

Chapter 1

Unifying Prolepsis and Cross-clausal Cliticization in Lubukusu

Lydia Newkirk

Rutgers University

Abstract goes here

1 Prolepsis in Lubukusu

The constructions I describe in this paper are analogous to prolepsis in English:

- (1) *John knows of Mary_i that she_i is an excellent linguist.*

These constructions differ from typical clause-embedding constructions in three ways: First, an extra nominal occurs to the left of the complementizer (the proleptic object); second, a pronoun that corefers with the proleptic object obligatorily occurs in the embedded clause; third, the embedded clause is interpreted as being about the proleptic object.

In Lubukusu there are three ways to license a proleptic object: first, a proleptic object can be introduced with a preposition (2a), as is the case in English, or equivalently with an applicative morpheme (2b).

- (2) a. John a-subil-a **khu** Bill_i a-li **o-mu-eene_i/niye_i** a-li
 John SM.c1-believe-FV PREP Bill c1-that c1-c1-own/him c1-be
 o-mu-miliyu
 c1-c1-smart
 ‘John believes of Bill_i that he_i is smart’ (Lubukusu)



a matrix (third person, non-reflexive) object marker rather than the reflexive marker, although it is degraded when the embedded object marker is in object position:

- (7) John a-a-mu-lom-a a-li o-mu-eene a-a-siim-a
 John SM.c1-PST-OM.c1-say-FV c1-that c1-c1-own SM.c1-PST-like-FV
 Mary
 Mary
 ‘John said about him_i that he_i likes Mary.’
- (8) John a-a-mu-lom-a a-li Mary a-a-lom-a khu
 John SM.c1-PST-OM.c1-say-FV c1-that Mary SM.c1-PST-say-FV of
 o-mu-eene
 c1-c1-own
 ‘John said about him_i that Mary speaks of him_i.’
- (9) ? John a-a-mu-lom-a a-li George a-mu-siima
 John SM.c1-PST-OM.c1-say-FV c1-that George SM.c1-OM.c1-like-FV
 ‘John say of him_i that George likes him_i.’

Further investigation of these constructions bears this this out. Constructions with *khu-mu-eene* in the matrix clause are insensitive to locality, whereas the construction with the RFM/OM cliticized to the matrix verb is sensitive to island boundaries:

With RFM:

- (10) John_i a-lom-a **khu-mu-eene_i** a-li Bill a-khaenj-a
 John SM.c1-say-FV PREP-c1-own c1-that Bill SM.c1-look.for-FV
 [o-mu-undu o-wa-mu-lip-a **o-mu-eene_i**]
 c1-c1-person wh-c1-OM.c1-PST-pay-FV c1-c1-own
 ‘John_i said about himself_i that Bill is looking for the person who paid himself_i.’
- (11) * John_i a-i_i-lom-a a-li Bill a-khaenj-a [o-mu-undu
 John SM.c1-RFM-say-FV c1-that Bill SM.c1-look.for-FV c1-c1-person
 o-w-a-mu-lip-a **o-mu-eene_i**]
 wh-c1-OM.c1-pst-pay-FV c1-c1-own
 ‘John_i said that Bill is looking for [the person who paid himself_i].’
- (12) * John_i a-i_i-lom-a a-li o-mu-eene_i a-rekukh-a [paata ya
 John SM.c1-RFM-say-FV c1-that c1-c1-own SM.c1-leave-FV after

- Mary khu-mu-khuu-p-a **o-mu-eene_i**
 Mary c15-OM.c1-c15?-hit-FV c1-c1-own
 ‘John_i said that he_i left [after Mary hit him_i].’ (Adjunct island)
- (13) * Jack a-i-many-il-e a-li George a-ch-a nge
 Jack SM.c1-RFM-know-APPL?-FV c1-that George SM.c1-leave-FV when
 a-mu-bon-a o-mu-eene
 SM.c1-OM.c1-see-FV c1-c1-own
 ‘Jack_i knows that George left when he saw himself_i’ (Adjunct island)
- (14) * Bill_i a-i_i-nyol-a [chilomo mbo John a-mu-lip-a
 Bill SM.c1-receive-FV information that John SM.c1-OM.c1-pay-FV
o-mu-eene_i]
 c1-c1-own
 ‘Bill_i heard [a rumor (about himself_i) that John paid him_i].’ (CNPC)
- (15) * John_i a-i_i-subil-a [likhuwa mbo Bill a-mu-bon-a
 John SM.c1-RFM-hope-FV claim that Bill SM.c1-OM.c1-see-FV
o-mu-eene_i]
 c1-c1-own
 ‘John_i believes [the claim that Bill saw himself_i].’ (CNPC)

