"Say"-complementation: re-analyzing agreeing complementizers as verbs

Travis Major (University of Southern California), Justine Sikuku (Moi University), and Michael Diercks (Pomona College)
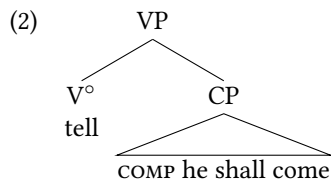
July 19, 2022

Abstract

Across the world's languages, one of the most common clausal complementation strategies involves a form of the verb "say". Most research in this area assumes "say" to be equivalent to the English complementizer "that." The present paper argues that an element in Lubukusu (Bantu, Kenya) that was formerly treated as an agreeing complementizer is actually a stative form of the verb "'say.'" This "say" predicate can take the appearance of a complementizer when it is the head of a clausal adjunct that merges at the VP level of the matrix clause. As a consequence of this analysis, an otherwise complex agreement system is reduced to run-of-the-mill subject-verb agreement both within the matrix clause and within a clausal adjunct. If correct, this proposal advances our understanding of complementation structures and solves a standing agreement puzzle. We introduce a broad range of new Lubukusu patterns relevant to complementation, serialization, and agreement.

1 Introduction

It has long been noted across the world's languages that many languages use some form of the verb "say" in clausal complementation structures.[1] The basic pattern is represented in (1) based on data from Twi (Kwa), where the verb *se* "say" functions as the main verb in (1a) and appears along with the predicate *kyerre* "tell" in (1b).

(1)   a.   o-n-**se**   biribi.
            he-NEG-say something
            'He said nothing.' (Lord, 1993, p. 176, ex: 304a)

       b.   ko ka-kyerre   no   **se**   ommere.
            go speak-show him say he-shall.come
            literally: 'Go, tell him, say, he shall come.' (Lord, 1993, p. 178, ex: 310)

Lord (1993) describes this as a diachronic development by which the verb "say" doubles as a main predicate and a complementizer within the language. Lord cites Riis (1854) as analyzing *se* in cases like (1b) as the verb "say" in a serial verb construction. Despite knowledge that many languages literally use the verb "say" in such structures, most syntactic literature would translate "say" as "that," yielding a translation such as "Go tell him that he should come" for cases like (1b). Given the distributional similarities between "say" elements and complementizers like "that," the most dominant approach in the literature has been to treat "say" elements as verbs that have undergone grammaticalization, resulting in speakers re-analyzing them as complementizers. In other words, most of the literature assumes (roughly) the structure in (2) for a sentence like (1).

(2)

```
              VP
             /  \
           /      \
         V°        CP
         tell     /  \
                /      \
              COMP he shall come
```

For simple cases like (1b), the structure in (2) appears to be sufficient, because "say" occurs in precisely the same position where one would expect a complementizer. However, in languages like Twi, this element similarly occurs in the position where one would expect to see the second verb in a serial verb construction (3).

---

[1] A CC BY license is applied to the Author Accepted Manuscript of this submission.

(3)

```
        VP
       /  \
     VP    XP
  tell (me) /\
          say CP
```
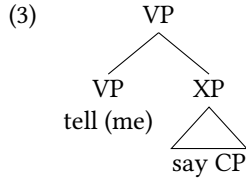
These two analyses make clear predictions with respect to constituency as well as different predictions with respect to semantics. In this paper, we revisit clausal complementation structures in Lubukusu (Bantu, Kenya), as depicted in (4).

(4) Alfredi ne nasi khw-a-lom-a [khu-li] ba-keni b-ool-ile
    1Alfred and I 1PL.SM-PST-speak-FV 1PL-SAY 2-guests 2SM-arrive-PFV

    'Alfred and I said that the guests arrived.' (Diercks, 2010, 298)

These constructions were initially discussed by Diercks (2010, 2013), with subsequent re-analyses appearing in Carstens (2016) and Diercks et al. (2020). All of this previous work treated the agreeing element -li (henceforth AGR-li) as a prototypical complementizer that heads a CP, as in (2), with phi-agreement as the exceptional and theoretically intriguing property. We instead argue in favor of the analysis in (3), in which AGR-li is the final verb in a serialization structure (5).

(5) a. Alfred and I say AGR-li guests arrived.
    b. [[Alfred and I]$_i$ [vP [VP speak (something)] [XP pro_i SAY [guests arrived]]]]

As (5b) shows, we argue that AGR-li should be analyzed as the verbal predicate "say" which agrees in a typical fashion in both main-clause uses and complementizer-like uses. We show that AGR-li 'say' in Lubukusu is a stative predicate and exhibits a subset of the properties of English "say". English "say" can function as either a dynamic activity predicate or a stative predicate. Lubukusu -li is only stative and is thus far more restricted in its functionality.
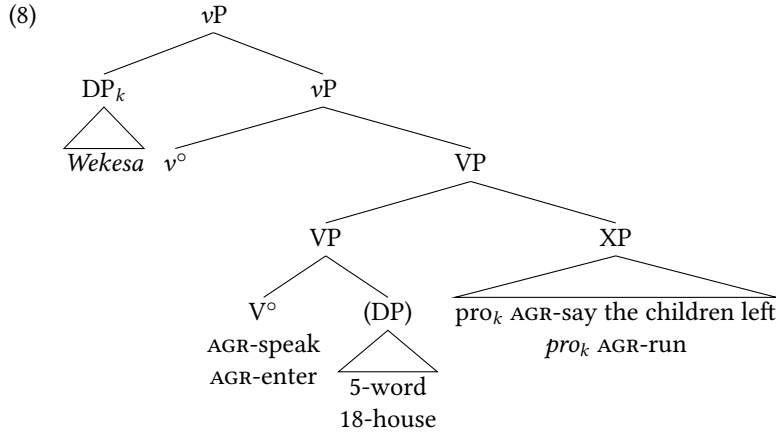
A novel empirical observation (to our knowledge) that this paper contributes to the literature on Lubukusu is that Lubukusu has a type of manner adjunct clause that modifies the matrix VP. We argue that the analysis in (5b) is not specific to complementation structures. Instead, it is the same grammatical mechanism used more generally in manner adjunct clauses, such as (6):

(6) Wekesa [a-kha-] engil-e mu-nju [a-] tim-a
    1Wekesa 1SM-FUT-enter-FUT 18-house 1SM-run-FV

    'Wekesa will enter the house running.'

In cases like (6), the matrix verb "enter" hosts both a subject marker (a) and tense marker (-kha-), while the manner modifying clause is truncated, where "run" only hosts a subject marker. This clause indicates the manner (i.e. "enter the house *running*"). This is the same morphological pattern observed in AGR-li constructions like (7); we will propose that these constructions receive parallel analyses.

(7) Wekesa [a-a-] lom-a li-khuwa [a-] li ba-ba-ana b-a-ch-a
    1Wekesa 1SM-PST-say-FV 5-word 1SM-SAY 2-2-children 2SM-PST-go-FV

    'Wekesa said that the children left.'

Notice that there is a single tense marker, which occurs on the matrix verb -lom- "speak," which is preceded by a SM. The second clause displays -li, which we argue to be a predicate, which hosts only a SM, similar to "run" in (6) . We propose that the cases in (6) and (7) share the same core syntactic structure, as illustrated in (8).

(8)

vP
DP$_k$ — *Wekesa* — $v°$ — VP
VP — XP
V° — (DP) — pro$_k$ AGR-say the children left
AGR-speak / AGR-enter — 5-word / 18-house — *pro$_k$* AGR-run

Our proposal (as sketched in (8)) is substantially different from prior accounts that have assumed *-li* to be an agreeing complementizer that heads a CP that is selected by embedded verbs (or nouns). Our analysis predicts a range of empirical effects that are not predicted on a complementizer analysis of AGR-*li*:

(9)   Predictions of a verbal analysis of Lubukusu AGR-*li*
    1.   AGR-*li* clauses do not form a constituent with N or V; instead, they form a constituent with VP (§2)
    2.   AGR-*li* can occur in unselected environments (does not need to be a complement clause) (§2.2)
    3.   AGR-*li* can occur as a matrix stative verb (§3)

In this paper we show that these predictions are upheld.

In §2, we tackle Prediction 1, which demonstrates that AGR-*li* clauses are not selected by V or N even in contexts involving an apparent CP-selecting verb or noun. This is especially clearly indicated in §2.2, showing that Prediction 2 is upheld. §3 demonstrates that Prediction 3 above holds, with the caveat that it requires rethinking the syntax (and lexical semantics) of "say". We demonstrate that *-li* corresponds to a subset of the uses of "say" in (e.g.) English. More specifically, we suggest it does not encode a dynamic speech event, but instead introduces the communicated content introduced by the subject, which is an underexplored property of English "say". The fact that Prediction 3 holds (i.e. *-li* can function as a main verb) is a particular example of Prediction 2, which also emphasizes the validity of the verbal analysis of these morphemes. In addition to this, §4 demonstrates that AGR-*li* clauses are VP modifiers. This includes detailed discussion of cases like (6) and cases where AGR-*li* exhibits similar behavior. §5 offers some relevant analytical and theoretical discussion that extends the conclusions of the previous sections.

There are several implications of this paper for both our understanding of Lubukusu and the theory of complementation and "complementizer" agreement. First, it provides a simpler explanation for the Lubukusu agreement facts than previous proposals (Diercks, 2010, 2013; Diercks et al., 2020; Carstens, 2016). Based on our analysis, all agreement is reduced to run-of-the-mill subject-verb agreement, making the Lubukusu facts compatible with most standard formulations of Agree (Chomsky, 2000, 2001). Second, it contributes to the typology of languages with "say" complementation structures, offering additional argumentation in favor of a verbal analysis, in line with analyses of Vata (Koopman, 1984), Abe (Koopman, 1989), Avatime (Major and Torrence, 2020; Major, 2021a), Fon Kinyalolo (1993), Nilo-Saharan (Driemel and Kouneli, 2020), Turkish (Ozyıldız, 2017), Uyghur (Major, 2021a,b, To Appear). Although there are precedents for this kind of approach it has been far more common for researchers to assume "say" elements to be complementizers roughly akin to English "that," given that they are generally translational equivalents on English "that" and also perform a very similar syntactic function (introducing an embedded clause).[2] Finally, this study has implications for the status of "say" as a semantically bleached verb. These overt "say" elements in languages like Lubukusu inspired (e.g.) Kratzer (2016) to posit a silent SAY modal in the left periphery of some English complement clauses or a silent light verb akin to "be" or "do" for Grimshaw (2015). Our findings are compatible with

---

[2]There has been considerable discussion of apparent agreeing complementizers in recent years. Diercks and Rao (2019) show that a similar pattern occurs in the southern Nilotic language Kipsigis, Idiatov (2010) shows a similar pattern in some Mande languages, and Torrence (2016) shows a parallel construction in Ibibio. The pattern is most widely documented for Bantu languages, including Chokwe, Luchazi, Lunda, and Luvale (central Bantu languages; Kawasha 2007), Ikalanga (southern Bantu; Letsholo and Safir 2019), and Kinande (Baker, 2008).

the main points brought up by Kratzer (2016) and Grimshaw (2015), but we motivate a more fine-grained syntactic structure.

2   Reasons to think that *AGR-li* clauses are not selected by V or N

We first demonstrate that a traditional analysis of clausal complementation (10) does not adequately account for (11).

(10)
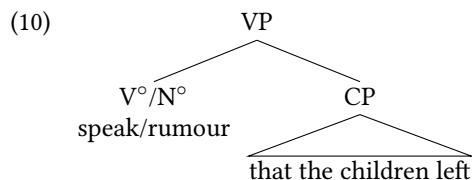
The structure that we introduced above in (8) shows stark differences with the traditional complementation structure in (10). Consider these differences in the context of the example in (11), repeated from (7):

(11)   Wekesa   a-a- lom-a   li-khuwa a- li   ba-ba-ana   b-a-ch-a
       1Wekesa 1SM-PST-say-FV 5-word   1SM-SAY 2-2-children 2SM-PST-go-FV

       'Wekesa said that the children left.'

On the approach we advocate for here, *AGR-li* is not an argument of a predicate, but rather is itself a predicate that introduces an argument. We can see this differentiation in (11), where there is a DP argument of *-loma* 'say,' but the *AGR-li* clause is nonetheless present.[3] (11) at least suggests that the *AGR-li* clause can combine with a VP that has already saturated its internal argument slot. Each of the subsections in this section offer reasons to think that the *AGR-li* constituent is not selected by a predicate.

2.1   "Scream" is unergative; *-li* takes an internal argument

The predicate *-jokel-* "scream" is not inherently a communicative predicate; it is an unergative activity predicate. The example in (12a) contains a question with *-jokel-* "scream;" this shows that *siina* "what" is interpreted as "what for" when there is no argument position to be "questioned."[4]

(12)   a. Nafula   a-n-jokel-el-a   siina.
          1Nafula 1SM.PST-1SG.OM-scream-APPL-FV 7what

          'What did Nafula scream at me for?' i.e. 'Why did Nafula scream at me?'

       b. Nafula   a-n-jokel-el-a   a-li   siina.
          1Nafula 1SM.PST-1SG.OM-scream-APPL-FV 1SM-SAY 7what

          'What did Nafula scream at me, saying?' (answer = *Wafula stole the red cow*)

Under an analysis of *AGR-li* as a complementizer, the contrast in (12) is mysterious. The English equivalent to (12b) with "that" is ungrammatical: **What did Nafula scream that ~~what~~?*. But (12) is expected if *AGR-li* is analyzed as the verbal element "say," and therefore *AGR-li* is what is responsible for introducing the communicative aspect of the event.[5] So instead, in a context like (12b), the matrix clause indicates that Nafula made loud noises directed at 'me,' while the *AGR-li* clause indicates that it was communicative and what was communicated. This is suggestive that *AGR-li* plays a more substantive role than just that of a complementizer, instead itself introducing the concept of communication and an argument (the content of that communication).

---

[3]We will address the alternative analysis (that this is a noun complement clause) in §2.4.

[4]This same alternation is observed across manner of speech predicates (e.g. "cry" or "whisper"). More generally, the interpretation of *siina* covaries with transitivity. For transitive predicates, *siina* is most naturally construed as questioning the internal argument of the predicate. For intransitives, *siina* is interpreted as a reason question.

[5]Kratzer (2016) suggests a silent "say" modal does similar work in English.

2.2 Unselected *AGR-li* on a purpose/reason reading

An additional argument that *AGR-li* is not a complementizer heading a complement clause is that *AGR-li* clauses can appear in environments where they are unselected. In (13), *AGR-li* bears no apparent selectional relationship with the matrix predicate but is acceptable nonetheless. The *AGR-li* clause in (13) receives a purpose/reason reading (in this instance, the reason that Wafula went outside).

(13) Wafula a-kha-ch-a   a-nje   a-li   a-nyol-e   e-m-beo.
      1Wafula 1SM-PFV-go-FV 16-outside 1SM-SAY 1SM-get-SUBJ 9-9-air

      'Wafula went outside to get fresh air.' (lit. '…saying he would get fresh air')

A sentence like (13) is very difficult to explain on an analysis that *AGR-li* is a declarative-embedding complementizer akin to English *that*. This is far less surprising under the present "saying" analysis, which is equally grammatical in English (i.e. "Wafula went outside saying he would get fresh air").

