent top-down evidence. With its great wealth of indirect objects and dative subjects, Georgian grammar makes it very easy to avoid parses that involve a human patient. Indeed, in the Georgian Reference Corpus (Gippert & Tandashvili 2015), verbs which morphologically license a dative indirect object outnumber verbs licensing dative direct objects (Foley 2022). Yet, overall, a dative noun is easiest to process when parsed as a direct object, in line with Skopeteas et al.'s (2011) generalizations.

Another perspective on high-animacy objects is that they increase a clause's prototypical Transitivity (Hopper & Thompson 1980). In the context of real-time sentence processing, a highly Transitive predicate can be conceived as one whose arguments can be held in sharper focus in the mind's eye: an event with highly distinct participants, which are easier to store together in memory (Bornkessel-Schlesewsky & Schlesewsky 2006's Distinctness). Even holding constant lexeme and valence (cf. 2a,e), Georgian shows that dative indirect objects are harder to process than dative direct objects. This follows naturally if the comprehenders are biased towards parses that are more Transitive — rather than more optimal/harmonic (Hoeks & Hendriks 2011, Skopeteas et al. 2011, Foley 2020), or less surprising given linguistic experience (Hale 2001, Levy 2008).

In what sense is Georgian a maladaptive verb-final language? Why does it through the comprehender a few more bones, by making word order stricter, or pronouns obligatorily overt? How has such an entropic morphosyntax evaded the forces of grammaticization so long? In fact, a case–role mapping system nearly identical to Mondern Georgian's is reconstructable to Proto South Caucasian (Harris 1985). And, across the 17 centuries of Georgian's written attestation, dative arguments have if anything gained subjecthood properties (like the ability to control plural agreement; Tuite 1998). Clearly, the comprehender can tolerate a certain degree of morphosyntactic ambiguity before encountering the verb.

The rest of this paper explores this observation and weighs its theoretical consequences. Section 2 gives Key background about Georgian and about crosslinguistic sentence processing. Section 3 synthesizes the background into a theoretical framework couched in Harmonic Grammar, articulating predictions for real-time sentence processing. Section 4 reports novel behavioral data. Section 5 summarizes and discusses findings from the three experiments. Section 6 concludes.

2. Background

This section digests key facts about Georgian morphosyntax (Subsection 2.1), previous sentence-processing studies (Subsection 2.2) and corpus research (Subsection 2.3) on Georgian, and the psycholinguistic/neurolinguistic literature on case–role processing crosslinguistically (Subsection 2.4).

2.1 Key grammatical facts about Georgian

Georgian (Shanidze 1980; Harris 1981; Aronson 1990; Hewitt 1995; Tuite 1998) is a member of the South Caucasian language family (Harris 1991, Boeder 2005, Testelets 2021). It has weakly head final syntax, null pronouns, flexible word order, and verbs with rich person–number agreement. Argument structure alternations productively increase or decrease the valence of a verb and affect its inflectional morphology. Case assignment of core nominal arguments depends on grammatical role and tense (Harris 1985, Nash 2017). Together, this makes the grammatical role of a large proportion of nominal arguments temporarily ambiguous; cues from finite verb's inflection are usually the key disambiguator.

Take a regular verb like /gaʧʰereba/ "stop:NMLZ [citation form; root = ʧʰer]". It can be inflected as a bivalent active verb, taking a transitive subject ('A') and a direct object ('P'). It also has a monovalent nonactive form; the intransitive subject ('U')[2] of this verb corresponds thematically to the direct object of the transitive version (see Cherchi 1997 or Gérardin 2016 on nonactive verbs in Georgian).

(3) ekʰim-ma msaχiob-i gaaʧʰera
doctor-**ERG** actor-**NOM** stop:**ACT**:**AOR**
"The doctor$_{[A]}$ stopped the actor$_{[P]}$."

[2] Note the divergence from the typical symbol 'S' representing intransitive subjects (Dixon 1979, et seq). We reserve 'S' to symbolize any subject, whatever its morphosyntax. Likewise, 'O' here symbolizes any object, direct or indirect.

(4) msaχiob-i gaʧʰerda
actor-**NOM** stop:**NACT:AOR**
"The actor$_{[U]}$ stopped."

Both active and nonactive verbs have applicative forms. Applicativization here refers to an alternation in verbal morphology that signals the addition of an indirect object ('G') into the argument structure. The applied indirect object can have multiple interpretations in relation to the event described, depending on context and semantics of all clausal elements. Throughout this paper applied objects will be translated as benefactees (i.e., "for NP"), though readings involving a malefactee, accidental causer, or external possessor of the P/U argument may be possible and even preferable (cf. Bosse et al. 2012).

(5) ekʰim-ma msaχiob-i gauʧʰera mʦ'eral-s
doctor-ERG actor-NOM stop:**ACT:APPL:AOR** writer-**DAT**
"The doctor$_{[A]}$ stopped the actor$_{[P]}$ for the writer$_{[G]}$."

(6) msaχiob-i gauʧʰerda mʦ'eral-s
actor-NOM stop:**NACT:APPL:AOR** writer-**DAT**
"The actor$_{[U]}$ stopped for the writer$_{[G]}$."

In the examples so far, case marking corresponds straightforwardly to arguments' thematic interpretation. The three distinct case categories ("ergative", "nominative", and "dative") each correspond to a coherent cluster of grammatical roles. Ergative marks the most prototypically agent-like arguments (transitive subjects); nominative marks the most prototypically patient-like arguments (direct objects and nonactive subjects); dative marks other arguments (applied and other indirect objects). However, this case alignment pattern is only observed in "Series II" tenses, which include the aorist (perfective past; also imperative) and optative (nonpast irrealis, especially in embedded clauses). In other tenses, the case–role mapping operates differently.

In "Series I" tenses (present, imperfective past, future, conditional, and two subjunctives), case tracks syntactic function more closely. Nominative marks subjects, both nonactive and active; dative marks objects, both direct and indirect. This means that ditransitive clauses can be globally ambiguous in Series I tenses, since both objects will be in the same case.

(7) ekʰim-i msaχiob-s gaaʧʰerebs/
doctor-**NOM** actor-**DAT** stop:**ACT:FUT**
"The doctor$_{[A]}$ will stop the actor$_{[P]}$."

(8) ekʰim-i msaχiob-s gauʧʰerebs mʦ'eral-s
doctor-**NOM** actor-**DAT** stop:**ACT:APPL:FUT** writer-**DAT**
"The doctor$_{[A]}$ will stop the writer$_{[P]}$ for the actor$_{[G]}$." [AGVP reading]
or "The doctor$_{[A]}$ will stop the actor$_{[P]}$ for the writer$_{[G]}$." [APVG reading]

(9)	msaχiob-i	gaʧʰerdeba		
	actor-**NOM**	stop:**NACT:FUT**		
	“The actor$_{[U]}$ will stop.”			

(10)	msaχiob-i	gauʧʰerdeba	mʦ’eral-s	
	actor-**NOM**	stop:**NACT:APPL:FUT**	writer-**DAT**	
	“The actor$_{[U]}$ will stop for the writer$_{[G]}$.”			

As for “Series III” tenses (the perfect, with past evidential readings; and the pluperfect, with past counterfactual readings), they are traditionally described as having an “inverse” morphosyntax (Harris 1981, Cherchi 1997). Active subjects are dative, and agree with the verb like erstwhile indirect objects. Direct objects and nonactive subjects are nominative, as in Series II. Only nonactive verbs have applicative forms in Series III. Indirect objects of nonactive verbs are dative; for active verbs, would-be indirect objects must be expressed as postpositional phrases.

(11)	ekʰim-s	msaχiob-i	gauʧʰerebia	
	doctor-**DAT**	actor-**NOM**	stop:**ACT:PERF**	
	“The doctor$_{[A]}$ must have stopped the actor$_{[P]}$.”			

(12)	ekʰim-s	msaχiob-i	gauʧʰerebia	mʦ’erl-is=tʰvis
	doctor-**DAT**	actor-**NOM**	stop:**ACT:PERF**	writer-**GEN=for**
	“The doctor$_{[A]}$ must have stopped the actor$_{[P]}$ for the actor$_{[G]}$.” (No applicative available)			

(13)	msaχiob-i	gaʧʰerebula		
	actor-**NOM**	stop:**NACT:PERF**		
	“The actor$_{[U]}$ must have stopped.”			

(14)	msaχiob-i	gasʧʰerebia	mʦ’eral-s	
	actor-**NOM**	stop:**NACT:APPL:PERF**	writer-**DAT**	
	“The actors$_{[U]}$ must have stopped for the actor$_{[G]}$.” (Applicative form distinct)			

The following table summarizes Georgian’s case alignment system. This description omits unergative subjects (i.e., subjects of atelic intransitives; Holisky 1981, Nash 2022), but they behave the same as active transitive subjects. It also sets aside a class of psych verbs: in all tenses, their experiencer subjects are dative and their stimuli objects nominative. A few irregular verbs also have idiosyncratic case marking patterns.

	Active Subject (A)	**Nonactive Subject (U)**	**Direct Object (P)**	**Indirect Object (G)**
Series I (FUT…)	NOM		DAT	
Series II (AOR…)	ERG	NOM		DAT
Series III (PERF…)	DAT	NOM		DAT / PP_{for}

Table 1: Summary of case marking of core arguments across three "Series" of morphological tenses categories. In Series III, indirect objects are dative in nonactive clauses, but PPs in active clauses (i.e., when there is a dative subject).

2.2 Previous sentence-processing work on Georgian

The first experiment on Georgian sentence processing was Skopeteas et al. (2011), in an acceptability judgement study. Their stimuli were all simple bivalent verb-final sentences, varying in case and thematic mappings. The two nominal arguments, one nominative and one dative, were presented together on a screen for 5000 ms. Then the verb appeared, along with a binary acceptability prompt. Critical stimuli belonged to two experiments, each with a 2×2 factorial design. (These stimuli were all grammatical; fillers included various morphosyntactic errors.) The first experiment focused on active clauses, manipulating word order (SOV vs. OSV) and tense/case alignment ($A_{NOM}/P_{DAT}/V_{FUT}$ vs. $A_{DAT}/P_{NOM}/V_{PERF}$) while holding lexical items constant across itemsets. Analyzing response latency, the authors found a significant effect of case alignment: future tense sentences were endorsed as grammatical significantly faster than perfect ones, indicating a processing penalty for the A_{DAT}/P_{NOM} mapping. Their second experiment focused on nonactive clauses, with itemsets comparing pairs of applied nonactives and psych verbs ($U_{NOM}/G_{DAT}/V_{NACT.APPL}$ vs. $Ex_{DAT}/St_{NOM}/V_{PSYCH}$) in either word order (SOV vs. OSV). The authors found main effects of verb type, word order, and a significant interaction of those factors: OSV order was endorsed more slowly than SOV order, but only in clauses with psych verbs.

In light of these results, Skopeteas et al. theorize the following: the Georgian comprehender's preference to parse nominative as S and dative as O is stronger than their preference for SO order. This predicts that future-tense active clauses and applied nonactives are relatively easy to process, whatever their word order. Dative-subject clauses ($A_{DAT}/P_{NOM}/V_{PERF}$ or $Ex_{DAT}/St_{NOM}/V_{PSYCH}$) all violate the hypothesized case–role preference. The St_{NOM}–Ex_{DAT}–V_{PSYCH} condition of their Experiment 2 additionally violates the weaker preference for SO, hence its slow RTs. Given Experiment 1's sole main effect of case marking, the SO preference seems to operate differently for perfect active verbs. The authors speculate that this is related to semantic and discourse–pragmatic properties of the Georgian perfect tense, a past evidential that is particularly felicitous when A is less relevant to or identifiable in an event than P. So, perhaps in perfect clauses P_{NOM} is interpreted by default with higher discourse prominence than A_{DAT} is. The A_{DAT}–P_{NOM}–V_{PERF} condition violates this secondary preference for initial high-prominence P, but it still satisfies the general SO preference; the P_{NOM}–A_{DAT}–V_{PERF} condition does the opposite. If the two costs related to linear order cancel out here, this explains the lack of main effect of word order in Experiment 1.

Building on Skopeteas et al. 2011, Foley (2020, chapter 2) ran two self-paced reading experiments. Both had a 2×2 design manipulating word order (SOV vs. OSV) and tense / case marking ($A_{NOM}/P_{DAT}/V_{FUT}$ vs. $A_{ERG}/P_{NOM}/V_{AOR}$) of simple active transitive clauses; arguments both referred to humans in Experiment 1, and both to inanimate objects in Experiment 2. For all-human sentences, an analysis of trimmed raw RTs found a significant main effect of word order and an case–order interaction at the verb region, reflected numerically as greater RTs in the P_{NOM}–A_{ERG}–V_{AOR} condition than the other three conditions. Foley interprets this a garden path effect: high-animacy nominative in initial position is preferentially parsed as A; ergative at N2 forces the comprehender to revise that parse, leading to slower RTs at the immediately following verb. The lack of a word-order effect in the A_{NOM}/P_{DAT} future conditions is compatible with Skopeteas et al.'s hypothesis that a dative argument is preferentially linked to objecthood, whatever its linear order or animacy. Foley (2020) offers an alternative explanation, linked to the availability of the G parse for datives. If the G role is more prominent than P (Belletti & Rizzi 1988, Dowty 1991), and if high-animacy arguments are parsed with the highest grammatically available role by default in Georgian, then $A_{NOM}/G_{DAT}/V_{ACT.APPL.FUT}$ (i.e., a ditransitive) is the most harmonic parse for a $N1_{HU.NOM}$–$N2_{HU.DAT}$ string. The observed future-tense monotransitive verb would necessitate a reparse. On the other hand, a $N1_{HU.DAT}$–$N2_{HU.NOM}$ string would be best parsed in a $A_{DAT}/P_{NOM}/V_{ACT.PERF}$ clause; encountering a future-tense monotransitive here again forces revision.