With OM:

- (16) * John a-a-mu_i-lom-a a-li George a-khaenj-a
 John SM.c1-PST-OM.c1-say-FV c1-that George SM.c1-look.for
 [o-muu-ndu o-w-a-mu-lip-a o-mu-eene]
 c1-c1-person wh-c1-pst-pay-FV c1-c1-own
 ‘John said of him_i that George is looking for [the person who paid
 him_i.]’
- (17) ? John a-**mu_i**-lom-a a-li o-mu-eene_i a-rekukh-a [paata ya
 John SM.c1-OM.c1-say-FV c1-that c1-c1-own SM.c1-leave-FV after
 Mary khu-mu-khuu-p-a **o-mu-eene_i**]
 Mary c15-OM.c1-c15?-hit-FV c1-c1-own
 John said of him_i that he_i left after Mary hit him_i.’

These correlate with the island/locality constraints for *wh*-movement in Lubukusu. The following are the corresponding island examples from Wasike (2006):

- (18) * Naanu ni-y-e Wafula a-kha-enj-a [o-muu-ndu
 who pred-c1-pron Wafula c1-prs-look.for-FV c1-c1-person
 o-w-a-kul-a]
 wh-c1-pst-buy-fv
 ‘What is it that Wafula is looking for [the person who bought]?’
- (19) * Naanu ni-y-e Nasike a-a-rekukh-a [paata ye *t* khu-khuup-a
 Who pred-c1-pron Nasike c1-pst-leave-FV after of inf-beat-FV
 Nanjala]
 Nanjala
 ‘Who is that Nasike left [after *t* hitting Nanjala]?’
- (20) ?? [Chi-lomo mbo Wafula a-a-ib-a si(ina) cha-a-chun-i-a]
 c7-report that Wafula c1-PST-steal-FV what c7-PST-hurt-CAUSE-FV
 Nafula ku-mw-oyo?
 Nafula PP-3-heart
 ‘What did [the report that Wafula stole] hurt Nafula?’

Based on the demonstrated island restrictions, I take the cliticization strategy to be movement of a pronoun from its argument position in the embedded clause to the matrix clause, and the applicative and prepositional phrase strategies to be base-generation of a pronoun or DP in the matrix clause. These same sentences are illicit without the appropriate embedded object marking, however:

- (21) * John_{*i*} a-i_{*i*}-lom-a a-li Mary a-siim-a o-mu-eene_{*i*}
 John SM.c1-RFM-say-FV c1-that Mary SM.c1-like-FV c1-c1-own
 ‘John_{*i*} said that Mary likes him_{*i*}.’
- (22) * John_{*i*} a-i_{*i*}-lom-a a-li George a-khaeknj-a
 John SM.c1-RFM-say-FV c1-that George SM.c1-look.for-FV
 o-muu-ndu o-wa-lip-a o-mu-eene_{*i*}
 c1-c1-person wh-c1-pay-FV c1-c1-own
 ‘John_{*i*} said that George is looking for the person who paid him_{*i*}.’

The ungrammaticality of (22) is unsurprising, given the general island sensitivity of this construction. (21) shows that the embedded object marker is obligatory, a fact I will return to later. If the cliticization strategy is movement from the embedded clause to the matrix clause, I will have to explain why the embedded OM remains obligatory.

Three main characteristics that are common across these constructions:

1. An “extra” nominal argument in the matrix clause, which the matrix verb does not ordinarily take
2. A aboutness relation between the extra argument and the embedded predicate
3. A specific (*de se*-like) acquaintance relation between the extra argument and the matrix attitude holder

I will conclude that characteristics 2 and 3 come about by the same process, and so I will consider them together. Characteristic 3 is a separate concern, so I will address it first.

