Relative (pro)nominals in Kirundi

Willie Myers

May 15, 2023

1 Introducing the puzzle

- Across the Bantu family, it is common for subject relative clauses to display 'anti-agreement effects' (AAEs), most apparently when the relativized subject is a singular human (for Lubukusu (Diercks 2009, 2010), Nande (Schneider-Zioga 2000, 2007), Abo (Burns 2013), and Bemba (Cheng 2006; Henderson 2007, 2009, 2013)).

- In a standard clause with a Class 1 (singular human) subject, the typical Class 1 agreement morpheme *a-* appears. When the same subject is relativized, standard agreement is not allowed (1b); instead, *u-* appears (1c).

(1) a. Umulendo **a**-ka-belenga   ibuku.
         1boy   **1.AGR**-FUT-read 5book
         'The boy will read the book.'
     b. *umulendo ú-**a**-ka-belenga   ibuku
         1boy   1.REL-**1.AGR**-FUT-read 5book
         'the boy who will read the book'
     c. umulendo ú-**u**-ka-belenga   ibuku
         1boy   1.REL-**1.AAE**-FUT-read 5book
         'the boy who will read the book'   (Bemba; Cheng 2006, 197)

- These so-called 'anti-agreement effects' (AAEs) are characterized by two properties: (i) distinct morphology, in comparison to standard subject-verb agreement marking and (ii) a limited distribution, appearing only when the highest clause-bound argument is relativized.

> **Puzzle**
> ⇒ What triggers the distinct morphology (*a-* vs. *u-*)?
> ⇒ Why does it only appear in some contexts (local subject relatives)?

- Many past analyses attribute the alternative morphology to the movement and/or features associated with relativization (see, e.g., Schneider-Zioga 1995, 2007; Cheng 2006; Henderson 2013; Baier 2018).

- The locality restrictions, however, are not otherwise expected in Ā-movement phenomena, necessitating a variety of complex secondary claims (e.g. anti-locality, strong chain repair, criterial freezing, etc.).

- Kirundi provides a fruitful context in which to (re-)assess past analyses because it shows agreement alternations akin to those analyzed as AAEs but in a notably *different* distribution.

- In Kirundi, *headed* relative clauses show typical agreement (*a-* for Class 1) (2a). It is only in *headless* relative clauses (HRCs) that alternative morphology *u-* appears (2b).

(2) a. Ndâzi [u-mu-gabo **a**-garúk-a].
         I.know AUG-1-man 1.AGR-return-IPFV
         'I know the man who will return.'
     b. Ndâzi [u-**u**-gáruk-a].
         I.know AUG-1.AAE-return-IPFV
         'I know the (person who) will return.'

- **Uh oh!** If AAEs are a result of Ā-movement, both headed and headless relative clauses should display AAEs.

2 Background on Kirundi

- Kirundi is an Eastern Bantu language, classified as Zone J (Interlacustrine/Great Lakes) (Bastin 2003). It is spoken by 12.5 million speakers in the country of Burundi where it is the national language.

- Kirundi is part of a dialect continuum with Kinyarwanda (Rwanda) and smaller contiguous communities in Tanzania, Uganda, and the DRC.

2.1 Nouns

- Nouns are composed of an initial "augment" vowel (D), a noun class prefix (NCP; *n*), and a root:

(3) Kirundi Noun Structure
    $D - n - \sqrt{}$

- The augment is situated in D (Ndayiragije et al. 2012).

- The NCP is a categorizing *n* head with gender and number features (Kramer 2014, 2015; Carstens 2008; Fuchs and van der Wal 2021).

- Noun class (gender) denotes 16 broad semantic categories.

(4) Variation in noun class with the root *-ntu*

Noun Class	Singular	Plural
People (1/2)	*u-mu-ntu* 'person'	*a-ba-ntu* 'people'
Things (7/8)	*i-ki-ntu* 'thing'	*i-bi-ntu* 'things'
Diminutive (12/13)	*a-ka-ntu* 'small thing'	*u-du-ntu* 'small things'
Non-tangible (14)	*u-bu-ntu* 'humanity'	
Places (16)	*a-ha-ntu* 'place'	*a-ha-ntu* 'places'

2.2 Verbs

- Verbs in Kirundi obligatorily display subject-verb agreement expressing number and gender/person. This agreement morpheme appears at the start of the verbal construction.