In Experiment 2, involving all-inanimate clauses, Foley finds a case–order interaction at N1: numerically, ergative inanimates in initial position are read slower than nominative or dative ones. This ergative-inanimate cost is expected given crosslinguistic observations that low-prominence proto-actors are associated with processing difficulty (Bornkessel-Schlesewsky & Schlesewky 2009b). At the verb region, there is another interaction, interpreted as a crossover effect: OS order is harder to process in the $A_{ERG}/P_{NOM}/V_{AOR}$ context, but SO is harder given $A_{NOM}/P_{DAT}/V_{FUT}$. The P_{NOM}–A_{ERG} penalty is likely another ergative-inanimate cost, rooted in animacy–role disharmony, spilling over onto verb region. To explain the A_{NOM}–P_{DAT} penalty, Foley again suggests that the availability of the G parse leads comprehenders down a garden path: a ditransitive parse (P_{NOM}–G_{DAT}) would allow them to avoid positing an inanimate A, but the monotransitive verb dispels that possibility.

2.3 Corpus frequencies

Foley 2022 estimates the relative proportion of grammaticality case–role mappings in Georgian by querying the morphologically parsed Georgian Reference Corpus (a 200-million word corpus, part of the GNC Project; Gippert & Tandashvili 2015). The following tables reproduces estimated counts of the various grammatical roles licensed across tense Series, and thus case. For example, there were enough verbs a Series II tense to license 61,928 transitive subjects (Series III × A_{TR}, Table 2). That figure corresponds to all the possible ergative transitive subjects in the corpus, or about 13% (ERG × A_{TR}, Table 3) of the 475,805 total licensed arguments.

	S (subjects)				O (objects)			
	A_{TR}	**A_{INTR}**	**U**	**Ex**	**P**	**Th**	**G**	
Series I	37,474	15,235	38,467	14,639	43,143	26,516	26,148	201,622
Series II	61,928	6,572	48,717	3,596	75,825	5,356	36,980	238,974
Series III	9,607	574	7,947	786	12,101	1,181	3,013	35,209
	109,009	22,381	95,131	19,021	131,069	33,053	66,141	

Table 2: Corpus counts adapted from Foley (2022). Figures give the maximum number of arguments across tense categories licensed by verbs of various argument structures in the GRC. Cell-shading indicates case (white for nominative, grey for dative, black for ergative).

	S (subjects)				O (objects)			
	A_{TR}	**A_{INTR}**	**U**	**Ex**	**P**	**Th**	**G**	
NOM	7.9%	3.2%	20.0%	0	18.5%	7.0%	0	56.6%
DAT	2.0%	0.1%	0	4.0%	9.1%	0	12.9%	28.1%
ERG	13.0%	1.4%	0	0	0	0	0	14.4%
	22.9%	4.7%	20.0%	4.0%	27.6%	7.0%	12.9%	

Table 3: Corpus proportions adapted from Foley (2022). Figures give the percentage of total arguments across case categories licensed by verbs of various argument structures in the GRC.

These findings suggest that in Georgian: (i) nominative-case proto-patients (i.e., U and P; internal arguments) very plentiful, more than nominative-case proto-agents (i.e., A_{TR} and A_{INTR}; external arguments); (ii) dative-case subjects are rather rare; (iii) proto-agents are mostly ergative; and (iv) dative-case indirect objects (i.e., G; applied arguments — participants which are neither prototypical agents or patients, perhaps also including Ex) are quite frequent, more so than dative-case direct objects. Future research, perhaps using a treebank, will be necessary to validate these estimates in a way that takes into account factors like word order, animacy, or null pronominal arguments.

2.4 Crosslinguistic processing of case–role ambiguities

For excellent overviews of the psycholinguistic–neurolinguistic literature on processing grammatical roles, see Bornkessel-Schlesewsky and Schlesewsky (2009b, 2012, 2014). This subsection summarizes a few crosslinguistic generalizations about case, word order, animacy, and verb-finality. It mostly sets aside issues related to the processing of verb-initiality or filler–gap dependencies; see the Conclusion (Section 6).

Its case syncretisms and quirky assignment patterns make German a good window to real-time role processing (Hemforth et al 1993, Schriefers et al. 1995, Bader and Meng 1999, Bornkessel et al. 2004). In that language, subject-initial orders are easier to process, and inanimate transitive subjects are difficult. Comparable results have been found for Turkish (Demiral et al. 2008), Chinese (Philipp et al. 2008, Wang et al. 2009), Estonian (Miljan et al. 2017), Tamil (Muralikrishnan et al. 2008), Italian (de Vincenzi 1991), Hindi (Choudhary et al. 2008, Bickel et al. 2015), etc. A bias towards subject-initial parses is a common theme across languages, but it is perhaps not universal — see, for example, Yasunaga et al. (2015) on Kaqchikel. A bias against inanimate transitive subjects is very plausibly universal, and plausibly related to cognitive principles more fundamental than grammar or language (Bornkessel-Schlesewsky & Schlesewsky 2014).

Studies on other languages verb-final ergative languages are few (on the interaction with verb initiality, see Clemens et al. 2015, Longenbaugh & Polinsky 2016, Tollan et al. 2019). Relevant is Polinsky et al. (2012), a reading-time study on Avar. There ergative case clearly increases processing cost, at least within prenominal relative clauses temporarily parsable as root clauses. In Georgian, postverbal ergative subjects are also relatively difficult to process, at least when they disambiguate object-relative clauses (Lau et al. 2013).

3. Predictions

Georgian's case system makes the relationship between arguments' case morphology and their grammatical function highly entropic (Foley 2022). Given the language's flexible verb-final word order and null pronouns, this poses a unique challenge for the active comprehender. While ergative is a very rich cue in and of itself (being found only on active subjects in Series II tenses), nominative and dative are highly ambiguous, at least without reliable cues to tense or argument structure. How does the comprehender navigate these ambiguities? This section synthesizes previous sentence-processing research into a theory to be tested in the current study.

3.1 Decomposing processing costs

Previous sentence processing research, into Georgian and beyond, implicates a set of constraints that guide comprehenders' expectations about a preverbal argument's grammatical role. This section develops a theory of role-processing couched in the formalisms of Harmonic Grammar (HG; Legendre et al. 1990, Smolensky & Legendre 2006). HG evaluates a set of competitors (here, 'parses': grammatical continuations for an observed input string of words, varying in argument structure and case mappings) relative to a set of violable constraints, each given a weight. The sum of those weighted violations is called a harmony score (*H*). Our key linking hypothesis (15) is that the processing cost associated with a particular parse — as reflected in the RT of a verb disambiguating to it — is directly proportional to its Harmony Score. This approach is similar to previous Optimality Theoretic approaches to incremental sentence processing (e.g., Hoeks & Hendriks 2011), but the gradient output of HG allows non-optimal parses to be meaningfully compared.

(15) **Linking Hypothesis: RT(cue) $\propto$ *H*(parse)**
The processing cost (e.g., reaction time) of a morphosyntactic cue (e.g., tense–argument structure morphology on a verb) is proportional to the harmony score of the parse that cue entails.

One family of constraints privileges the harmonic alignment of prominence scales, ranked values of linguistic categories like those in 16. In naturalistic discourse, values along these scales are highly correlated (Hopper & Thompson 1980): the more Proto-Agent properties an argument has by virtue of its licensing verb's lexical semantics, the more likely it will be an animate, clause-initial, topical subject in an unmarked case; the more Proto-Patient properties, the more likely the argument is inanimate, non-initial, non-topical object in a marked case. Misaligned scales lead systematically to processing costs (Bornkessel Schlesewsky & Schlesewky 2009b, 2014), and in many languages certain misalignments are altogether ungrammatical (Aissen 1999).

(16) Thematic role: More Proto-Agent properties > More Proto-Patient properties
Syntactic role: Subject > Object
Animacy role: Human > Animal > Inanimate
Linear order: First > Second
Discourse role: Topical, Given, Definite > Focused, New information, Indefinite
Case: Nominative, Absolutive > Ergative, Accusative > Dative, Oblique

(17) **AlignAnimacy:** For every argument in a parse, assign violations according to the alignment of its animacy and thematic prototypicality (see table below).

	Proto-Agent	Neither	Proto-Patient
Human		*	**
Inanimate	**	*	

Table 4: Violations assigned by AlignAnimacy

(18) **AlignOrder:** For every pair of arguments, assign a violation if the argument with the more prominent thematic role linearly follows the argument with the less prominent role.

These Align constraints evaluate an argument on its absolute prominence. It is also useful to evaluate arguments on their prominence relative to each other. Processing multivalent clauses involves holding several discourse entities in memory — a task which becomes more difficult as the entities share more linguistic features (cf. encoding and retrieval interference; e.g., Villata et al. 2018). Focusing on thematic similarity, Bornkessel-Schlesewsky & Schlesewsky (2006, 2009b) call this constraint Distinctness. It is empirically motivated by the processing disadvantage of transitive clauses with experiencer-object verbs relative to more canonical transitive verbs (Bornkessel et al. 2005). Because experiencer objects have few canonical Proto-Patient properties, they are less thematically distinct from their subject than is grammatically possible.

(19) **Distinctness:** For every pair of arguments in a parse, assign violations according to how far from each other they are on the thematic prominence scale. (Monovalent parses receive no violations.)
Proto-Agent & Proto-Patient: No violations
Proto-Agent & Neither: *
Proto-Patient & Neither: *

Finally, I posit that comprehenders favor more economical valence structures: all things being equal, a clause with fewer arguments will be easier to process than one with more (cf. Polinsky et al. 2012).

(20) **Economy:** For every argument in a parse, assign a violation.

3.2 Harmony as a predictor of processing cost

To illustrate how these constraints interact, consider the space of grammatical parses in Georgian compatible with a single nominative noun ($N1_{NOM}$), a nominative noun immediately followed by a dative noun ($N1_{NOM}$–$N2_{DAT}$), or a dative–nominative sequence ($N1_{NOM}$–$N2_{DAT}$). The sentences below illustrate four parses compatible with those strings (21–23), differing in argument structure and tense / case mapping. These particular parses are compared directly in Experiments I–III; discussion here serves to motivate the design of the present study and ground discussion of its results.

As sentences in 32 show, a single high-animacy nominative noun ($N1_{NOM}$) could in principle be parsed as an active transitive subject (A), a nonactive subject (U), or a direct object (P). Being a human noun in clause initial position, $N1_{NOM}$ is unambiguously prominent on animacy and linearity scales. Therefore Align constraints will favor continuations where it is parsed as a Proto-Agent (A) over ones where it is a Proto-Patient (U or P). Argument prominence is orthogonal to valence, though: we predict independent processing costs associated with bivalent parses (21a/AVP, 21c/UVG, 21d/PVX) relative to the more economical monovalent one (21b/UVX), and costs associated with thematically distinct active transitives (21a/AVP, 21d/PVX) relative to the less-distinct applied nonactive (21c/UVG, where the U argument has many Proto-Patient properties but the G argument does not have many Proto-Agent properties).

(21) a.

ActSubj	**Verb:Tr.SeriesI**	**DirObj**		**(AVP)**
msaχiob-i	gaatʃʰerebs	mʦ’eral-s	ezo-ʃi	
actor-**NOM**	stop:**ACT:FUT**	writer-**DAT**	garden-in	

"The actor$_{[A]}$ will stop the writer$_{[P]}$ in the garden."

b.

NactSubj	**Verb:Nact**	**Adjunct**		**(UVX)**
msaχiob-i	gatʃʰerdeba	mʦ’erl-is	ezo-ʃi	
actor-**NOM**	stop:**NACT:FUT**	writer-**GEN**	garden-in	

"The actor$_{[U]}$ will stop in the writer's$_{[X]}$ garden."

c.

NactSubj	**Verb:Nact.Appl**	**IndObj**		**(UVG)**
msaχiob-i	gautʃʰerdeba	mʦ’eral-s	ezo-ʃi	
actor-**NOM**	stop:**NACT:APPL:FUT**	writer-**DAT**	garden-in	

"The actor$_{[U]}$ will stop for the writer$_{[G]}$ in the garden."

d.	**DirObj**	**Verb:Tr.SeriesII/III**	**Adjunct**		**(PVX; null A)**
	msaχiob-i	gavaʧʰere	mʦ’erl-is	ezo-ʃi	
	actor-**NOM**	stop:**ACT.AOR.1SGA**	writer-**GEN**	garden-in	

“I$_{[A]}$ stopped the actor$_{[P]}$ in the writer’s$_{[X]}$ garden.”