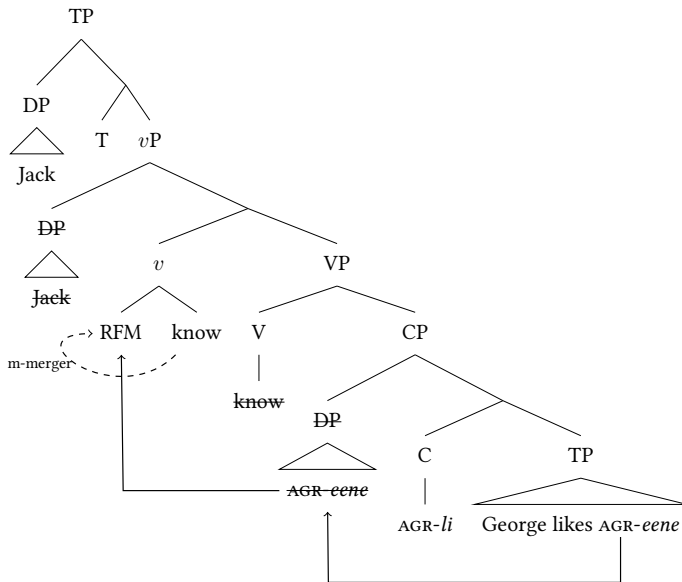
2 Nominal licensing

In analyzing the island-sensitive clitic-licensed prolepsis, I generally follow analyses of cross-clausal agreement in Polinsky & Potsdam (2001); Bruening (2001); Branigan & MacKenzie (2002). The embedded DP A' -moves to the embedded left periphery. In Lubukusu, that pronoun can then undergo further A' -movement to cliticize to the matrix verb. I follow the analysis of clitics as incorporated pronouns from Matushansky (2006); Baker & Kramer (2016).

On this analysis, (5) has the preliminary structure in (23).

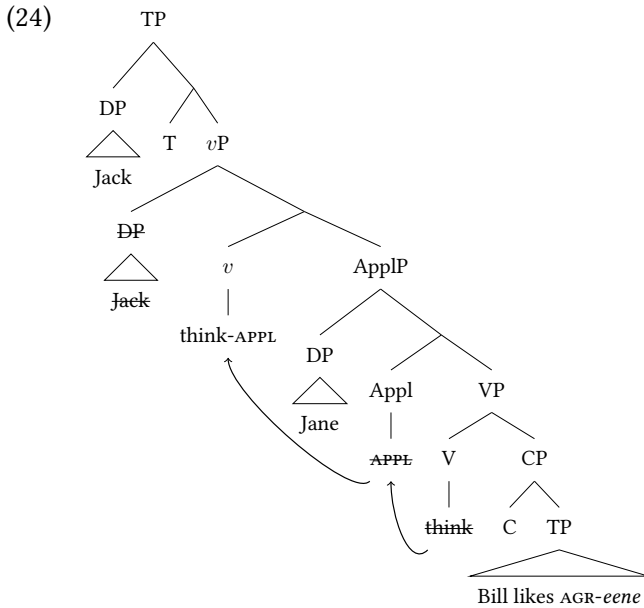
- (5) Jack_i a-i_i-many-il-e a-li George a-mu_i-siim-a
Jack SM.c1-RFM-knows-TNS-FV c1-that George SM.c1-OM.c1-like-FV
o-mu-eene_i
c1-c1-own
'Jack_i knows that George likes him_i.'

(23)



The preposition-licensed and applicative-licensed cases, on the other hand, have a base-generated proleptic object, introduced by a preposition or applicative, and then are related to the embedded pronoun by binding.

- (2b) John a-kanakan-il-a Jane_i a-li Bill a-mu-siim-a
 John SM.c1-think-APPL-FV Jane c1-that Bill SM.c1-OM.c1-like-FV
o-mu-eene_i/niye_i
 c1-c1-own/her
 ‘John thinks of Jane_i that Bill likes her_i.’