(5) a. N-a-bón-ye   i-gi-tabo.
       1SG.AGR-PST-read-PFV AUG-7-book
       'I read a book (yesterday).'
    b. U-mu-gabo a-tēk-a   u-mu-ceri.
       AUG-1-man 1.AGR-cook-IPFV AUG-3-rice
       'The man cooks rice.'

- Tone can be lexical (*umuryāngo* 'family' vs. *umuryângo* 'door') or grammatical (*nasomye* 'I read (earlier today)' vs. *nasómye* 'I read (before today)' (Ntahokaja 1994, 162)).

3 Evidence against Ā-movement approach

- Variation in the morphology of headed vs. headless relative clauses in Kirundi places doubt on an Ā-driven approach to AAEs.

3.1 Headed vs. headless clauses

- Standard headed relative clauses have no overt relative complementizer but are nonetheless distinguishable because they exhibit common properties of dependent (non-matrix) clauses.

- The most apparent of these properties is the addition of a high tone on the second syllable of the embedded verb root. This is found in all headed relative clauses and embedded complement clauses.

- In the example below, this dependent high tone differentiates a matrix clause (6a) from its relative clause counterpart (6b)

(6) a. U-mu-gabo a-tēk-**a** u-mu-ceri.
AUG-1-man 1.AGR-cook-IPFV AUG-3-rice

'The man cooks rice.' (Matrix)

b. Ndakûnda [u-mu-gabo ____ a-tēk-**á** u-mu-ceri].
I.like [AUG-1-man 1.AGR-cook-IPFV AUG-3-rice]

'I like the man (who) cooks rice.' (Relative)

- Headed relative clauses can also be formed with objects and across clauses.

(7) Ndakûnda [u-mu-ceri a-tēk-**á** u-mu-gabo ____].
I.like AUG-3-RICE 1.AGR-cook-IPFV AUG-1-man

'I like the rice the man cooks.'

(8) Ndakûnda [u-mu-ceri tūzí [kó u-mu-gabo a-tēk-**á** ____]].
I.like AUG-3-RICE we.know C AUG-1-man 1.AGR-cook-IPFV

'I like the rice we know that the man cooks.'

- HRCs are formed from an augment, a class-denoting morpheme, and a verbal projection. The class morpheme is where we see irregular morphology.

(9) Ndakûnda [u-**u-tê**k-a u-mu-ceri].
I.like AUG-1-cook-IPFV |sc aug-3-rice

'I like the (one (person) who) cooks rice.'

- In terms of distribution, Kirundi HRCs pattern exactly like previously documented AAE relative clause constructions – they must be local and clause-bound.

- Attempts to create object or long-distance HRCs are illicit. In fact, it's not even clear what it would look like to have an object or long-distance HRC.

(10) I like...

a. *[u-**u**-kûnd-a]
AUG-1-like-IPFV

'the (one who) is liked'

b. *[u-**u**-a-kûnd-a Muco]
AUG-1-1.AGR-like-IPFV Muco

'the (one who) Muco likes'

c. *[u-**u**-tūzí Muco a-kûnd-a]
AUG-1-we.like Muco 1.AGR-like-IPFV

'the (one who) we know Muco likes'

- Instead, passivization (11) or a pronoun head (12) is required.

(11) Ndakûnda [u-**u**-kûnd-u-a].
I.like AUG-1-like-PASS-IPFV

'I like the one (who) is liked.'

(12) a. Ndakûnda [**u-u-ó** Muco a-kūnd-á.
I.like AUG-1-ONE Muco 1.AGR-like-IPFV

'I like the one Muco likes.'

b. Ndakûnda [**u-u-ó** tūzí [kó Muco a-kūnd-á]].
I.like AUG-1-ONE we.know that Muco 1.AGR-like-IPFV

'I like the one we know that Muco likes.'

Recap

– Kirundi HRCs pattern exactly like AAE constructions as documented in Bemba and other Bantu languages.

1. distinct morphology

2. limited distribution

– However, Kirundi *headed* relative clauses do *not* pattern in this way.

(13) Headed vs. headless relative clauses

	Headed	Headless
Morphology	*a-*	*u-*
Distribution	everywhere	highest local argument
Tone	σσ́	σ́σ

– In addition, HRCs differ from headed relative clauses in their tonal pattern – a corollary of dependent clause syntax.