The next set of sentences illustrate parses compatible with the preverbal sequence $N1_{NOM}$–$N2_{DAT}$: a S_{NOM}–O_{DAT} monotransitive in Series I (22a), an S_{NOM}–O_{DAT} ditransitive in Series I (22b), an O_{NOM}–S_{DAT} monotransitive in Series III (22c), or a O_{NOM}–O_{DAT} ditransitive in Series II (22d). The monotransitive parses fair better than the ditransitive ones both in terms of Economy and Distinctness: positing a trivalent argument structure after encountering just two arguments is not a conservative parsing strategy, and the applied G argument will be relatively difficult to distinguish from both the A and P arguments. Since both preverbal nouns are human, AlignAnimacy will favor parsing them in the most prominence possible grammatical roles: A_{NOM}–G_{DAT} (22b) being more harmonic than either A_{NOM}–P_{DAT} (22a) or P_{NOM}–A_{DAT} (22c), which in turn are more harmonic than P_{NOM}–G_{DAT} (22d). Finally, AlignOrder prefers parses where nouns with more prominent roles precede nouns with less prominent ones: A–P (22a) and A–G (22b) parses satisfying that constraint better than P–A (22c) and P–G (22d).

(22) a.	**ActSubj**	**DirObj**	**Verb:Tr.SeriesI**	**Adjunct**		**(APVX)**
	ekʰim-i	msaχiob-s	gaaʧʰerebs	mʦ’erl-is	ezo-ʃi	
	doctor-**NOM**	actor-**DAT**	stop:**ACT:FUT**	writer-**GEN**	garden-in	

“The doctor$_{[A]}$ will stop the actor$_{[P]}$ in the writer’s$_{[X]}$ garden.”

b.	**ActSubj**	**IndObj**	**Verb:Ditr.SeriesI**	**DirObj**		**(AOVO)**
	ekʰim-i	msaχiob-s	gauʧʰerebs	mʦ’eral-s	ezo-ʃi	
	doctor-**NOM**	actor-**DAT**	stop:**ACT:APPL:FUT**	writer-**DAT**	garden-in	

“The doctor$_{[A]}$ will stop the writer$_{[P]}$ for the actor$_{[G]}$ in the garden.” (AGVP reading)
or “The doctor$_{[A]}$ will stop the actor$_{[P]}$ for the writer$_{[G]}$ in the garden.” (APVG reading)

c.	**DirObj**	**ActSubj**	**Verb:Tr.SeriesIII**	**Adjunct**		**(PAVX)**
	ekʰim-i	msaχiob-s	gauʧʰerebia	mʦ’erl-is	ezo-ʃi	
	doctor-**NOM**	actor-**DAT**	stop:**ACT:PERF**	writer-**GEN**	garden-in	

“The actor$_{[A]}$ must have stopped the doctor$_{[P]}$ in the writer’s$_{[X]}$ garden.”

d.	**DirObj**	**IndObj**	**Verb:Ditr.SeriesII**	**Adjunct**		**(PGVX; null A)**
	ekʰim-i	msaχiob-s	gavuʧʰere	mʦ’erl-is	ezo-ʃi	
	doctor-**NOM**	actor-**DAT**	stop:**ACT:APPL:AOR:1SGA**	writer-**GEN**	garden-in	

“I$_{[A]}$ stopped the doctor$_{[P]}$ for the actor$_{[G]}$ in the writer’s$_{[X]}$ garden.”

Given Georgian’s flexible word order, all parses compatible with $N1_{NOM}$–$N2_{DAT}$ are also compatible with $N1_{DAT}$–$N2_{NOM}$. Thus dative-initial monotransitive (23a, 23d) and ditransitive parses (23b, 23c) will violate Economy, Distinctness, and AlignAnimacy just as their nominative-initial counterparts do. The reversed word order, though, does change how AlignOrder is evaluated: the $A_{DAT}P_{NOM}$ (23a) and $G_{DAT}P_{NOM}$ (23c) parses do not violate this constraint, but $G_{DAT}A_{NOM}$ (23b) and $P_{DAT}A_{NOM}$ (23d) parses do.

(23)

a.

ActSubj	**DirObj**	**Verb:Tr.SeriesIII**	**Adjunct**		**(APVX)**
ekʰim-s	msaχiob-i	gauʧʰerebia	mʦ’erl-is	ezo-ʃi	
doctor-**DAT**	actor-**NOM**	stop:**ACT:PERF**	writer-**GEN**	garden-in	

“The doctor$_{[A]}$ must have stopped the actor$_{[P]}$ in the writer’s$_{[X]}$ garden.”

b.

IndObj	**ActSubj**	**Verb:Ditr.SeriesI**	**DirObj**		**(OAVO)**
ekʰim-s	msaχiob-i	gauʧʰerebs	mʦ’eral-s	ezo-ʃi	
doctor-**DAT**	actor-**NOM**	stop:**ACT:APPL:FUT**	writer-**DAT**	garden-in	

“The actor$_{[A]}$ will stop the writer$_{[P]}$ for the doctor$_{[G]}$ in the garden.”
or “The actor$_{[A]}$ will stop the doctor$_{[P]}$ for the writer$_{[G]}$ in the garden.” (PAVG reading)

c.

IndObj	**DirObj**	**Verb:Ditr.SeriesII**	**Adjunct**		**(GPVX)**
ekʰim-s	msaχiob-i	gavuʧʰere	mʦ’erl-is	ezo-ʃi	
doctor-**DAT**	actor-**NOM**	stop:**ACT:APPL:AOR:1SGA**	writer-**GEN**	garden-in	

“We$_{[A]}$ stopped the actor$_{[P]}$ for the doctor$_{[G]}$ in the writer’s$_{[X]}$ garden.”

d.

DirObj	**ActSubj**	**Verb:Tr.SeriesI**	**Adjunct**		**(PAVX)**
ekʰim-s	msaχiob-i	gaaʧʰerebs	mʦ’erl-is	ezo-ʃi	
doctor-**DAT**	actor-**NOM**	stop:**ACT:PERF**	writer-**GEN**	garden-in	

“The actor$_{[A]}$ will stop the doctor$_{[P]}$ in the writer’s$_{[X]}$ garden.”

The following tableaux summarize the constraint evaluation for parses discussed in this section. And given our linking hypothesis (15), these also convey a possible set of predictions for Experiments I–III. The ‘inputs’ refer to previously observed linguistic arguments, so the harmony scores here correspond to processing of the verb. (For some parses, the verb’s argument structure morphology entails a yet-unencountered argument, represented in parentheses; the existence of that argument suffices to evaluate Economy and Distinctness, but bottom-up evidence of its lexical semantics and position/overtness are necessary to evaluate AlignAnimacy and AlignOrder.)

Calculating harmony scores (*H*, rightmost column) requires assigning each constraint a weight. Given previous experimental research on Georgian, we might speculate that AlignAnimacy is weighted relatively high (given the dramatic inanimate-ergative costs Foley 2020 finds) while AlignOrder is weighted relatively low (since Skopeteas et al. 2011 found no main effect of word order); Economy and Distinctness are given weights arbitrarily. At this point, illustrating the logic of HG is more important than precisely predicting harmony scores.

Note that the Series I ditransitive clauses (22b, 23b) correspond to two possible parses, since both P and G arguments receive dative case in those tenses. Thus the verb’s inflection cannot completely disambiguate the case mapping, like it does elsewhere. Because P and G have different thematic interpretations, how the dative argument is parsed will affect the evaluation of the alignment constraints. We assume that, for strings whose structure–mapping is not fully disambiguated (like $N1_{NOM}$–$N2_{DAT}$–$V_{FUT.APPL}$), processing cost will be proportional to the average of all compatible parses’ harmony scores.

Input: $N1_{HU.NOM}$	AlignAnim $w = 3$	Economy $w = 2$	Distinctness $w = 2$	AlignOrder $w = 1$	*H*
A_{NOM} V_{FUT} [P_{DAT}]		**			4
U_{NOM} V_{NACT}	**	*			8
U_{NOM} $V_{NACT.APPL}$ [G_{DAT}]	**	**	*		12
P_{NOM} V_{AOR} [A_{ERG}]	**	**			10

Table 5: HG tableau for parses compatible with a single preverbal nominative noun (21)

Input: $N1_{HU.NOM}$ $N2_{HU.DAT}$	AlignAnim $w = 3$	Economy $w = 2$	Distinctness $w = 2$	AlignOrder $w = 1$	*H*
A_{NOM} P_{DAT} V_{FUT}	**	**			10
A_{NOM} G_{DAT} $V_{FUT.APPL}$ [P_{DAT}]	*	***	* + *		13
A_{NOM} P_{DAT} $V_{FUT.APPL}$ [G_{DAT}]	**	***	* + *		16
P_{NOM} A_{DAT} V_{PERF}	**	**		*	11
P_{NOM} G_{DAT} $V_{AOR.APPL}$ [A_{ERG}]	** + *	***	* + *	*	20

Table 6: HG tableau for parses compatible with a nominative noun followed by a dative noun (22)

Input: $N1_{HU.DAT}$ $N2_{HU.NOM}$	AlignAnim $w = 3$	Economy $w = 2$	Distinctness $w = 2$	AlignOrder $w = 1$	*H*
A_{DAT} P_{NOM} V_{PERF}	**	**			10
G_{DAT} A_{NOM} $V_{FUT.APPL}$ [P_{DAT}]	*	***	* + *	*	14
P_{DAT} A_{NOM} $V_{FUT.APPL}$ [G_{DAT}]	**	***	* + *	*	17
G_{DAT} P_{NOM} $V_{AOR.APPL}$ [A_{ERG}]	* + **	***	* + *		19
P_{DAT} A_{NOM} V_{FUT}	**	**		*	13

Table 7: HG tableau for parses compatible with a dative noun followed by a nominative one (23)

4. L-Maze study

The present study comprises three experiments designed to test hypotheses laid out in the previous section about the comprehension of incrementally ambiguous mono-, bi-, and tri-valent clauses with nominative and dative arguments in Georgian. 208 total itemsets were constructed: 24 for Experiment I; 32 for Experiment II; 32 for Experiment III; and 120 additional fillers of comparable length and complexity.

These were divided evenly into two lists, which each served as the stimuli for two experimental sessions of 104 trials each.

The experimental methodology used was the Lexicality Maze (L-Maze; Freedman & Forster 1985, Boyce et al. 2020), a variant of self paced reading that has participants make a series of incremental forced-choice lexicality decisions. Each word of each itemset was paired with a nonce word of equal orthographic length. The nonce words were generated automatically from orthographic trigram frequencies extracted from a collection of texts written in Georgian and the closely related language Megrelian.

The real and nonce words appeared side by side, in a random order. Participants were instructed to choose the real Georgian word of the pair, and that together the real Georgian words would form a sentence. They used the [e] and [i] keys to input their lexicality decision. A correct judgement would lead automatically to the next pair of words. A feedback message would appear after an incorrect judgement; participants would remake the lexicality decision and continue on through the rest of the item. The mock-up of one trial, translated into English, follows.

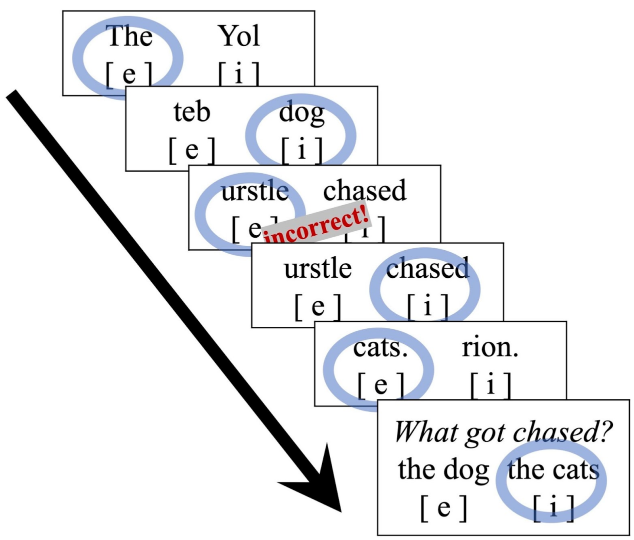

Figure 1: Illustration of the L-Maze methodology. Participant selections, input with the keyboard, are indicated with blue circles.

Hosted on PCIbex (Zehr & Schwartz 2018), the study was conducted entirely remotely. Each experimental session began with a demographic form, instructions, and three practice items. The 104 target items of each session were presented in a random order, with breaks built in at about one-third and two-thirds through. Each participant saw one version of each itemset, according to the Latin Square manner. Each session ended with a few optional debriefing questions.

56 native speakers of Georgian residing in Georgia were recruited to participate in the study. 44 participants took both experimental sessions. Among them, 38 took Session A first and 6 took Session B first. Another 8 participants took only Session A, and 6 took only Session B. All participants were paid 30 GEL for each experimental session they participated in.

All data from two participants with lexical decision accuracies lower than 60% were excluded from analysis. (Average accuracy for the other 54 participants was 97%.) Typos were found in five items across Experiments I–III; observations at or after misspelled words were set aside. Finally, we also chose to exclude within a given trial all correct lexicality decisions after any error, assuming that the error feedback message was likely to impede participants' comprehension of the whole sentence. These exclusion decisions left 88% of all collected data for analysis.

There were no comprehension questions. This design decision decreased the length and difficulty of the task, but not including a direct measure of whole-sentence understanding is somewhat unusual for an online processing experiment. Comprehension questions incentivize participants to pay closer attention to stimuli, and they give the experimenter another dependent variable. Without them, it is likely that some trials included in our analyses were in fact miscomprehended, for systematic or incidental reasons. It may even be that some participants made lexicality decisions without attending to grammatical dependencies between the words they chose. However, the clear and coherent results of this study give us some confidence that the L-Maze task alone was a good measure of incremental syntactic processing.

