

Constraining long-distance *wh*-movement: the case of Georgian*

Tanya Bondarenko (Harvard University)

October 10, 2025 • SCCC4

1 Introduction

- **Question:** Why do we see variation in availability of long-distance *wh*-movement?
- E.g., why is it available from the complement of *think* in English, but for many Georgian speakers a similar sentence with *pikrobs* is ungrammatical?

- (1) **Who**₁ did Mariam think that this task upset **t**₁? *English*
- (2) ***vin**₁ ipikra mariam-ma, [rom am davaleba-m **t**₁ gaanac'qena]? *Georgian*
who.NOM thought Mariam-ERG [ROM this task-ERG upset] (some speakers)
Intended: ‘Who did Mariam think that this task upset?’

- **Today’s answer:** the variation comes from (at least) the following differences:
 - the size of the embedded clause that the *wh*-phrase vacates;
 - the height of the landing site of the *wh*-phrase;
 - the presence and overtness of the nominal structure on the embedded clause;
- **Preview of the proposal:**
 - constraints like (2) are the so-called Williams Cycle effect, which can be accounted by the Level Embedding Approach (Meadows 2023; Poole 2023; Williams 2003, 2011).
 - some cases of impossible long-distance *wh*-extraction are due to the case-matching requirement (Berulava 2025b)—movement becomes available once it’s met;
 - nominalized clauses with an overt demonstrative *is* are islands for extraction due to how properties of D interact with linearization.

*I thank Tamari Berulava, Jacob Fernandes, Kaden Holladay, Richard Luo, and the WOLF-lab at Harvard for productive discussions about various aspects of this topic. I am extremely grateful to our Georgian consultants in Tbilisi and in Boston for their intuitions and insightful comments. All errors are my own.

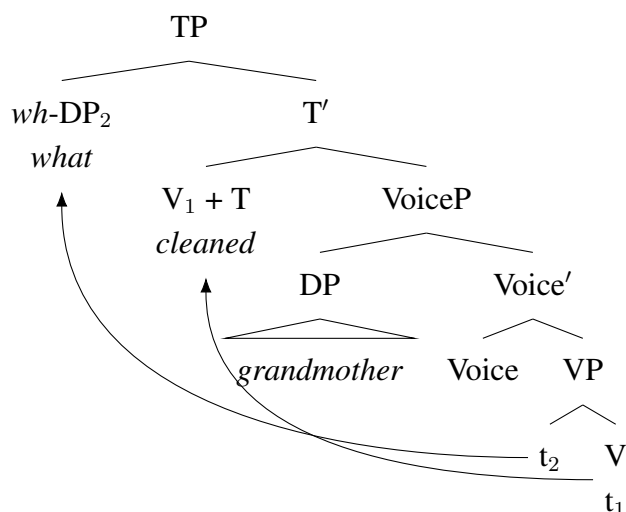
2 Georgian *Wh*-Movement

- Following Borise (2023), I will assume that Georgian VP is head final, but higher projections are head initial, and there's optional V-to-Voice movement which derives SVO.
- *Wh*-words in Georgian simple clauses must occur in an immediately pre-verbal position (Borise 2023; Harris 1981). Other material can precede or follow them quite freely.
- E.g., while SOV, SVO, OSV and OVS are generally available word orders in Georgian, only SOV and OVS are possible when the DO is a *wh* phrase, (3).

- (3) a. ✓ SOV
 bebia ras
 grandma.NOM what.ACC
 alagebda?
 clean.IPFT.3SG
 'What did grandma clean?'
 (Borise 2023: 184)
- b. ✗ SVO
 *bebia alagebda
 grandma.NOM clean.IPFT.3SG
 ras?
 what.ACC
 'What did grandma clean?'
 (Borise 2023: 185)
- c. ✗ OSV
 *ras bebia
 what.ACC grandma.NOM
 alagebda?
 clean.IPFT.3SG
 'What did grandma clean?'
 (Borise 2023: 184)
- d. ✓ OVS
 ras alagebda
 what.ACC clean.IPFT.3SG
 bebia?
 grandma.NOM
 'What did grandma clean?'

- I adopt Borise (2023)'s proposal that *wh*-movement in Georgian targets a position lower than CP and higher than VoiceP. I will assume that the relevant projection is TP.

- (4) Structure of the *wh*-question (3d)



3 Georgian Embedded Clauses

3.1 External syntax

- Georgian has two main strategies of clausal embedding:
 - (i) finite complements with the complementizer *rom*; (ii) nominalizations (*masdars*).
- In (Bondarenko *et al.* 2025) we've proposed an analysis of the external syntax of *rom*-clauses that I will adopt in this talk.
- **Key claim:** some CPs are base-generated within a nominal structure (DP shell), others are not.
- The argument structure of the verb determines which structure is available:
 - transitive verbs (class I) always embed CPs within a DP shell.
 - intransitive verbs (class II) never embed CPs within a DP shell.

(5) Demonstrative-on-CP Generalization

The demonstrative *is* 'that' can occur on the embedded *rom*-clause iff the verb can combine with DPs, and the interpretation that CP receives is the same that a DP would receive.

- Cf. verbs *gaipikrebs* (class I) and *šepikriandeba* (class II) 'will think', both of which are derived from the root *pikr* 'think', but differ in transitivity.

(6) nino [am idea-s]_{DP} ga-i-pikr-eb-s /*še-pikr-ian-d-eb-a.
Nino.NOM this idea-ACC PVB-PV-think-TS-PRS.3SG /PVB-think-IAN-D-TS-PRS.3SG
'Nino will think this idea.'

(7) nino ga-i-pikr-eb-s /še-pikr-ian-d-eb-a,
Nino.NOM PVB-PV-think-TS-PRS.3SG /PVB-think-IAN-D-TS-PRS.3SG
[rom šota-m moigo čemp'ionat'i].
ROM Shota-ERG won competition
'Nino will think that Shota won the competition.'

(8) nino ga-i-pikr-eb-s /*še-pikr-ian-d-eb-a,
Nino.NOM PVB-PV-think-TS-PRS.3SG /PVB-think-IAN-D-TS-PRS.3SG
[imas rom šota-m moigo čemp'ionat'i].
DEM.ACC ROM Shota-ERG won competition
'Nino will think that Shota won the competition.'

- The data in (6)-(8) could be thought of in two ways:
 - **Demonstrative = Nominalization:** both *gaipikrebs* and *šepikriandeba* can combine with bare CPs, only *gaipikrebs* can combine with nominalized CPs.
 - **Demonstrative ≠ Nominalization; Optional D:** *gaipikrebs* only combines with DPs and nominalized CPs, *šepikriandeba* only combines with bare CPs. Dem is optional/droppable.

- We suggest that the second view (*Demonstrative ≠ Nominalization, Optional D*) is correct.
- (Bondarenko 2022): cross-linguistically, nominalized CPs always have the same interpretations as DPs, bare CPs never do—they describe content associated with the attitudinal/speech event.
- We can see from verbs like ‘deny’ that clauses with transitive verbs in Georgian must have the same interpretation as DPs→*nominalized structure*. “Content” readings are impossible.

(9) šota-m uarqo [es azri].
 Shota-ERG denied this thought.NOM
 ‘Shota denied this thought.’