Notably, (13) contains no factive presupposition (in contrast to English *because*); instead, the *AGR-li* clause introduces the stated/purported reason, rather than (necessarily) the true reason. (14) demonstrates this lack of a factive presupposition with a purpose-clause use of *AGR-li*.[6]

(14) Wekesa a-ch-a   mu-n-ju   mw-ewe a-li   a-lw-iile,   nekakhali eny-a
      1Wekesa 1SM.PST-go-FV 18-9-house 18-his   1SM-SAY 1SM-be.sick-PFV but   1SM.want-FV
      a-lol-e   e-TV.
      1SM-watch-SUBJ 9-TV

      'Wekesa went to his house saying he was sick/tired, but (I know that) he wanted to watch TV.' (not a contradiction)

Again, this suggests that the unselected *AGR-li* is simply describing the stated rationale for the main-clause action, not a presuppositional reason adjunct clause. So in effect, as we will show, *AGR-li* here is serving a very similar role to what it serves in instances of apparent complementation.

2.3 *AGR-li* is obligatory to introduce communicative content

Another argument that suggests that *AGR-li* is more than a simple complementizer is that it is obligatory in almost all environments where a clausal complement is introduced. This clausal complement can represent either a direct or an indirect speech report. In other words, a case like (15) is ambiguous between a direct or indirect speech report.[7]

(15) a. Watulo a-a-loma   *(a-li) ba-ba-ana b-a-ch-a
        1Watulo 1SM-PST-say-FV 1SM-SAY 2-2-children 2SM-PST-go-FV

        'Watulo said that the children left.'
        'Watulo said "the children left."'
     b. Watulo a-n-jokel-el-a   *(a-li) ba-ba-ana b-a-ch-a.
        1Watulo 1SM.PST-1SG.OM-scream-APPL-FV 1SM-SAY 2-2-children 2SM-PST-go-FV

        'Watulo screamed to me, saying the children left.'
        'Watulo screamed to me, saying, "the children left."'

Complementizers like English "that" are not able to introduce direct quotations in English, but they are able to introduce indirect speech:

(16) a. He said (*that), "I'm going".
     b. He said (that) I'm going.
        (Munro, 1982, p. 302, ex: 1)

---

[6]Similar facts have been reported for Turkic (Ozyıldız, 2017; Major, 2021a), Mongolic (Bondarenko, 2020), and beyond.

[7]In earlier work on this issue, there was not careful control of whether a predicate/circumstance was specifically introducing communicative content or not, so (for example) Diercks (2010) lists a null complementizer as simply another option alongside *AGR-li*.

Notice that when the complementizer is absent, however, the verb "say" is able to introduce either direct or indirect quotation. Under our analysis, it is unsurprising that *AGR-li* introduces direct and indirect quotation, because it is a general property of "say" cross-linguistically. In other words, it is not "that" at all; it is a verb, for which reason it does not distribute like "that".

There is a small set of other predicates that can directly take a finite CP complement such as "want." We are currently investigating whether there are any syntactic/semantic differences introduced by the presence/absence of *-li*. But as (17) shows, *enya* "want" allows but does not require *AGR-li*.[8]

(17) Watulo eny-a   (a-li)   ba-ba-ana   ba-ch-e
     1Watulo 1SM.want-FV 1SM-SAY 2-2-children 2SM-go-SUBJ
     'Watulo wants the children to go.'

2.4 Apparent Noun Complement Clauses

Both *-loma* 'speak/say' and *AGR-li* can independently function as matrix verbs, with different selectional properties. (18a) shows that *-loma* can take a DP such as *bubeyi* 'lie' as its complement, but it is incapable of selecting for the finite CP "the children left," shown in (18b).

(18) a. Wekesa a-a-lom-a   bu-beyi
       1Wekesa 1SM-PST-say-FV 14-falsehood
       'Wekesa told a lie.'
    b. *Wekesa a-a-lom-a   ba-ba-ana   b-a-ch-a
       1Wekesa 1SM-PST-say-FV 2-2-children 2SM-PST-go-FV
       Intended: 'Wekesa said the children left.'

AGR-li 'say' is able to introduce a clausal complement like "the children left" (19a) in a matrix usage, unlike *-loma*. (We explore the matrix use of *AGR-li* in depth in §3.2.) But *AGR-li* 'say' is unable to introduce a DP complement like 'lie' as its complement (19b), again unlike *-loma* (18a).

(19) a. Wekesa a-li   ba-ba-ana   b-a-ch-a
       1Wekesa 1SM-SAY 2-2-children 2SM-PST-go-FV
       'Wekesa said the children left.'
    b. *Wekesa a-li   bu-beyi
       1Wekesa 1SM-SAY 14-falsehood
       'Wekesa told a lie.'

Consider the full sentence containing both a *-loma* 'speak/say' clause and *AGR-li* 'say' clause (20), where each introduces a different kind of argument. *-loma* 'speak/say' introduces the DP *bubeyi* 'lie,' and *AGR-li* 'say' introduces the finite CP or a wh-expression (*siina* 'what') that corresponds to that finite CP.

(20) Wekesa a-a-lom-a   bu-beyi   *(a-li)   ba-ba-ana   b-a-ch-a
     1Wekesa 1SM-PST-say-FV 14-falsehood 1SM-SAY 2-2-children 2SM-PST-go-FV
     'Wekesa told a lie, saying the children left.'

There are two possible analyses here: First, these may be consecutive verb phrases, each with verbal elements taking internal arguments. Second, this could be a noun complement clause (NCC): a nominal *bubeyi* 'lie' which takes a CP complement headed by *AGR-li*. Diercks (2010, 2013) uncritically adopted the NCC analysis for such examples, seeking an explanation for how a complementizer embedded within a DP can agree with the matrix subject. The linear order above is identical to a noun-complement construction in English, which understandably could lead to one positing that this is a case where the noun "lie" takes a CP complement, as is traditionally assumed for English.

---

[8]Complements to any predicate other than *AGR-li* cannot be construed as direct quotes.

Recall (12), which demonstrated that 'what' receives two distinct interpretations, depending on whether it is directly embedded under 'scream' ('why did they scream?') or under *AGR-li* ('what did they scream?'). The same kind of alternation is observed in these constructions that resemble N-Comp constructions, where "what" can be embedded under *AGR-li* (21a) or directly under "lie" (21b).

(21) a. Wekesa a-a-lom-a   bu-beyi   a-li   siina
         1Wekesa 1SM-PST-say-FV 14-falsehood 1SM-SAY 7what

      'What did Wekesa say, telling a lie?' (Answer: *that the children left*)

     b. Wekesa a-a-lom-a   bu-beyi   siina
         1Wekesa 1SM-PST-say-FV 14-falsehood 7what

      'What lie did Wekesa tell?' (Answer: the one he heard from Wafula)

Notice that when embedded under *AGR-li* (21a), *siina* corresponds to the content communicated by Wekesa (i.e. what was said). When *AGR-li* is dropped, as in (21b), the question refers to the particular lie that is under discussion. If this were simply an instance of complementizer drop, we would not expect an interpretive difference at all.[9] We claim that the verbal analysis better captures the full range of facts, as we will make clear in what follows.

2.5 *AGR-li* clauses modify the matrix VP

The verb *-reeba* "ask" can select a DP internal argument such as *lireeba* 'question' (22a), but it cannot directly embed an interrogative CP (22b). (22c) illustrates that *AGR-li* 'say' is able to directly introduce the question, "Where did Nafula go?". Finally, (22d) demonstrates that the "ask" and "say" clauses can combine to form an embedded question. This follows straightforwardly from a VP-modifier analysis—*AGR-li* clauses do not require an open argument slot from the matrix predicate to be merged into the structure.

(22) a. Wekesa a-a-reeb-a   li-reeba
         1Wekesa 1SM-PST-ask-FV 5-question

      'Wekesa asked a question.'

     b. *Wekesa a-a-reeb-a   Nafula a-ch-a   wae(ena)
         1Wekesa 1SM-PST-ask-FV 1Nafula 1SM.PST-go-FV where

      Intended: 'Wekesa asked where Nafula went.'

     c. Wekesa a-li   "Nafula a-ch-a   wae(ena)?"
         1Wekesa 1SM-SAY 1Nafula 1SM.PST-go-FV where

      'Wekesa said/asked, "Where did Nafula go?"'

     d. Wekesa a-a-reeb-a   li-reeba   a-li   "Nafula acha   wae(ena)?"
         1Wekesa 1SM-PST-ask-FV 5-question 1SM-SAY 1Nafula 1SM.PST-go-FV where

      'Wekesa asked a question saying, "Where did Nafula go?"'

It is surprising that (22d) is grammatical if we assume *-li* to be a complementizer. For example, **Wekesa asked a/the question that/whether the children left* is ungrammatical regardless of the complementizer that is chosen in English. If we assume that both *-reeba* "ask" and *-li* "say" are verbal, both of which independently introduce internal arguments (*-reeba* 'ask' selects *lireeba* 'question' and *AGR-li* 'say' selects "Where did Nafula go?"), the grammaticality of (22d) is far less surprising.

(23a) shows another construction where the first verb takes "word" as its complement and where *-li* takes the clausal complement "the children left". (23b) demonstrates that the complement of *-li* can be wh-questioned to the exclusion of *-li* itself.

---

[9]Under a Noun-Complement analysis of (21a), one could argue that it should induce a Complex Noun Phrase Constraint violation (Ross, 1967). However, Wasike (2007) illustrates that wh-objects can be extracted from adjuncts or complex noun phrases with marginal acceptability. Given this marginal acceptability, attributing (un)grammaticality to islandhood is confounded.

(23) a. Wafula a-a-lom-a   li-khuwa a-li   ba-ba-ana   b-a-ch-a
          1Wafula 1SM-PST-say-FV 5-word   1SM-SAY 2-2-children 2SM-PST-go-FV
          'Wafula said the news that the children left.'

      b. Wafula a-a-lom-a   li-khuwa a-li   siina?
          1Wafula 1SM-PST-say-FV 5-word   1SM-SAY 7what
          'What was the news that Wafula told?' (lit. 'What was the word Wafula said?')

On a complementizer analysis of *AGR-li*, it is wholly unexpected that a *wh*-word commonly used for nominal(ized) elements could replace the complement of C°. If *-li* is a "say" verb, this wh-question simply targets the internal argument of *-li*. This would predict the meaning difference observed above (i.e. (23b) requests the content that was 'said').

2.6 Distinct subjects of matrix predicate and *AGR-li*

There are some environments where the agreement on *AGR-li* does not match the phi-features on the matrix subject. For the second author, agreement with a non-subject source is acceptable. (24) is an instance of *-ulila* 'hear,' which specifies a source of the information. In this instance, agreement with the information source (and *not* the matrix subject) is acceptable.

(24) N-a-ulil-a   khu-Wekesa a- li ba-ba-ana   b-a-ch-a
      1SG.SM-PST-hear-FV 17-1Wekesa 1-SAY 2-2-children 2SM-PST-leave-FV
      'I heard from Wekesa that the children left.'

Likewise, even an information source that is pragmatically accessible but not explicitly mentioned in the matrix clause can trigger agreement on *AGR-li*.

(25) a. N-a-ulil-a   o- li   w-a-siim-a   ka-ma-beele
          1SG.SM-PST-hear-FV 2SG-SAY 2SG.SM-PST-like-FV 6-6-milk
          'I heard that you like milk.'
          (*source = addressee: context must make this clear for sentence to be acceptable*)

      b. n-a-ulil-a   a- li Wekesa   a-a-siim-a ka-ma-beele
          1SG.SM-PST-hear-FV 1-SAY 2SG.SM-PST-like-FV 6-6-milk
          'I heard that Wekesa likes milk.'
          (*source = Wekesa: context must make this clear for sentence to be acceptable*)

This is in contrast with what was reported by Diercks (2010, 2013), where agreement with a non-subject source of reported information was documented as being ungrammatical. It's not clear to us whether this difference from what is reported by Diercks (2013) is a genuine grammatical difference, or if the pragmatic conditions were not set appropriately for the consultants at that point. If the latter (which is what we suspect), the previous unacceptability should have been marked as "#" instead of "*" (we understand the conditioning environments better now). For the most part, the second author's judgments align with the patterns reported in Diercks (2013), so these are not obviously distinct grammars. We consider it likely that Diercks (2013) was not controlling for the pragmatic constraints in the way that we have here: it requires pragmatic precision to get agreement with the non-matrix subject. For example, even though examples like (24) and (25) are acceptable, is not as simple as specifying *AGR-li* as always "agreeing with the source of information," because even for the second author, agreement with a passive by-phrase is ungrammatical, as was also reported by Diercks (2010, 2013). Instead, the context must be established such that there is a single event of information transfer and the particular manner of transfer is communicated by the *AGR-li* phrase. We discuss the dissociation of sources in 'hearing' verbs from passive by-phrases in §5.2 (the former can trigger agreement on *AGR-li*; the latter cannot).

Nonetheless, (24) and (25) make clear that it is possible for *AGR-li* to agree with something that is not the matrix subject, and even not represented in the matrix clause at all, in contrast to what has been generally reported/assumed about *AGR-li*. In these instances, the subject of the hearing predicate is distinct from the subject of the 'saying' predicate *AGR-li* (on the analysis that we set forward here).

3   Lubukusu *-li* is stative "say"

The previous section provided argumentation that AGR-*li* should not be treated as equivalent to standard clausal complementation structures. This section probes deeper into what the internal structure of AGR-*li* clauses actually are, illustrating that this element is not semantically vacuous and even occurs in environments that lack an alternative, clause-selecting predicate, which we first saw in §2.4 above.[10] (26a-26b) below both demonstrate that AGR-*li* bears strong resemblance to the predicate "say" in English, which relates content to its source, whether the subject is animate (26a) or inanimate (26b).

(26) a. Wekesa a-li   ba-ba-ana   b-a-ch-a
          1Wekesa 1SM-SAY 2-2-children 2SM-PST-leave-FV
          'Wekesa says that the children left.'

      b. e-barua e-li   ba-ba-ana   b-a-ch-a
          9-letter 9SM-SAY 2-2-children 2SM-PST-leave-FV
          'The letter says that the children left.'

The goal of this section is to demonstrate that *-li* is a stative form of the verb "say," which allows us to understand its role in more complex structures when it combines with another predicate (e.g. "tell," "scream," or "ask"). First, in §3.1 we show that a range of recent research argues for an approach that proposes an abstract light verb SAY that introduces communicative content. We then show in §3.2 that Lubukusu AGR-*li* shows a large number of the same properties, arguing for a similar treatment of English "say" and Lubukusu AGR-*li*.

(27)
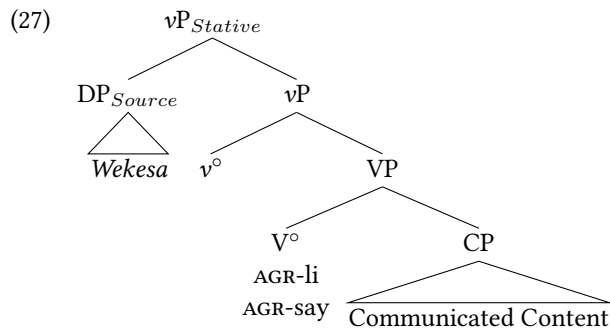
Due to AGR-*li* being a stative predicate, we predict that this structure should lack some of the properties of verbs that encode a dynamic event, such as "tell," "ask," or "scream." We show that this is the case for AGR-*li* with respect to argument structure (e.g. no goal is permitted), event structure, negation, and tense.