4.1 Experiment I (Nom–Verb)

Experiment I investigated how immediately preverbal nominative arguments are processed in active transitive and nonactive argument structures with various word orders.

4.1.1 Materials

Twenty-four itemsets were constructed in a four-condition design, similar to the quartet of examples in example 21 above. Key regions were a preverbal noun (N1), the verb, and a postverbal noun (N2). N1 was always in the nominative case, but functioned as a transitive subject (in the AVP condition), nonactive subject (UVX or UVG), or direct object (PVX). In the AVP and UVG conditions, N2 was a dative argument of the verb; in the UVX and PVX conditions, it was the first word of a postverbal adjunct constituent (usually a possessor, inflected genitive). Both N1 and N2 always referred to humans.

N1 was preceded by a single word (wd1), either an adjective or adverb. After N2 were two words (wd5 and wd6) that formed an adjunct constituent (usually a PP). In some conditions of some itemsets, the verb was immediately preceded by a grammatical particle (part). This was either the modal particle /unda/ "must, should" or the ordinary negation marker /ar/ "NEG". The particles were included to increase an item's global plausibility, generally by facilitating particular readings of the various tenses used. Of the 96 items (24 itemsets with 4 conditions each), 9 had one of these preverbal particles.

Table 8 summarizes the syntactic template for target stimuli in Experiment I.

<table>
<tr><th></th><th>wd1</th><th>noun1</th><th>(part)</th><th>verb</th><th>noun2</th><th>wd5 wd6</th></tr>
<tr><td>(Ia) AVP</td><td rowspan="4">Modifier</td><td>ActSubj-NOM</td><td rowspan="4">(MODAL, NEG)</td><td>Trans, SeriesI</td><td>DirObj-DAT</td><td rowspan="4">Adjunct XP</td></tr>
<tr><td>(Ib) UVX</td><td>NactS-NOM</td><td>Nonactive</td><td>Adjunct-GEN</td></tr>
<tr><td>(Ic) UVG</td><td>NactS-NOM</td><td>Nonact Appl</td><td>IndObj-DAT</td></tr>
<tr><td>(Id) PVX</td><td>DirObj-NOM</td><td>Trans, SeriesII, 1/2SubjAgr</td><td>Adjunct-GEN</td></tr>
</table>

Table 8: Syntactic template for target stimuli in Experiment I, with each word's grammatical function and inflection across the four conditions.

4.1.2 Analysis

Mixed linear effects models were run on log-transformed RTs, using the R package *lme4* (Bates et al. 2015); interaction effects were investigated with *emmeans* (Lenth et al. 2020). The design of Experiment 1 was not binary-factorial per se, but it was assimilated into one with the following sum-coding scheme. There were factors, *Voice* and *Markedness*. *Voice* was coded to compare the mean of the conditions with active verbs (Ia/AVP, Ib/PVX = –1/2) to the mean of conditions with nonactive verbs (Ib/UVX, Ic/UVG = +1/2). *Markedness* was coded to compare the mean of the subject initial, non-applicativized conditions (Ia/AVP, Ib/UVX = –1/2) to the mean of the object initial or applicativized conditions (Ic/UVG, Id/PVX = +1/2;). This latter contrast is artificial, conflating several grammatical factors. Log RTs were analyzed at the critical verb region, the postverbal noun, and a spillover region. The complexity of the random-effect structure was decreased bit by bit (Barr et al. 2013) until models converged without singularity errors or other issues.

4.1.3 Results

Figure 2 shows mean RTs, in raw ms, for each word region. Mean RTs are higher than in a typical self paced reading experiment. This is unsurprising given the L-Maze methodology, a hybrid lexical decision / SPR task. But, it may also be that properties of the Georgian script or Georgian morphology make the language inherently slower to read, or that the participants were less accustomed to tasks like a psycholinguistics experiment conducted online. Some standard errors are also quite large. This too is unsurprising, since RTs were only trimmed according to accuracy on the lexical decision task, not by their absolute magnitude. Some very large RTs, probably attributable to temporary distraction or loss of focus, remain in the dataset. Log transformation helps correct for these extreme observations.

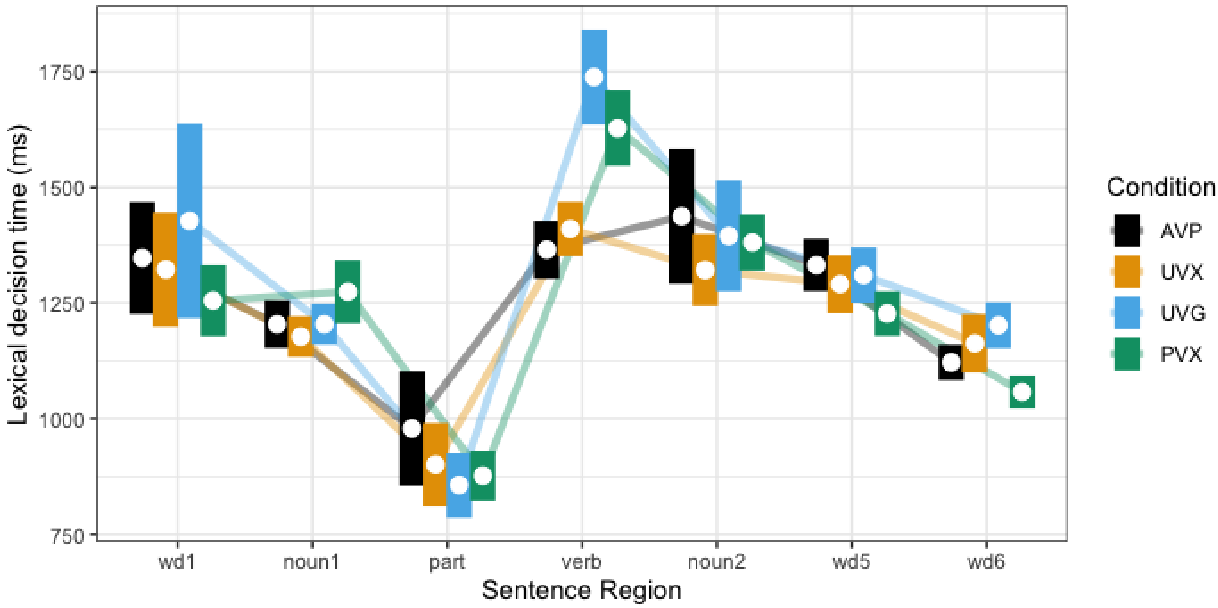

Figure 2: RTs by region for Experiment I, in raw milliseconds. The dots within colored bars indicate mean RTs ± one standard error, calculated by participant.

Results of linear models are summarized in the following tables. First, consider the verb region (Table 9). Using the first coding scheme, we find a significant effect of the factor *Markedness*: verbs in the relatively marked Ic/UVG and Id/PVX conditions were read reliably slower than verbs in the relatively unmarked conditions Ia/AVP and Ib/UVX.

Verb region lmer function: LogRT ~ Voice * Markedness + (0 + Voice * Markedness | Participant) + (0 + Voice * Markedness | Itemset)					
	Estimate	SE	df	*t*	*p*
Intercept	7.20	0.040	63.4	179	< 0.001
Voice	0.0376	0.038	25.6	0.969	0.341
Markedness	0.159	0.031	26.0	5.07	< 0.001
Voice:Markedness	–0.0054	0.076	21.5	–0.072	0.944

Table 9: Results of linear mixed effect modeling of log RTs at the verb region of Experiment I.

Next, the postverbal noun region (‘noun2’). There were no significant effects.

Postverbal noun region lmer function: LogRT ~ Voice * Markedness + (1 | Participant) + (1 + Voice + Markedness | Itemset)					
	Estimate	SE	df	*t*	*p*
Intercept	7.08	0.039	63.5	180	< 0.001
Voice	–0.0278	0.022	18.5	–1.22	0.236
Markedness	0.0423	0.027	20.7	1.56	0.134
Voice:Markedness	–0.0313	0.044	952	–0.711	0.477

Table 10: Results of linear mixed effect modeling of log RTs at the postverbal noun region of Experiment I.

Finally, the spillover region ('wd5'). There was a significant interaction of *Voice* and *Markedness*; pairwise comparisons revealed that this word was recognized significantly slower in the 'unmarked' version of the active transitive pair (i.e., Ia/AVP $>_{RT}$ Id/PVX; Est. = 0.0531, SE = 0.0267, $t(881) = 1.98$; $p < 0.05$), but 'marked' and 'unmarked' nonactive conditions (Ib/UVX and Ic/UVG) were not significantly different in this region (Est. = –0.0386, SE = 0.0275, $t(887) = -1.40$, $p = 0.16$).

Spillover region lmer function: LogRT ~ Voice * Markedness + (1 + Voice | Participant) + (1 | Itemset)					
	Estimate	SE	df	*t*	*p*
Intercept	7.06	0.0485	45.3	145	< 0.001
Voice	0.0135	0.0212	50.2	0.639	0.52
Markedness	–0.00725	0.0191	891	–0.379	0.70
Voice:Markedness	0.0916	0.0383	885	2.392	< 0.05

Table 11: Results of linear mixed effect modeling of log RTs at the spillover region of Experiment I.

4.1.4 Discussion

The starkest effect is how slow RTs are at verbs in the applicativized nonactive and object-initial transitive conditions (Ic/UVG and Id/PVX) compared to RTs there for the subject-initial transitive and simple nonactive conditions (Ia/AVP and Ib/UVX). The UVG cost is not surprising. Its verb eliminates the possibility of parsing $N1_{NOM}$ as a transitive subject, thereby violating Align(Animacy,Role). It also forces them to accommodate a yet unencountered argument, violating Economy. And, since that anticipated argument is an indirect object, which is only slightly more prominent than the observed Proto-Patient ($N1_{NOM}$), Distinctness is also violated.

As for the difficulty associated with the PVX verb, that could be because it violates Economy (forcing a bivalent parse) and Align(Animacy,Role) (forcing a high-animacy direct object), but not Distinctness (since the accommodated argument is a transitive subject). A design confound, however, prevents us from confidently attributing the effect here to processing difficulty associated with the $N1_{HU.NOM}$→DirObj parse per se. Unlike any other condition, Id/PVX had a first or second person subject, signaled solely by verbal agreement. If there is a significant processing cost of processing first/second agreement morphology registering a null pronoun, or accommodating a first/second person discourse referent, it may be that these costs are inflating the RTs observed at the verb in condition Id/PVX.

Finally, consider the AVP verb (which disambiguates to a parse violating Economy) and the UVX verb (which violates Align(Animacy,Role)). These constraints seem to be weighted about equally, given the lack of significant difference between RTs in these conditions.

At the spillover region, the *Voice–Markedness* interaction seems driven by the relative ease of processing this region in the 1d/PVX condition. Perhaps this is related to the fact that the previous word (N2) was an argument of the verb in the Ia/AVP and Ic/UVG conditions, but part of an adjunct in Ib/UVX and Id/PVX. It could also be related to the first/second verbal agreement in Id/PVX: after the processing inhibition it caused in the previous two regions, the comprehender 'bounces back'.

4.2 Experiment II (Nom–Dat–Verb)

Experiment II investigated how nominative–dative sequences are processed across monotransitive and ditransitive argument structures with various word orders.

4.2.1 Materials

Thirty-two itemsets were constructed in a four-condition design, similar to the quartet of examples above (22). Key regions were two preverbal nouns (N1 and N2), the verb, and a postverbal noun (N3). N1 was always in the nominative case, functioning either as an active subject (IIa/APVX, IIb/AOVO) or a direct object (IIc/PAVX, IId/PGVX). N2 was always dative, functioning as a direct object (IIa/APVX), indirect object (IIb/AOVO, IId/PGVX), or active subject (IIc/PAVX). N3 was a dative direct object in one condition (IIb/AOVO), and a genitive possessor within a clausal adjunct in the other conditions. Both N1 and N2 always referred to humans; N3 also referred to a human in almost every itemset.

As discussed above (Section 3.3), P and G are not distinguishable by case marking in Series I; both are dative. Therefore, at least in principle, condition IIb is globally ambiguous between an AGVP parse and an APVG parse. In order to bias comprehenders towards the AGVP parse, in almost every itemset N3 was a relational noun like "cousin" or "colleague". In general, an applied indirect object can be readily interpreted as an external possessor of the direct object in Georgian. If relational nouns are most felicitous with thematic possessors, and sentence-internal possessors are easier to accommodate than implicit ones, then comprehenders should be more likely to parse N3 as P than G in this condition (IIb); insofar as N2 is preferentially parsed as P than G, interpreting N3 as P will force reanalysis of N2 to G. The other applied condition (IId/PGVX) does not have this ambiguity, since P and G have different case marking in Series II.

N1 was preceded by a single word (wd1), either an adjective or adverb. After N3 were two words (wd6 and wd7) that formed an adjunct constituent (usually a PP). In some conditions of some itemsets, the verb was immediately preceded by a grammatical particle (part). This was either the modal particle /unda/ “must, should”, the ordinary negation marker /ar/ “NEG”, or the modal negation marker /ver/ “can’t”. The particles were included to increase an item’s global plausibility, generally by facilitating particular readings of the various tenses used. Of the 128 items (32 itemsets with 4 conditions each), 26 had one of these preverbal particles.

Table 12 summarizes the syntactic template for target stimuli in Experiment II.