(10) **Context for the Content Reading:**

Someone claimed that Shota /the speaker broke into the house and robbed it. At the trial, Shota denies this allegation. He says: I am /the speaker is innocent.

(11) šota-m (is) uarqo [rom me var udanašaulo]
 Shota-ERG (DEM.NOM) denied ROM I am innocent

Theme Reading: OK ‘Shota denied that I (the speaker) am innocent.’

Content Reading: *‘Shota denied some claim by saying “I am/the speaker is innocent”.’

Bare CPs (intransitive verbs)

- (Bondarenko *et al.* 2025): bare CPs are base-generated in the complement of the verb, and then extrapose to the right in order to obey the Final-over-Final (FOFC) constraint (Biberauer, Holmberg, *et al.* 2009; Biberauer, Newton, *et al.* 2009; Biberauer and Sheehan 2012).

(12) **Final-Over-Final Constraint (FOFC)**

“Head-final heads can’t take head-initial complements”

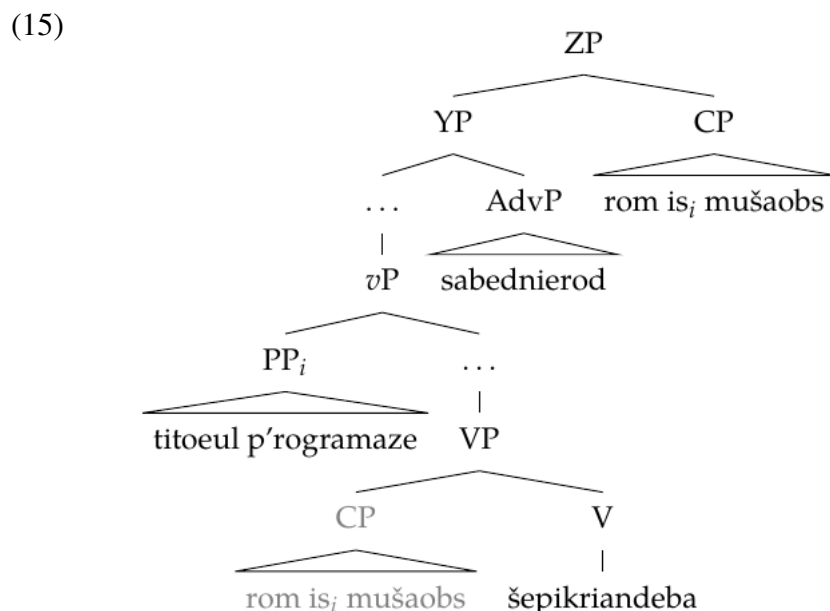
Within a clausal or nominal extended projection, if a phase head bears $\wedge$, then all the categorially alike heads in its complement domain (i.e., those making up the “spine” of the projection in question) must have $\wedge$. (Biberauer, Newton, *et al.* 2009: p. 711)

- Bare clauses (as well as clauses without overt demonstratives) cannot remain in situ:

(13) *nino [rom šota-m moigo čemp’ionat’i] še-pikr-ian-d-eb-a
 Nino.NOM ROM Shota-ERG won competition PVB-think-IAN-D-TS-PRS.3SG
 /ga-i-pikr-eb-s.
 /PVB-PV-think-TS-PRS.3SG
 Intended: ‘Nino will think that Shota won the competition.’

- The clauses must extrapose, and (i) they extrapose to a position above high adverbs in the matrix clause; (ii) they reconstruct for the purposes of binding, (14).

- (14) mariami [titoeul p'rograma-ze]₁ še-pikr-ian-d-eb-a
 Mariam.NOM each program PVB-think-IAN-D-TS-PRS.3SG
 /ga-i-pikr-eb-s sabednierod [rom is₁ muša-ob-s].
 /PVB-PV-think-TS-PRS.3SG fortunately ROM it work-TS-PRS.3SG
 'Mariam will fortunately think about [every program]₁ that it₁ works.'



Nominalized CPs (transitive verbs)

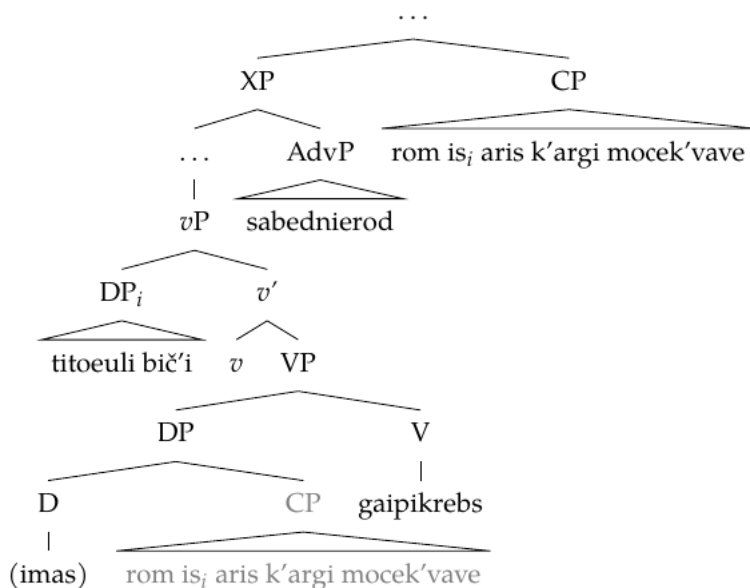
- (Bondarenko *et al.* 2025): nominalized CPs are CPs in a DP shell, where the DP is base-generated as the complement of the verb.
- The CP in these structures obligatory extraposes to the right if the demonstrative is omitted, as we can see from ungrammaticality of *gaipikrebs* in (13). Examples like (14) with *gaipikrebs* suggest that such CPs extrapose high, and reconstruct low, just like CPs in transitive structures.
- If the demonstrative is overt, the CP can remain in its base-generated position:

- (16) nino [imas rom šota-m moigo čemp'ionat'i] ga-i-pikr-eb-s.
 Nino.NOM DEM.ACC ROM Shota-ERG won competition PVB-PV-think-TS-PRS.3SG
 'Nino will think that Shota won the competition.'

- **Implication:** FOFC really only “counts” overt heads. When the demonstrative is phonologically vacuous, it doesn’t count, and extraposition is needed to avoid FOFC. When the demonstrative is overt, it creates a barrier and FOFC violation is not incurred.
- While it doesn’t have to, bare CP can undergo rightwards extraposition even when the demonstrative is overt. It targets a high position, and reconstructs low, (17)-(18).

- (17) [titoeuli bič'i]₁ (**imas**) ga-i-pikr-eb-s sabednierod,
 each boy (**that.ACC**) PVB-PV-think-TS-PRS.3SG fortunately,
 [rom is₁ aris k'argi mocek'vave]
 ROM 3SG is good dancer
 '[Every boy]₁ will fortunately think that he₁ is a good dancer.'

(18)



3.2 Internal Syntax

- **Q:** ‘What’s the size of the embedded clauses?’
 ≈ ‘What’s the amount of *verbal projections* in the embedded clauses?’
- **Proposal:** there are (at least) 4 different sizes of embedded clauses (table 1).
- **NB#1:** For some speakers, the same verb might be able to take several sizes of complements.
- **NB#2:** There is variation among speakers as to which verbs take which size(s) of complements.