3.1   "Say" in English

Before turning to the Lubukusu facts, we first demonstrate that stative "say" exists in English and exhibits similar properties to Lubukusu AGR-*li*. The difference is that English allows "say" to encode either stative or eventive semantics in most environments, whereas Lubukusu AGR-*li* is unambiguous. Recent investigations of English "say" demonstrate that although the prototypical representation most speakers assume for "say" treats it as a transitive predicate that encodes a dynamic speech event involving an Agent (the speaker), some content (what was said), and often a Goal (the person the speech was directed at), this is not the only representation. Grimshaw (2015) introduces the following subcategorizaton frame for this proto-typical use of "say" (28).

(28)   The teacher said to the students that the exam was easy. (Grimshaw, 2015, 80:2)

In (28), Grimshaw refers to "the teacher" as an Agent, "the students" as a Goal, and the clausal complement "that the exam was easy" as *Linguistic Material* (henceforth LM). We adopt the following characterization of LM:

---

[10]For similar patterns in languages with similar say-complementation, see also Kawasha (2007), Diercks and Rao (2019), and Diercks et al. (2020).

(29) **Linguistic Material**: Direct quotation or any argument that can be substituted with direct quotation. (Major 2021a, interpreting a concept from Grimshaw 2015)

Referring back to (28), the clausal complement is a LM argument because it can be replaced by direct quotation (i.e. "The teacher said to the students, 'the exam was easy'."). We use this term because direct quotation is possible in all (and only in) *AGR-li* environments. For present purposes, LM can be construed as content resulting from a communicative act (or an abstract nominal anaphorically referring to the communicated content, e.g., 'a few words' or 'nonsense').

Grimshaw argues that the presence of a LM argument serves as evidence for the presence of the "light verb" SAY. The main verb "say" is proposed to be the overt realization of the light verb SAY, which otherwise goes unpronounced as a result of it incorporating into another predicate, such as "scream". In other words, cases such as (30) are underlyingly (31), but "say" is only pronounced when no other predicate is merged into the structure.

(30) The teacher screamed, "The exam was easy."

(31) The teacher screamed SAY "The exam was easy."

Grimshaw observes one other property that is crucial to this paper; "say" sometimes functions as a stative verb when it occurs with certain inanimate subjects, such as "sign" or "poster," which she refers to as *Locations*, as illustrated in (32). More specifically, rather than "the sign" functioning as an Agent in (32), it instead indicates the Location of LM.

(32) The sign/poster/book/article said that the park was closed.

Grimshaw demonstrates that cases like (32) are stative by illustrating incompatibility with the progressive aspect (33a) and Goals (33b).[11]

(33) a. ??The sign/poster/book/article was saying that the park was closed.
b. ??The sign/poster/book/article said to the tourists that the park was closed.
(Grimshaw, 2015, 87: ex 30)

Major (2021a) argues that "say" can function as a stative verb even in contexts with animate subjects, but it is often hard to observe due to the ease at which the dynamic (activity) reading of "say" can be coerced when the subject is a viable agent, which is not the case for the "signs" examples above. Major (2021a) demonstrates that the use of present tense to report a previous speech act is an instance where stative properties of "say" emerge. First, note that it is possible to use present tense form of "say" to describe an instance of communication that happened in the past.[12]

(34) I ran into Kayla yesterday. She says/said that she is coming to the conference tonight.

Here, context makes clear that the communication occurred in the past, but this occurs naturally either with past tense or present tense on "say." Without delving into how the present tense form is possible here, relevant to our point is that these tense differences show distinctions with respect to the eventivity of the predicate. The availability of manner adverbials and goal arguments are both typical properties of eventive predicates (and not of stative predicates). As (35) illustrates, both manner adverbials and goal arguments are unacceptable in this present tense usage of "say." Finally, this stative use of "say" closely resembles the behavior of other stative attitude predicates, such as 'think,' 'believe,' and 'know' (35c).

---

[11]Major and Stockwell (2021) demonstrate that "say" is also compatible with an expletive subject ("It says 'the park is closed' on the sign."). Major and Stockwell argue that Agents are introduced by an Agent Voice head and that Locations are introduced by Holder Voice heads (Kratzer, 1996), while expletives are introduced when there is no Voice head merged.

[12]Major (2021a) demonstrates that this pattern needs to be distinguished from the so-called "narrative present" in English: see that work for details.

(35) I saw Katie yesterday.

    a. She (#cheerfully) says (#to me) that Kayla will be visiting!

    b. She cheerfully said to me that Kayla will be visiting!

    c. She says/thinks/believes/knows that Kayla will be visiting!

It appears, then, that the use of present tense "say" to report a previous communicative event is a stative use. As is evident in (36), these stative instances of "say" lack a range of typical eventive properties beyond what is noted in (35a).

(36) I met Katie for the first time yesterday.

a.	She says that Kayla will be visiting!	
b.	# She **cheerfully** says that Kayla will be visiting!	*No manner modification*
c.	# She says **to me** that Kayla will be visiting!	*No goal*
d.	# It **is said** that Kayla will be visiting!	*No passivization*
e.	# She **doesn't say** that Kayla will be visiting!	*No negation*
f.	# She says **the news (that Kayla will be visiting)**!	*No content nouns*

These judgments can be tricky: "say" can be construed as dynamic under historical present, generic, or habitual readings, but the diagnostic context in (36) does help to isolate the properties we are concerned with here. The lead-in sentence (36) largely mitigates these concerns, because it is most natural under a reading in which there was one unspecified communicative act carried out by "Katie," which makes a habitual reading highly unlikely. It is not embedded within a narrative, which eliminates the narrative present reading. Our interpretation is that "say" has a stative form, like other attitude verbs (e.g. (35c)). Unlike other attitude verbs, the same root can be construed as an activity predicate. For this reason, coercion is often simple if the context does not restrict it.[13] Most important for our current concerns is that we show that the list of restrictions presented in (36) almost entirely matches the set of restrictions on AGR-*li* in Lubukusu. In this way, we will suggest that AGR-*li* is the lexicalization of a stative version of 'say'.

3.2 Matrix AGR-*li*

As we mentioned above in (26), repeated here in (37), -*li* "say" can function as the sole predicate of sentence, which can occur with an external argument that is animate (37a) or inanimate (37b). Just as stative "say" in English presupposes a communicative act but does not directly denote the act itself, the same is true in Lubukusu.

(37) a. Wekesa a-li   ba-ba-ana b-a-ch-a
         1Wekesa 1SM-SAY 2-2-children 2SM-PST-leave-FV

         'Wekesa says that the children left.'

    b. e-barua e-li   ba-ba-ana b-a-ch-a
         9-letter 9SM-SAY 2-2-children 2SM-PST-leave-FV

         'The letter says the children left.'

Unlike English, Lubukusu AGR-*li* corresponds to truncated, stative "say" even when the subject is a capable agent. In other words, the contexts where AGR-*li* is felicitous are precisely those where the communicated content and the source are introduced independent of the actual communicative act (e.g. screaming/telling).

(38) *Context: The speaker says, "Hey Wekesa—Nafula screamed some lies at me yesterday. You won't believe it … "*

a-li   Wafula eb-a   e-khaafu e-m-besemu
1SM-SAY 1Wafula 1SM.PST.steal-FV 9-cow   9-9-red

'She said that Wafula stole a red cow!'

---

[13]We refer the interested reader to Major (2021a) for a discussion of these eventive properties of verbs of communication, including interesting subpatterns regarding nominals that contain linguistic material as opposed to content nouns.

In any context where one refers to the communicative act itself (because it cannot be presupposed from context), a second predicate (e.g. *-lom-* "say/tell") is required:

(39) Q: w-a-ch-a   mw-i-duka   khubela si(ina)?
       2SG.SM-PST-go-FV 18-9-bodega reason   7what
       'Why did you go to the store?'

      A: Nafula   #(a-lom-a)   a-li   eny-a   ka-ma-bele
       1Nafula 1SM.PST-speak-FV 1SM-SAY 1SM.PST.want-FV 6-6-milk
       'Nafula said that she wants milk.'

When *AGR-li* is used as the sole predicate, it is restricted to contexts where the communicative content holds present relevance and the communicative act is either irrelevant or already salient in the discourse.

3.3 Evidence for *AGR-li* as a stative 'say'

To illustrate that Lubukusu *AGR-li* exhibits the same properties as stative "say" in English, we show that *AGR-li* cannot take manner modification or a goal argument, undergo passivization, occur under negation, or take a content nominal as complement. In following with Major and Stockwell (2021) and Major (2021a), we attribute these restrictions to the absence of the syntactic structure responsible for introducing an agent/agentive semantics and eventive semantics.[14]

3.3.1 No manner modification

As we saw for the English stative uses of 'say,' *AGR-li* is also incompatible with manner modification (40).[15]

(40) Wekesa **a-li**   [(*kalaa)] ba-ba-ana   b-a-ch-a
      1Wekesa 1SM-SAY slowly   2-2-children 2SM-PST-go-FV
      'Wekesa said (*quietly) that the children left.'

Lexical predicates of communication that combine with *AGR-li* often encode manner of speech; in these instances, manner modification becomes possible:

(41) Wekesa **a-a-lom-a**   [kalaa] **a-li**   ba-ba-ana   b-a-ch-a
      1Wekesa 1SM-PST-say-FV slowly   1SM-SAY 2-2-children 2SM-PST-go-FV
      'Wekesa said quietly that the children left.'

3.3.2 *AGR-li* cannot take a Goal argument

Unlike "say" as an activity predicate, stative "say" does not encode a communicative act. For this reason, a Goal argument to which the communicative act was directed cannot be introduced unless a verb denoting the communicative act is present (either dynamic "say" or another predicate). Under our proposal for *AGR-li*, it corresponds to only stative "say" and thus cannot host applicative morphology (i.e. introduce a Goal):

(42) *Wekesa a-n-li-li   ba-ba-ana   b-a-ch-a
      1Wekesa 1SM-1SG.OM-say-APPL 2-2-children 2SM-PST-go-FV
      attempted: 'Wekesa said to me that the children left.'

However, when *AGR-li* combines with another predicate that denotes a communicative event, that predicate can bear applicative morphology and introduce a Goal argument (43):

---

[14]We believe our approach is compatible with many modern approaches to the syntax of event structure, such as Ramchand (2008). For our purposes, we simply assume that only the lowest VP projection is introduced in Lubukusu *AGR-li* constructions.

[15]No alternative positioning of the manner adverb makes the sentence in (40) acceptable with the manner adverb: **Wekesa ali kalaa …; *Wekesa kalaa ali …; *kalaa Wekesa ali … .*

(43) Wekesa a-a-m-bol-el-a   a-li   ba-ba-ana   b-a-ch-a
      1Wekesa 1SM-PST-1SG.OM-say-APPL-FV 1SM-SAY 2-2-children 2SM-PST-go-FV
      'Wekesa said to me (i.e. told me) that the children left.'

We interpret this as additional evidence that AGR-*li* is a stative predicate with a restricted distribution, in contrast to other lexical predicates which contribute the manner of a communicative event.

3.3.3 AGR-*li* cannot be passivized

Again similar to English, despite its matrix verbal usage AGR-*li* cannot be passivized. Expletive constructions in Lubukusu typically take a class 6 subject marker; (44) is what an impersonal passive of AGR-*li* would look like if it were possible (it is not).

(44) *ka-l-w-a babaana ba-a-ch-a
      6SM-say-PASS-FV 2-2-children 2SM-PST-go-FV
      Attempted: 'It was said that the children left.'

As with these other properties, this continues to replicate the English facts on the stative present-tense use of "say" that we reported in (36).

3.3.4 AGR-*li* cannot be negated

Despite its available use as a main clause predicate, AGR-*li* cannot be negated on its own.

(45) *Wekesa se-a-li   ba-ba-ana ba-ch-a   ta.
      1Wekesa NEG-1SM-SAY 2-2-children 2SM-go-FV NEG
      Attempted: 'Wekesa didn't say that the children left.'

This is again reminiscent of English use of "say" present tense forms, which is also restricted in the same way (see (36e) above). This property is not tethered to AGR-*li*'s status as a stative predicate, as there is no prohibition on negation of stative predicates more generally. However, given that matrix AGR-*li* requires a salient communicative act in the discourse and is responsible for reporting the content of that act, negating it would result in the content that was not asserted as part of the relevant communicative act, which is entirely uninformative. This is roughly equivalent to English (36e).

3.3.5 Nominal complements to AGR-*li*

As discussed for English "say" in Major (2021a), AGR-*li* cannot select a so-called content nominal as its complement (e.g. rumor/myth/story), unlike most attitude predicates (see Moulton, 2009).[16] "say" is able to select nominals that stand in for CPs (46a), not that occur in addition to CPs (46b).

(46) a.   Kayla said a few words/something interesting (*that the children left).
      b. * Kayla said the rumor/myth/story (that the children left).

Notice the contrast below between AGR-*li* and -*loma* in (47): AGR-*li* is able to directly take a propositional DP that stands in for communicated content (*mbao*), but -*loma* cannot (on its own).

(47) a. Wekesa a-li   mbao   si-layi ta
         Wekesa 1SM-say nothing 7-good not
         'Nafula says nothing interesting.'

---

[16]The taxonomy of nominals selected by clause-selecting predicates has received some attention. The types of nominals introduced by predicates like "say" are deemed *Words-NPs* in Moltmann (2017). Elliott (2016) introduces a relevant discussion between the explanans and explanandum readings of the relevant complement types. We leave this discussion to future research.

b. *Nafula a-lom-a   mbao   si-layi ta
   1Nafula 1SM.PST-speak-FV nothing 7-good not

   Intended: 'Nafula said nothing interesting.'

On the other hand, if *a-li* is present in the structure, it is possible to introduce the propositional DP in a construction with the matrix predicate *-loma*; this follows if *mbao* is in fact an argument of AGR-*li* in this instance.

(48) Nafula a-lom-a   a-li   mbao   si-layi ta
     1Nafula 1SM.PST-speak-FV 1SM-say nothing 7-good not

     'Nafula said nothing interesting.'

On the other hand, acceptability flips for content nouns such as *bu-beyi* 'lie,' which has content associated with it but does not stand in for content itself (in the terms of Major 2021a, it is not a Linguistic Material Nominal).

(49) a. *Wekesa a-li   bu-beyi
        Wekesa SM1-say 14-lie

        Intended: 'Nafula says the lie.'
     b. Nafula   a-lom-a   bu-beyi
        1Nafula 1SM.PST-speak-FV lie

        'Nafula said the lie.'

The primary conclusion to take from these data is that the purpose of AGR-*li* is to link communicated content to its source, which aligns with stative "say" in English. It indicates that content was communicated and does not introduce any information about the manner in which the communicative act was carried out. In order to describe the manner, the discourse participants, or other aspects of the communicative act itself, a different predicate is needed.

3.3.6 Brief Summary

We've seen that there is an apparent near-synonymity between clauses that solely use AGR-*li* to express a communicative event and those that use a predicate communicating the manner of the communication, like *-loma* 'speak' together with AGR-*li*.

(50) a. Nafula   a-li   eny-a   ka-ma-bele
        1Nafula 1SM.PST-speak-FV 1SM-SAY 1SM.PST.want-FV 6-6-milk

        'Nafula said that she wants milk.'
     b. Nafula   a-lom-a   a-li   eny-a   ka-ma-bele
        1Nafula 1SM.PST-speak-FV 1SM-SAY 1SM.PST.want-FV 6-6-milk

        'Nafula said that she wants milk.'

As the evidence noted above made clear, however, there are stark differences in the properties of (matrix) AGR-*li* and other predicates like *-loma*, as summarized in (51).