<table>
<tr><th></th><th>wd1</th><th>noun1</th><th>noun2</th><th>(part)</th><th>verb</th><th>noun3</th><th>wd6 wd7</th></tr>
<tr><td>(IIa) APVX</td><td rowspan="4">Adj / Adv</td><td>ActSubj-NOM</td><td>DirObj-DAT</td><td rowspan="4">(MOD, NEG)</td><td>Monotrans, SeriesI</td><td>Adjunct-GEN</td><td rowspan="4">Adjunct XP</td></tr>
<tr><td>(IIb) AOVO</td><td>ActSubj-NOM</td><td>IndObj-DAT</td><td>Ditr, SeriesI</td><td>DirObj-DAT</td></tr>
<tr><td>(IIc) PAVX</td><td>DirObj-NOM</td><td>ActSubj-DAT</td><td>Monotrans, Series III</td><td>Adjunct-GEN</td></tr>
<tr><td>(IId) PGVX</td><td>DirObj-NOM</td><td>IndObj-DAT</td><td>Ditr, Series II, 1/2SubjAgr</td><td>Adjunct-GEN</td></tr>
</table>

Table 12: Syntactic template for target stimuli in Experiment II, with each word’s grammatical function and inflection across the four conditions.

4.2.2 Analysis

Mixed linear effects models were run on log-transformed RTs, using the R package *lme4* (Bates et al. 2015). Conditions were coded with two contrast factors, *ArgStrux* and *NomTheta*. The former factor compares the average of monotransitive conditions (IIa/APVX, IIc/PAVX = –1/2) to the average of ditransitive ones (IIb/AOVO, IId/PGVX = +1/2). The latter factor compares the conditions where $N1_{NOM}$ was a Proto-Agent (IIa/APVX, IIb/AOVO = –1/2) to the conditions where $N1_{NOM}$ was a Proto-Patient (IIc/PAVX, IId/PGVX = +1/2). Log RTs were analyzed at the critical verb region, the postverbal noun, and a spillover region. The complexity of the random-effect structure was decreased bit by bit (Barr et al. 2013) until models converged without singularity errors or other issues.

4.2.3 Results

Figure 3 shows mean RTs, in raw milliseconds, for each word region.

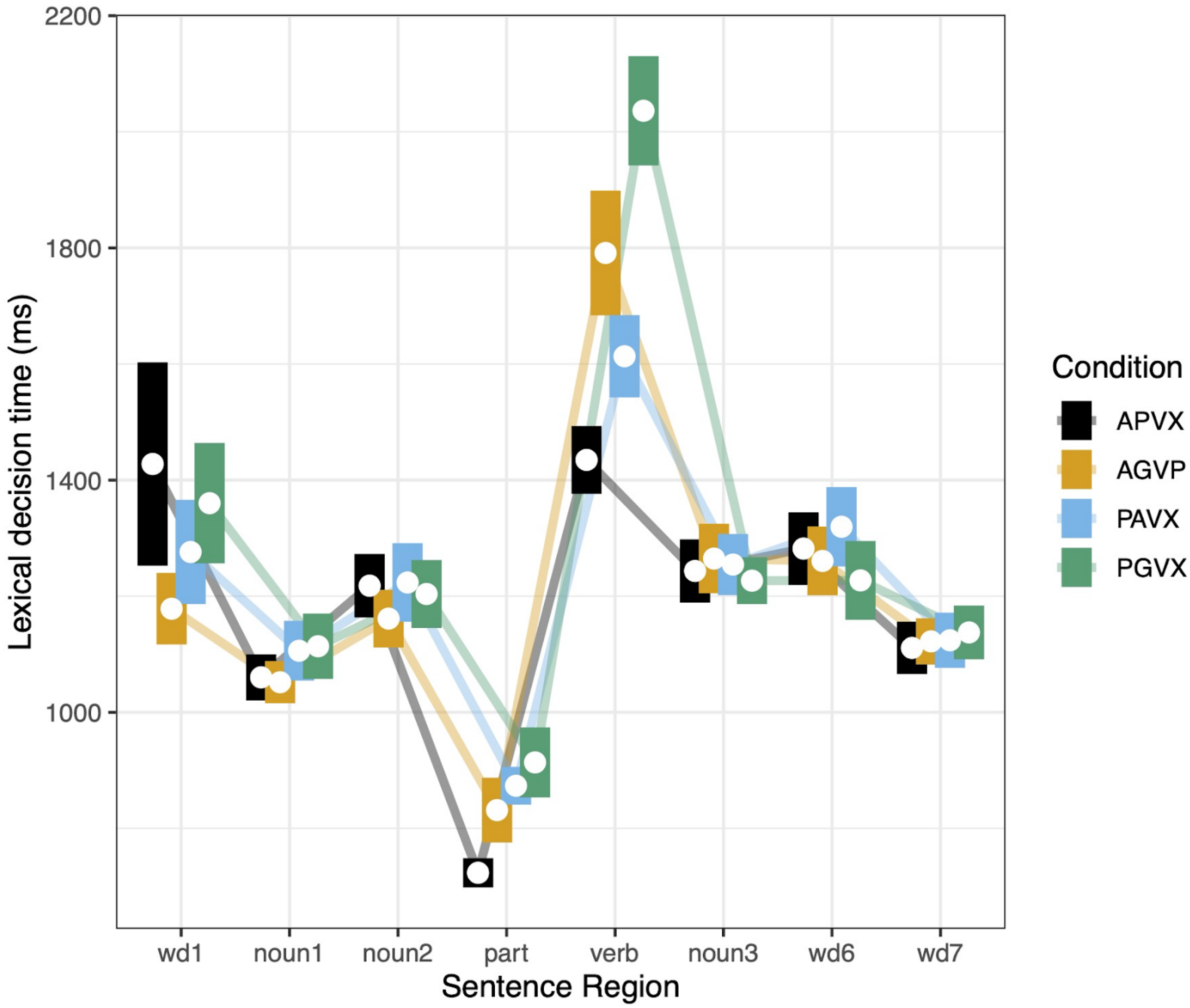

Figure 3: RTs by region for Experiment II, in raw milliseconds. The dots within colored bars indicate mean RTs ± one standard error, calculated by participant.

Results of linear models of RTs at the verb region are summarized in the following table. There were significant main effects of *ArgStrux* and *NomTheta*: ditransitive verbs were recognized more slowly than monotransitive ones, and verbs licensing a nominative Proto-Patient were slower than ones with nominative Proto-Agents.

Verb region lmer function: LogRT ~ ArgStrux*NomTheta + (1|Participant) + (1+ArgStrux|Itemset)					
	Estimate	SE	df	*t*	*p*
Intercept	7.30	0.0398	60.5	183	< 0.001
ArgStrux	0.187	0.0257	30.2	7.27	< 0.001
NomTheta	0.132	0.0207	1,350	6.37	< 0.001
ArgStrux:NomTheta	0.0641	0.0415	1,350	1.54	0.123

Table 13: Results of linear mixed effect modeling of log RTs at the verb region of Experiment II.

4.2.4 Discussion

The main effect of *NomTheta* shows that sentence-initial high-animacy nominative arguments are more easily parsed as A than P. This result is consistent with a relatively high weight for at least one of AlignAnimacy, AlignOrder, or AlignCase — that is, a relatively strong advantage for events with high-animacy Proto-Patients, sentences with Proto-Agents that precede Proto-Patients, or Proto-Agents in the nominative rather than dative case.

The main effect of *ArgStrux* shows that ditransitive verbs are harder to process than monotransitive ones. This difficulty could have several sources. It could be connected to violations of Economy or Distinctness, parsing constraints that favor fewer arguments with maximally distinct thematic roles. But it also be related to an ambiguity inherent to condition Ib/AOVO, which is also parseable as APVG; perhaps verbs whose morphological cues do not totally disambiguate their arguments' grammatical roles are harder than verbs that do. (Recall that N3 was usually a relational noun, hypothesized to bias the comprehender towards the former parse; insofar as this design feature was successful, the lack of significant effects at N3 and the following spillover region seems to indicate it does not lead to a reanalysis cost.) The first/second person agreement morphology on the verb in Id/PGVX might also be a source of processing difficulty independent of valence. In any case, this main effect shows that high-animacy dative arguments are not preferentially parsed as indirect objects, as Foley (2020) hypothesized; harmonic alignment constraints do not seem to countenance the full gamut of grammatical roles.

4.3 Experiment III (Dat–Nom–Verb)

Experiment III investigated how dative–nominative sequences are processed across monotransitive and ditransitive argument structures with various word orders.

4.3.1 Materials

Thirty-two itemsets were constructed in a four-condition design, similar to the quartet of examples above (23). Key regions were two preverbal nouns (N1 and N2), the verb, and a postverbal noun (N3). N1 was always in the dative case, functioning either as an active subject (IIIa/APVX), applied indirect object (IIIb/OAVO, IIIc/GPVX), or direct object (IIId/PAVX). N2 was always nominative, functioning as direct object (IIIa/APVX, IIIc/GPVX) or active subject (IIIb/OAVO, IIId/PAVX). N3 was a dative direct object in one condition (IIIb/OAVO), and a genitive possessor within a clausal adjunct in the other conditions. Both N1 and N2 always referred to humans, except in one itemset where they were animals; N3 was also almost always human.

Just like as in Experiment II, the Series I ditransitive condition (IIIb/OAVO) has an alternative parse: PAVG. To bias comprehenders towards the former reading, N3 was typically a relational noun. (See Section 4.2.1.)

N1 was preceded by a single word (wd1), either an adjective or adverb. After N3 were two words (wd6 and wd7) that formed an adjunct constituent (usually a PP). In some conditions of some itemsets, the verb was immediately preceded by a grammatical particle (part). This was either the modal particle /unda/ "must, should", the ordinary negation marker /ar/ "NEG", or the modal negation marker /ver/ "can't". The particles were included to increase an item's global plausibility, generally by facilitating particular readings of the various tenses used. Of the 128 items (32 itemsets with 4 conditions each), 42 had one of these preverbal particles.

Table 14 summarizes the syntactic template for target stimuli in Experiment III.

<table>
<tr><th></th><th>wd1</th><th>noun1</th><th>noun2</th><th>(part)</th><th>verb</th><th>noun3</th><th>wd6 wd7</th></tr>
<tr><td>(IIIa) APVX</td><td rowspan="4">Adj / Adv</td><td>ActSubj-DAT</td><td>DirObj-NOM</td><td rowspan="4">(MOD, NEG)</td><td>Monotr, SeriesIII</td><td>Adjunct-GEN</td><td rowspan="4">Adjunct XP</td></tr>
<tr><td>(IIIb) OAVO</td><td>IndObj-DAT</td><td>ActSubj-NOM</td><td>Ditr, SeriesI</td><td>DirObj-DAT</td></tr>
<tr><td>(IIIc) GPVX</td><td>IndObj-DAT</td><td>DirObj-NOM</td><td>Ditr, Series II, 1/2SubjAgr</td><td>Adjunct-GEN</td></tr>
<tr><td>(IIId) PAVX</td><td>DirObj-DAT</td><td>TrSubj-NOM</td><td>Monotr, SeriesI</td><td>Adjunct-GEN</td></tr>
</table>

Table 14: Syntactic template for target stimuli in Experiment 3, with each word's grammatical function and inflection across the four conditions.

4.3.2 Analysis

Mixed linear effects models were run on log-transformed RTs, using the R package *lme4* (Bates et al. 2015); interaction effects were investigated with *emmeans* (Lenth et al. 2020). Conditions were coded with two contrast factors, *ArgStrux* and *NomTheta*. The former factor compares the average of monotransitive conditions (IIIa/APVX, IIId/PAVX = –1/2) to the average of ditransitive ones (IIIb/OAVO, IIIc/GPVX = +1/2). The latter factor compares the conditions where $N2_{NOM}$ was a Proto-Agent (IIIb/OAVO, IIId/APAVXGVP = –1/2) to the conditions where $N2_{NOM}$ was a Proto-Patient (IIIa/APVX, IIIc/GPVX = +1/2). Log RTs were analyzed at the critical verb region, the postverbal noun, and a spillover region. The complexity of the random-effect structure was decreased bit by bit (Barr et al. 2013) until models converged without singularity errors or other issues.

4.3.3 Results

Figure 4 shows mean RTs, in raw milliseconds, for each word region.