The movement strategy is restricted to pronouns due to independent facts about Lubukusu object marking. The object markers are clitics, and these clitics can only be doubled by pronouns, and not by full DPs:

- (25) N-a-**mu_i**-bon-a (#Wekesa_i)
 1sgS-PST-OM.c1-see-FV Wekesa
 ‘I saw him.’ (Diercks & Sikuku 2015: 2)
- (26) Wekesa a-a-**mu_i**-p-a (**niye_i**)
 Wekesa SM.c1-PST-OM.c1-beat-FV **him**
 ‘Wekesa beat him.’ (Afranaph ID: 3734/5039)
- (27) Yòháná_i á-á-**i_i**-bon-a (o-mu-eene_i)
 Yohana SM.c1-PST-RFM-see-fv c1-c1-own
 ‘John_i saw himself_i’ (Afranaph ID:1248/1249)

In principle, a full DP could undergo movement to the matrix clause, but Lubukusu has no way of licensing it there with a clitic. At the same time, although prepositions and applicatives can both provide licensing to an additional matrix argument, they are not viable landing sites for movement, and so preclude movement of an embedded argument into their complement.

I can now offer a tentative explanation for why the embedded OM remains obligatory even in the movement cases. The embedded pronoun begins by re-

ceiving a theta role in the embedded clause, but while it is then syntactically licensed in the matrix clause via cliticization, it is not semantically licensed there. So the embedded clitic contains information about where (and from what) the embedded pronoun received semantic licensing, while the matrix clitic contains information about its syntactic licensing in the proleptic construction. Since the two copies contain different information, they both must be pronounced.

Since the distinction between movement-based and base-generated prolepsis ultimately rests on the particular nominal licensing strategies in Lubukusu, we should expect cross-linguistic variation along the lines of what types of nominals can be licensed in what position, and what that licensing strategy is.

3 Acquaintance Relations

There are still several questions left to address, however. The obligatory binding relationship between base-generated proleptic objects and the embedded pronoun is so far unexplained, as is the identical interpretation for all three types of prolepsis.

An important fact on the way to addressing these issues is that proleptic objects must always be read transparently (Quine 1956; Salzmann 2006; to appear).

(28)

Context:

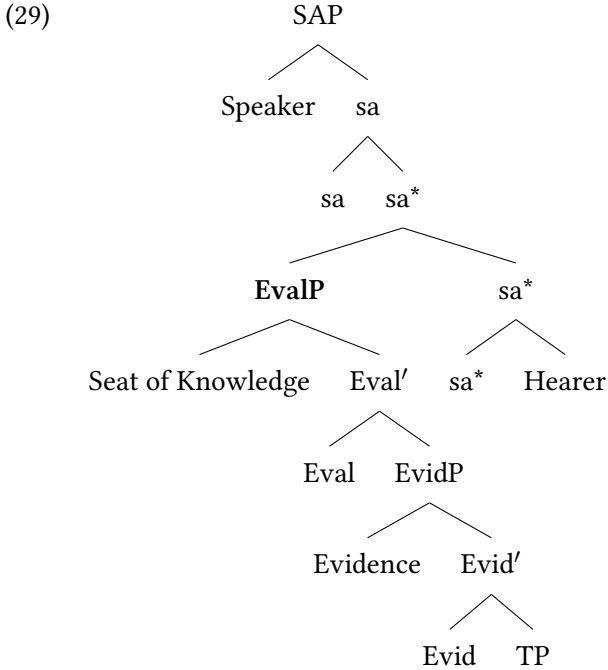
Bill is walking down the street. He glances down a dark alley and sees a man in a trench coat talking into his watch. Bill, who reads too many thrillers, immediately thinks to himself “That man is a spy.” In reality, the man in the alley is Bill’s friend Wayne, although Bill didn’t recognize him.

- a. # Bill thinks of Wayne_i that he_i is a spy.
- b. Bill thinks that Wayne is a spy.