3.2 Analysis: HRCs are non-Ā relative clauses

⇒ Proposal: HRCs in Kirundi are not created by Ā-movement and therefore, their distinct morphology is not driven by Ā-driven AAEs.

- Only HRCs exhibit the properties associated with Bantu AAEs. Headed relative clauses do not.

- The properties of HRCs do not *obviously* support an Ā-construction.

1. only the highest nominal is targeted

2. HRCs cannot be long distance (a basic test of Ā-movement)

3. there is no wh-element or relative pronoun

- All other expected Ā-constructions in the language exhibit standard subject-verb agreement and not *u*-morphology.

(14) Distribution of agreement/concord morphemes

a-	*u-*
matrix clauses	obj. markers
dependent clauses	demonstratives
adjunct clauses	quantifiers
cleft/*wh*-constructions	numerals
headed relatives	pronouns
light-headed relatives	HRCs

- Are HRCs relative clauses at all? Newman 2022 argues for the need to disassociate the descriptive label "relative clause" from the theoretical concept Ā-dependency:

  a relative clause is "any kind of modification of an NP which contains an extended projection of a verb root"

- An example of a non-Ā relative clause in English is found in modifying gerunds. Like Kirundi HRCs, these constructions lack overt wh-morphology and can only modify the highest clause-bound nominal.

(15) English non-Ā relative clause
    a.   the [___ girl-stinging] bee
    b.   *the [bee ___-stinging] girl

- Non-Ā constructions with relative clause interpretations have similarly been proposed for Northern Paiute and Hiaki (Toosarvandani 2014; Harley 2020).

- An additional piece of evidence against HRCs being Ā-movement phenomena is that the distinct morphology associated with Bantu AAEs is found elsewhere in Kirundi, outside of the verbal domain.

- The same morphology and a similar distribution occurs with a subset of pronouns and noun modifiers. In these forms, the *u-* morpheme is clearly an *n* head, like the NCPs of standard nouns.

(16)   a.   Ndakûnda [u-**u**-ŭndi].
              I.like   AUG-1-other
              'I like the other (one).'

    b.   Ndakûnda [u-**u**-â-njé].
              I.like   AUG-1-of-me
              'I like mine (my one).'

- However, similar to HRCs, what appears after the NCP can be more complex than a bare noun root (see for instance the possessive in (16b)).

- This overlap in morphology and distribution opens the door for an account of Kirundi HRCs in the nominal and not verbal domain.

4   HRCs in the nominal domain

- Having argued against an Ā-movement approach, I now argue for a nominalization approach to explaining HRCs.

4.1   *u-* as a NCP

- As noted, all nouns in Kirundi have a noun class prefix (NCP) expressing gender and number (note: this is also true of adjectives).

(17)   [U-**mu**-âna **mu**-īzá]   a-ra-za.
        AUG-1-child 1-beautiful 1.AGR-DJ-come.IPFV
        'The beautiful child is coming.'

- If HRCs are nominal constructions, we must also assume that they possess an *n* head akin to *mu-* above.

- However, HRCs never appear with a *mu-* morpheme; instead, they have a *u-* morpheme in this position.

(18)   [U-**u**-têka] a-ra-za.
        AUG-1-cook 1.AGR-DJ-come.IPFV
        'The one who cooks is coming.'

- While Ā-movement accounts claim that this morpheme is subject-verb (anti-)agreement, I will present evidence that it is a noun class prefix.

(19)   AAE HRC Structure
        D – AGR – verbal projection

(20)   Nominalization HRC Structure (Repeated)
        D – *n* – verbal projection

- First, HRCs pattern like nouns in their co-occurrence with the augment. The augment is obligatory for all nouns unless preceded by the negative operator *nta* or another determiner. The same is true of HRCs.

  (21) a. Nta   (*u-)mu-hǔngu.
            there.is.no (*AUG)-1-boy
            'There is no boy.'
        b. Nta   (*u-)u-têka.
            there.is.no (*AUG)-1-cook
            'There is no boy.'

- Second, the *u-* NCP appears in other words that occupy nominal positions.

- Third, these other words lack verbal elements that could take part in agreement.

  (22) a. U-kūnd-a   [u-mu-gabo]?
            2s-love-IPFV AUG-1-man
            'Do you love the man?'
        b. U-kūnd-a   [u-û-hé]?
            2s-love-IPFV AUG-1-which
            'Which one do you love?'