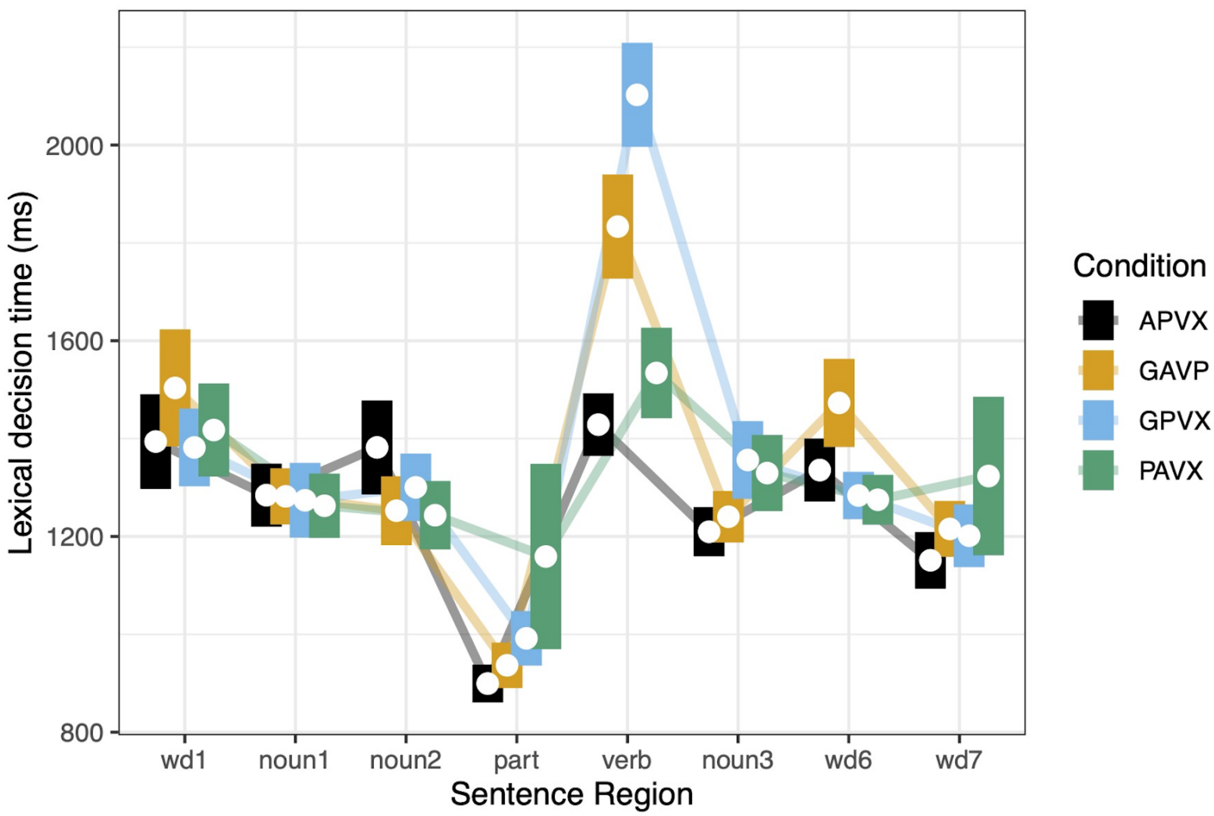

Figure 4: RTs by region for Experiment III, in raw milliseconds. The dots within colored bars indicate mean RTs ± one standard error, calculated by participant.

Results of linear models are summarized in the following tables. First, at the verb region (Table 15), there was a significant main effect of *ArgStrux*: ditransitive verbs were recognized more slowly than monotransitive ones. There was also a marginal *ArgStrux:NomTheta* interaction: for ditransitive conditions (IIIb/OAVO, IIIc/GPVX), the $N2_{NOM}$→P parse is harder than the $N2_{NOM}$→A parse (Est. = –0.105, SE = 0.0430, $t(38.9) = –2.43$, $p < 0.05$); for the monotransitive conditions (IIIa/APVX, IIId/PAVX) there is no reliable effect of $N2_{NOM}$'s grammatical role (Est. = 0.019, SE = 0.0411, $t(28.9) = 0.464$, $p = 0.646$).

Verb region lmer function: LogRT ~ ArgStrux*NomTheta + (1|Participant) + (1+ArgStrux*NomTheta|Itemset)					
	Estimate	SE	df	*t*	*p*
Intercept	7.29	0.0460	74.0	158	< 0.001
ArgStrux	0.242	0.0334	27.9	7.25	< 0.001
NomTheta	0.0427	0.0287	25.9	1.48	0.149
ArgStrux:Nom Theta	0.123	0.0608	29.3	2.03	0.0514

Table 15: Results of linear mixed effect modeling of log RTs at the verb region of Experiment III.

At the immediately postverbal noun region (Table 16), we find a marginal main effect of *ArgStrux* (N3 was marginally slower after ditransitives), and a significant *ArgStrux:NomTheta* interaction: when $N2_{NOM}$→P, the ditransitive is slower than the monotransitive (Est. = 0.0879, SE = 0.0281, $t(48.9) = 3.123$, $p < 0.01$), but not when $N2_{NOM}$→A (Est. = –0.0222, SE = 0.0285, $t(49.9) = –0.781$, $p = 0.438$).

Postverbal noun region lmer function: LogRT ~ ArgStrux*NomTheta + (1+ArgStrux*NomTheta|Participant) + (1+NomTheta|Itemset)					
	Estimate	SE	df	*t*	*p*
Intercept	7.03	0.0396	74.4	177	< 0.001
ArgStrux	0.0328	0.0188	97.2	1.73	0.0855
NomTheta	0.00667	0.0258	27.7	0.258	0.798
ArgStrux:Nom Theta	0.110	0.0417	50.7	2.64	< 0.05

Table 16: Results of linear mixed effect modeling of log RTs at the verb region of Experiment III.

In the spillover region (Table 17), there was a main effect of *ArgStrux*, with ditransitive conditions being slower than monotransitive ones. There was also an *ArgStrux:NomTheta* interaction — for conditions where $N2_{NOM}$→A (IIIb/OAVO, IIId/PAVX), the ditransitive parse is significantly slower than the monotransitive (Est. = 0.0877, SE = 0.0267, $t(1180) = 3.28$, $p < 0.05$), but the conditions where $N2_{NOM}$→P (IIIa/APVX, IIIc/GPVX) are not significantly different (Est. = –0.00317, SE = 0.0258, $t(1180) = –0.123$, $p = 0.902$); for ditransitive conditions (IIIb/OAVO, IIIc/GPVX), the $N2_{NOM}$→A parse is slower than the $N2_{NOM}$→P parse (Est. = 0.0720, SE = 0.0274, $t(775.5) = 2.630$, $p < 0.05$), but the monotransitive conditions (IIIa/APVX, IIId/PAVX) are not significantly different (Est. –0.0189, SE = 0.0265, $t(69.3) = –0.713$, $p = 0.478$).

Postverbal noun region lmer function: LogRT ~ ArgStrux*NomTheta + (1+NomTheta|PartID) + (0+NomTheta|Itemset)					
	Estimate	SE	df	*t*	*p*
Intercept	7.07	0.0453	65.8	156	< 0.001
ArgStrux	0.0422	0.0185	1220	2.28	< 0.05
NomTheta	–0.0265	0.0194	198	–1.36	0.173
ArgStrux:Nom Theta	–0.0909	0.0370	1220	–2.45	< 0.05

Table 17: Results of linear mixed effect modeling of log RTs at the verb region of Experiment III.

4.3.4 Discussion

As in Experiment II, the ditransitive verbs were harder to process than monotransitive ones — perhaps because of the Economy and Distinctness violations they incur by virtue of their argument structure per se. For condition IIIb/OAVO, we cannot rule out the possibility of an independent cost associated with lingering role ambiguity ($N1_{DAT}$ here being parsable as G or P). And given the very long RTs in IIIc/GPVX, there does seem to be a processing cost associated with first/second person verbal agreement.

Particularly notable is the fact that verbs in the IIIa/APVX and IIId/PAVX conditions are recognized at about the same speed. In other words, given $N1_{DAT}$–$N2_{NOM}$, it is not any harder to process cues disambiguating to $O_{DAT}S_{NOM}$ order (IIId) than cues disambiguating to $S_{DAT}O_{NOM}$ order (IIIa). This contrasts with the clear $O_{NOM}S_{DAT}$ cost in Experiment II: verbs in condition IIc/PAVX were slower than verbs in IIa/APVX.

In postverbal regions, the first/second agreement cost in IIIc/GPVX seems to linger into the processing of N3; the slow RTs of the following spillover region in the IIIc/OAVO condition is likely related to the fact that N3 is a verbal argument, rather than a possessor within an adjunct as in the other conditions.

4.4 Exploratory analyses

Here we report a few exploratory findings related to quirks of the design of Experiments I–III and the task / procedure of the study as a whole.

Within any experiment's itemsets, word order was not manipulated directly. However, conditions between experiments correspond to each other, identical in design except for word order and lexical items. Table 18 reports log RTs in all 12 conditions of this study, grouping together conditions with comparable arguments structures and case mappings. Arranging data this way suggests a few additional important findings. First, for SOV sentences, the A_{DAT}/P_{NOM} mapping of Series III tenses is not markedly harder to process than the A_{DAT}/P_{NOM} mapping of Series III. Second, OSV order is harder to process than SOV order for Series III clauses, but not Series I clauses. Third, for Series II ditransitives, where P and G get distinct cases, PGV order is not harder than GPV order. In other words, neither a disharmonic word order (O before S, or P before G) nor a disharmonic case mapping (with A_{DAT} rather than A_{NOM}) is difficult on its own to process in Georgian, but the two disharmonic properties in combination (as in the $P_{NOM}A_{DAT}V_{PERF}$ condition) do in fact seem to inhibit processing.

A post-hoc analysis pooling the monotransitive conditions from Experiment II and III lend credence to this generalization. The linear model[3] finds no main effect of case mapping (Est. = 0.0394, SE = 0.0260, $t(56.2)$, $p = 0.135$) but does find a main effect of word order (Est. = 0.0607, SE = 0.0260, $t(56.2) = 2.33$, $p < 0.05$). Pairwise comparison finds an OSV cost for the clauses with A_{DAT}/P_{NOM} mapping (Est. = 0.0993, SE = 0.0413, $t(59.8) = 2.40$, $p < 0.05$), but no effect of word order for the A_{NOM}/P_{DAT} mapping (Est. = 0.0222, SE = 0.0455, $t(59.6)$, $p = 0.626$).

[3] lmer function: LogRT ~ CaseMap*WordOrder + (1|PartID) + (1+CaseMap|Itemset); The sum-coded factors were *CaseMap* (–1/2 for A_{NOM}/P_{DAT} and +1/2 for A_{DAT}/P_{NOM}) and *WordOrder* (–1/2 for SOV and +1/2 for OSV).

	SV	SOV	OSV
Active transitive, Series I tense	$A_{\text{NOM}} - \mathbf{V}_{\textbf{TR.FUT}}$ [– P_{DAT}] Ia: **7.11 log ms (0.040)**	$A_{\text{NOM}} - P_{\text{DAT}} - \mathbf{V}_{\textbf{TR.FUT}}$ IIa: **7.16 log ms (0.036)**	$P_{\text{DAT}} - A_{\text{NOM}} - \mathbf{V}_{\textbf{TR.FUT}}$ IIId: **7.18 log ms (0.042)**
Active transitive, Series III tense	$A_{\text{DAT}} - V_{\text{TR.PERF}}$ [– P_{NOM}]	$A_{\text{DAT}} - P_{\text{NOM}} - \mathbf{V}_{\textbf{TR.PERF}}$ IIIa: **7.16 log ms (0.039)**	$P_{\text{NOM}} - A_{\text{DAT}} - \mathbf{V}_{\textbf{TR.PERF}}$ IIc: **7.26 log ms (0.040)**
Nonactive, Any tense	$U_{\text{NOM}} - \mathbf{V}_{\textbf{NACT}}$ Ib: **7.15 log ms (0.038)**	———	———
Nonactive appl., Any tense	$U_{\text{NOM}} - \mathbf{V}_{\textbf{NACT.APPL}}$ [– G_{DAT}] Ic: **7.30 log ms (0.051)**	$U_{\text{NOM}} - G_{\text{DAT}} - V_{\text{NACT.APPL}}$	$G_{\text{DAT}} - U_{\text{NOM}} - V_{\text{NACT.APPL}}$
Active appl., Series I tense	$A_{\text{NOM}} - V_{\text{DITR.FUT}}$ [– P/G_{DAT} …]	$A_{\text{NOM}} - G/P_{\text{DAT}} - V_{\text{DITR.FUT}}$ [– P/G_{DAT}] IIb: **7.32 log ms (0.045)**	$G/P_{\text{DAT}} - A_{\text{NOM}} - V_{\text{DITR.FUT}}$ [– P/G_{DAT}] IIIb: **7.34 log ms (0.050)**

	OV	GPV	PGV
Active transitive, Series-II, Agr-A	$P_{\text{NOM}} - V_{\text{TR.AOR.1A}}$ [– $A_{\text{1.ERG}}$] Id: **7.25 log ms (0.041)**	———	———
Active appl., Series-II, Agr-A	$G_{\text{DAT}} - V_{\text{DITR.AOR.1A}}$ [– P_{NOM}…] $P_{\text{NOM}} - V_{\text{DITR.AOR.1A}}$ [– G_{DAT}…]	$G_{\text{DAT}} - P_{\text{NOM}} - V_{\text{DITR.AOR.1A}}$ [– $A_{\text{1.ERG}}$] IIIc: **7.47 log ms (0.043)**	$P_{\text{NOM}} - G_{\text{DAT}} - V_{\text{DITR.AOR.1A}}$ [– $A_{\text{1.ERG}}$] IId: **7.48 log ms (0.043)**

Table 18: Summary of mean log recognition times (by-participant standard errors in parentheses) at the verb region across Experiments I–III, arranged by morphosyntactic features and word order. Corresponding stimuli are schematized above each observation; postverbal arguments which can be predicted from agreement or argument-structure morphology are shown in brackets. In grey are grammatically possible word orders not included in this study.

The experiment was rather long; the median completion time of one session was 49 minutes, including two breaks. It could be that the same structures elicit qualitatively different processing behavior towards the beginning or the end of an experiment, or between the first and second experimental session completed. Table 19 reports RTs at the verb region across experiments, broken up grossly by trial order and also pooled. (Recall that the experiments were intermixed, the stimuli for each session being a list of half of each experiment's itemsets.)