Properties	<i>rom</i> -CPs	<i>rom</i> -TPs	<i>rom</i> -ModPs	VoiceP-NMNs
“high” adverbs	✓	✗	✗	✗
unrestricted TAM	✓	✓	✗	✗
TAM present	✓	✓	✓	✗
“low” adverbs	✓	✓	✓	✓

Table 1: The sizes of embedded clauses.

- **Implication: presence of *rom* ≠ size!** Not all *rom*-clauses are full CPs.
 Cf. *hogy*-clauses in Hungarian (Egressy 2025), *ki*-clauses in Hindi-Urdu (Manetta 2011).

Examples of Verbs

- With intransitive verbs (e.g., *šepikriandeba* ‘think’, *aris bednieri* ‘be happy’), clauses are always CPs (see Bondarenko 2022, ch. 4, for a semantic account of this fact and the exceptions).
- With transitive verbs, clauses come in different sizes; a given verb can be compatible with more than one “size”. In our data, all TP-selecting verbs can take CPs too.

(19) Examples of verbs which can take:

- CPs:** *pikrobs*, *gaipikrebs*, *šepikriandeba* ‘think’, *hgonia* ‘think’, *ečvobs* ‘suspect’, *sjera* ‘believe’, *ambobs* ‘say’, *čaičurčulebs* ‘whisper’, *k’amatobs* ‘argue’, *dainaxavs* ‘see’, *aris bednieri* ‘be happy’;
- TPs:** for all: *hgonia* ‘think’, *sjera* ‘believe’, for some speakers: *pikrobs* ‘think’, *ečvobs* ‘suspect’, *gaigonebs*, *gaigebbs* ‘hear’, *dainaxavs* ‘see’;
- ModPs:** *unda* ‘want’, *scada* ‘try’, *ganizraxa* ‘intend’, *gadac’qvet’s* ‘decide’;
- VoicePs:** *daic’qebbs* ‘begin’, *daasrulebs* ‘finish’, *šec’qvet’s* ‘stop’.

High Adverbs

- “High” adverbs like *sabednierod* ‘fortunately’, *savaraudod* ‘allegedly’, *albat* ‘probably’ can appear in embedded clauses and take embedded scope only with verbs that can take CPs:

- (20) mariam-ma tkva [rom **albat** nene-m daxat’a st’udent’i] CP only
 Mariam-ERG said ROM **probably** Nene-ERG painted student.NOM
 ‘Mariam said that probably, Nene painted the student.’
- (21) mariam-s hgonia [rom **albat** nene-m daxat’a st’udent’i] CP ∨ TP
 Mariam-DAT thinks ROM **probably** Nene-ERG painted student.NOM
 ‘Mariam thinks that probably, Nene painted the student.’
- (22) *mariam-s unda [rom **albat** am davaleba-m shota gaanac’qenos] ModP
 Mariam-DAT wants ROM **probably** this task-ERG Shota.NOM upset.3SG.OPT
 ‘Mariam wants that, probably, this task would upset Shota.’
 CC: not possible to understand “*she wants it to be probable that this task upsets Shota.*”
- (23) *luk’a-m daic’qo [ğob-is **albat** a-šen-eb-a]. VoiceP
 Luka-ERG started fence-GEN **probably** PVB-build-TS-NMN.NOM
 Intended: ‘Luka started probably building a fence.’
 (Cannot mean: initiated the state of affairs that building a fence is probable).

- The difference between CP only verbs and CP ∨ TP verbs arises when we look at how *wh*-extraction interacts with the presence of “high adverbs”.
- Unlike CP only verbs, which prohibit *wh*-extraction, (24), CP ∨ TP verbs are verbs that allow *wh*-extraction on its own, and “high adverbs” on its own, but don’t allow them together, (25).

- (24) *vis ambobs mariami [rom xat'avš šota]?
 who.ACC say.PRS.3SG Mariam.NOM ROM draw.PRS.3SG Shota.NOM
 'Who does Mariam say that Shota is painting?' (*judgements: from * to 2.5 /3 out of 5*)
- (25) a. vin hgonia mariam-s [rom am davaleba-m gaabraza]?
 who.NOM thinks Mariam-DAT ROM this task-ERG upset
 'Who does Mariam think that this task upset?'
 b. *vin hgonia mariam-s [rom albat am davaleba-m gaabraza]?
 who.NOM thinks Mariam-DAT ROM **probably** this task-ERG upset
 Intended: 'Who does Mariam think that probably this task upset?'

- **Hypothesis:** *wh*-extraction is possible out of TP, but not CP complements.
- Verbs like *ambobs* 'say' can take only CPs → "high adverbs" good, *wh*-extraction bad.
- Verbs like *hgonia* 'think' can take both CPs and TPs → "high adverbs" good, *wh*-extraction good, but they cannot co-occur in one sentence.

Tense–Aspect–Modality

- Only verbs that take CPs and/or TPs are unrestricted in the T(enst)-A(spect)-M(odality) of the embedded clause. Clauses of smaller sizes show restrictions or lack TAM completely.
- Verbs that combine with CPs only, or with CPs and TPs, are free to have any TAM forms in their complements (provided they are compatible with their semantics), just like unembedded clauses:

- (26) *CP only, CP ∨ TP verbs*
- a. mariam-s u-tkv-am-s /hgonia
 Mariam-DAT pv-say-TS-PERF.3SG /think.PRS.3SG
 [rom šota-m c'erili da-c'er-a].
 [ROM Shota-ERG letter.NOM PVB-write-AOR.3SG]
 'Mariam has said/thinks [that Shota wrote a letter].'
- b. mariam-s u-tkv-am-s /hgonia
 Mariam-DAT pv-say-TS-PERF.3SG /think.PRS.3SG
 [rom šota-m c'erili unda da-c'er-os].
 [ROM Shota-ERG letter.NOM must PVB-write-OPT.3SG]
 'Mariam has said/thinks [that Shota should write a letter].'

- ModP clauses on the other hand are restricted: they need modality, e.g. *optative* marking, to be expressed inside of the embedded clause, and so their choice of TAM forms is restricted.

- (27) *ModP verbs*
- mariam-s u-nda
 Mariam-DAT pv-want.PRS.3SG
 [rom šota-m c'erili da-c'er-os /*da-c'er-a].
 [ROM Shota-ERG letter.NOM PVB-write-OPT.3SG /PVB-write-AOR.3SG]
 'Mariam wants Shota to write a letter.'

- Verbs with CP only and CP ∨ TP complements allow tense mismatches between the matrix and the embedded clauses, as evidenced by incompatible matrix and embedded adverbs in (28):

- (28) a. mariam-ma **gušin** ipikra
 Mariam-ERG **yesterday** thought
 [rom šota **xval** mankana-s iqidis]
 [ROM Shota.NOM **tomorrow** car-ACC buy.FUT.3SG]
 ‘Mariam thought yesterday that Shota will buy a car tomorrow.’
- b. mariami **xval** ipikrebs
 Mariam.NOM **tomorrow** think.FUT.3SG
 [rom šota-m **gušin** c’erili dac’era]
 [ROM Shota-ERG **yesterday** letter.NOM write.AOR.3SG]
 ‘Mariam will think tomorrow that Shota wrote a letter yesterday.’

