Epistemic Parallels between Nouns and Clauses

John Gluckman

1. Introduction

While much previous work has illustrated syntactic similarities between the nominal and clausal domain (Abney 1987, among many others) I provide semantic evidence for the parallelism. Starting from the observation that there is strong empirical support for a "perspectival projection" in the CP domain (Koopman & Sportiche, 1989; Speas, 2004; Sudo, 2012; Sundaresan, 2018). I introduce data from the Luhia, Bantu language Nyala East which motivates that a similar projection exists in the DP domain (cf, Charnavel 2019).

In (1), I sketch the distribution of epistemic marking (EM) in Nyala East. EM is ostensibly inflectional morphology which appears in three contexts, prefixed to verbs, to complementizers, and to nouns in the following configurations.

(1) a. VERBAL-EM : EM-Verb … CP
   b. CLAUSAL-EM : EM-Complementizer
   c. NOMINAL-EM : EM-Noun

I argue below that the three configurations in (1) provide evidence for a perspectival operator in two syntactic loci: the clausal and nominal domains. I equate verbal-EM with CP-linked (null) expletives, and follow Gluckman (to appear) in proposing that they are derived by agreement with an embedded clause. The data thus underscore the inherent and well-noted categorial/distributional connection between the clausal and nominal domains—to the exclusion of the verbal domain—and provide novel evidence for a perspectival element in the nominal domain.

2. Verbal-EM

Impersonal constructions in Nyala East allow four different morphemes to occupy the subject agreement slot. I refer to this as verbal-EM. The markers are morphologically identical to subject markers for the noun class 8, 6, 5, 9.¹

(2) a. vi-on-ekh-an-a [ Masika ka-chi-e Nairobi ]
   8EM-see-NEUT-FV Masika 1SM-go-FV Nairobi
   ‘It looks like Masika went to Nairobi (based on a lot of evidence).’

b. ka-on-ekh-an-a [ Masika ka-chi-e Nairobi ]
   6EM-see-NEUT-FV Masika 1SM-go-FV Nairobi
   ‘It looks like Masika went to Nairobi (based on a few pieces of evidence).’

¹ Like all other Eastern Bantu languages, Nyala East has an extensive noun class system. By convention, the class nouns are numbered. Up until class 10, odd numbers are generally singulars of the following even number. The classes have broad semantic properties (e.g., 1/2 consists of humans, 3/4 contains plants and round things; etc), though there are numerous exceptions as well.

Verbal-EM is possible on any predicate that selects for a CP complement (or “extraposed” CP Stowell 1981) and does not select for a thematic subject. It is not possible in any other context.2

(3) a. Perception verbs: -onekha- ‘seem/appear,’ -ulekha- ‘be heard/sound’
   b. Verbs of saying: -orekhana-, ‘be said,’ -monyekhana- ‘be whispered,’ …
   c. Modal verbs: -nyalekhan-, ‘be possible,’ ’-enyekhan- ‘should, ought,’
   d. Emotives: -sindus-, ‘be surprising,’ -esiver- ‘be sad,’ …
   e. Verbs of thought: -paarekhan- ‘be thought,’ -manyekhan- ‘be known,’ -suvirekhan- ‘be believed,’ …

There are no other expletive constructions in Nyala East—no weather-/ambient-expletives, no existentials, no subject-in situ (cf Xhosa, Zulu data in Carstens & Mletshe 2015; Halpert 2016). Verbal-EM is possible when the CP is a subject (though not all predicates allow for CP-subjects, e.g., -onekha- ‘seem/appear’).

(4) [ mbo Masika ka-chi-e Nairobi ]
   \[
   \text{vi-ka-li-e-vi-ka-li-e-a-man-ikh-an-a EM-TNS-know-NEUT-REC-FV}
   \]
   ‘That Masika went to Nairobi is known.’

Briefly, the choice in verbal-EM conveys subtly different meanings depending both on lexico-semantic and contextual factors (cf, Gluckman & Bowler 2016 for the related Logoori).

- **vi-** (Class 8) indicates general, or universal knowledge. It can mean either “based on a lot of available information,” or “according to everyone contextually relevant.”
- **ka-** (Class 6) indicates relatively restricted knowledge. It can mean either “based on a few pieces of evidence,” or “according to a few contextually relevant people.”
- **li-** (Class 5) indicates a specific knowledge source. It can mean either, “based on a specific piece of evidence,” or “according to a specific person.”
- **e-** (Class 9) indicates a non-specific knowledge source. It can mean either, “based on some non-specific evidence,” or “according to some non-specific person.”

Class 9 is the weakest meaning, and its use is fairly restricted across all three occurrences (verbal-, clausal-, and nominal-EM). Finally, we note that class 8, 6, 5, and 9 are not otherwise “quantificational,” and subject agreement regularly (and obligatorily) appears with such noun classes without the meanings indicated above.

\footnote{Many of these verbs are morphologically complex appearing with the “neuter/middle/stative” suffix -ekh and the reciprocal -an. The combination of these two affixes in this particular context is well-known in Bantu linguistics (Seidl & Dimitriadis, 2003).}
3. Clausal-EM

Like the closely related Lubukusu (cf, Diercks 2013; Carstens 2016), Nyala East has an agreeing complementizer, -chi.

As the head of an embedded clause, the agreeing complementizer in Nyala East (like Lubukusu) is “subject-oriented,” reflecting the phi-features of the subject of the most local embedding verb.