	Session 1		**Session 2**		**Across sessions**
	First half	**Second half**	**First half**	**Second Half**	
Ia/AVP	7.32 (0.065)	7.03 (0.073)	7.04 (0.074)	6.92 (0.053)	7.11 (0.040)
Ib/UVX	7.08 (0.065)	7.24 (0.066)	7.17 (0.060)	6.94 (0.060)	7.15 (0.038)
Ic/UVG	7.29 (0.090)	7.44 (0.078)	7.21 (0.092)	7.20 (0.080)	7.30 (0.051)
Id/PVX	7.34 (0.083)	7.34 (0.083)	7.20 (0.063)	7.14 (0.070)	7.25 (0.041)
IIa/APVX	7.27 (0.050)	7.12 (0.053)	7.02 (0.045)	7.05 (0.059)	7.16 (0.036)
IIb/AOVO	7.39 (0.071)	7.45 (0.079)	7.19 (0.061)	7.16 (0.060)	7.32 (0.045)
IIc/PAVX	7.38 (0.058)	7.25 (0.056)	7.13 (0.052)	7.09 (0.057)	7.26 (0.040)
IId/PGVX	7.53 (0.063)	7.50 (0.067)	7.50 (0.068)	7.28 (0.061)	7.48 (0.043)
IIIa/APVX	7.26 (0.054)	7.11 (0.049)	7.06 (0.049)	7.05 (0.051)	7.16 (0.039)
IIIb/OAVO	7.47 (0.072)	7.26 (0.062)	7.24 (0.070)	7.18 (0.067)	7.34 (0.050)
IIIc/GPVX	7.63 (0.067)	7.47 (0.068)	7.36 (0.068)	7.27 (0.083)	7.47 (0.043)
IIId/PAVX	7.28 (0.053)	7.18 (0.074)	7.16 (0.062)	7.00 (0.066)	7.18 (0.042)

Table 19: Mean RTs at the verb region for each condition of each experiment, in log ms (with standard errors). Each session had 104 trials, including non-experimental fillers, in a random order; each half-session is thus a block of 52 sentences.

In general, RTs decrease as trial order increases: participants become better at the L-Maze methodology over time. Effects across conditions of each experiment are rather consistent across as exposure increases, though it is notable how unstable the relative processing cost of SV transitive vs. nonactive verbs (Ia/AVP vs. Ib/UVX) seems.

Finally, a methodological note: the lack of comprehension questions in this remote-run reading-time study casts some minor doubt over the results as a whole — see Stewart et al. (2007) on the effect comprehension questions have on participants' depth of processing. High accuracy in the Maze task shows that participants were at least attending carefully to the lexicality of words across the stimuli. And the quite coherent RT results are hard to explain if participants were not also attending to morphosyntactic dependencies between those words, at least to some degree. Future research, though, should include a comprehension task to explicitly encourage deeper attention and filter out data from misunderstood trials.

5. General discussion

Three L-Maze experiments investigated how Georgian comprehenders parse preverbal nominative and dative arguments, which are radically ambiguous for grammatical role due to a complex split-ergative case mapping grammar sensitive to tense and argument structure. To recap key findings: (i) verbs with an applied G argument are harder to process than verbs with simple nonactive or active argument structures; (iii) lone nominative arguments are about as easy to parse as the subject of a simple nonactive (U) as they are the subject of an active transitive (A); (iii) verbs agreeing with a first- or second-person null subject are harder to process than ones with an overt preverbal third-person subject; and, by post-hoc reasoning, (iv) disharmonic word orders (e.g., OSV) impede processing only for verbs with A_{DAT}/P_{NOM} case mapping, triggered by Series III tenses. Table 20 summarizes our predictions and empirical findings.

	AlAn w_1**=3**	**Econ** w_2**=2**	**Dist** w_3**=2**	**AlOr** w_4**=1**	***1/2** w_5**=?**	***S**$_{DAT}$ w_6**=?**	H	**RT**$_{Verb}$
(Ia) A_{NOM} V_{FUT} [P_{DAT}]		**					6	7.11
(Ib) U_{NOM} V_{NACT}	**	*					8	7.15
(Id) P_{NOM} V_{AOR} [$A_{1/2}$]	**	**			*		10+w_5	7.25
(Ic) U_{NOM} $V_{NACT.AP}$[G_{DAT}]	**	**	*				12	7.30
(IIa) A_{NOM} P_{DAT} V_{FUT}	**	**					10	7.16
(IIc) P_{NOM} A_{DAT} V_{PERF}	**	**		*		*	11+w_6	7.26
(IIb) A_{NOM} G_{DAT} $V_{FUT.AP}$ [P_{DAT}]	*	***	**				13	7.32
(IIb) A_{NOM} P_{DAT} $V_{FUT.AP}$ [G_{DAT}]	**	***	**				16	
(IId) P_{NOM} G_{DAT} $V_{AOR.AP}$ [$A_{1/2}$]	***	***	**	*	*		20+w_5	7.48
(IIIa) A_{DAT} P_{NOM} V_{PERF}	**	**				*	10+w_6	7.16
(IIId) P_{DAT} A_{NOM} V_{FUT}	**	**		*			11	7.18
(IIIb) G_{DAT} A_{NOM} $V_{FUT.AP}$ [P_{DAT}]	*	***	**	*			14	7.34
(IIIb) P_{DAT} A_{NOM} $V_{FUT.AP}$ [G_{DAT}]	**	***	**	*			17	
(IIIc) G_{DAT} P_{NOM} $V_{AOR.AP}$ [$A_{1/2}$]	***	***	**		*		19+w_5	7.47

Table 20: HG tableau for parses disambiguated by a clause medial/final verb, with corresponding RTs from Experiments I–III (in log ms). Constraints are evaluated relative to parses compatible with the sentence up to the verb; yet-unencountered arguments entailed by verbal morphology are in brackets. Shaded columns give constraint evaluations for two speculative constraints; see text.

The first finding is remarkable because all incrementally ambiguous arguments in these experiments referred to humans; if, as Foley (2020) predicted, comprehenders prioritize the harmonic alignment of animacy and role scales, then a high-animacy G_{DAT} (having intermediate thematic prominence) should be preferable to a high-animacy P_{DAT} (have low thematic prominence). The fact that evidence of applicativization instead impedes processing suggests the relative importance of Economy (a constraint favoring as few arguments as possible) or Distinctness (disfavoring any multivalent parse with a medium-prominence argument). Something like AlignAnimacy still seems important to explain Foley's

(2020) inanimate-ergative costs — perhaps the constraint penalizes low-prominence Proto-Agents more than it does high-prominence Proto-Patients. Crosslinguistic processing evidence indeed suggests such asymmetry: inanimate transitive subjects routinely challenge comprehenders more than animate direct objects do (Bornkessel-Schlesewky & Schlesewsky 2009b). Moreover, the second finding — no significant processing difference between $U_{NOM}V_{NACT}$ and $A_{NOM}V_{FUT}$ verbs in Experiment I — also suggests that human Proto-Patients are not particularly penalized, at least not not enough to outweigh the extra Economy violation incurred by the active transitive.

The third finding indicates some processing cost associated with $P_{NOM}V_{AOR}$ (Id), $P_{NOM}G_{DAT}V_{AOR}$ (IId), and $G_{DAT}P_{NOM}V_{AOR}$ (IIIc) verbs. This effect cannot be confidently attributed to the role disambiguation of preverbal arguments per se, given a design confound: unlike any other stimuli in Experiments I–III, these conditions all had first- or second-person null pronominal subjects, indicated by verbal agreement. Perhaps processing this agreement morphology, or accommodating a speech-act participant argument, is in itself taxing. (That speculative processing cost is expressed in Table 20 with the markedness constraint "*1/2".) Were it not for this confound, the role of Economy could be better isolated: $P_{NOM}V_{FUT}$ and $U_{NOM}V_{NACT}$ verbs both assign a Proto-Patient role to their nominative argument, but only the former parse also has a Proto-Agent argument. Future studies could very easily change these conditions to have all third-person arguments, perhaps including an overt subject in postverbal position. The role Georgian's complex agreement morphology plays in sentence processing is its own can of worms deserving systematic research.

The fourth finding is surprising, given presupposed constraints contributing to parse-harmony (Section 3). Post-hoc analysis finds no compelling evidence that either OSV order or A_{DAT}/P_{NOM} mapping is difficult to process, yet the factors together seem to combine superadditively in the $P_{NOM}A_{DAT}V_{PERF}$ condition (IIc). This contrasts with Skopeteas et al.'s (2011) findings: measuring response latencies of binary acceptability judgements, they find a main effect of case mapping for active transitives (with A_{DAT}/P_{NOM} verbs endorsed more slowly than A_{NOM}/P_{DAT} verbs), and a OSV cost for $Exp_{DAT}/Stim_{NOM}$ verbs (i.e., psych predicates in any tense) but not A_{DAT}/P_{NOM} verbs (i.e., canonical transitives in Series III). Skopeteas et al. suggest an independent processing cost associated with dative subjects (their "Locality" constraint), represented in Table 20 as "$*S_{DAT}$". However, the relatively fast recognition times of $A_{DAT}P_{NOM}V_{PERF}$ verbs (IIId) means this cannot be the whole story. Something like constraint conjunction (Shih 2017) is descriptively necessary, with the constraint "$*S_{DAT}$&AlignOrder" having a much higher weight than either $*S_{DAT}$ or AlignOrder.

I offer two possible interpretations of a conjoined constraint like this. First, it could be that initial nominative arguments ($N1_{NOM}$ of Experiment II) simply evoke stronger expectations for an A parse than initial dative arguments do ($N1_{DAT}$ of Experiment III). This might reflect an asymmetry in the distribution of nominative and dative case in the language: A_{NOM} is found in the six tenses of Series I (present, imperfect, present subjunctive, future, conditional, and future subjunctive), but A_{DAT} is only found in the two tenses of Series III (perfect and pluperfect). Future corpus research and Cloze studies should verify the relative frequency of A_{NOM} and A_{DAT} in Georgian, but it could be that a theory which incorporates parse frequency more adequately models case–role processing.

Another possibility is that $P_{NOM}A_{DAT}V_{PERF}$ parses are difficult to process due to the interaction of semantic–pragmatic factors. Recall that Series III tenses, especially the perfect, have robust inferential evidential uses. According to Skopeteas et al. (2011), the evidential in Georgian is particularly felicitous when the A argument is pragmatically backgrounded. It has also been observed that focused constituents prefer the immediately preverbal position in the language (Skopeteas et al. 2009). If comprehenders in this study accommodated OSV order by interpreting the subject as focal, and accommodated evidential semantics of Series III verbs by pragmatically backgrounding A, and if focus and this notion of pragmatic backgrounding are infelicitous in combination, then this could explain the superadditive processing penalty for $P_{NOM}A_{DAT}V_{PERF}$ verbs. Future studies could provide context sentences which facilitate scrambled word orders or evidential readings.

6. Conclusion

The present study, building on Skopeteas et al. (2011), was designed to better understand how comprehenders of a verb-final split-ergative language process morphosyntactic ambiguity. Incremental processing behavior evidences clear costs associated with grammatical features lowering the clause's prototypical Transitivity (Hopper & Thompson 1980). These results contribute to a growing comparative literature in sentence processing (e.g., Wagers et al. 2015, Keshev & Meltzer-Asscher 2017, Pizarro-Guevara & Wagers 2020, Bhatia & Dillon 2022, Hammerly et al. 2022, Tollan & Heller 2022, Fuchs 2023; see Polinsky 2023 and Sauppe et al. 2023 for overviews), and in many ways they validate crosslinguistic processing predictions of the extended Argument Dependency Model (eADM; Bornkessel-Schlesewsky & Schlesewsky 2006, 2009a, 2009b, 2014, 2016).

We conclude by returning to the typological observation we began with: verb-final word orders are very frequent in the world's languages, and frequently associated with case marking. Georgian's complex split-ergative alignment system is prima facie evidence that case need not streamline incremental identification of preverbal arguments' grammatical roles; comprehension pressures do not seem to be a strong force driving the development of case systems (Foley 2022; cf. Bickel et al. 2015). The way that Georgian comprehenders have adapted to their non-optimized grammar is highly revealing: the general strategy for ambiguous arguments seems to be "if nominative, posit subject; if dative, posit direct object". This suggests that prototypically Transitive (Hopper & Thompson 1980) parses, with arguments that are highly distinct (Bornkessel-Schlesewsky & Schlesewsky 2006), are the ones privileged during comprehension — rather than parses with harmonically aligned scales.

References

Aissen, Judith. 1999. Markedness and subject choice in Optimality Theory. *Natural Language and Linguistic Theory*, 17: 673–711.

Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory*, 21: 435–483.

Aronson, Howard I. 1990. *Georgian: A Reading Grammar*. Slavica Publishers.

Bader, Markus and Michael Meng. 1999. Subject-object ambiguities in German embedded clauses: An across-the-board comparison. *Journal of Psycholinguistic Research*, 28. 121–143.

Barr, Dale, Roger Levy, Christoph Sheepers, and Harry Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3): 255–278. doi:10.1016/j.jml.2012.11.001.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using *lme4*. *Journal of Statistical Software*, 67(1): 1–48. DOI: https://doi.org/10.18637/jss.v067.i01

Belletti, Adriana, and Luigi Rizzi. 1988. Psych verbs and Θ-Theory. *Natural Language and Linguistic Theory*, 6(3): 291–352.