- This is not possible with optative complements of verbs like *unda* ‘want’:

- (29) a. *mariam-s **gušin** undoda
 Mariam-DAT **yesterday** wanted
 [rom šota-m **xval** mankana iqidos]
 [ROM Shota-ERG **tomorrow** car.NOM buy.OPT.3SG]
 Intended: ‘Mariam wanted yesterday that Shota will buy a car tomorrow.’
- b. *mariam-s **xval** endomeba
 Mariam-DAT **tomorrow** will.want
 [rom šota-m **gušin** c’erili da-c’er-os]
 [ROM Shota-ERG **yesterday** letter.NOM PVB-write-OPT.3SG]
 ‘Mariam will want tomorrow that Shota wrote a letter yesterday.’

- **NB:** This mismatch becomes possible if a plurperfect form is used in the embedded clause (*da-e-cer-a* PVB-pv-write-PLURPERF.3SG), so such complements might have a bit more structure to them (perhaps, an *impoverished TP* that’s “smaller” than TP with verbs like *hgonia*).
- Masdars don’t have any ability to showcase temporal distinctions:
 - they do not contain any TAM morphemes¹;
 - they cannot host temporal adverbs in their structure, (30).

- (30) a. *mariam-ma **gušin** [c’eril-is **xval** c’era] daic’qo
 Mariam-ERG **yesterday** letter-GEN **tomorrow** write.NMN.NOM began
 ‘Mariam yesterday started writing the letter tomorrow.’
- b. *mariami **xval** [c’eril-is **gushin** c’era-s] daic’qeb
 Mariam.NOM **tomorrow** letter-GEN **yesterday** write-ACC begin.FUT.3SG
 ‘Mariam will begin tomorrow to write a letter yesterday.’

¹ Besides the aspectual preverbs, which I consider to originate vP-internally and specify lexical aspect (telicity/atelicity), rather than correspond to viewpoint aspect and the TAM part of the clausal spine.

“Low” Adverbs

- Despite their lack of TAM, masdar complements in Georgian admit VoiceP-level adjuncts like *ganzrax* ‘on purpose’ and *jibr-ze* ‘spitefully’, just like bigger complements do, (31)-(32), suggesting the presence of VoiceP in their structure (see also Bondarenko 2017 and Berulava 2025a).²

- (31) luk’a-m tkva /luk’a-s hgonia
Luka-ERG said /Luka-DAT thinks
[rom šota-m c’igni **ganzrax** /**jibr-ze** da-c’er-a].
ROM Shota-ERG book.NOM **on.purpose** /**spite-on** PVB-write-AOR.3SG
‘Luka said/thinks that Shota wrote the book on purpose/spitefully.’
- (32) luk’a-s unda [rom šota-m c’igni **ganzrax** /**jibr-ze** da-c’er-a].
Luka-DAT wants ROM Shota-ERG book.NOM **on.purpose** /**spite-on** PVB-write-AOR.3SG
‘Luka wants Shota to write the book on purpose/spitefully.’
- (33) luk’a-m daic’qo [c’ignis **ganzrax** /**jibr-ze** c’era].
Luka-ERG started c’ign-GEN **on.purpose** /**spite-on** write.NMN.NOM
‘Luka started writing a book on purpose/spitefully.’

- **Licensing of arguments** also distinguishes ModP complements and masdars:
 - in ModP clauses, case licensing is more restricted than in CP and TP clauses (not all existing case alignments are available), which goes hand-in-hand with TAM restrictions, but DPs can receive cases that exist in independent sentences (e.g., ERG and NOM for optative clauses);
 - in masdars, only 1 DP can be expressed without postpositions, and it receives Genitive case, other arguments will become PPs (e.g., *tvīs*-PPs for indirect objects).

Upshot:

- Masdars don’t have any TAM, case/licensing depends on TAM, thereby even though they include VoiceP structure, arguments won’t be able to be expressed as they are in tensed clauses.
- Optative clauses (ModPs) have some TAM, but it’s restricted → they can have case/licensing of arguments, but it is restricted in accordance with the TAM they have.

4 Constraints on long-distance *wh*-movement

- Recall the types of embedded clauses we have:

1. Bare: CPs;
2. Nominalized: CPs, TPs, ModPs, VoicePs x overt/covert D (*is*).

- These parameters matter when it comes to *wh*-extraction.

² That masdars in Georgian are VP-level nominalizations has been proposed in (Wier 2014) and (Berikashvili 2024).

4.1 Extraction from bare CPs

- Extraction from bare CPs is never good, (34)-(35). This is expected on the view that such complement clauses are adjuncts, and are thus subject to the *Adjunct Island* (but see section 5).³

- (34) a. nino še-pikr-ian-d-eb-a [rom mariami šota-s naxavs]
Nino.NOM PVB-think-IAN-D-TS-PRS.3SG ROM Mariam.NOM Shota-ACC sees
‘Nino will think that Mariam sees Shota.’
b. *vis₁ še-pikr-ian-d-eb-a nino [rom mariami t₁ naxavs]?
who.ACC PVB-think-IAN-D-TS-PRS.3SG Nino [ROM Mariam.NOM sees]
Intended: ‘Who will Nino think that Mariam sees?’
- (35) a. mariami bednieri aris, [rom šota nene-s xat’avs]
Mariam.NOM happy is ROM Shota.NOM Nene-ACC draws
‘Mariam was happy that Shota is painting Nene.’
b. *vis₁ aris bednieri mariami, [rom šota t₁ xat’avs]?
who.ACC is happy Mariam.NOM ROM Shota.NOM draws
Intended: ‘Who is Mariam happy that Shota is painting?’

4.2 Extraction from nominalized CPs

- Table 2 summarizes behavior of 4 types of nominalized clauses when it comes to *wh*-extraction.

	<i>rom</i> -CPs	<i>rom</i> -TPs	<i>rom</i> -ModPs	VoiceP-NMNs
∅ (no demonstr.)	✗	✓ (case-matching)	✓ (case-matching?)	✓ (pied-piping)
DEM (<i>is</i> ‘that’)	✗	✗	✗	✗

Table 2: Possibility of *wh*-movement out of nominalized embedded clauses.

Wh-extraction from nominalized CPs

- When a verb combines exclusively with CP-level clauses, *wh*-extraction is ungrammatical, regardless of the presence of the demonstrative:

- (36) *vis ambobs mariami, (is) [rom xat’avs šota]?
who.ACC say.PRS.3SG Mariam.NOM (DEM) ROM draw.PRS.3SG Shota.NOM
‘Who does Mariam say that Shota is painting?’
- (37) *vin čaičurčula mariam-ma, (is) [rom am davaleba-m gaabrazā]?
who.NOM whisper.AOR.3SG Mariam-ERG (DEM) ROM this task-ERG upset.AOR.3SG
‘Who did Mariam whisper that this task upset?’

- Other things that don’t matter*: what kind of a phrase is being extracted (e.g., subject vs. object), what is the case of the moving phrase vs. the case alignment in the matrix clause.

³ Cross-linguistically, not all bare CPs are islands for extraction, as are not all adjuncts in general. See (Bondarenko 2022) for discussion of extraction out of bare CPs cross-linguistically. See (Johnson 2003; Privoznov 2022; Truswell 2007, 2011; Uriagereka 1999), a.o., for theories of extraction from adjuncts.