(51) AGR-*li* vs *-loma* "say"/"tell" (and others)

	AGR-*li*	*-loma* 'speak'
Agreement	✔	✔
Matrix use	✔	✔
Introduces LM	✔	✗
Takes Goal	✗	✔
Manner modification	✗	✔
Negation	✗	✔
Tense/Aspect	✗	✔
Content Noun (e.g. *rumor*)	✗	✔
Indefinite (e.g. *something*)	✗	✔

Our claim is that *-loma* 'speak' is an eventive predicate taking an agent and describing the manner of communication and it is therefore compatible with the eventive properties noted above (Goals, manner modification, negation, etc.). AGR-*li*, in contrast, is not. So despite its availability for use in a matrix context, AGR-*li* does not share the same full distribution as most lexical predicates of communication.

One property of AGR-*li* that distinguishes it from other verbs is its inability to inflect for tense/aspect/mood.

(52) Wekesa a-li   ba-ba-ana   b-a-ch-a
      1Wekesa 1SM-SAY 2-2-children 2SM-PST-go-FV
      'Wekesa said that the children left.'

AGR-*li* in its non-matrix use is compatible with any tense/aspect, in principle. When AGR-*li* combines with another predicate, that predicate can be inflected for tense/aspect. We offer further discussion about these restrictions in 4.2.

3.4   Notes on typology of "say"

Munro (1982) discusses the morpho-syntax of "say" cross-linguistically, showing that "say" often behaves (morpho-syntactically) like an intransitive predicate cross-linguistically. This is often shown with respect to object marking and case (e.g. Samoan does not allow ergative case with "say," in contrast to other predicates), and "say" often is restricted with respect to the complements it can take (especially nominal complements). For instance, transitive verbs in Chickasaw generally require their objects to be marked with the morpheme *-ã*, which is unavailable when "say" introduces quotation (53).

(53)   a.   "Ihoo"   (*-ã) aachi
            woman OBJ   say
            'He said, "woman."'(Munro 1982, 303)
       b.   "Hilha" (*-ã) aachi
            dancing OBJ   say
   'He says, "She's dancing."' or 'He says that she's dancing.'(Munro 1982, 303)

In Cahuilla, third person singular objects trigger the agreement prefix *pe-* on the verb (54a)-(54b), but this is not true of "say" (54c):

(54)   a.   Pe-n-'ayaw-qa mansaana-y.
            it-I-want-pres apple-obj
            'I want an apple.' (Munro 1982, 306)
       b.   Pe-n-'ayaw-qa hen-hichi-ka.
            it-I-want-pres I-go-incomp
            'I want to go.' (Munro 1982, 306)
       c.   Ni-ya-qa   "Hen-hichi-ka"
            I-say-PRES I-go-INCOMP
            'I say, "I'm going."' (Munro 1982, 306)

Our interpretation of Munro's findings is that the alternation between eventive/dynamic "say" and stative "say" is realized morpho-syntactically in some languages. The same may serve as an explanation for similar restrictions in other languages (e.g. Tigrinya (Spadine, 2020), Ewe (Clements, 1975; Spadine, 2020), Amharic (Ruth Kramer, p.c.), Avatime (Major and Torrence, 2020; Major, 2021a)). Beyond these observations regarding "say" predicates, "say" predicates are often used to introduce complement-clause-like material cross-linguistically, as we showed in the introduction for Twi. We argue that Lubukusu *-li* lexicalizes stative "say," not only in the matrix usages commented on above, but also in the instances where it appears to be serving the function of a subordinating complementizer. We discuss how this happens in the next section.

4 How does *AGR-li* combine with other predicates?

We have shown that *AGR-li* exhibits properties of an independent predicate, suggesting that the previous treatments of *AGR-li* as a canonical complementizer (that bears agreement) were on the wrong track. Under the present analysis, *AGR-li* is a stative "say" verb that introduces communicative content and its source. We assume the structure of an *AGR-li* clause that is shown in (55).

(55)

X°
khu-
1plsM

v°

pro$_k$

V
SAY
-li

LM
ba-keni b-ool-ile
the guests arrived
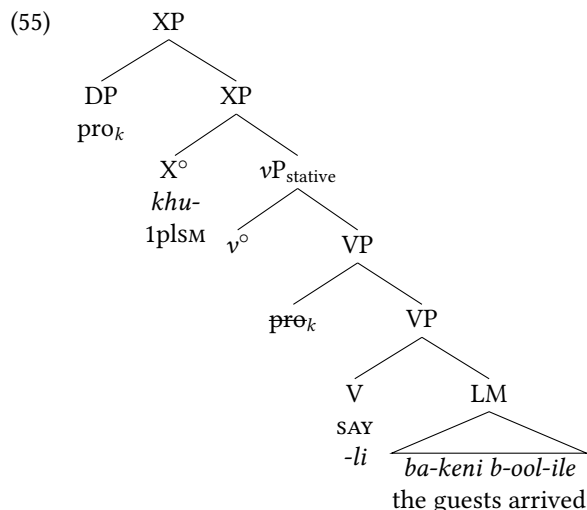
Following Borer (2004), Folli and Harley (2007), Ramchand (2008), and Rothmayr (2009) (among others), we assume states to involve a truncated structure compared to eventive structures.[17] However, notice that the structure above does not explain how the *AGR-li* clause (XP above) combines with another predicate superficially resembling clausal complementation structures (e.g. English "that" or Spanish "que" clauses). As we've discussed previously, we argue that these XPs are not actually complements to V (or N), but are instead modifiers, which is in line with recent literature on English and beyond (Aboh, 2005; Bochnak and Hanink, 2021; Bondarenko, 2020; Elliott, 2020; Kayne, 2014; Kratzer, 2006, 2016; Major, 2021a; Moulton, 2009; Ozyıldız, 2017; Potts, 2002).

The structure we propose shares properties with recent analyses of a series of unrelated languages, such as Washo (Bochnak and Hanink, 2021) and Buryat (Bondarenko, 2020). Unlike Washo, however, Lubukusu does not restrict *AGR-li* clauses to unergative predicates (it occurs with 'ask,' "tell," etc.). Unlike Buryat, we suggest that *AGR-li* introduces a CP; it does not head a CP. We argue that *AGR-li* XPs (55) adjoin to the matrix VP where it indicates (roughly) the manner of the matrix VP, as illustrated in (56).

(56)
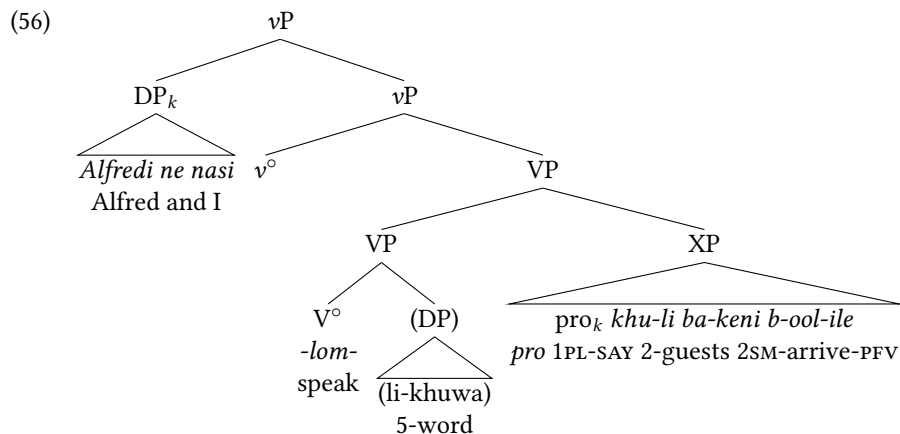
It may seem surprising that an XP containing an agreeing verbal element could freely merge as a VP adjunct,

---

[17]The structure in (55) should be translatable into any of the aforementioned theories without affecting the core contributions of our analysis.

but we illustrate that this is a general process in the language. Agreeing verbal elements can merge as manner modifiers more generally in the language; these manner modifiers are morpho-syntactically identical to *AGR-li* clauses.[18] We motivate the link between verbal adjunct structures and our analysis of *AGR-li* clauses in the next section.

4.1 Truncated verbal adjunction

We analyze *AGR-li* as equivalent adjunct clauses to (57), which are highly restricted manner modifiers of the matrix VP.

(57) [Alfredi] engil-a   mu-nju [a-] tim-a.
     1Alfred 1SM.enter-FV 18-house 1SM-run-FV

     'Alfred entered the house running.'

The form of these truncated adjunct clauses is SM-verb.root-a; the subordinated predicate has extremely limited morphological structure, bearing no tense, aspect, negation, etc. As we will see below, the subordinated predicate is interpreted as having a tight interpretive link to the main predicate, e.g. manner modification or a contrastive counterpoint (e.g. *Wekesa ate rice despite hating it*).

We give some more empirical detail about these constructions below, but to start, the analysis that we assume for these constructions is sketched in (58). We will refer to these as 'truncated verbal adjunction' structures, because the adjunct clause is itself highly truncated, and we assume low (manner) adjunct status of the structure.[19]

(58)
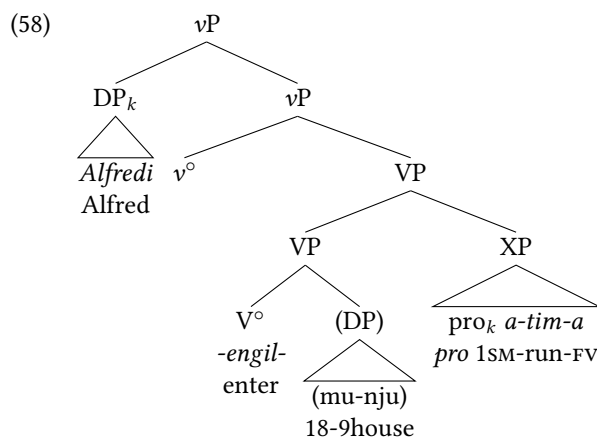
Within the XP adjunct clause we assume a highly truncated structure that consists only of a *v*P and the XP projection that generates the SM, as we showed above in (55). Therefore there is a single T in the sentence (matrix T). Due to its status as a manner adjunct, the entire resulting complex predicate is interpreted relative to matrix T (and Asp).

4.1.1 Truncated adjunction as manner modification

There is a clear distinction between truncated adjunct clauses and tensed adjunct clauses, shown by comparing the truncated adjunct clause in (59) with the tensed adjunct clause in (60):

(59) Wekesa a-kha-engil-e   mu-nju a-tim-a
     Wekesa 1SM-FUT-enter-FUT 18-house 1SM-run-FV

     'Wekesa will enter the house running.'

---

[18]The distribution and function of *AGR-li* XPs is represented morphologically as (say + converb) in Buryat (Bondarenko, 2020) and Uyghur (Major, 2021a) and as a clause chaining suffix in Washo (Bochnak and Hanink, 2021).

[19]This analysis is roughly equivalent to manner modifying converbial constructions in Uyghur and nuclear serial verb constructions in Avatime (Major, 2021a).

(60) Wekesa a-kha-engil-e   mu-nju [$_{TP}$ a-kha-tim-e   ]
Wekesa 1SM-FUT-enter-FUT 18-9house   1SM-FUT-run-FUT

'Wekesa will enter the house and then will run.'

(59) is necessarily construed as a manner modifier of the matrix clause that temporally overlaps with it, necessarily describing a single event. (60) encodes two separate temporally related events, interpreted sequentially.

The bare truncation strategy is only available when the XP modifier can be construed as the manner in which something happened. In (61) only *dancing* (61a) can be considered an actual way of entering the house. A similarly simultaneous activity of greeting people in (61b) is nonetheless not itself the manner of entering, and is less acceptable using bare truncation, requiring the NE subordinator.

(61) a. Wafula engil-a   mu-nju (??n-)a-khin-a   ka-ma-beka
1Wafula 1SM.enter-FV 18-house (??NE-)1SM-dance-FV 6-6-shoulders

'Wafula entered the house dancing.' (*dancing shoulders = a particular kind of dance*)

b. Wafula engil-a   mu-nju ??(n-)a-khesi-a   buli mundu
1Wafula 1SM.enter-FV 18-house ??(NE-)1SM-greet-FV every 1person

'Wafula entered the house (while) greeting everyone.'

4.1.2   Same restriction on Truncated Adjunction as for *AGR-li*

Central to our proposal here, the bare truncation strategy for manner adjunct clauses is subject to similar kinds of restrictions that we discussed above for *AGR-li*. For example, it is impossible to have goals/applicatives added to a bare truncated adjunct clause:

(62) *Alfred engila   munju a-tim-a   khu-ese
1Alfred 1SM.enter-FV 18-house 1SM-run-FV 17-me

Intended:'Alfred entered the house running toward me.'

(63) a. Wafula a-a-khin-il-a   Nafula
1Wafula 1SM-PST-dance-APPL-FV 1Nafula

'Wafula danced for Nafula.'

b. *Wafula engil-a   mu-nju a-khin-il-a   Nafula
1Wafula 1SM.enter-FV 18-house 1SM-dance-APPL-FV 1Nafula
Wafula entered the house dancing-for Nafula

Likewise, there is no negation with truncated adjunct clauses; any attempt to do so requires the NE- subordinator, a distinct clausal adjunction strategy:

(64) Wafula engil-a   mu-nju *(ne)-a-kha-khin-a   ta
1Wafula 1SM.enter-FV 18-house NE-1SM-NEG2-dance-FV NEG

'Wafula entered the room while not dancing.'

(65) *Wafula engil-a   mu-nju se-a-khin-a   ta.
1Wafula 1SM.enter-FV 18-house NEG-1SM-dance-FV NEG

Intended:'Wafula entered the house not-dancing.'

Therefore, we see significant overlap in the properties of truncated adjunct clauses and *AGR-li*; it is our claim that this is due to them being instances of the same grammatical construction, a truncated clause adjoined as an instance of manner modification.[20]

---

[20]We have yet been unable to diagnose whether it is possible to perform manner modification inside an truncated adjunct clause: to show whether this is possible (or not) would require a construction where the attempted manner modification inside the truncated adjunct clause could not be possibly interpreted as manner modification of the main clause. Given that the truncated adjunct clauses are themselves manner modifiers of the main clause, it is not clear if there is ever such a context that would allow us to diagnose this question.

(66) Truncated adjunct clauses are restricted in ways similar to AGR-*LI*:

	Manner XPs	AGR-*li*
Agreement	✔	✔
Takes applicative/associative	✗	✗
Negation	✗	✗
Tense	✗	✗
Manner adverb	unclear	✗

4.2   An overview of clausal adjunct modifiers

Cross-linguistically, the morphology involved in "say" complementation structures uses the (or a) syntactic mechanism used more generally in the language to link clauses. In some languages, this involves no overt morphology, as is the case for languages with serial verb constructions. In other languages, this is accomplished via an overt linker (Turkic/Monglic converbs, switch reference markers, etc.). In Lubukusu, the most restricted clause-linking mechanism (the truncated verbal adjunction strategy discussed above) is used to combine an AGR-*li* clause with the matrix clause as a manner modifier. In this structure, both the adjunct clause and the matrix VP are embedded under the tense/aspect/mood of the matrix clause, which restricts the adjoined predicate to the same spatio-temporal specifications as the matrix clause.

Let's begin with the particle NE. The morphologically defining feature of NE-adjunction is that the the verb of the modifying clause displays *n-/ne-/na-* subordinating morphology as a prefix to the verbal form. Whereas truncated adjunct clauses are highly restricted, NE-adjunction can link any two clauses that can occur simultaneously (essentially equivalent to English *while*-constructions).

(67) Wekesa osiy-a   e-nyama n-a-nyw-a   ka-ma-lwa
      1Wekesa 1SM.PST.roast-FV 9-meat   NE-1SM-drink-FV 6-6-beer
      'Wekesa roasted meat while drinking beer.'