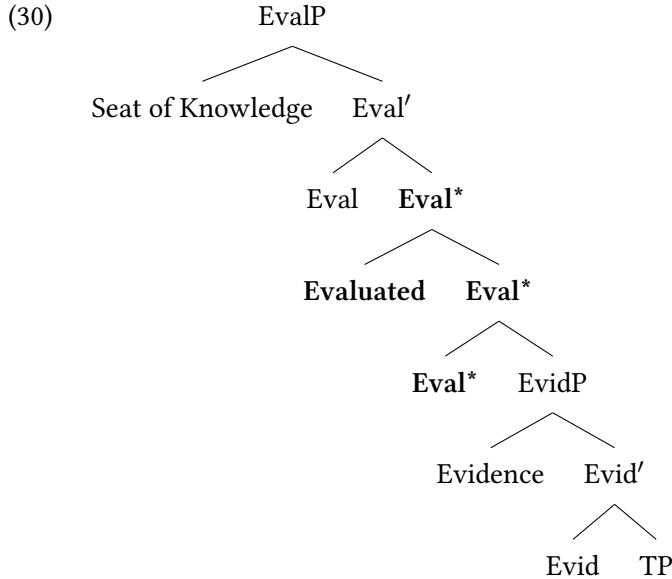
Saying that the embedded clause is “about” the proleptic object is not sufficient to account for this data. The matrix attitude holder has to **knowingly ascribe** the embedded predicate to the proleptic object, and properly identify the proleptic object as well.

The framework I will use to account for these facts is from Speas & Tenny (2003). They propose a set of projections in the left periphery to account for various perspectival phenomena. The projections include a Speech-Act Phrase (SAP), Evaluative Phrase (EvalP), and Evidential Phrase (EvidP). The projections host various null nominals that have a perspectival semantics, and can both bind

embedded pronouns and be bound by higher nominals to force coreference. A sketch of their left periphery is in (29).

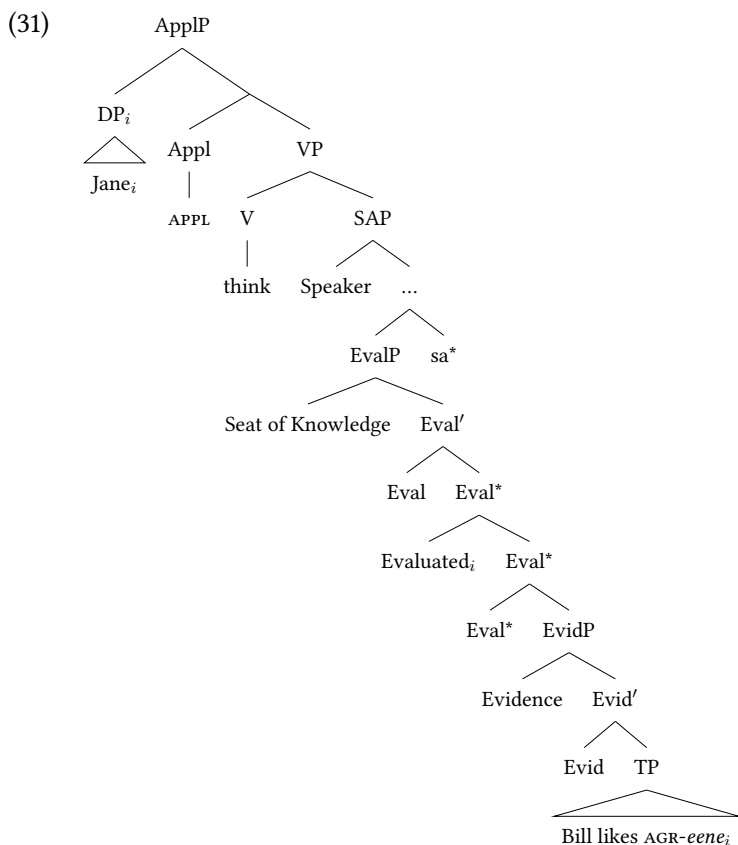


All of these positions are inherently perspectival, however. Accordingly, they won't work for a proleptic object (which doesn't even have to be sentient). But within their system, there is space to add one more position, for an **evaluated object**. Speas & Tenny derive an extended SAP by head movement of the speech act head. The same movement can apply to the evaluative head, creating an additional position for the evaluated object. Rather than having a perspective-taking semantics, the evaluated object can be non-sentient, so long as it is the object perceived by the seat of knowledge evaluating the embedded propositional content.