Wh-extraction from nominalized TPs

- *Wh*-extraction is grammatical provided case-matching is obeyed: the case on the moved phrase is the same as the case this phrase would have gotten in the matrix clause (Berulava 2025b).
- The exact details of this phenomenon are to be determined (ongoing work with Tamari Berulava), but here’s a tentative generalization of the case-matching restriction:

Matrix \ Embedded	NOM	ERG	DAT	ACC
	×	×	✓	✓
NOM	×	variation	×	×
ERG	✓	variation	×	×
DAT	✓	variation	×	×

Table 3: Case-Matching (preliminary findings).

- **NB:** there’s variation among the speakers which verbs belong to this class, so e.g. for some *pikrobs* ‘think’ never allows *wh*-extraction (“CP only”). The judgments presented below come from those speakers who allow TP complements for these verbs.
- Subject/object-hood and animacy do not seem to play a role: it really is about structural Case.
- Clauses cannot co-occur with an overt demonstrative in cases of *wh*-extraction even in otherwise grammatical configurations (i.e., when the case-matching is obeyed).

- (38) a. mariam-s hgonia (is), [rom šota mankana-s qidulobs].
 Mariam-DAT thinks (DEM.NOM) ROM Shota.NOM car-ACC buy.PRS.3SG
 ‘Mariam thinks that Shota is buying a car.’
 b. **vin**₁ hgonia mariam-s (*is), [rom **t**₁ mankana-s qidulobs]?
who.NOM thinks Mariam-DAT (DEM.NOM) ROM car-ACC buy.PRS.3SG
 ‘Who does Mariam think is buying a car?’
 c. ***ra-s**₁ hgonia mariam-s (is), [rom šota **t**₁ qidulobs]?
what-ACC thinks Mariam-DAT (DEM.NOM) ROM Shota.NOM buy.PRS.3SG
 ‘What does Mariam think Shota is buying?’

- (39) a. mariam-s hgonia, [rom šota-s mankana uqidia].
 Mariam-DAT thinks ROM Shota-DAT car.NOM buy.PERF.3SG
 ‘Mariam thinks that Shota has bought a car.’
 b. ***vis**₁ hgonia mariam-s (is), [rom **t**₁ mankana uqidia]?
who.DAT thinks Mariam-DAT (DEM.NOM) ROM car.NOM buy.PERF.3SG
 ‘Who does Mariam think has bought a car?’
 c. **ra**₁ hgonia mariam-s (*is), [rom šota-s **t**₁ uqidia]?
what.NOM thinks Mariam-DAT (DEM.NOM) ROM Shota-DAT buy.PERF.3SG
 ‘What does Mariam think that Shota has bought?’

- For more detailed data on the case-matching in *wh*-extraction, see the appendix.

Wh-extraction from nominalized ModPs

- For all speakers, extraction from these clauses is in general possible.
- Inserting an overt demonstrative *is* ‘that’ leads to ungrammaticality:

(40) **vin**₁ unda mariam-s (*is), [rom masc'avlebel-ma **t**₁ šeakos]?
who.NOM wants Mariam-DAT (DEM.NOM) ROM teacher-ERG praise.OPT.3SG
 ‘Who does Mariam want the teacher to praise?’

- *Case-matching*: in general seems to be present, e.g. (41).
- But there are also examples that we found to be more permissive: e.g., (42) has been accepted on multiple occasions as grammatical, even though it violates the generalization in table 3.

(41) ??**vis**₁ gadac'qvít'a mariam-ma [rom **t**₁ nene uqvardes]?
who.DAT decided Mariam-ERG ROM Nene.NOM love.OPT.3SG
 ‘Who did Mariam decide that will love Nene?’

(42) %**vis**₁ unda mariam-s [rom shota **t**₁ da-e-xat'os]
who.DAT wants Mariam-DAT ROM Shota.NOM PVB-pv-paint.OPT.3SG
 ‘For whom did Mariam want Shota to be painted?’

Wh-extraction from nominalized VoicePs

- *Wh*-extraction out of masdar complements is possible for all speakers.
- When the extracted phrase is an argument, the nominalization must pied-pipe to the preverbal position, (43). When it is an adjunct, e.g. a PP, the PP is extracted without pied-piping, (44).

- (43) a. mariam-ma [c'ign-is k'itxva] daasrula.
 Mariam-ERG book-GEN read.NMN.NOM finished
 ‘Mariam finished reading a book.’
 b. [**ris k'itxva**] daasrula mariam-ma?
what.GEN read.NMN.NOM finished Mariam-ERG
 ‘What did Mariam finish reading?’
 c. ***ris**₁ daasrula mariam-ma [**t**₁ k'itxva]?
what.GEN finished Mariam-ERG read.NMN.NOM
 ‘What did Mariam finish reading?’
 d. *[**ris k'itxva**] mariam-ma daasrula?
what.GEN read.NMN.NOM Mariam-ERG finished
 ‘What did Mariam finish reading?’
- (44) a. mariam-ma [k'at'is-tvis tevz-is č'meva] daasrula.
 Mariam-ERG cat.GEN-for fish-GEN feed.NMN.NOM finished
 ‘Mariam finished feeding the cat the fish.’

- b. **ris-tvis**₁ daasrula mariam-ma [**t**₁ tevz-is č'meva]?
what.GEN-for finished Mariam-ERG fish-GEN feed.NMN.NOM
 ‘What did Mariam finish feeding the fish to?’
- c. ***[ris-tvis tevz-is č'meva]** daasrula mariam-ma?
what.GEN-for fish-GEN feed.NMN.NOM finished Mariam-ERG
 ‘What did Mariam finish feeding the fish to?’

Q: Why do we see the constraints on the *wh*-extraction that we see in Georgian?

5 Proposal

- I propose an account of the size-sensitivity of Georgian *wh*-movement that follows the “*Position Intuition*” (Meadows 2024) and suggest that we are looking at a *Williams Cycle* effect.

(45) **The Position Intuition**

The location of the final landing site constrains the location of the launching site.

(46) **The Williams Cycle**

Movement must not land lower in the clausal functional sequence than any projections crossed along the way.

- **Functional sequence:** CP > TP > ModP > VoiceP.
- **Hypothesis:** *Wh*-movement lands in Spec,TP.
 - explains the word order in *wh*-questions (phrases can precede *wh*-words);
 - explains the case-matching effect: movement lands in a domain where case is assigned (TP), and the DP is evaluated for case again in the matrix clause (→ the “*matching*” requirement).

Landing Site	<i>rom</i> -CPs	<i>rom</i> -TPs	<i>rom</i> -ModPs	VoiceP-NMNs
Spec,TP	×	✓ (case-matching)	✓ (case-matching?)	✓ (pied-piping)

Table 4: Possibility of *wh*-movement out of embedded clauses with a null D.

- If it’s true that bare clauses without nominal structure are always CPs (Bondarenko 2022), this hypothesis would also explain why extraction out of bare CPs is not possible in Georgian (unlike in other languages, e.g. *Buryat*, *Korean*, *Russian*). We don’t need to appeal to *Adjunct Island*.
- **Q:** But what derives the Williams Cycle?