Notice here that 'drinking beer' is not a conceivable manner in which 'roasting meat' is done. Thus while enforcing simultaneity, NE-adjunction does not impose the same restrictions on event structure or clausal relations as truncated verbal adjunction.

Turning to more *and*-like conjunctions, the coordinators *mala/lundi* allow any two clauses to be conjoined regardless of their relationship or event structure (68).

(68) Wafula ech-a   likolooba mala / lundi a-kh-ach-e
      1Wafula 1SM.PST.come-FV yesterday CONJ / CONJ 1SM-FUT-go-FUT
      'Wafula came yesterday and will leave.'

As in many other languages, particularly languages with serial verb constructions (e.g. Baker, 1989; Collins, 1997), there is a clausal coordination construction in Lubukusu equivalent to (68) in all senses other than the fact that the conjunction is covert (69). In these structures, it is possible for the clause-linking to occur below T (at roughly Asp), but there is remains no requirement that the clauses be closely related, able to link 'coming' and 'leaving,' for example.

(69) a. Wafula ech-a   a-a-ch-a
          1Wafula 1SM.PST.come-FV 1SM-NAR-go-FV
          'Wafula came and (then) went.'
      b. Wafula a-kh-ech-e   a-ch-e
          1Wafula 1SM-FUT-come-FUT 1SM-go-SUBJ
          'Wafula will come and (then) leave.'

Again, none of the cases above are as restricted as truncated adjunct clauses, which are restricted in size (only AGR+VP) and function (only manner modification).

It is worth discussing examples like (69) in more depth in order to distinguish them from our discussion of truncated verbal adjunction; we will refer to these as narrative clause chains. These clause chains bear many of the hallmarks of clausal adjunct modifiers, with the exception that they contain independent mood marking. This mood-marking is the only morphological difference between truncated adjunct clauses and narrative clause chains. Predicates in these clauses are either in a subjunctive form or in what we refer to as the *narrative tense*.[21]

(70) a. SM-a- verb.root-a:
      Narrative tense is used for past tense events that have actually occurred at the narrative time (71a).

   b. SM-verb.root-e :
      Subjunctive is used for sequences of future (irrealis) events unrealized at the narrative time (71b).

This is illustrated in (71), where (71a) is past tense (and the chained clause uses the narrative -a-) and (71b) is a future tense construction where the chained clause uses the subjunctive.

(71) a. Wekesa   osiy-a   e-nyam-a   a-a-(ki-)ly-a
         1Wekesa 1SM.PST.roast-FV 9-meat   1SM-NAR-9OM-eat-FV
         'Wekesa roasted meat and (then) ate it.'
      b. Wekesa   a-kh-osiy-e   e-nyama   a-(ki-)ly-e
         1Wekesa 1SM-FUT-roast-FUT 9-meat   1SM-9OM-eat-SUBJ
         'Wekesa will roast meat and eat (it).'

There is no overt conjunction in these examples, but there is also no tense specified after the first verb. The interpretation of non-initial, mood-marked predicates is dependent on the tense established by the first predicate (or the discourse in some cases). This is not unlike the narrative present in English, where the chain of events is obligatorily interpreted as a progression (for overview and discussion, see Anand and Tooservandani, 2020; Anand and Toosarvandani, to appear; Pancheva and Zubizarreta, 2020). We suggest that the cases in (71a) and (71b) correspond to (72) and (73) respectively.

(72) So Wekesa walked in, roasts the meat, puts it on a platter, eats some of it, and gives us the rest.

(73) Wekesa will walk in, roast the meat, put it on a platter, eat some of it, and then give us the rest.

These cases differ from NE-adjunction: with NE-adjunction, simultaneity is required. The examples in (71) can be compared with the NE-form in (74), wherein the roasting and the eating have to be simultaneous:

(74) Wekesa   osiy-a   e-nyama   n-a-ki-ly-a
      1Wekesa 1SM.PST.roast-FV 9-meat   NE-1SM-9OM-eat-FV
      'Wekesa roasted meat while eating it.'
      (*i.e. roasting it and eating bits as he went*)

Given that truncated adjunct clauses are obligatorily simultaneous, yet lack NE, and are unable to host mood marking, we argue that truncated adjunct clauses are a highly restricted type of clause-linking mechanism. Notably, these are all properties of *AGR-li*, and therefore we argue that it is precisely this mechanism that is implemented in *AGR-li* constructions.

4.3   Truncated verbal adjunction and *AGR-li*

Recall from §2, that we argue that *AGR-li* clauses function as manner modifiers. In a case such as 'Wekesa scream + say the children left,' we suggest that the "say" (*AGR-li*) clause describes the manner in which the screaming was carried out. In other words, 'Wekesa screamed in such a way as to say that the children left.' Truncated clausal adjuncts are all restricted in this way. The cases below can only be construed such that running describes manner, not as a sequence of events.

---

[21]The narrative tense here is the same form that Safir et al. (2020) previously analyzed as an 'actual' clause.

(75) Wekesa engil-a   mu-nju   a-tim-a
1Wekesa 1SM.PST.enter-FV 19-9house 1SM-run-FV
'Wekesa entered the house running.'

(76) Wekesa a-kha-engil-e   mu-nju   a-tim-a
1Wekesa 1SM-FUT-enter-FUT 18-9house 1SM-run-FV
'Wekesa will enter the house running.'

When narrative tense is prefixed to the second predicate, the interpretation changes (77).

(77) Wekesa a-kha-engil-e   mu-nju   a-a-tim-a
1Wekesa 1SM-FUT-enter-FUT 18-9house 1SM-NAR-run-FV
'Wekesa ran when he entered the house.'

(78) a. Wafula e-cha   e-e-khala.
          Wafula 1SM-enter 1SM-NAR-sit
          'Wafula entered and sat down.'
      b. ??Wafula e-cha   e-khala.
          Wafula   1SM-enter 1SM-sit
          'Wafula entered as he sat (e.g. kept sitting in the process of entering).'

If we take AGR-*li* clauses to be a truncated verbal adjunction clauses it enables us to make some predictions about the distribution of AGR-*li* clauses. Just as 'running' is able to better describe the manner in which one 'enters' than 'sitting,' the same should hold of AGR-*li* clauses. In other words, the matrix VP that encodes the communicative event has to be compatible with the communicated content.

(79)
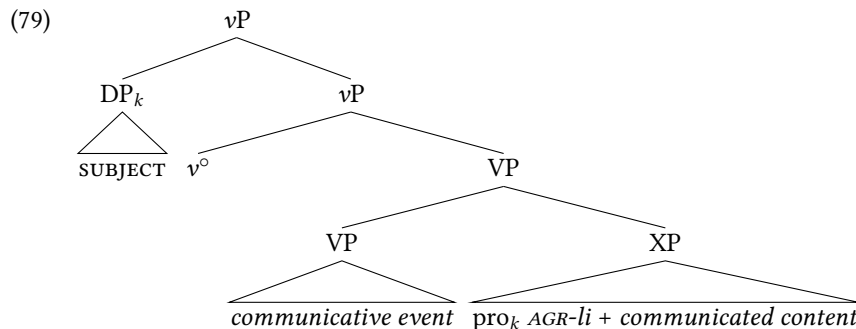
Because we argue that AGR-*li* is a stative predicate, the communicative event need not involve the physical production of sounds. If this were the case, we might only expect it to occur with predicates that describe (for example) manner of speech, such as 'scream'. This is clearly not the case, as we have shown that AGR-*li* can occur with "say," "tell," "ask (a question)," "scream," "go outside," and on its own as the main predicate.

What we are suggesting here is not far from the pattern we observe for English. First, 'go outside' does not take a clausal complement:

(80) *Wekesa went outside that he would get fresh air.

However, it is compatible with a manner modifier, including a "saying" modifier:

(81) a. Wekesa went outside {saying he was leaving, dancing, screaming}.
      b. Wekesa cried saying he was sorry.
      c. Wekesa silently sat in the corner saying he was sorry.
      d. Wekesa was thinking (about it) saying he was sorry.

As mentioned throughout this paper, even English "say" is a semantically lightweight predicate. For this reason, it is possible to use "say" as a manner modifier with a wide range of unrelated predicates, as shown above. It is uncontroversial that there is variation with respect to the lexical semantics of speech/attitude predicates. For this reason, we suggest it is not unreasonable that in a language such as Lubukusu, "say" is semantically lighter than English, allowing *AGR-li* to occur in an even wider range of environments.[22]

4.4 *AGR-li* modifies complex VPs

If we begin with a verb like *-lila* 'cry' in (82a) without *AGR-li*, it is not construed as a communicative predicate at all. This is encoded only by the presence of *AGR-li*, as in (82b). We suggest this should receive a paraphrase, such as 'Wekesa cried in such a way that he communicated his mother passed away'.

(82) a. Wekesa a-a-lil-a
      1Wekesa 1SM-PST-cry-FV

      'Wekesa cried.'

   b. Wekesa a-a-lil-a   *(a-li)   maayi   a-a-tib-a
      1Wekesa 1SM-PST-cry-FV 1SM-SAY 1mother 1SM-PST-get.lost-FV

      'Wekesa cried, saying his mother passed away. (lit. got lost, fig. died)'

We suggest that this is the same reason that purpose/reason readings are possible with *AGR-li* clauses, such as (13), repeated below as (83), where Wafula's going outside indicates ('says') to the speaker that he would get fresh air.

(83) Wafula a-kha-ch-a   a-nje   a-li   a-nyol-e   e-m-beo.
      1Wafula 1SM-PFV-go-FV 16-outside 1SM-SAY 1SM-get-SUBJ 9-9-air

      'Wafula went outside to get fresh air.' (lit. '…saying he would get fresh air')

Recall from earlier that there are multiple mechanisms that allow clauses to be chained together as sequences or simultaneous events. In (84), the *AGR-li* clause specifies what was communicated ("said") via a series of actions 'say/speak,' 'cry,' and 'sing.'

(84) Wekesa a-lom-ile   a-a-lil-a   e-emb-a   a-li   maawe a-a-tib-a.
      1Wekesa 1SM-speak-PFV 1SM-NAR-cry 1SM.NAR-sing-FV 1SM-SAY 1mother 1SM-PST-get.lost-FV

      'Wekesa spoke, cried, and sang saying that his mother passed away.' (lit. got lost, fig. died)

In (84) the entire matrix clause has a single TAM marker (the *-ile* perfective), and the interpretation of the *-li* clause is such that there are 'saying/speaking,' 'crying,' and 'singing' events and the result of all of these events is the communication of the message that Wekesa's mother passed away. This arrangement would be extremely surprising if it were the complement of only 'speak' or only "sing." As is evident in (85), the verbs in the clause chaining construction may themselves take arguments, but this is not disruptive to the availability of the *AGR-li* light verb in the chain:

(85) Wekesa a-ulil-a   li-khuwa   a-a-lil-a   eemb-a   lu-lw-imbo a-li   maawe
      1Wekesa 1SM.PST-hear-FV 5-word/news 1SM-NAR-cry-FV 1SM.NAR.sing-FV 11-11-song 1SM-SAY 1mother
      a-a-tib-a.
      1SM-PST-get.lost-FV

      'Wekesa heard the news, cried, sang the song, saying that his mother passed away.' (lit. got lost, fig. died)
      (all interpreted as part of a single grieving event)

---

[22]In other words, we do not find it particularly problematic that a predicate like 'know' in English cannot occur with a 'saying' modifier, while it is perfectly natural in Lubukusu. We assume this variation to be akin to the fact that doxastic predicates like 'know' can occur in the progressive in Japanese, Korean, and Turkic, but not in English. Japanese, Korean, Turkic, and Lubukusu are all more flexible with respect to object drop than English, as well. For this reason, the possibility that there is a silent internal argument to these predicates with the entire VP being modified by the "say" clause is also a possible solution.

As with many patterns mentioned in this paper, this is largely unexpected if *AGR-li* is analyzed as a complementizer heading a complement clause argument of the main verb communication predicate. It is much easier to understand as a modifier of the VP that embeds the entire chain of events that indicates both that there was content communicated as part of this series of events and what that content was. This analysis also accommodates the possibility that, in cases like (84), the communicated content was the result of only the 'singing,' while the 'speaking' and 'crying' are independent, which is a possible interpretation of the string.

4.5   CPs without *AGR-li*

As reported by Diercks (2010, 2013), there are a variety of contexts where CPs are available in Lubukusu but the *AGR-li* construction is not. None of these effects have received wholly satisfactory explanations in the previous work, but all of them follow very naturally if *AGR-li* is a predicate which introduces Linguistic Material, because not all CPs comprise Linguistic Material arguments. Diercks (2010) notes three main restrictions on the distribution of *AGR-li* generally that diverge from other non-agreeing complementizers: non-agreeing complementizers are used in because-clauses (86), if-clauses (87), and with complements of emotive factive predicates (88).[23] In each instance, a non-agreeing complementizer form must be used, and using *AGR-li* is unacceptable.

(86) Mikaeli a-likho a-tekh-a   sy-akhulia sikila ⌈mbo⌉   (*a-li)   a-likho   a-ulil-a   e-njela
      1Michael 1SM-PROG 1SM-cook-FV 7-food   reason COMP/COMP (*1SM-SAY) 1SM-PROG 1SM-hear-FV 9-hunger

      'Michael is cooking because he is hungry.' (Diercks, 2010, 322)

(87) Alfred ka-reb-a   Sammy nekaba ⌈mbo⌉ (*a-li)   ba-keni ba-ach-e
      1Alfred 1SM.PST-ask-FV 1Sammy if   COMP (*1SM-SAY) 2-guests 2SM-go-SUBJ

      'Alfred asked Sammy if the guests left.' (Diercks, 2010, 323)

(88) N-a-beelele   ⌈mbo⌉ (*n-di)   si-n-a-ch-ile   Bungoma ta.
      1sgSM-PST-regret COMP (*1sgSM-SAY) NEG-1sgSM-PST-go-PFV Bungoma NEG

      'I regretted that I didn't go to Bungoma.' (Diercks, 2013, 398)

Each of these instances are examples of dependent/complement CPs that are most reasonably *not* Linguistic Material, per the defining characteristics proposed by Grimshaw (2015) and Major (2021a). Given that they are not LM, we would have no reason to expect that *AGR-li* would be used to introduce them. Previous work offered no principled reason for these effects (Diercks, 2010, 2013; Carstens, 2016; Diercks et al., 2020), but these effects follow quite directly from the proposal here. We save comprehensive discussion of these additional elements for future research, but we find it important to point out that *AGR-li* is not found in CP structures that lack some sort of communicated content, be it externalized or internal to the mind of the source.[24]

5   Additional analytical and theoretical discussion

5.1   More on tense in root *AGR-li* contexts

We offered an explanation for the absence of tense/aspect on *AGR-li* in truncated clausal adjunct structures; namely, that truncated clausal adjuncts in general are too small to host tense/aspect morphology. However, it remains a bit mysterious that tense is similarly unable to be expressed in root *AGR-li* constructions. It is worth taking a moment to explore this in more depth, given that "say" clauses have been shown to be similarly truncated in Avatime (Major, 2021a), Ewe (Clements, 1975; Spadine, 2020), and Tigrinya (Spadine, 2020).

First, root *AGR-li* constructions are not truncated in the C domain. For instance, there is no prohibition on matrix questions in *AGR-li* constructions. Yes/no questions are indicated by intonation, so the segmental material

---

[23]The second author disprefers the *mbo* complementizer, instead preferring *bali*. There seems to be variation across Lubukusu speakers in which of these (non-agreeing) complementizers are preferred.