  (23) a. Hari   [u-mū-ntu]   u-kūnd-á.
            there.is AUG-1-person 2s-love-IPFV
            'There's the person you love.'
        b. Hari   [u-u-ó]   ukūnd-á.
            there.is AUG-1-one 2s-love-IPFV
            'There's the one you love.'

Quick aside: *u-* vs. *mu-*?

- Why do HRCs have an *u-* NCP and not a *mu-* one?

- In examining their distribution, the *u-* words in Kirundi fall into three clear groupings which map exactly onto the nominal proform typology proposed by Déchaine and Wiltschko 2002.

Type	Proform	Syntax	Distribution	Semantics
demonstratives	pro-DP	D syntax	argument	definite
numerals	pro-ϕP	neither	argument/predicate	variable
other	pro-NP	N syntax	predicate	constant

- **Proposal**: Kirundi proforms (like HRCs) always use *u-* morphology.

- The claim that HRCs are pro-NPs supports the difference in the interpretation of HRCs and DPs (24). It also touches upon the proposal by Little et al. 2023 for a null nominal head anaphoric to a discourse salient set ($\emptyset_{\text{dom}}$).

  (24) A-bā-ntu   ba-tatu.
        AUG-2-people 2-three
        'There are three people.'
        a. Hari   u-u-urírīmba.
            there.is AUG-1-sing
            'There is one (of them) who sings.'
        b. Hari   u-mū-ntu   a-rirîmba.
            there.is AUG-1-person 1.AGR-sing
            'There is (another) person who sings.'

4.2 HRCs as nominalizations

- How do HRCs fit into Kirundi nominalization strategies?

- The NCP, as a categorizing *n* head, is often used to nominalize a verb root.

  (25) a. *-rim-* 'to farm' → *u-**mu**-rim-a* 'a field'
        b. *-rim-* 'to farm' → *u-**mu**-rim-yi* 'a farmer'

- Let's look specifically at this second example because like HRCs, it denotes an entity. In Kirundi, the agent-introducing suffix *-yi* is adjoined to verb roots to derive agent nominalizations. I locate this suffix in Voice (Myers 2023).

  (26) *u-mu-rim-yi*
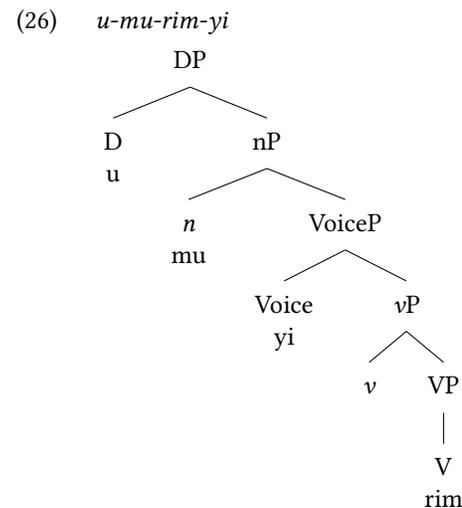
- In entity-denoting nominalizations, it is always the highest argument which is expressed (similar to Baker and Vinokurova 2009 on Kikuyu). Thus, an object nominalization can only be done with a passivized verb, marked by the *-u* suffix.

  (27) a. *-tum-* 'to send' → *u-**mu**-tum-u-a* 'a messenger'
  b. *-yobor-* 'to lead' → *u-**mu**-yobor-u-a* 'an employee'

- Notably, these are the same locality restrictions that characterize HRCs and other proposed AAE constructions in Bantu.

- There are a several differences between *mu-* nominalizations and HRCs.

  1. HRCs are productive; other nominalizations are limited in the lexicon.

     (28) a. *u-u-rím-a* 'the one who farms' – *u-mu-rim-yi* 'a farmer'
     b. *u-u-bón-a* 'the one who sees' – **u-mu-bon-yi* 'a seer'

  2. HRC nominalizations can include tense/aspect/mood. Others cannot.

     (29) a. u-u-a-rím-ye
        AUG-1-PST-cook-PFV
        'the one who cooked'
     b. *u-mu-a-rim-i-ye
        AUG-1-PST-cook-VOICE-PFV
        Intended: 'the former(?) farmer'

 3. HRCs can include objects and adverbs; non-HRC nominalizations cannot.

     (30) a. u-u-sóm-a   i-bi-tabo
        AUG-1-read-IPFV AUG-8-book 8-many
        'the one who reads lots of books'
     b. *u-mu-som-yi   i-bi-tabo
        AUG-1-read--VOICE-PFV
        'the book-reader'

- Based on these data, it is clear that HRCs require a more complex structure than *mu-* nominalizations: the verbal projection is larger, likely up to TP.