(47) **The Level Embedding Approach** (Meadows 2023; Poole 2023; Williams 2003, 2011)

- Level Embedding Conjecture (LEC):** The bigger an embedded clause is, the later in the derivation it is introduced into a matrix clause.
- Cyclic Requirement:** Syntactic operations must target the periphery of the current clausal extended projection.

- I will adopt Meadows (2023)’s version of this approach, which has the following ingredients:

(48) **The Featural Extension Requirement (FER)**

Featural triggers (e.g. [urel]) on a clausal F can only be representationally altered (valued, checked, satisfied etc.) at the cycle at which they are introduced. Each clausal F is introduced in its own cycle.

- Every clausal FP is a cyclic node, i.e. denotes a domain of rule application.
- Some operations are featurally-triggered, others are driven by the legibility at the interfaces.

(49) **Parallel Derivation of Main and Embedded Clauses**

The Merge of a component of clausal functional sequence applies in parallel across all workspaces.

(50) **Embedding as Adjunction**

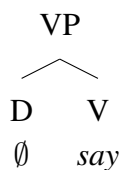
Clausal Embedding involves adjunction: it does not check/value features.

- **Nominal Projections:** Nominal structure is built “out-of-sync” with the verbal structure. E.g., all the nominal structure of the direct object will be inserted at the VP cycle.

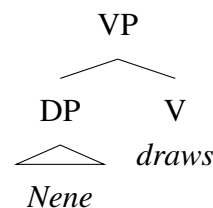
Example:

- (51) mariami ambobs, [rom šota nene-s xat’avs].
 Mariam.NOM is.saying ROM Shota.NOM Nene-ACC is.drawing
 ‘Mariam is saying that Shota is drawing Nene.’

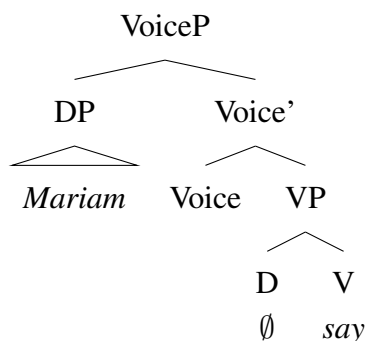
(52) **Matrix VP**



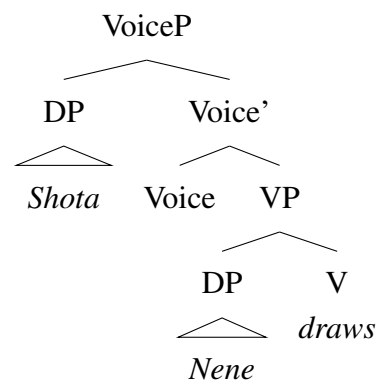
(53) **Embedded VP**

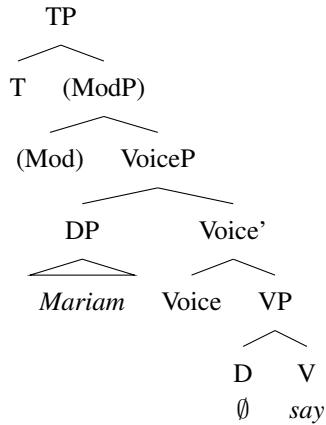
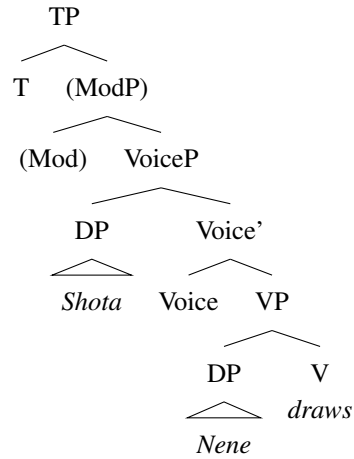
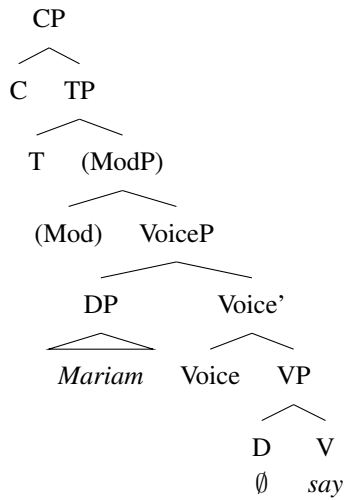
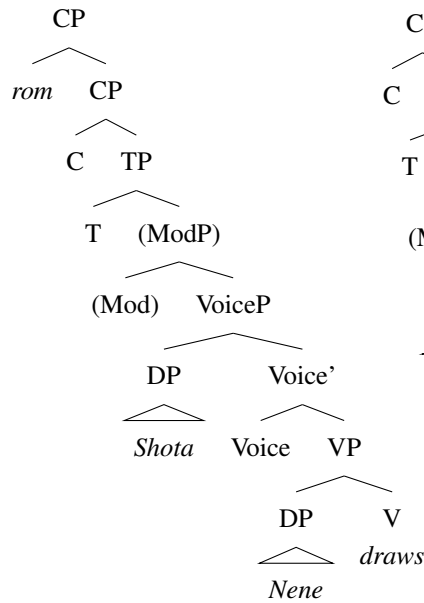
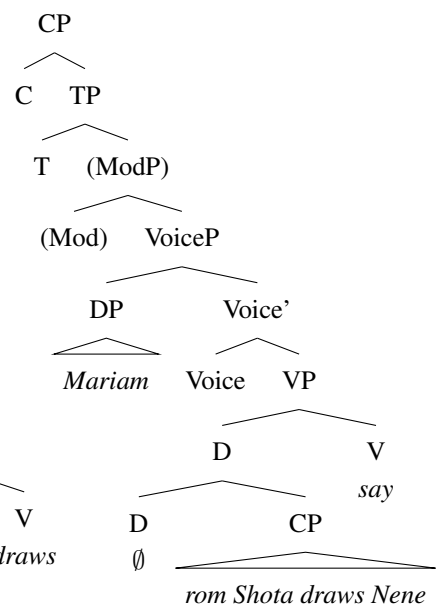


(54) **Matrix VoiceP**



(55) **Embedded VoiceP**



(56) Matrix TP(57) Embedded TP(58) Matrix CP(59) Embedded CP(60) →Single Clause

- This structure will be semantically well-formed: e.g., if the D head has the meaning of a definite article, (61), it would take the meaning of the CP, which is a predicate, (Bondarenko 2022; Elliott 2020; Kratzer 2006; Moulton 2009) as its argument.
- If there is more nominal structure in nominalized CPs—at least an NP/nP layer, then the CP will combine with the null noun as its modifier by Predicate Modification.

$$(61) \quad \llbracket D \rrbracket = \lambda f \in D_{et}. \iota x(f(x) = 1)$$

$$(62) \quad \llbracket CP \rrbracket = \lambda x \in D_{et} \cap D_{Cont}. \text{CONT}(x) = \{w: \text{Shota draws Nene in } w\}$$

- **Upshot:** CP is a restrictor/modifier in the nominal structure, which can be merged late.