[24]In the contexts where Bungoma speakers used *mbo* forms as non-agreeing complementizers, the second author uses a non-agreeing *bali* form. This *bali* form is homophonous with the third plural form of *AGR-li*: it remains to be seen whether these are truly distinct forms, or some kind of impersonal use of the "say" predicate that allows it to be used in these kinds of contexts.

below is identical to the declarative equivalent (89). If we assume that polar questions are formed based on a feature in the CP domain (or perhaps Force), this suggests that the left periphery is present in root *AGR-li* constructions.

(89) Wekesa a-li   ba-ba-ana   b-a-ch-a?
      1Wekesa 1SM-SAY 2-2-children 2SM-PST-go-FV
      'Does Wekesa say that the children left.'

Spadine (2020) suggests that there is only one clausal periphery present in similar data in Tigrinya. However, it is possible to embed a question as well, suggesting that both the matrix and embedded clause consist of CPs (90).

(90) Wekesa a-li   naanu b-a-ch-a?
      1Wekesa 1SM-SAY 2who 2SM-PST-go-FV
      'Does Wekesa say who left?'

We interpret this as evidence that there are in fact two clauses in these structures (i.e. *AGR-li* is not a left-peripheral modal element). For Spadine, this analysis offered an explanation for why tense morphology and temporal adverbs are not permitted. At first glance, the restriction on temporal adverbs seems to hold in Lubukusu, as well. As was shown for Ewe and Tigrinya, 'yesterday' is not permitted (91).

(91) #Likolooba Wekesa a-li   ba-ba-ana   b-a-ch-a
      yesterday   1Wekesa 1SM-SAY 2-2-children 2SM-PST-go-FV
      Intended: 'Yesterday, Wekesa said that the children left.'

However, the adverb 'now' is felicitous, which suggests that the restriction cannot be characterized as a prohibition on temporal adverbs in *AGR-li* constructions; instead, it seems there is a prohibition on past temporal adverbs.

(92) Luno Wekesa a-li   ba-ba-ana   b-a-ch-a
      now 1Wekesa 1SM-SAY 2-2-children 2SM-PST-go-FV
      'Now, Wekesa says that the children left.'

We interpret this as evidence that there is tense in these constructions, which is likely present tense. If we take null tense in Lubukusu to share properties with the English present, this data is perhaps not that surprising. Building on the discussion in Section 3.1, notice that the same property holds of English 'says,' where 'yesterday' is not possible, but 'now' is.

(93) a. I saw Katie yesterday - she says/said she will visit next week.
      b. #Yesterday, Katie says she will visit next week.
      c. (I saw Katie yesterday.) Now she says she will visit next week (she changed her plan).

This is yet another instance of *AGR-li* clauses behaving like this particular use of English present tense constructions. While this issue requires more work (as do the semantics of Lubukusu tense/aspect more generally), this at least begins to help us understand that the stative *-li* "say" predicate may simply be restricted to a present tense usage, albeit one that is highly dependent on its immediate discourse context.

One could question why only a single verb (*-li* 'say') shows these restrictions. Perhaps offering some credence to the possibility that "say" should be treated as a light verb (Grimshaw, 2015), signs of defectiveness in light verbs are not so uncommon. For instance, the Welsh verb *gwneud* 'do' is sensitive to the stativity/eventivity contrast and shows similar inflectional restrictions, but is quite clearly not a simple auxiliary (Rouveret, 2011). Landau (2006) demonstrates that [n,g,d] 'say'/'tell' in Modern Hebrew is only permitted in future or infinitival forms. Furthermore, if we are on the right track, we might conclude the Ewe and Tigrinya constructions (along with *AGR-li*) are verbs that are defective in much the same way.

If we take the next step in assuming that *AGR-li* is in the (silent) present tense, we might expect it to behave like the English cases in (72), where the temporal properties of the simple present in English are determined by the context in which they are embedded. When there is no local tense, the present is interpreted as 'now' (at the utterance time). If a reference time is established, the present is interpreted relative to that time. This is precisely what we see when *AGR-li* occurs as a modifier. Notice in (94), for instance, that when modifying the matrix VP in the future, it is interpreted in the future.

(94)　*Context: A speechwriter explains to the press what the president will say in his speech later:*

o-mw-aami a-la-lom-a　　a-li　ba-a-sikari ba-la-yukh-a　　e-ngo.
1-1-leader　1SM-FUT-speak-FV 1SM-SAY 2-2-soldiers 2SM-FUT-return-FV 9-home
'The president/leader will say that the soldiers are returning home.'

This is only a preliminary discussion, and a more in-depth treatment of the tense properties of *AGR-li* clauses (and of Lubukusu tense/aspect more broadly) is certainly necessary. For now, we leave this as an initial direction of investigating the apparently-exceptional properties of the *AGR-li* predicate with respect to tense.

5.2　Agreement mismatches

A central piece of Diercks' (2010; 2013) analysis is that the agreement on *AGR-li* is simply subject-oriented, rather than semantically controlled. As we have showed in this work, finer control of the pragmatic context reveals that conclusion to have been incorrect; nonetheless, a number of the original motivating patterns still hold, which do raise interesting questions for how to explain *AGR-li* on the account we advanced in this paper. For example, despite the frequent assumption that the agreement trigger for *AGR-li* as it introduces embedded clauses is the communicative source, there are regular mismatches in this regard. For example, in a verb of hearing the *AGR-li* form naturally agrees with the matrix subject, as is the norm.

(95)　Khw-a-ulil-a　　khukhwama khu Sammy { khu-li　/ %a-li　} ba-limi　ba-a-fun-a
　　　1PL.SM-PST-hear-FV from　　LOC 1Sammy 1PL-SAY 1SM-SAY 2-farmers 2SM-PST-harvest-FV
　　　ka-ma-indi.
　　　6-6-maize
　　　'We heard from Sammy that the farmers harvested the maize.' (Diercks, 2013, 366)

In the earlier work, Diercks reports *a-li* class 1 agreement as ungrammatical. In this paper we reported a genuine empirical difference here, as for the second author, class 1 agreement in this construction is acceptable.[25] Nonetheless, agreement with the subject (i.e. the hear-er, not the say-er) is unproblematic with verbs of hearing.

(96)　Khw-a-ulil-a　　khu-li　ba-limi　　2-farmers ba-a-fun-a ka-ma-indi.
　　　1PL.SM-PST-hear-FV 1PL-SAY 2SM-PST-harvest-FV 6-6-maize
　　　'We heard that the farmers harvested the maize.' (Diercks, 2013, 366)

Diercks (2010, 2013) reports a similar scenario for derived subjects of passives: given the appropriate predicate (one with an overt matrix object), the derived subject of the passive can trigger agreement on *AGR-li*.

(97)　Sammy ka-bol-el-w-a　　a-li　ba-keni b-ol-a.
　　　1Sammy 1SM-say-APPL-PASS-FV 1SM-SAY 2-guests 2SM-PST.arrive-FV
　　　'Sammy was told that the guests arrived.'

These facts (among others) led Diercks (2013) to argue for a binding-theoretic approach to the *AGR-li* puzzle where agreement on *AGR-li* is mediated by a null subject-oriented self-anaphor, as opposed to agreement with the source of information.

Similarly, Diercks (2010, 2013) reported that despite the ability to agree with derived subjects and passives and subjects of verbs of hearing, the inclusion of sources/attitude holders can at times disrupt the canonical agreement

---

[25]We commented above in §2.6 on the possible differences between the previous reports and this one.

on *AGR-li* targeting the matrix subject. This is evident in subjects of passives (among other contexts). In (98), we can see that including a passive by-phrase is highly disruptive to *AGR-li* agreeing with the derived subject of the passive:

(98) Nelsoni ka-a-bol-el-w-a   nende ese { mbo / *n-di   / *a-li   } ba-keni b-a-ach-a.
      1Nelson 1SM-PST-say-APPL-PASS-FV by   me   COMP *1SG.SM-SAY 1SM-SAY 2-guests 2SM-PST-go-FV

      'Nelson was told by me that the guests left.' (Diercks, 2013, 380)

Therefore, we can see that there are many instances where *AGR-li* naturally agrees with an apparent non-say-er (subjects of passives and verbs of hearing), canonically agreeing with the matrix subject. Even the early work on the issue (Diercks, 2010, 2013) showed that there is influence of the presence of SAY-ers in the syntax, especially in disrupting the canonical subject-orientation of the agreement on the *AGR-li* construction.

      None of this had an explanation previously, but the proposed analysis of *AGR-li* introduced in this paper can largely account for these facts. As we've noted previously in this paper, for the second author it is readily available for *AGR-li* to agree with the source of information in a verb of hearing.

(99) N-a-ulil-a   khu⌈Wekesa⌉ *pro* ⌈a-⌉ li ba-ba-ana   b-a-ch-a
      1SG.SM-PST-hear-FV 17-1Wekesa   1-SAY 2-2-children 2SM-PST-leave-FV

      'I heard from Wekesa that the children left.'

This follows naturally from the analysis advanced in this paper. The controller of agreement on *AGR-li* on our account is a pronominal subject of *-li* 'say,' which in a sentence like (99) can readily occur with *Wekesa* as an antecedent. Not just any source is eligible to trigger agreement on *AGR-li*, however: Diercks (2013) reported that passive by-phrases cannot trigger CA, and that remains true for the second author as well:

(100) ??N-a-bol-el-w-a   ne-⌈Wekesa⌉ ⌈a-⌉ li ba-ba-an-a   b-a-ch-a.
       1SG.SM-PST-say-APPL-PASS-FV by-1Wekesa 1-SAY 2-2-children 2SM-PST-leave-FV

       'I was told by Wekesa that the children left.'

This is somewhat surprising, given that non-subject sources of hearing predicates readily trigger agreement on *AGR-li*. And in a non-passive parallel to (100), the agentive subject readily triggers agreement in *AGR-li*:

(101) ⌈Wekesa⌉ a-a-m-bol-el-a   ⌈a-⌉ li ba-ba-an-a   b-a-ch-a.
       Wekesa   1SM-PST-1sgOM-say-APPL-PASS-FV 1-SAY 2-2-children 2SM-PST-leave-FV

       'Wekesa told me that the children left.'

      So why can the source in a verb of hearing trigger agreement on *AGR-li*, but the source of the reported information in a passive sentence cannot? We suggest that the issue here is in fact a familiar one regarding antecedence of pronominals. While pronominals can refer to any discourse referent (that does not trigger a binding violation), salient/activated referents are highly preferred, and potential referents of lower salience are less likely (e.g. Cardinaletti and Starke, 1999). Likewise, passivization is an operation that promotes a non-subject (and non-topic) to subject position. We would suggest, then, that the main problem with *AGR-li* agreeing with the argument in the passive by-phrase in (100) is that the act of promoting the passivized subject and demoting the by-phrase agent makes the passive subject an aboutness topic, and of higher salience, and therefore the highly preferred antecedent to *pro*.

      We can see this effect in play in instances that have nothing to do with *AGR-li*. Each example in (102) has two sentences, the second of which has a *pro* subject. In (102a) the initial sentence is active, and the subject *Nafula* is readily interpreted as the antecedent of the *pro* subject in the second sentence. In (102b), in contrast, the first sentence is a passive (based off the active sentence in the first example); here, the judgments for the antecedent of the subsequent *pro* subject of the second clause are inverted. In this instance, the passivized subject *Nelsoni* is the preferred antecedent for *pro*.

(102) a. Nafula a-p-a Nelson~i~. Mala *pro* a-a-ch-a.
1Nafula 1SM.PST-hit-FV 1Nelson Then 1SM-NAR-go-FV

'Nafula hit Nelson. Then *pro* left.'
* Nelson left.
✓Nafula left.

b. Nelson~i~ a-p-w-a ne Nafula. Mala *pro* a-a-ch-a.
1Nelson 1SM.PST-hit-PASS-FV by 1Nafula Then 1SM-NAR-go-FV

'Nelson was hit by Nafula. Then *pro* left.'
✓Nelson left.
* Nafula left.

Presumably, the discourse function of a passive sentence is what is at play here: by demoting a subject and promoting an object to the subject role in (102b), a speaker makes the derived subject the topic of discussion and therefore of higher salience than the by-phrase agent, which affects the preferred antecedent for the *pro* subject of the subsequent sentence. We take this as evidence that the dispreference for passive by-phrases to control agreement on *AGR-li* has to do with general principles of pronominal antecedence and not anything particular to the *AGR-li* construction.

Passives and verbs of hearing pose a different problem on the analysis we have advanced, however. On the account advanced here, the agreement trigger on *AGR-li* is the *pro* subject of the *AGR-li* 'say' predicate. If this is the case, however, how is it possible for *AGR-li* to agree with the subject of 'hear' in (103a) and with the derived subject of a passivized verb of speech in (103b)?

(103) a. Khw-a-ulil-a khu-li ba-limi 2-farmers ba-a-fun-a ka-ma-indi.
1PL.SM-PST-hear-FV 1PL-SAY 2SM-PST-harvest-FV 6-6-maize

'We heard that the farmers harvested the maize.' (Diercks, 2013, 366)

b. Sammy ka-bol-el-w-a a-li ba-keni b-ol-a.
1Sammy 1SM-say-APPL-PASS-FV 1SM-SAY 2-guests 2SM-PST.arrive-FV

'Sammy was told that the guests arrived.'

In both of these instances the *pro* subject of *AGR-li* 'say' has an antecedent that is not the original speaker of the Linguistic Material in the *AGR-li* clause, but instead was the recipient/Goal/target of that communicated information. Data like these were among what originally motivated Diercks' (2010; 2013) analysis of *AGR-li* as controlled by the grammatical subject and not by any thematic role or strict semantic effect. So how are these accounted for on the proposal advanced here?

We find the most likely explanation to be centered on the fact that the speaker of a sentence is the one who is asserting something to the effect that "X communicated LM" when they use an *AGR-li* clause, and the speaker of the sentence uses a *pro* subject of *AGR-li* that refers to their own source of the information, i.e. who they attribute the Linguistic Material too. In most situations where someone can report "Sammy was told LM" or "Wekesa heard LM," the way that a speaker can report this is if Sammy or Wekesa are the ones that told the speaker of this event. As we've reported, it is in fact possible to agree with the original source of the LM as in (99), but this would only be possible when the utterer of (99) themselves can assert with appropriate evidence (for typical Gricean reasons) that the source was someone *else*. This might be the case if they were a part of the original exchange, or if they have some other kind of reliable evidence. This more restricted pragmatic context may also help explain why speakers at times rule out agreement with a non-subject source, as was originally reported by Diercks (2010, 2013).[26]

---

[26]This also has potential to explain some additional outstanding data patterns from the original work on *AGR-li*. As reported by Diercks (2010, 2013), there are instances where speakers can use a non-*AGR-li* clause-embedding strategy instead of *AGR-li*. Without delving into details here (for the sake of space), the consistent judgment that speakers had in response to these sentences is that the speaker of the sentence was creating an inference for the audience that the subject of the sentence was unreliable, and that the reported information (inside the *AGR-li* clause) was less than certain. A more full exploration of these patterns will take significant empirical work based on our new proposals here, which can't be done in the context of this paper, but the proposal here offers much promise for explaining the evidential-like interpretive effects of the presence/absence of *AGR-li* in some contexts.

5.3 Apparent expletive agreement on *AGR-li*

Diercks (2013) does point out that the agreement trigger for *AGR-li* does in fact appear to need to be *capable* of being an attitude-holder. A sentence like (104) sets up a scenario where a non-animate cause is the probably trigger of agreement on *AGR-li*, but agreement in this instance is anomalous.