  (31)   [DP u- [NP u- [TP -ríma]]]

- However, there is no overt subject or overt subject-verb agreement to hint at what is projected in subject position.

(32)   *u-u-rim-a*
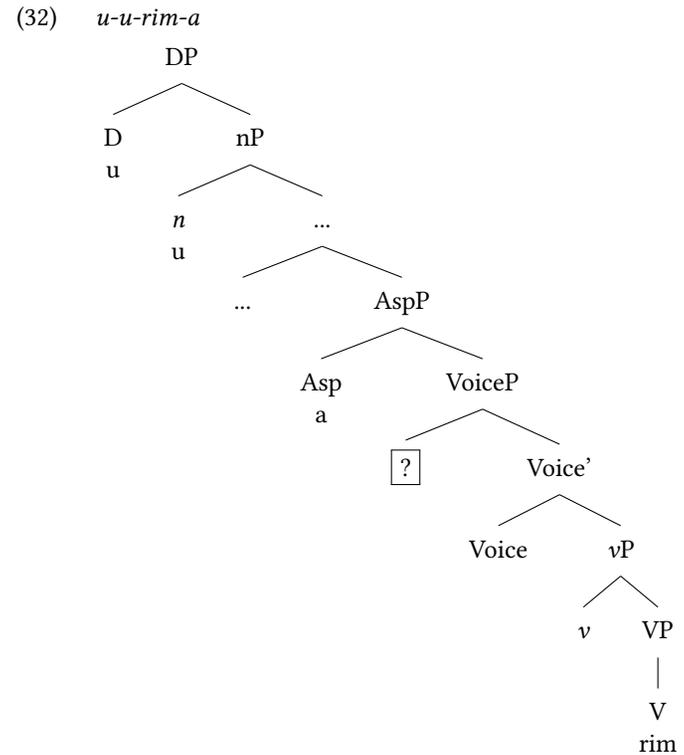
5   Interpreting HRCs

- How are we getting an entity interpretation from the nominalization of such a high verbal projection?

  – Unlike with *u-mu-rim-yi* which can be nominalized when the predicate is still unsaturated, the presence of higher projections require an alternate analysis.

- What is in the subject position introduced by Voice?

  – If a null subject is projected, we may expect agreement and a saturated predicate (<t>). The latter would require type gymnastics of some sort to result in an NP of type <e,t>.

  – If there is no subject projected, how does the derivation continue to project higher?

- Potential options for interpretation:

  1. No subject – the verbal projection remains unsaturated throughout, leaving a TP of type <e,t>.

  2. Null subject + type-shifting

     (a) A-movement – phrase-internal subject movement à la quantification into an XP (Heim and Kratzer 1998)
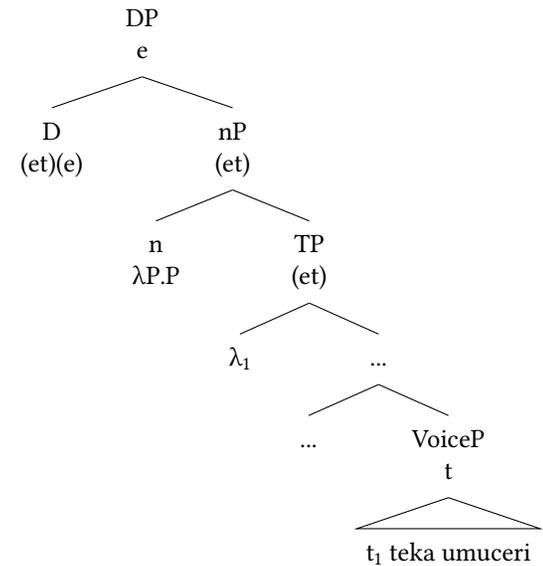
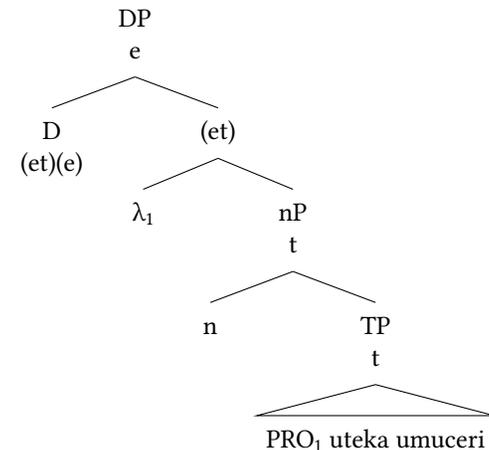   (33) *u-u-teka u-muceri* à la Heim & Kratzer

```
              DP
              e
          ┌────┴────┐
          D         NP
        (et)(e)     (et)
                 ┌───┴───┐
                 n       TP
                        (et)
                    ┌────┴────┐
                   PRO
                       ┌───┴───┐
                       1       TP
                               t
                            ┌───┴───┐
                         t₁ uteka umuceri
```
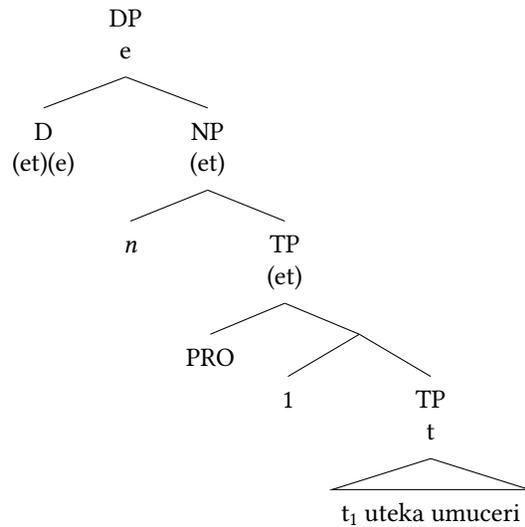
     (b) Operator at VoiceP – an operator is projected in subject position and then binds its own trace after moving to the highest pre-nominalization projection (Harley 2020)
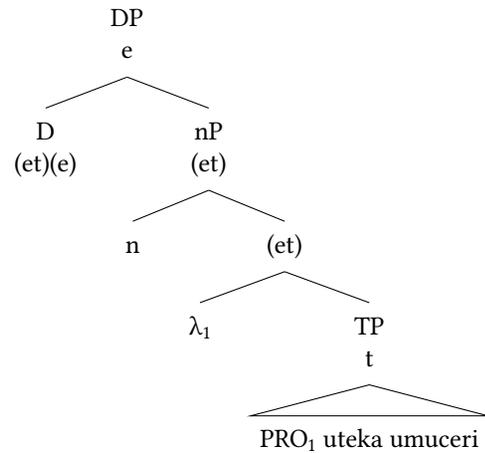
         (34) *u-u-teka u-muceri* à la Harley

```
                    DP
                    e
                ┌───┴───┐
                D       nP
              (et)(e)   (et)
                     ┌───┴───┐
                     n       TP
                    λP.P     (et)
                         ┌────┴────┐
                         λ₁       ...
                               ┌───┴───┐
                              ...      VoiceP
                                         t
                                      ┌───┴───┐
                                   t₁ teka umuceri
```

     (c) Operator at determiner – operator inserted with D because a determiner requires a set of entities as its argument (Toosarvandani 2014)

         (35) *u-u-teka u-muceri* à la Toosarvandani

```
                DP
                e
            ┌───┴───┐
            D       (et)
          (et)(e)
                 ┌───┴───┐
                 λ₁      nP
                          t
                      ┌───┴───┐
                      n       TP
                               t
                            ┌───┴───┐
                         PRO₁ uteka umuceri
```

     (d) Operator at nominalization (riff on Toosarvandi's D operator; for Kirundi, we need the NP to take in type <e,t>) – operator is inserted with *n* because a nominalizer requires <e,t>

(36)   *u-u-teka u-muceri* with *n* operator

```
                    DP
                    e
                   / \
                  D   nP
               (et)(e) (et)
                     /  \
                    n   (et)
                       /  \
                      λ₁   TP
                           t
                          / \
                    PRO₁ uteka umuceri
```

(36) *u-u-teka u-muceri* with *n* operator

DP — e
D — (et)(e)   nP — (et)
n   (et)
λ_1 TP — t
PRO$_1$ uteka umuceri

Next steps

- Pros/Cons for these structures – which are theoretically more attractive?