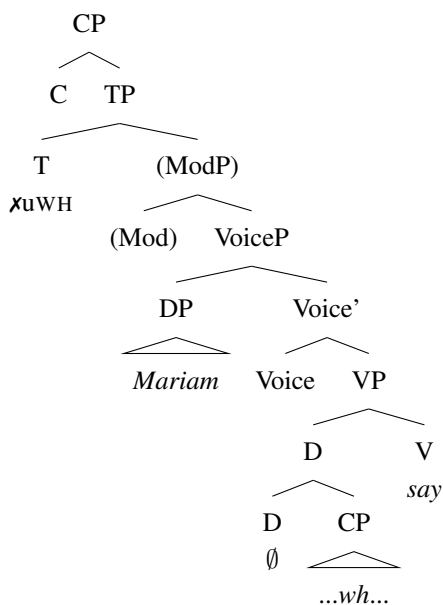
- (63) **TP-Cycle: embedded clause visible**

(64) **TP-Cycle: *wh*-movement to Spec,TP**

- (65) TP-cycle: embedded CP inaccessible

(66) TP-cycle: embedded clause not yet built

(67) CP-cycle: too late for T to probe



- The embedded clause is embedded at the CP-cycle.
- But by then, it's too late for the uWH on T to probe!
- **The Featural Extension Requirement (FER):** featural triggers can be representationally altered only at the cycle at which they are introduced.
- *Wh*-movement from CP clauses would violate FER.
- **NB:** DP layer is irrelevant for this restriction, so the explanation extends to the bare CPs.
- *Why Georgian's different:* because *wh*-movement targets a lower position in the functional sequence.
- This raises questions about phases & intermediate movement, see (Meadows 2023) for discussion.

Why does overttness of D matter?

- We've assumed that in all cases with transitive verbs, clauses are nominalized: this is because whenever it's possible to distinguish the content vs. DP readings, the CPs have the exact same meanings as corresponding DPs (recall the example in (11)).
- Having an overt D bleeds *wh*-movement, both from *rom*-clauses but also from masdars:

- (68) [$\langle *is \rangle$ **r-is-i** $\langle *is \rangle$ gadac'era]₁ $\langle *is \rangle$ daasrula mariam-ma $\langle *is \rangle$ t₁?
 $\langle that \rangle$ what-GEN-NOM $\langle that \rangle$ rewriting $\langle that \rangle$ finished Mariam-ERG $\langle that \rangle$
 'What did Mariam finish rewriting?' (*What did Mariam finish that rewriting of?)

Proposal:

- Georgian D does not have an EPP feature that allows it to attract phrases to its specifier.
 → We don't find cases of subextraction from DPs—*Left Branch Condition* is obeyed.
- (69) *mariam-ma is daasrula [c'erilis gadac'era].
 Mariam-ERG DEM finished letter.GEN rewriting
 Intended: 'Mariam finished that rewriting of the letter.'
- DPs are a Spell-Out domain for *Cyclic Linearization* (Fox and Pesetsky 2005), and thus at DP, D_{obj} < CP order is established: D must precede the CP's material. This means that D_{obj} < wh.
- If a *wh*-phrase moves to Spec,TP without passing through DP edge (due to the lack of EPP), then at the CP domain of linearization, we will get ordering: $C_m < wh < T_m < \dots < DP_{subj} < DP_{obj} V$, and thus wh < D_{obj}. This contradicts the ordering established at D_{obj}P → **ungrammaticality!**

(70) **Cyclic Linearization** (after Fox and Pesetsky 2005)

a. ***Spell-Out Domains:***

At each Spell-Out Domain (*for us here: CP, DP*), linearization happens:
linear order between syntactic units is established.

b. ***Order Preservation:***

Information about linearization, once established at the end of a given Spell-Out domain, is never deleted in the course of a derivation.

- **NB:** If this account is correct, *only overt terminals matter for linearization*—if a D is null, no violating ordering will be established. The *wh*-phrase would then move to the edge of the verbal projections of the embedded clause, and thus be the topmost element in the linearized DP.

6 Conclusion

- We have observed several constraints on long-distance *wh*-extraction in Georgian:
 - ★ ***Size Constraint:*** *wh*-phrases cannot be extracted out of CPs.
 - ★ ***Overt D Constraint:*** clauses with the overt demonstrative (*is*) prohibit extraction.
 - ★ ***Case-matching Constraint:*** *wh*-DPs extracted from TPs and ModPs must have Case that they could have gotten in the matrix clause.
 - ★ ***Pied-piping Constraint:*** *wh*-DPs extracted from masdars (VoiceP NMNS) must pied-pipe them.
- I’ve suggested that the *Size Constraint* is a Williams Cycle effect, which can be explained e.g. by the Level Embedding Approach (Meadows 2023).
- I’ve proposed that the *Overt D Constraint* arises because Georgian Ds don’t have EPP features, and movement out of nominals with overt demonstratives would violate Cyclic Linearization.

Open Questions:

- Does the provided WC-explanation for the *Size Constraint* align with other locality restrictions on syntactic dependencies in the language (*check out Jacob Fernandes’s poster!*)?
- How exactly does the *Case-matching Constraint* work, and how can it inform theories of case?
- Alternative ways to account for the *Overt D* requirement? Explanation of pied-piping behavior?

References

- Berikashvili, Svetlana (2024). *Differential Subject Marking in Georgian*. PhD thesis. Georg August Universität Göttingen / Ilia State University.
- Berulava, Tamari (2025a). “Georgian Masdars”. Squib, MIT.