Bhatia, Sakshi, and Brian Dillon. 2022. Processing agreement in Hindi: When agreement feeds attraction. *Journal of Memory and Language*, 125: 104322. DOI: https://doi.org/10.1016/j.jml.2022.104322

Bickel, Balthasar, AlenaWitzlack-Makarevich, Kamal Choudhary, Matthias Schlesewsky, and Ina Bornkessel-Schlesewsky. 2015. The neurophysiology of language processing shapes the evolution of grammar: Evidence from case marking. *PLoS ONE*, 10(8): e0132819.

Boeder, Winfried. 2005. The South Caucasian languages. *Lingua*, 155: 5–89.

Bornkessel, Ina, Brian McElree, Matthias Schlesewsky, and Angela Friederici. 2004. Multi-dimensional contributions to garden path strength: Dissociating phrase structure from case marking. *Journal of Memory and Language*, 51. 495–522.

Bornkessel, Ina, and Matthias Schlesewsky. 2006. The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review*, 113(4): 787–821.

Bornkessel, Ina, Stefan Zysset, Angela Friederici, D. Yves von Cramon, and Matthias Schlesewsky. 2005. Who did what to whom? The neural basis of argument hierarchies during language comprehension. *NeuroImage*, 26. 221–233.

Bornkessel-Schlesewsky, Ina and Matthias Schlesewsky. 2009a. *Processing syntax and morphology: A neurocognitive perspective*. Oxford University Press.

Bornkessel-Schlesewsky, Ina and Matthias Schlesewsky. 2009b. The role of prominence information in the real time comprehension of transitive constructions: A cross-linguistic approach. *Language and Linguistics Compass*, 3. 19–58.

Bornkessel-Schlesewsky, Ina and Matthias Schlesewsky. 2014. Scales in real-time language comprehension: A review. In *Scales and Hierarchies: A cross-disciplinary perspective*, eds. Ina Bornkessel-Schlesewsky, Andrej Malchukov, and Marc Richards. Trends in Linguistics: Studies and Monographs, 277. De Gruyter Mouton. 321–352.

Bornkessel-Schlesewsky, Ina and Matthias Schlesewsky. 2016. The Argument Dependency Model. In *Neurobiology of Language*, eds. Gregory Hickok and Steven L. Small. Academic Press. 357–369.

Bosse, Solveig, Benjamin Bruening, and Masahiro Yamada. 2012. Affected experiencers. *Natural Language and Linguistic Theory*, 30. 1185–1230. DOI: 10.1007/s11049-012-9177-1

Boyce, Veronica, Richard Futrell, and Roger Levy. 2020. Maze made easy: Better and easier measurement of incremental processing difficulty. *Journal of Memory and Language*, 111: 104082. DOI: https://doi.org/10.1016/j.jml.2019.104082

Clemens, Laura 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.

Cherchi, Marcello. 1997. *Modern Georgian Morphosyntax: A grammatico-categorial hierarchy-based analysis with special reference to indirect verbs and passives of state*. Harrassowitz Verlag.

Choudhary, Kamal K. et al. 2010. An Actor-preference in a split-ergative language: Electrophysiological evidence from Hindi. Poster presented at the 23rd Annual Meeting of the CUNY Conference on Human Sentence Processing.

de Vincenzi, Marica. 1991. *Syntactic Parsing Strategies in Italian*. Dordrecht: Kluwer.

Demiral, Şükrü Barış, Matthias Schlesewsky, and Ina Bornkessel-Schlesewky. 2008. On the universality of language comprehension strategies: Evidence from Turkish. *Cognition* 106: 484–500.

Dixon, R. M. W. 1979. Ergativity. *Language*, 55. 59–138.

Dowty, David. Thematic proto-roles and argument selection. *Language*, 67(3): 547–619.

Foley, Steven. 2020. Case, agreement, and sentence processing in Georgian. Dissertation, University of California Santa Cruz.

Foley, Steven. 2022. Quantifying performance constraints on case-alignment typology with information theory. Poster presented at the 35th Annual Conference on Human Sentence Processing (HSP 2022). University of California Santa Cruz, Santa Cruz, CA. March 24–26.

Freedman, Sandra, and Kenneth Forster. 1985. The psychological status of overgenerated sentences. *Cognition*, 19: 101–131.

Fuchs, Zuzanna. 2023. Processing of grammatical gender agreement morphemes in Polish: Evidence from the Visual World Paradigm. *Morphology*. https://doi.org/10.1007/s11525-023-09418-3

Gippert, Jost, and Manana Tandashvili. 2015. Structuring a diachronic corpus: The Georgian National Corpus project. In *Historical Corpora: Challenges and perspectives*, eds. Jost Gippert and Ralf Gehrke. Narr Verlag. 305–322.

Gérardin, Hélène. 2016. Les verbes intransitifs primaires et dérivés en Géorgien: Description morphosyntaxique, sémantique, et dérivationnelle [Primary and derived intransitive verbs in Georgian: Morphosyntactic, semantic, and derivational description]. Dissertation, Institut National des Langues et Civilisations Orientales (Inalco).

Greenberg, Joseph. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of Language*, ed. Joseph Greenberg. MIT Press. 73–113.

Hale, John. 2001. A probabilistic Earley parsey as a psycholinguistic model. In the *Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics. 159–166.

Hammerly, Christopher, Adrian Staub, and Brian Dillon. 2022. Person-based prominence guides incremental interpretation: Evidence from obviation in Ojibwe. *Cognition*, 225: 105122. https://doi.org/10.1016/j.cognition.2022.105122

Harris, Alice. 1981. *Georgian Syntax: A study in Relational Grammar.* Cambridge University Press.

Harris, Alice. 1985. *Diachronic Syntax: The Kartvelian Case.* Academic Press.

Harris, Alice (ed). 1991. *The Indigenous Languages of the Caucasus, Vol. 1: The Kartvelian Languages*. Caravan Books.

Haspelmath, Martin. 2021. Explaining grammatical coding asymmetries: Form–frequency correspondences and predictability. *Journal of Linguistics*, 57. 605–633.

Hemforth, Barbara, Lars Konieczny, and Gerhard Strube. 1993. Incremental syntax processing and parsing strategies. In the *Proceedings of the 15th Annual Conference of the Cognitive Science Society*. Erlbaum. 539–545.

Hewitt, B. G. 1995. *Georgian: A Structural Reference Grammar*. John Benjamins Publishing Company.

Hoeks, John, and Petra Hendriks. 2011. Optimality Theory and human sentence processing: Towards a cross-modular analysis of coordination. *Journal of Cognitive Science*, 12: 83–128.

Holisky, Dee Ann. 1981. *Aspect and Georgian Medial Verbs*. Academic Resources Corp.

Hopper, Paul, and Sandra Thompson. 1980. Transitivity in grammar and discourse. *Language*, 56(2): 251–299.

Inoue, Atsu, and Janet Dean Fodor. 1995. Information-paced parsing of Japanese. In *Japanese Sentence Processing*, eds. Reiko Mazuka and Noriko Nagai. Psychology Press. 19–48.

Kalin, Laura. 2018. Licensing and differential object marking: The view from Neo-Aramaic. *Syntax*, 21(2): 112–159.

Keshev, Maayan, and Aya Meltzer-Asscher. 2017. Active dependency formation in islands: How grammatical resumption affects sentence processing. *Language*, 93(3): 549–568.

Lau, Ellen, Rusudan Asatiani, Nancy Clarke, Michaela Socolof, and Maria Polinsky. 2023. A subject relative clause preference in a split-ergative language: ERP evidence from Georgian. *Brain and Language 2022*, 236.

Legendre, Géraldine, Yoshiro Miyata, and Paul Smolensky. 1990. Harmonic Grammar: A formal multi-level connectionist theory of linguistic well-formedness: an application. In *Proceedings of the 12th Annual Conference of the Cognitive Science Society*. Erlbaum. 884–891.

Lenth, Russell, Ben Bolker, Paul Buerkner, Iago Giné-Vázquez, Maxime Herve, Maarten Jung, Jonathan Love, Fernando Miguez, Hannes Riebl, and Henrik Signman. 2020. *emmeans*: estimated marginal means, aka least-squares mean. R package version 1.5.2-2.

Levy, Roger. 2008. Expectation-based syntactic comprehension. *Cognition*, 106. 1126–1177.

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.

Miljan, Merilin, Elsi Kaiser, and Virve-Anneli Vihman. 2017. Interplay between case, animacy and number: Interpretations of grammatical role in Estonian. *Finno-Ugric Languages and Linguistics*, 6(1). 55–77.

Muralikrishnan, R., Matthias Schlesewsky, and Ina Bornkessel-Schlesewsky. 2008. Universal and cross-linguistic influences on the processing of word order and animacy: Neurophysiological evidence from Tamil. Poster presented at the 21st Annual CUNY Conference on Human Sentence Processing.

Nash, Léa. 2017. 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. 175–200.

Nash, Léa. 2022. Nonunitary structure of unergative verbs in Georgian. *Natural Language & Linguistic Theory*, 40. 1201–1267

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*, 205: 112–133.

Pizarro-Guevara, Jed and Matt Wagers. 2020. The predictive value of Tagalog voice morphology in filler–gap dependency formation. *Frontiers in Psychology*, 11(517). DOI:10.3389/fpsyg.2020.00517

Polinsky, Maria. 2023. Experimental syntax and linguistic fieldwork. In *The Oxford Handbook of Experimental Syntax*, ed. John Sprouse. Oxford University Press. 97–126.

Polinsky, Maria, Carlos Gómez Gallo, Peter Graff, and Ekaterina Kravtchenko. 2012. Subject preference and ergativity. *Lingua*, 122. 267–277.

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.

Sauppe, Sebastian, Caroline Andrews, and Elisabeth Norcliffe. 2023. Experimental research in cross-linguistic psycholinguistics. In *Routledge Handbook of Experimental Linguistics*, eds. Sandrine Zufferey and Pascal Gygax. Routledge. 156–172. doi: 10.4324/9781003392972-12

Schriefers, Herbert, Angela Friederici, and Katja Kühn. 1995. The processing of locally ambiguous relative clauses in German. *Journal of Memory and Language*, 34. 499–520.

Shanidze, Akaki. 1980. Kartuli enis gramat'ik'is sapudzvlebi [Fundamentals of Georgian grammar]. In *Txzulebani* [Collected Works], vol. 3. Scientific Academy, Tbilisi State University.

Shih, Stephanie. 2017. Constraint conjunction in weighted probabilistic grammar. *Phonology*, 24: 243–268. doi:10.1017/S0952675717000136

Skopeteas, Stavros, Gisbert Fanselow, and Rusudan Asatiani. 2011. 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. Studies in Theoretical Psycholinguistics, 40. Springer. 145–172.

Skopeteas, Stavros, Caroline Féry, and Rusudan Asatiani. 2009. Word order and intonation in Georgian. *Lingua*, 119. 102–127.

Smolensky, Paul, and Géraldine Legendre (eds.) 2006. *The Harmonic Mind: From neural computation to optimality-theoretic grammar*. 2 vols. MIT Press.

Stewart, Andrew, Judith Holler, and Evan Kidd. 2007. Shallow processing of ambiguous pronouns: Evidence for delay. *The Quarterly Journal of Experimental Psychology*, 60(12): 1680–1696.

Testelets, Yakov. 2021. Kartvelian (South Caucasian) Languages. In *The Oxford Handbook of Languages of the Caucasus*, ed. Maria Polinsky. 490–528.

Tollan, Rebecca and Daphna Heller. 2022. Pronoun resolution and ergativity: Effects of subjecthood and case in Niuean. *Language*, 98(1): 157–183.

Tollan, Rebecca, Diane Massam, and Daphna Heller. Effects of case and transitivity on processing dependencies: Evidence from Niuean. *Cognitive Science*, 46(3). 1–14. DOI: 10.111/cogs.12736.

Tuite, Kevin. 1998. Kartvelian morphosyntax: Number agreement and morphosyntactic orientation in the south caucasian languages. Dissertation, University of Montreal.

Villata, Sandra, Whitney Tabor, and Julie Franck. 2018. Encoding and Retrieval Interference in Sentence Comprehension: Evidence from Agreement. *Frontiers in Psychology*, 19. https://doi.org/10.3389/fpsyg.2018.00002

Wagers, Matt, Manuel Borja, and Sandra Chung. 2015. The real-time comprehension of *wh*-dependencies in a *Wh*-Agreement language. *Language*, 91(1): 109–144.

Wang, Luming, Matthias Schlesewsky, Balthasar Bickel, and Ina Bornkessel-Schlesewsky. 2009 Exploring the nature of the 'subject'-preference: Evidence from the online comprehension of simple sentences in Mandarin Chinese. *Language and Cognitive Processes*, 24(7–8). 1180-1226. DOI: 10.1080/01690960802159937

Yasunaga, Daichi, Masataka Yano, Yoshiho Yasugi, and Masatoshi Koizumi. 2015. Is the subject-before-object preference universal? An event-related potential study in the Kaqchikel Mayan language. *Language, Cognition and Neuroscience*, 30(9): 1209–1229. DOI: 10.1080/23273798.2015.1080372

Zehr, Jeremy, and Florian Schwarz. 2018. PennController for Internet Based Experiments (IBEX). https://doi.org/10.17605/OSF.IO/MD832