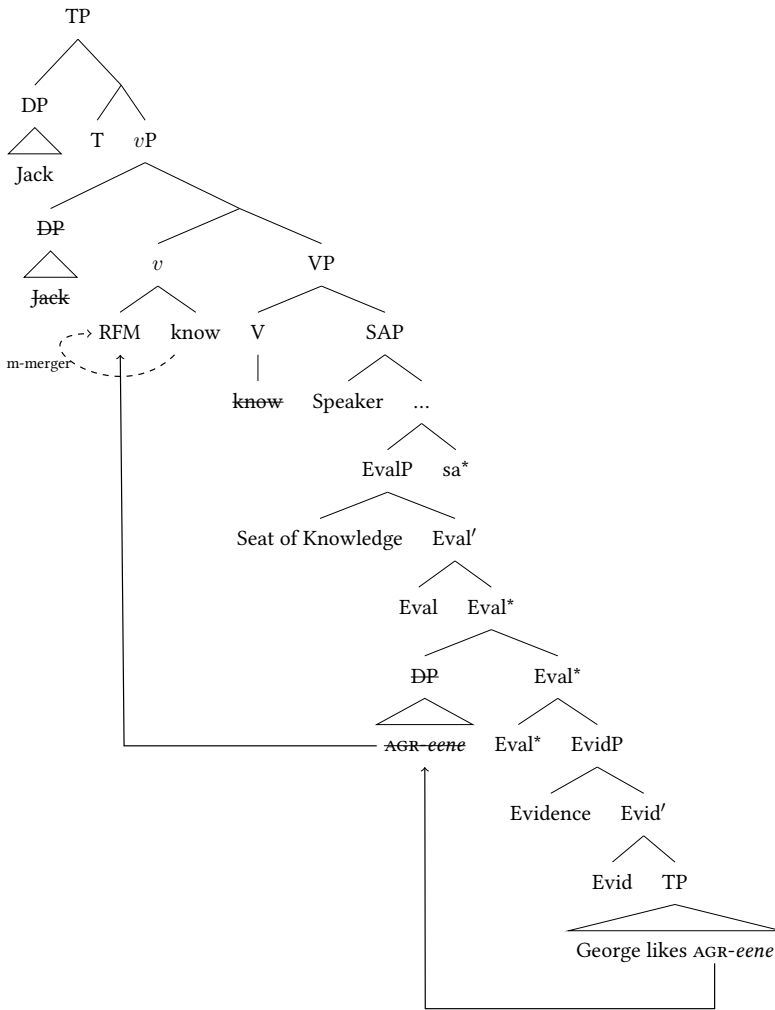
In base-generated prolepsis, the evaluated object binds the embedded *AGR-ene*, and in turn the evaluated object is bound by the proleptic object in the matrix clause. Therefore the modified tree for (2b) is in (31).

- (2b) John a-kanakan-il-a Jane_i a-li Bill a-mu-siim-a
 John SM.c1-think-APPL-FV Jane c1-that Bill SM.c1-OM.c1-like-FV
o-mu-eene_i/niye_i
 c1-c1-own/her
 ‘John thinks of Jane_i that Bill likes her_i.’



The movement-based prolepsis construction is much as it was before, but now we can pinpoint the left-peripheral location that serves as an escape hatch for the moved pronoun: it passes through the site of the evaluated object, and thereby receives its proleptic semantics. Then *AGR-eene* moves further upward to cliticize to the matrix verb for its syntactic licensing.

(32)



Since both constructions involve the same projection in the left periphery, they get the same interpretation from the Eval head. Despite their disparate syntax, a common left periphery allows them to get the same attested semantics.

4 Cross-linguistic predictions

Turning our attention to other languages, we can see that the difference between movement-based and base-generated prolepsis is how the nominal in the matrix clause is syntactically licensed, and whether that licensing position is eligible

for movement or base-generation. For Passamaquoddy (Bruening 2001), Innu-Aimûn (Branigan & MacKenzie 2002), and Tsez (Polinsky & Potsdam 2001), agreement can reach to the CP domain and license the nominal there. But the nominal can only surface in the matrix clause if it is licensed by an agreeing matrix verb. If the verb surfaces in the non-agreeing (TI) voice, the nominal must stay in-situ, and there is no topicality:

- (33) Innu-Aimûn (Branigan & MacKenzie 2002):
- a. N-uî-tshissenit-en tshetshî mûpishtâshkuenit **kassinu**
1-want-know-TI if visited-2/INV every
kâuâpikueshit.
priest
'I want to know if every priest visited you.'
 - b. *N-uî-tshissenit-en_i [**kassinu kâuâpikueshit**]_i tshetshî mûpishtâshkuenit.