- Are there tests that could help decide between one or the other?

References

Baier, Nico. 2018. Anti-Agreement. Ph.D. dissertation, University of California, Berkeley.

Baker, Mark C., and Nadya Vinokurova. 2009. Agent nominalizations and why they are not like event nominalizations. *Language* 85:517–556.

Bastin, Yvonne. 2003. The Interlactrine Zone (Zone J). In *The Bantu Languages*, eds. Derek Nurse and Gerard Philippson, 501–528. London: Routledge.

Burns, Roslyn. 2013. Abo optional anti-agreement. In *Selected Proceedings of the 43rd Annual Conference on African Linguistics*, eds. Ólaniké Óla Orie and Karen W. Sandersq. Cascadilla Press.

Carstens, Vicki. 2008. DP in Bantu and Romance. In *The Bantu-Romance Connection*, eds. Katherine Demuth and Cécile de Cat, 131–165. Amsterdam: John Benjamins.

Cheng, Lisa L.-S. 2006. Decomposing Bantu relatives. In *North East Linguistics Society (NELS) 36*, 197–216. GLSA.

Déchaine, Rose-Marie, and Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33:409–442.

Diercks, Michael. 2009. Subject extraction and (so-called) anti-agreement effects in Lubukusu: A criterial freezing approach. In *CLS 45*. Chicaigo Linguistics Society.

Diercks, Michael. 2010. Agreement with subjects in Lubukusu. Ph.D. dissertation, Georgetown University.

Fuchs, Zusanna, and Jenneke van der Wal. 2021. The locus of parametric variation in Bantu gender and nominal derivation. *Linguistic Variation* 22:267–323.

Harley, Heidi. 2020. Relative nominals and event nominals in Hiaki. In *Nominalization: 50 years on from Chomsky's Remarks*, eds. Artemis Alexiadou and Hagit Borer. Oxford: Oxford University Press.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell Publishers.

Henderson, Brent. 2007. Anti-agreement and [person] in Bantu. In *Selected Proceedings of the 38th Annual Converence on African Linguistics*, eds. Masangu Matondo, Fiona McLaughlin, and Eric Potsdam, 173–181. Cascadilla Proceedings Project.

Henderson, Brent. 2009. Anti-agreement: Locality of movement or agreement? In *CLS 45*, 89–102. Chicago Linguistic Society.

Henderson, Brent. 2013. Agreement and person in anti-agreement. *Natural Language and Linguistic Theory* 31:453–481.

Kramer, Ruth. 2014. Clitic doubling or object agreement: An Amharic investigation. *Natural Language and Linguistic Theory* 32:593–634.

Kramer, Ruth. 2015. *The morphosyntax of gender*. Oxford University Press.

Little, Carol-Rose, Scott AnderBois, and Jessica Coon. 2023. Type-shifters in headless relative clauses. Manuscript.

Myers, Willie. 2023. Unpacking -ish: a bimorphemic account of Kirundi causatives. In *Proceedings of BaTOM 1: Toronto-Montreal Bantu Colloquium*, volume 45. Toronto Working Papers in Linguistics.

Ndayiragije, Juvenal, Emmanuel Nikema, and Parth Bhatt. 2012. The augment in Kirundi: when syntax meets phonology. In *Proceedings of the 42nd Annual Conference of African Linguistics*, ed. Michael Marlo et al., 112–121. Somerville, MA: Cascadilla Proceedings Project.

Newman, Elise. 2022. Probing for the closest DP: a reply to Branan et al. 2022. Unpublished manuscript, from personal website.

Ntahokaja, Jean-Baptiste. 1994. *Grammaire Structurale du Kirundi*. Bujumbura, Burundi: Faculté de Philosophie et Lettres.

Schneider-Zioga, Patricia. 1995. Specifier/head agreement in Kinande. *Cahiers linguistiques d'Ottowa* 23.

Schneider-Zioga, Patricia. 2000. Anti-agreement and the fine structure of the left periphery. In *University of California Irvine Working Papers in Linguistics*, volume 6.

Schneider-Zioga, Patricia. 2007. Anti-agreement, anti-locality and minimality. *Natural Language and Linguistic Theory* 25:403–446.

Toosarvandani, Maziar. 2014. Two types of deverbal nominalization in Northern Paiute. *Language* 90:786–833.