(104) Chi-sale khu-mesa chy-a-subi-sy-a   Alfredi (mbo/ ??chi-li/ *a-li)   chi-mbeba chi-li
      10-marks 17-table 10SM-PST-believe-CAUS-FV 1Alfred COMP/ ??10SM-SAY/ *1SM-SAY 10-rats   10-be
      mu-nju
      18-house
      'The marks on the table made Alfred believe that rats were in the house.' (Diercks, 2013, 400)

This appears to be in conflict with other patterns reported by Diercks, because it is possible for apparent (null) expletives to control agreement on *AGR-li*.

(105) a. Ka-lolekhan-a ka-li   Tegani ka-a-kw-a
          6SM-seems-FV 6SM-SAY 1Tegan 1SM-PST-fall-FV
          'It seems like Tegan fell.'
       b. Li-lolekhan-a li-li   Sammy a-likho a-lwal-a
          5SM-seems-FV 5SM-SAY 1Sammy 1S-PROG 1SM-be.sick-FV
          'It seems like Sammy is sick.' (Diercks, 2010, 381)

Carstens (2016) takes the facts in (105) as central to adjudicating between her standard Agree-based approach to *AGR-li* and the anaphoric approach advocated for by Diercks (2013) and Diercks et al. (2020), as we would not expect an expletive to be the antecedent of an anaphor, though it may readily be the target of an Agree operation. And these data raise a similar question for the approach we advance here: if the agreement on *AGR-li* is the result of agreeing with the *pro* subject of *AGR-li* 'say,' how is it possible for an expletive to be the antecedent for *pro*, and for that referent to be a viable subject of the "say" predicate?

We don't have a final answer to this question as it requires a lot of additional research, but work on related Luyia languages that has occurred in the interim provides some insight that will be useful for our concerns here. Consider the following data from Gluckman and Bowler (2017), which shows that different expletive-type agreements yield different interpretations in Logoori (Luyia, Kenya). The main distinction has to do with the relative (in)directness of the evidence the speaker has at hand: if it is more direct evidence, the class 6 *ga-* subject marker is more appropriate. If the evidence is more indirect, the class 9 *e-* subject marker is more appropriate.

(106) Logoori
      Context: It's flu season, and Imali didn't come to school. The speaker says:
      a. e-fan-a   kuresa Imali a-saal-a
         9SM-seem-FV like   1Imali 1SM-be.sick-FV
         'It seems like Imali is sick.'
      b. #ga-fan-a   kuresa Imali a-saal-a
         6SM-seem-FV like   1Imali 1SM-be.sick-FV
         'It seems like Imali is sick.' (Gluckman and Bowler, 2017, 1065)

(107) Logoori
      Context: The speaker sees Imali coughing and sneezing. They say:
      a. ?e-fan-a   kuresa Imali a-saal-a
         9SM-seem-FV like   1Imali 1SM-be.sick-FV
         'It seems like Imali is sick.'
      b. ga-fan-a   kuresa Imali a-saal-a
         6SM-seem-FV like   1Imali 1SM-be.sick-FV
         'It seems like Imali is sick.' Speaker's comment: "(4b) is only appropriate if you're looking at Imali."
         (Gluckman and Bowler, 2017, 1065)

While work is ongoing on this issue, we can report similar patterns in at least Tiriki and Wanga (nearby Luyia languages). And though Diercks (2010) didn't investigate this question, consultants in that original work frequently offered the intuition that the distinction between the singular class 5 and the plural class 6 "expletive" agreements in (105) had to do with the amount/degree of evidence available regarding the reported information, suggesting that those Bungoma Lubukusu speakers in that work shared a similar evidential-like effect of different expletive agreements in (105).

The details of Gluckman and Bowler's (2017) analysis are not central to our concerns here. What is crucial is the observation that these pseudo-expletive agreements are not in fact contentless pleonastic pronouns: they in fact have semantic content that interacts with modal contexts (Gluckman and Bowler, 2017; Gluckman, 2021). Specifically, if these pseudo-expletives (or whatever generates them) have semantics that are capable of referencing evidential properties of a proposition, they are plausibly likewise capable of serving in the stative subject position of AGR-*li* 'say' (something like "the evidence says …"). This suggests to us that the availability of AGR-*li* agreeing with apparent expletives in Lubukusu is not an immediate counter-argument to the claims advanced here. Instead, we conclude that these pseudo-expletives require additional investigation, investigation which should take into account their ability to serve as the Holder/Location subjects of AGR-*li* "say."

A similar explanation could be at play with another puzzle introduced by Diercks (2010, 2013). Diercks showed that while non-raised versions of raising constructions could occur with AGR-*li*, it is impossible for AGR-*li* to occur when agreeing with a raised subject.

(108) Sammy a-lolekhan-a   mbo (*a-li)   a-likho   a-lwal-a
       1Sammy 1SM-appears-FV COMP (*1SM-SAY) 1SM-PROG 1SM-be.sick-FV

       'Sammy appears to be sick.' (lit. "Sammy seems that is sick") (Diercks, 2010, 382)

The answer here may well be the same as the discussion above: whereas the apparent null expletives are in fact capable of being the Holder/Location of the propositional content, the raised subjects cannot be, as raised subjects in raising constructions are themselves arguments of the reported proposition.

This discussion is not final, of course: detailed work on hyper-raising in Lubukusu is still pending (initial discussion is available in Carstens and Diercks 2013). What we mean to point out, however, is that several apparent challenges to the account presented here quite reasonably have alternative explanations for Lubukusu, though more work is certainly necessary.

5.4 Implications for Bantu languages and beyond

The analysis presented here is situated within a growing theoretical discussion about clausal complementation across Bantu languages. Halpert (2019) introduces discussion of "say" complementation in Zulu, illustrating that the morphological makeup of the complementizer plays a crucial role in determining both its syntactic and semantic properties. In particular, Halpert differentiates between constructions containing the nominalized form of 'say,' *ukuthi*, compared to a form that is more verb-like, *sengathi*. In Pietraszko (2019), it is argued that *ukuthi* clauses in Ndebele form DPs. We find that our approach and overall analysis is compatible with most of the findings in these papers.

Take the two nominalized verbs from Ndebele in (109), for instance. Pietraszko argues that both forms consist of the augument, *u*, a noun class marker, and a verb root. Both 'cook' and "say" host class 15 prefixes and the augment *u-*.

(109) a. u-ku-pheka
          15aug-15-cook

          'to cook/cooking' (Pietraszko, 2017, 210)

       b. u-ku-thi
          15aug-15-say

          'to say/saying/COMP' (Pietraszko, 2017, 210)

Pietraszko suggests that the complementizer use of *ukuthi* provided in (110) has not grammaticalized the

augment *u*, treated as a D head layered on the complementizer element. But she suggests that *kuthi* is a pure complementizer.[27]

(110) a. Ngicabanga [*DP* u-   [*CP* kuthi usukile]].
         1sg.thought   aug-   comp 1.left
         'I thought that she left.' (Pietraszko, 2019, 68)

A nontrivial aspect of how one approaches the decomposition of an element like *ukuthi* is dependent on the null hypothesis or set of assumptions that one makes leading into the analysis. Pietraszko clearly demonstrates that *u* has not fused with *kuthi*. However, the possibility that *-thi* remains a verbal root is not really entertained. The question is whether the decomposition of *ukuthi* should be the same in both (109b) and (110a). If one chooses to treat both (109a) and (109b) as equivalent, the result is that the augment selects a deverbal nominalized complement in both cases. If this is the case, it is not necessarily the case that the augment *u-* takes a CP complement; instead, it would take a deverbal nominal complement.

This leads to a second argument against the presence of an NP in Ndebele clausal complementation: namely, the fact that an overt nominal element such as *indaba* '9.news' is prohibited in these structures.

(111) a. Ngi-zwe   u-kuthi   u-ya-m-thanda.
         1sg.sbj-hear-.pst 15.aug-15.comp 1.sbj-tam-1.obj-like
         'I heard that she likes him.' (Pietraszko, 2019, 78)
     b. *Ngi-zwe   indaba (u)-kuthi   u-ya-m-thanda.
         1sg.sbj-hear-.pst 9.news 15.aug-15.comp 1.sbj-tam-1.obj-like
         Intended: 'I heard the news that she likes him.' (Pietraszko, 2019, 78)

If we apply the decompositional analysis to the cases above, the augment *u* selects for a deverbal nominal complement in (111a); as mentioned above, the NP in this structure is the nominalized verb. Applying the same logic to (111b), the ungrammaticality would result from 'news' being unable to select for a DP complement, which would presumably hold if *-thi* "say" were replaced with e.g. *-pheka* 'eat,' as well. In other words, this illustrates a problem with a bare nominal selecting a DP headed by the augment, but it does not show that the augment does not select for an NP complement.

One argument in favor of decomposition comes from the existence of other elements containing the root *-thi*, that have similarly been treated as simplex C heads, such as *sengathi*. Contrasts between *ukuthi* and *sengathi* are discussed at length in Halpert (2019). Halpert illustrates that *ukuthi* clauses distribute like DPs, which is expected given that this clausal argument is headed by the augment. However, Halpert points to a range of environments where *sengathi* seems to merge as a low VP adjunct, which is fully in line with our analysis of AGR-*li* presented in this paper. Under a decompositional analysis, we predict that these differences arise from the decomposition; namely, differences between the morphosyntactic and semantic properties of *uku-* and *senga-*. In other words, cross-linguistic differences between "say" complementation structures arise from differences in the morphosyntactic linking mechanisms and idiosyncratic properties of the particular "say" verb. Lubukusu has a single mechanism for merging AGR-*li* clauses: truncated clausal adjunction. There is no evidence of nominalization in any instance of AGR-*li* in Lubukusu, but we do see it in both Ndebele and Zulu.

One language that shows further evidence that these C-like elements do contain verbal roots is Kipsigis (Driemel and Kouneli, 2020). Interestingly, the element *-le* in Kipsigis can take applicative morphology, agreement, and also aspect/mood, like a full TP (unlike Bukusu). Interestingly, this "say" element is able to be modified by adverbs (e.g. 'slowly') and occur with an explicit addressee (introduced by the applicative). We interpret the differences between Lubukusu and Kipsigis to arise from the fact that Kipsigis "say" is not obligatorily stative (likely ambiguous like its English equivalent), and the fact that the linking strategy is not as truncated as the Lubukusu adjunction strategy.[28] More work is necessary on Kipsigis, though, to examine the precise method of incorporating *-le* clauses into matrix clauses (e.g. complementation, adjunction, etc.).

---

[27]Equivalent augments in Zulu were treated as K(ase) heads in (Halpert, 2012). This distinction does not have implications for the present discussion, however.

[28]In Uyghur, it is argued that "say" complementation structures derive from the converbial linker *-(I)p* and the main verb *de-* 'say,' which has both eventive and stative uses, like English and Kipsigis (Major, To Appear, 2021a).

We hope that our re-analysis of AGR-*li* constructions in Lubukusu can help open up debate with respect to the grammatical mechanisms implicated in clausal complementation cross-linguistically. Perhaps the most common reason that these elements are not treated as verbs, comes from the fact that they often do not denote 'saying out loud to somebody'. The present paper has demonstrated that this is not the right conclusion. As a matter of fact, as suggested in Major (2021a), it remains entirely plausible that the lexical semantics of "say" could undergo bleaching to the point that it is hard to find any meaningful contribution. This does not allow us to conclude that there is no longer a verbal head; it may just be that its primary role as a verb is to participate in whatever the verbal-linking structure is in the language and to introduce a clause. This seems to be a productive way to compare the differences between Lubukusu and Kipsigis, for instance.

5.5   Broader Implications, in brief

There are a variety of broader theoretical implications of our proposal which we only mention here. Specifically, the new structural proposal for Lubukusu AGR-*li* changes the theoretical implications of the construction. Most preceding work on the construction accepted relatively uncritically the assumption that it is a complementizer: this was reasonable, given that it serves a complementizer-like function and is the translational equivalent of a complementizer. On that structural analysis of AGR-*li*, the agreement properties of AGR-*li* posed the largest theoretical puzzle, as reflected by the existing work on the issue (Diercks, 2010; Carstens, 2016; Diercks et al., 2020). On the account we propose here, agreement is no longer problematic: agreement is handled in the same local way that all agreement is handled in truncated adjunct clauses, presumably an Agree relation from a local functional head with the *pro* subject of "say."

Instead of agreement being the central question, our proposal induces a new set of empirical questions about clause chaining mechanisms in Lubukusu, as well as broader theoretical questions about the nature of say-complementation cross-linguistically.

In addition to these theoretical contributions, it is worth noting more explicitly some important aspects of the empirical investigation that led to the theoretical conclusions. First, this paper benefited considerably from the insights of the second author, a native speaker of Lubukusu and a linguist. Pinpointing these subtle differences would have been impossible without this, and the more access that we can create for native speaker linguists, the better off the field will be.

Our findings emphasize another methodological note, which is by no means new. If one restricts their investigation to sentences elicited mainly from direct translations in the absence of robust context, we risk missing important generalizations. It is of course part of the normal course of research for initial ideas/analyses to be revisited as more information is discovered: there are still robust research communities working on European languages, for example, despite the vast amount of knowledge we have generated on those languages. We can't learn everything at once. But findings like these emphasize the centrality of context in accurate empirical analyses. Again, we don't want to imply that elicitation methods are problematic in themselves: the extraordinary level of grammatical detail we've learned about this one construction in Lubukusu is only possible via elicitation methods. But more and more it is becoming clear that sentence-based analysis in the absence of carefully-controlled pragmatic context can lead to incomplete empirical generalizations. [29]

5.6   Issues for future research

As productive of a research area that these say-complementation constructions (formerly known as agreeing complementizer constructions) are proving to be, there are nonetheless important additional questions that we haven't explained in this work. First, as we noted above, the nature of the null pseudo-expletives in Lubukusu (and Luyia more broadly) is a deeply interesting question that intersects with our puzzle here. But given that there are extensive questions yet to be answered about those constructions, we cannot fully engage their relevance for AGR-*li* until that parallel work is complete.

Second, there are several empirical domains that have been previously reported that we are not able to cover in this paper. As we mentioned above, Diercks (2010, 2013) shows that there are interpretive distinctions between using

---

[29]Some of these authors' own work on object marking in Lubukusu reflects this same progression: cf. Sikuku et al. (2018), Sikuku and Diercks (2022), Lippard et al. (2022).

AGR-*li* complementation and using a non-agreeing complementizer (*bali*). Specifically, AGR-*li* is used in instances where the speaker has confidence in the veracity of the information being reported, but using non-agreeing *bali* instead creates a sense that the information is not certainly true, or specifically that the subject of the sentence is an unreliable reporter of the information. We think there may be a quite natural account of these facts under the proposal advanced here, but addressing this would extend this paper to a bloated state in ways that are not necessary to argue for our proposal, so we leave it to future work.

In §4.2 we gave an initial outline of clause-chain/serialization-type constructions in Lubukusu, a previously unexplored area in the language. These are also under-researched in narrow Bantu languages more generally, at least from a theoretical perspective. What we report here, however, can only be considered an initial description, as each of the constructions require their own detailed investigation (NE-subordination, narrative/subjunctive clause chains, and truncated verbal adjunction). These are an area of growing crosslinguistic research from a generative perspective, and Lubukusu clearly has contributions to offer in those areas.