- Berulava, Tamari (2025b). “Long-distance Questions in Georgian”. Squib, MIT.
- Biberauer, Theresa, Anders Holmberg, and Ian Roberts (2009). “Linearization and the architecture of grammar: A view from the Final-over-Final Constraint”. In V. Moscati and E. Servidio, eds., *Proceedings of the Incontro di Grammatica Generativa XXXV, Studies in Linguistics*, pp. 77–89. Siena: CISCL.
- Biberauer, Theresa, Glenda Newton, and Michelle Sheehan (2009). “Limiting synchronic and diachronic variation and change: The Final-over-Final Constraint”. *Language and Linguistics* 10.4: 701–743.
- Biberauer, Theresa and Michelle Sheehan (2012). “Disharmony, antisymmetry, and the Final-over-Final Constraint”. In Myriam Uribe-Etxebarria and Vidal Valmala, eds., *Ways of structure building*, pp. 106–244. Cambridge: Cambridge University Press.
- Bondarenko, Tatiana (2022). *Anatomy of An Attitude*. PhD thesis. MIT.
- Bondarenko, Tatiana, Jacob Fernandes, Kaden Holladay, and Richard Luo (2025). “Interpreting clausal movement”. Manuscript under review.
- Bondarenko, Tatiana I (2017). “Two aspectual subsystems of Georgian”. *Voprosy Jazykoznanija* 4: 74–100.
- Borise, Lena (2023). “The syntax of wh-phrases, narrow foci, and neg-words in Georgian”. *The Linguistic Review* 40.2: 173–215.
- Egressy, János (2025). “Size- and Position-dependent opacity in Hungarian”. Manuscript. UCLA.
- Elliott, Patrick D (2020). *Elements of Clausal Embedding*. PhD thesis. UCL (University College London).
- Fox, Danny and David Pesetsky (2005). “Cyclic Linearization of Syntactic Structure”. *Theoretical Linguistics* 31.1-2: 1–45.
- Harris, Alice (1981). *Georgian syntax: A study in relational grammar (Cambridge Studies in Linguistics)*. Vol. 33. Cambridge: Cambridge University Press.
- Johnson, Kyle (2003). “Towards an etiology of adjunct islands”. Ed. by Anne Dahl, Kristine Bentzen, and Peter Svenonius. *Nordlyd* 31.1. Tromsø, Norway: University of Tromsø’s Working Papers in Linguistics: 187–215.
- Kratzer, Angelika (2006). “Decomposing Attitude verbs”. Talk given at the Hebrew University of Jerusalem.
- Manetta, Emily (2011). *Peripheries in Kashmiri and Hindi-Urdu: The Syntax of Discourse-Driven Movement*. Amsterdam: John Benjamins.
- Meadows, Tom (2023). *Size Matters: Clause structure and locality constraints in Swahili relatives*. PhD thesis. Queen Mary University of London.
- (2024). “Advanced Syntax: The Williams Cycle”. EGG Summer School Slides.
- Moulton, Keir (2009). *Natural Selection and the Syntax of Clausal Complementation*. PhD thesis. UMass Amherst.
- Poole, Ethan (2023). “Improper case”. *Natural Language & Linguistic Theory* 41.1: 347–397.
- Privoznov, Dmitry (2022). “Adjunct islands are configurational”. *Glossa: a journal of general linguistics* 7.1.
- Truswell, Robert (2007). “Extraction from adjuncts and the structure of events”. *Lingua* 117.8: 1355–1377.
- (2011). *Events, Phrases, and Questions*. 33. Oxford University Press.
- Uriagereka, Juan (1999). “Multiple Spell-out”. In S. Epstein and N. Hornstein, eds., *Working Minimalism*, pp. 251–282. Cambridge, MA: MIT Press.
- Wier, Thomas (2014). “Nonconfigurationality and argumenthood in Georgian”. *Lingua* 145: 36–64.
- Williams, Edwin (2003). *Representation Theory*. Cambridge, MA: MIT Press.

7 Appendix: Data for Case-Matching *Wh*-Movement

• NOM Matrix Subject

(71) NOM extracted *wh*-phrase

- a. ***vin**₁ pikrobs mariami, [rom **t**₁ mankana-s qidulobs]?
who.NOM thinks Mariam.NOM ROM car-ACC buy.PRS.3SG
 ‘Who does Mariam think that is buying a car?’
- b. ***vin**₁ pikrobs mariami, [rom am davaleba-m **t**₁ gaabraza]?
who.NOM thinks Mariam.NOM ROM this task-ERG upset.AOR.3SG
 ‘Who does Mariam think that this task upset?’

(72) ERG extracted *wh*-phrase

- a. ***ra-m**₁ pikrobs mariami, [rom **t**₁ šota gaabraza]?
what-ERG thinks Mariam.NOM ROM Shota.NOM upset.AOR.3SG
 ‘What does Mariam think that upset Shota?’
- b. ***vin**₁ pikrobs mariami, [rom **t**₁ mankana iqida]?
who.ERG thinks Mariam.NOM ROM car.NOM buy.AOR.3SG
 ‘Who does Mariam think that bought a car?’

(73) DAT extracted *wh*-phrase

- a. **vis**₁ pikrobs mariami, [rom **t**₁ nino uqvars]?
who.DAT thinks Mariam.NOM ROM Nino.NOM love.PRS.3SG
 ‘Who does Mariam think that loves Nino?’

(74) ACC extracted *wh*-phrase

- a. **vis**₁ pikrobs mariami, [rom es sakme **t**₁ agižebbs]?
who.ACC thinks Mariam.NOM ROM this task.NOM make.mad.PRS.3SG
 ‘Who does Mariam think that this task is making mad?’

• ERG Matrix Subject

(75) NOM extracted *wh*-phrase

- a. **vin**₁ ipikra mariam-ma, [rom **t**₁ mankana-s qidulobs]?
who.NOM think.AOR.3SG Mariam-ERG ROM car-ACC buy.PRS.3SG
 ‘Who did Mariam think that is buying a car?’
- b. ?**vin**₁ ipikra mariam-ma, [rom am davaleba-m **t**₁ gaanac’qena]?
who.NOM think.AOR.3SG Mariam-ERG ROM this task-ERG upset.AOR.3SG
 ‘Who did Mariam think that this task upset?’

- (76) ERG extracted *wh*-phrase
- a. %**vin**₁ ipikra mariam-ma, [rom **t**₁ mankana iqida]?
who.NOM think.AOR.3SG Mariam-ERG ROM car.NOM buy.AOR.3SG
 ‘Who did Mariam think that bought a car?’
- b. %**ra-m**₁ ipikra mariam-ma, [rom **t**₁ šota gaanac’qena]?
what-ERG think.AOR.3SG Mariam-ERG ROM Shota.NOM upset.AOR.3SG
 ‘What did Mariam think that upset Shota?’
- (77) DAT extracted *wh*-phrase
- ***vis**₁ ipikra mariam-ma, [rom **t**₁ nino uqvars]?
who.DAT think.AOR.3SG Mariam-ERG ROM Nino.NOM love.PRS.3SG
 ‘Who did Mariam think that loves Nino?’
- (78) ACC extracted *wh*-phrase
- ***vis**₁ ipikra mariam-ma, [rom es sakme **t**₁ agižebš]?
who.ACC think.AOR.3SG Mariam-ERG ROM this task.NOM make.mad.PRS.3SG
 ‘Who did Mariam think that this task is making mad?’
- **DAT Matrix Subject**
- (79) NOM extracted *wh*-phrase
- a. **vin**₁ upikria mariam-s, [rom **t**₁ mankana-s qidulobš]?
who.NOM think.PERF.3SG Mariam-DAT ROM car-ACC buy.PRS.3SG
 ‘Who has Mariam thought that is buying a car?’
- b. **vin**₁ upikria mariam-s, [rom am davaleba-m **t**₁ gaabraza]?
who.NOM think.PERF.3SG Mariam-DAT ROM this task-ERG upset.AOR.3SG
 ‘Who has Mariam thought that this task upset?’
- (80) ERG extracted *wh*-phrase
- a. %**vin**₁ upikria mariam-s, [rom **t**₁ mankana iqida]?
who.NOM think.PERF.3SG Mariam-DAT ROM car.NOM buy.AOR.3SG
 ‘Who has Mariam thought that bought a car?’
- b. %**ra-m**₁ upikria mariam-s, [rom **t**₁ šota gaabraza]?
what-ERG think.PERF.3SG Mariam-DAT ROM Shota.NOM upset.AOR.3SG
 ‘What has Mariam thought that made Shota mad?’
- (81) DAT extracted *wh*-phrase
- ??**vis**₁ upikria mariam-s, [rom **t**₁ nino uqvars]?
who.DAT think.PERF.3SG Mariam-DAT ROM Nino.NOM love.PRS.3SG
 ‘Who has Mariam thought that loves Nino?’
- (82) ACC extracted *wh*-phrase
- ***vis**₁ upikria mariam-s, [rom es sakme **t**₁ agižebš]?
who.ACC think.PERF.3SG Mariam-DAT ROM this task.NOM make.mad.PRS.3SG
 ‘Who has Mariam thought that this task is making mad?’