For Middle Dutch the matrix nominal is licensed by case marking, but on the analysis van Koppen, Seuren & de Vries (2016) it's in spec,CP, although it hasn't been moved there. In German, prolepsis often feeds further movement that would otherwise be degraded:

- (34) Von [**welchem Maler**]_i glaubst du, dass Maria **ihn**_i mag?
of which.DAT painter think.2SG you that Mary him like.3SG
'Of which painter_i do you think that Mary likes him_i?'

If prolepsis is used when A'-movement is degraded, then it comes as no surprise that the proleptic object in those constructions would not be moved into that position, since movement out of the embedded clause is impossible in the first place. And similar to the base-generation strategy in Lubukusu, the complement of a preposition is not an eligible landing site for A'-movement. If German only licenses extra matrix clause nominals with a preposition, then those extra nominals will necessarily be base-generated there. Once again, the particulars of a given language condition which of the movement and base-generation strategies are available, and under which circumstances.

These considerations bring to the fore an important distinction between semantic and syntactic licensing. Semantically, the evaluated object head provides a viable semantic interpretation for the extra matrix nominal, so long as the context supports that interpretation. Thus the left periphery is identical in both types of structure. The syntactic licensing requirements, however, differ by construction (and by language), as independently established. It is precisely these syntactic facts that derive the differences between prolepsis types.

Acknowledgements

Special thanks to Jastino Sikuku, Ken Safir, Mark Baker, Vivian Deprez, and the attendees of Rutgers ST@R and SURGE reading groups. Some of the data for this project are from the Afranaph Project (NSF BCS 1324404).

References

- Baker, Mark & Ruth Kramer. Doubling Clitics are Pronouns: Reduce and Interpret. Unpublished Ms.
- Branigan, Phil & Marguerite MacKenzie. Altruism, \bar{A} -movement, and object agreement in Innu-aimûn. *Linguistic Inquiry* 33(3). 385–407. <http://www.mitpressjournals.org/doi/abs/10.1162/002438902760168545>, accessed 2017-10-14.
- Bruening, Benjamin. Raising to object and proper movement. *ms.*, *University of Delaware*.
- Diercks, Michael & Justine Sikuku. Pragmatic Effects of Clitic Doubling: Two Kinds of Object Markers in Lubukusu. Pomona College and Moi University.
- Matushansky, Ora. Head movement in linguistic theory. *Linguistic inquiry* 37(1). 69–109.
- Polinsky, Maria & Eric Potsdam. Long-distance agreement and topic in Tsez. *Natural Language & Linguistic Theory* 19(3). 583–646. <http://link.springer.com/article/10.1023/A:1010757806504>, accessed 2017-10-14.
- Quine, W.V. Quantifiers and propositional attitudes. *Journal of Philosophy* 53. 177–187.
- Safir, Ken & Justine Sikuku. Lubukusu Anaphora Sketch version 3.1. In Ken Safir (ed.), *Afranaph Website*, <http://www.africananaphora.rutgers.edu/>. <http://www.africananaphora.rutgers.edu/>.
- Salzmann, Martin. Prolepsis. In Martin Everaert & Henk van Riemsdijk (eds.), *The Wiley Blackwell Companion to Syntax*. John Wiley & Sons, Inc.
- Salzmann, Martin. *Resumptive prolepsis: a study in indirect A'-dependencies*. OCLC: 84611597. Utrecht: LOT.
- Speas, P & C Tenny. Configurational properties and point of view roles. In A.-M. Di Sciullo (ed.), *Asymmetry in Grammar*, 315–344. Amsterdam: John Benjamins.
- Van Koppen, Marjo, Lucas Seuren & Mark de Vries. The Proleptic Accusative as an exceptional Exceptional Case Marking construction. *ms.*
- Wasike, Aggrey. *The Left Periphery, Wh-in-situ and A-bar Movement in Lubukusu and Other Bantu Languages*. Ithaca, New York: Cornell University Ph.D. 393 pp.