On this new approach, Lubukusu fits into a broader typology of say-complementation. A broad variety of recent research (and some older research) is showing that say-complementation often has verbal properties across a broad variety of language familes: Bantu (Güldemann, 2008; Halpert, 2019; Kawasha, 2007; Letsholo and Safir, 2019); Nilo-Saharan (Driemel and Kouneli, 2020); Kwa (Kinyalolo, 1993; Koopman, 1984; Koopman and Sportiche, 1989; Major and Torrence, 2020); Turkic (Baker and Vinokurova, 2010; Major, To Appear, 2021a,b; Ozyıldız, 2017; Predolac, 2017); Mongolic (Knyazev, 2015; Bondarenko, 2020); Semitic (Spadine, 2020); Sinitic (Chappell, 2008); Dravidian and Indo-Aryan (Balusu, 2020; Bayer, 1999); Englishes like AAVE (Martin and Wolfram, 1998), Nigerian Pidgin (Mfon Udoinyang, pc), and colloquial American English (Major, 2021a). It is clear, however, that there are interesting differences between the extent of the verbal properties that are allowed on the "say" predicates in these "say"-complementation structures: Ikalanga (Letsholo and Safir, 2019), Kipsigis (Diercks and Rao, 2019; Diercks et al., 2020; Driemel and Kouneli, 2020), and Lubukusu make an interesting comparison set in this regard. Ikalanga shows many properties similar to Lubukusu, but also shows some degree of voice morphology on the "say"-based "complementizer." Meanwhile, Kipsigis even allows object agreement on "say" in its complementizer-like usage. Clarifying the say-complementation structures cross-linguistically will open doors to exploring how/why these constructions vary cross-linguistically.

6   Conclusions

We have argued that analyzing Lubukusu AGR-*li* as an agreeing complementizer was a mistake. AGR-*li* does introduce apparent complement clauses in many instances, but we have argued that AGR-*li* should instead be analyzed as a verbal element that occurs in serialization-type constructions together with a main predicate. On this analysis, AGR-*li* is the stative predicate "say" (Grimshaw, 2015; Major, 2021a), which introduces Linguistic Material and a subject argument that is the Holder/Location of that material. We argue that this explains a broader range of the previously reported facts than preceding analyses do, and we have shown that it makes a large range of additional predictions which are supported by the novel Lubukusu evidence reported here. The result is a novel analysis of Lubukusu AGR-*li* which also solves some standing theoretical issues, namely the atypical agreement configuration it stood in (on a complementizer analysis of AGR-*li*).

Acknowledgements

References

Aboh, Enoch Oladé. 2005. Object shift, Verb movement and Verb reduplication. In *The Oxford handbook of comparative syntax*, ed. Guglielmo Cinque and Richard S. Kayne, 138–177. Oxford: Oxford University Press.

Anand, Pranav, and Maziar Toosarvandani. to appear. Narrative and point of view. In *Linguistics meets philosophy*. Cambridge: Cambridge University Press.

Anand, Pranav, and Maziar Tooservandani. 2020. Embedded presents and the structure of narratives. In *Proceedings of SALT 30*, ed. Joseph Rhyne, Kaelyn Lamp, Nicole Dreier, and Chloe Kwon, 801–820. Linguistic Society of America. https://journals.linguisticsociety.org/proceedings/index.php/SALT/issue/view/174.

Baker, Mark. 2008. *The syntax of agreement and concord*. Cambridge: Cambridge University Press.

Baker, Mark, and Nadya Vinokurova. 2010. Two modalities of Case assignment: case in Sakha. *Natural Language and Linguistic Theory* 28:593–642.

Baker, Mark C. 1989. Object sharing and projection in serial verb constructions. *Linguistic Inquiry* 20:513–554.

Balusu, Rahul. 2020. The quotative complementizer says "I'm too Baroque for that.". In *Proceedings of Formal Approaches to South Asian Languages (FASAL) 8*, ed. Bhamati Dash and Gurmeet Kaur, volume 3.

Bayer, Josef. 1999. Final complementizers in hybrid languages. *Journal of Linguistics* 35:233–271. Publisher: Cambridge University Press.

Bochnak, M. Ryan, and Emily A. Hanink. 2021. Clausal embedding in Washo: Complementation vs. modification. Published online 7 Dec 2021.

Bondarenko, Tatiana. 2020. Factivity from pre-existence: Evidence from Barguzin Buryat. *Glossa: a journal of general linguistics* 5, 1:1–35.

Borer, Hagit. 2004. The grammar machine. In *The unaccusativity puzzle: Explorations of the syntax-lexicon interface*, ed. Artemis Alexiadou, Elena Anagnostopoulou, and Martin Evereart, 288–331. Oxford: Oxford University Press.

Cardinaletti, Anna, and Michal Starke. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In *Clitics in the languages of europe*, ed. Henk C. van Riemsdijk, 145–234. Berlin: Mouton de Gruyter.

Carstens, Vicki. 2016. Delayed valuation: a reanalysis of "upwards" complementizer agreement, and the mechanics of Case. *Syntax* 19:1–42. Lingbuzz/001432.

Carstens, Vicki, and Michael Diercks. 2013. Parameterizing Case and activity: Hyperraising in Bantu. In *Proceedings of the 40th Annual Meeting of the North East Linguistic Society*, ed. Seda Kan, Claire Moore-Cantwell, and Robert Staubs, 99–118. Amherst, MA: GLSA.

Chappell, Hilary. 2008. Variation in the grammaticalization of complementizers from verba dicendi in Sinitic languages. *Linguistic Typology* 12:45–98.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–156. Cambridge: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in linguistics*, ed. Michael Kenstowicz, 1–52. Cambridge: MIT Press.

Clements, George N. 1975. The logophoric pronoun in ewe: its role in discourse. *Journal of West African Languages* 10:141–177.

Collins, Chris. 1997. Argument sharing in serial verb constructions. *Linguistic Inquiry* 28:461–497.

Diercks, Michael. 2010. Agreement with subjects in Lubukusu. Doctoral Dissertation, Georgetown University.

Diercks, Michael. 2013. Indirect agree in Lubukusu complementizer agreement. *Natural Language and Linguistic Theory* 31:357–407.

Diercks, Michael, and Meghana Rao. 2019. Upward-oriented complementizer agreement with subjects and objects in Kipsigis. In *Theory and description in African Linguistics: Selected papers from the 47th Annual Conference on African Linguistics*, ed. Emily Clem, Peter Jenks, and Hannah Sande, 369–393. Berlin: Language Science Press. DOI:10.5281/zenodo.3367166.

Diercks, Michael, Marjo van Koppen, and Michael Putnam. 2020. Agree probes down: Anaphoric feature valuation and phase reference. In *Agree to agree: Agreement in the Minimalist Programme*, ed. Peter W. Smith, Johannes Mursell, and Katharina Hartmann, 347–390. Berlin: Language Science Press.

Driemel, Imke, and Maria Kouneli. 2020. C-Agree is local subject-verb agreement in Kipsigis. Ms, Humboldt University and University of Leipzig.

Elliott, Patrick D. 2016. Explaining DPs vs. CPs without syntax. *Proceedings of CLS* 52:171–185.

Elliott, Patrick D. 2020. Elements of clausal embedding. Doctoral Dissertation, UCL (University College London).

Folli, Raffaella, and Heidi Harley. 2007. Causation, obligation, and argument structure: On the nature of little v. *Linguistic Inquiry* 38:197–218.

Gluckman, John. 2021. Null expletives and embedded clauses in logoori. *Syntax* 24:334–375.

Gluckman, John, and Margit Bowler. 2017. Expletive agreement, evidentiality, and modality in logooli. *Proceedings of SALT* 26:1063–1082.

Grimshaw, Jane. 2015. The light verbs say and say. *Structures in the Mind: Essays on Language, Music, and Cognition in Honor of Ray Jackendoff* 79.

Güldemann, Tom. 2008. *Quotative Indexes in African Languages: A Synchronic and Diachronic Survey*, volume 34 of *Empirical Approaches to Language Typology*. Mouton de Gruyter.

Halpert, Claire. 2012. Argument licensing and agreement in Zulu. Doctoral Dissertation, MIT.

Halpert, Claire. 2019. How to be an embedded clause: say complementizers in Bantu. In *Proceedings of triplea 5: Fieldwork perspectives on the semantics of African, Asian and Austronesian languages*, ed. M. Ryan Bochnak, Miriam Butt, Erlinde Meertens, and Mark-Matthias Zymla. https://publikationen.uni-tuebingen.de/xmlui/handle/10900/93789: Universitätsbibliothek Tübingen, Publikationssystem.

Idiatov, Dmitry. 2010. Person-number agreement on clause linking markers in Mande. *Studies in Language* 34:832–868.

Kawasha, Boniface. 2007. Subject-agreeing complementizers and their functions in Chokwe, Luchazi, Lunda, and Luvale. In *Selected Proceedings of the 37th Annual Conference on African Linguistics*, ed. Doris Payne and Jaime Peña, 180–190. Somerville, MA: Cascadilla Proceedings Project.

Kayne, Richard. 2014. Why isn't *this* a complementizer? In *Functional structure from top to toe: The cartography of syntactic structures*, ed. Peter Svenonius, volume 9, 188–231. Oxford University Press.

Kinyalolo, Kasangati K. W. 1993. The logophoric pronoun émi as a LF operator/anaphor. In *Proceedings of the North East Linguistic Society 23*, ed. Amy J. Schafer, 223–237. University of Ottawa: GRADUATE LINGUISTIC STUDENT ASSOCIATION.

Knyazev, Mikhail. 2015. Verbal complementizers in Kalmyk: V, c, or both? In *Proceedings of the 9th workshop on altaic formal linguistics (wafl9)*, volume 76, 145–158. MIT Working Papers in Linguistics.

Koopman, Hilda. 1984. *The syntax of verbs*. Dordrecht, The Netherlands: Foris Publications.

Koopman, Hilda. 1989. The structure of VP in Dutch. Presented at the Summer Syntax Workshop, University of California, Irvine.

Koopman, Hilda, and Dominique Sportiche. 1989. Pronouns, logical variables, and logophoricity in Abe. *Linguistic Inquiry* 20:555–588.

Kratzer, Angelika. 1996. Severing the external argument from its verb. In *Phrase structure and the lexicon*, ed. Johan Rooryck and Laurie Zaring, 109–137. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Kratzer, Angelika. 2006. Decomposing attitude verbs. Talk given at The Hebrew University of Jerusalem, July 2006.

Kratzer, Angelika. 2016. Evidential Moods in Attitude & Speech Reports (unpublished slides). Slides from talk at UConn, Sept 9.

Landau, Idan. 2006. Severing the distribution of PRO from case. *Syntax* 9:153–170.

Letsholo, Rose, and Ken Safir. 2019. Complement clause c-agreement beyond subject phi- agreement in Ikalanga. In *African linguistics across the disciplines: Selected papers from the 48th Annual Conference on African Linguistics*, ed. Samson Lotven, Silvina Bongiovanni, Phillip Weirich, Robert Botne, and Samuel Gyasi Obeng, 159–186. Berlin: Language Science Press.

Lippard, Hannah, Justine Sikuku, Crisofia Lánga da Câmara, Franco Colantes, Madelyn Liu, and Michael Diercks. 2022. Emphatic interpretations of object marking in bantu languages. Ms, Pomona College, Moi University, Universidade Eduardo Mondlane, and Cambridge University. Available at https://research.pomona.edu/mjkd/files/2022/02/EmphaticOMs.pdf.

Lord, Carol. 1993. *Historical change in serial verb constructions*, volume 26. John Benjamins Publishing.

Major, Travis. 2021a. On the nature of "say" complementation. Doctoral Dissertation, UCLA.

Major, Travis. 2021b. What uyghur "says" about dependent case theory. Ms, UCLA.

Major, Travis. To Appear. Revisiting the syntax of monsters in Uyghur. *Linguistic Inquiry* .

Major, Travis, and Richard Stockwell. 2021. "Say"-ing without a Voice. *Proceedings of NELS 51* Place: UQAM.

Major, Travis, and Harold Torrence. 2020. "say"-chains, not "say"-complementation. To appear in the Proceedings of WCCFL38.

Martin, Stefan, and Walt Wolfram. 1998. The sentence in african-american vernacular english. In *African american english: structure, history, and use*, ed. John R. Rickford Mufwene, Salikoko, Guy Bailey, and John Baugh, 11–36. Routledge New York, NY.

Moltmann, Frederike. 2017. Levels of linguistic acts and the semantics of saying and quoting. In *Interpreting austin: Critical essays*, 34–59. Cambridge.

Moulton, Keir. 2009. Natural selection and the syntax of clausal complementation. Doctoral Dissertation, University of Massachusetts-Amherst, Amherst, MA.

Munro, Pamela. 1982. On the transitivity of say verbs. In *Studies in transitivity*, volume 15 of *Syntax and Semantics*, 301–318. Brill.

Ozyıldız, Deniz. 2017. Attitude reports with and without true belief. In *Semantics and Linguistic Theory*, volume 27, 397–417.

Pancheva, Roumyana, and Maria Luisa Zubizarreta. 2020. Temporal reference in the absence of tense in Paraguayan Guarani. In *Proceedings of nels 50*, ed. Mariam Asatryan, Yixiao Song, and Ayana Whitmal, 265–278. GLSA, University of Massachusetts.

Pietraszko, Asia. 2017. Clause size and transparency in Ndebele. Talk given at the 91st Annual Meeting of the LSA, Austin TX.

Pietraszko, Asia. 2019. Obligatory CP nominalization in ndebele. *Syntax* 22:66–111.

Potts, Christopher. 2002. The syntax and semantics of as-parentheticals. *Natural Language and Linguistic Theory* 20:623–689.

Predolac, Esra. 2017. The syntax of sentential complementation in Turkish. Doctoral Dissertation, Cornell University.

Ramchand, Gillian. 2008. *Verb Meaning and the Lexicon: A First-Phase Syntax*. Cambridge: Cambridge University Press.

Riis, H. N. 1854. *Grammatical outline and vocabulary of the oji-language, with especial reference to the akwapim-dialect, together with a collection of proverbs of the natives.* Basel: Bahnmaier's Buchhandlung.

Ross, John. 1967. Constraints on variables in syntax. Doctoral Dissertation, Massachusetts Institute of Technology.

Rothmayr, Antonia. 2009. *The structure of stative verbs*, volume 143. John Benjamins Publishing.

Rouveret, Alain. 2011. VP ellipsis, phases and the syntax of morphology. *Natural Language & Linguistic Theory* 30:897–963.

Safir, Ken, Mark Baker, and Justine Sikuku. 2020. Actual clauses in Lubukusu. In *Festschrift for Akin Akinlabi*, ed. Oluseye Adesola. Ibadan, Nigeria: Scholarship Publishing and Educational Consult.

Sikuku, Justine, and Michael Diercks. 2022. Object marking in Bukusu: At the interface of pragmatics and syntax. Ms, Moi University and Pomona College.

Sikuku, Justine, Michael Diercks, and Michael Marlo. 2018. Pragmatic effects of clitic doubling: Two kinds of object markers in Lubukusu. *Linguistic Variation* 18:359–429.

Spadine, Carolyn. 2020. The structure of attitude reports: Representing context in grammar. Doctoral Dissertation, Massachusetts Institute of Technology.

Torrence, Harold. 2016. Subject oriented complementizer agreement in Ibibio. Talk given at the 47th Annual Conference on African Linguistics at the University of California, Berkeley.

Wasike, Aggrey. 2007. The left periphery, wh-in-situ and A-bar movement in Lubukusu and other Bantu languages. Doctoral Dissertation, Cornell University.