(5) a. **Masika** ka-ol-iy-e a-**chi** Ingwe ya-vich-ir-e
**Masika** 1SM-say-ASP-FV 1SM-COMP 9Ingwe 9SM-win-ASP-FV

‘Masika said that Ingwe won.’

b. **na-khu-ol-iy-e n-**chi / *o-**chi** Nekesa ka-**chi-e** Nairobi
1SG-2SG.OM-say-APPL-FV 1SG-COMP 2SG.AGR-COMP 1Nekesa 1SM-go-FV Nairobi

‘I told you that Nekesa went to Nairobi.’

c. **e-li-kazeti** li-ol-iy-e li-**chi** Ingwe ya-vich-ir-e
**5AUG-5NC-newspaper** 5SM-say-ASP-FV 5SM-COMP 9Ingwe 9SM-win-ASP-FV

‘The newspaper said that Ingwe won.’

The agreeing complementizer is available for all (selected) embedded clauses. Additionally, Nyala East permits unembedded complementizers. When unembedded, the agreeing complementizer may be inflected with the four epistemic markers. I refer to this use as clausal-EM.

(6) a. **vi-**chi Masika ka-**chi-e** Nairobi
**8EM-COMP** 1Masika 1SM-go-FV Nairobi

‘Masika went to Nairobi (according to a lot of evidence).’

b. **ka-**chi Masika ka-**chi-e** Nairobi
**6EM-COMP** 1Masika 1SM-go-FV Nairobi

‘Masika went to Nairobi (according to a few pieces of evidence).’

c. **li-**chi Masika ka-**chi-e** Nairobi
**5EM-COMP** 1Masika 1SM-go-FV Nairobi

‘Masika went to Nairobi (according to a specific piece of evidence).’

d. **e-**chi Masika ka-**chi-e** Nairobi
**9EM-COMP** 1Masika 1SM-go-FV Nairobi

‘Masika went to Nairobi (according to a non-specific piece of evidence).’

3 Agreeing complementizers are a not-uncommon property of Bantu languages (Güldemann, 2000, 2008; Kawasha, 2006, 2007). Besides the non-agreeing complementizer mbo shown in the previous section, Nyala East also has the non-agreeing complementizers vachi ‘that’ and nga ‘like.’ Unlike Bukusu, “whether/if” clauses do not have an analytic complementizer (oli in Bukusu), but rather use the prefix ni- which synthetically attaches to a finite verb.

4 At least one speaker allows demoted subjects in a by-phrase (neende-phrase) to control this agreement, though this appears to be the exception, rather than the rule. To the extent that this is possible, it confirms the proposal in Diercks (2013) (adopted below) that the goal for this agreement is syntactically independent of the subject; see also Sundaresan (2018) for corroborating evidence from Tamil.

5 Ingwe is the name of the Luhia football/soccer team. The word ingwe means ‘leopard.’

6 This is called desubordination in Boye & Kehayov (2016) and insubordination in Evans & Watanabe (2016).

7 The unembedded agreeing complementizer may also inflect for 1st/2nd (singular) person, but not 3rd. The non-agreeing complementizers mbo and vachi are also permitted in unembedded contexts, cf Kawasha (2007). They uniquely convey “hearsay” evidentiality: mbo indicates that you heard from one person, vachi that you heard from many people. We also note that not all Nyala East speakers accept unembedded agreeing complementizers with clausal-EM—though all speakers accept unembedded non-agreeing complementizers, as well as unembedded agreeing complementizers with 1st/2nd agreement. Finally, we note that historically, the agreeing complementizer -chi is derived from the proto-Bantu word for ‘say’ *-ti, but it cannot be synchronically analyzed as a verb. (It is “defective” in the sense of Güldemann 2000.) It does not inflect for tense/aspect/negation, cannot stand as the main predicate of a clause, cannot be inflected for object marking, and cannot take a lexical subject.
Unembedded clausal-EM contributes the same information that verbal-EM does, again forming a scale. (And again, class 9 is largely degraded in clausal-EM.)

- *vichi* $P$ means “according to a lot/all evidence/people, $P$”
- *kachi* $P$ means “according to a few pieces of evidence/people, $P$”
- *lichi* $P$ means “according to a specific piece of evidence/person, $P$”
- *echi* $P$ means “according to a non-specific piece of evidence/person, $P$”

Finally, verbal-EM and clausal-EM (i.e., an embedded agreeing complementizer with EM) can co-occur, in which case they *obligatorily* match. Note that I gloss the markers as EM in the following examples, though I will eventually conclude that only the marker on the complementizer is “true” epistemic marking.

(7) a. *vi*-on-ekh-an-a [ *vi-*chi Ingwe ya-vich-ir-e ]
    $8$EM-see-NEUT-REC-FV $8$EM-COMP 9Ingwe 9SM-win-ASP-FV
    ‘It looks like Ingwe lost (according to a lot of evidence)’

    b. *ka*-on-ekh-an-a [ *ka-*chi Ingwe ya-vich-ir-e ]
    $6$EM-see-NEUT-REC-FV $6$EM-COMP 9Ingwe 9SM-win-ASP-FV
    ‘It looks like Ingwe lost (according to a few pieces of evidence)’

    c. *li*-on-ekh-an-a [ *li-*chi Ingwe ya-vich-ir-e ]
    $5$EM-see-NEUT-REC-FV $5$EM-COMP 9Ingwe 9SM-win-ASP-FV
    ‘It looks like Ingwe lost (according to a specific piece of evidence)’

    d. *e*-on-ekh-an-a [ *e*-chi Ingwe ya-vich-ir-e ]
    $9$EM-see-NEUT-REC-FV $9$EM-COMP 9Ingwe 9SM-win-ASP-FV
    ‘It looks like Ingwe lost (according to a non-specific piece of evidence)’

4. Nominal-EM

As noted above, nouns in Nyala East indicate noun class using an obligatory prefixal noun class marker. The noun class may be preceded by a sometimes omitted augment (cf, Halpert 2017). The augment in Nyala East is only overt on CV noun classes.

(8) e-khuwa
    5NC-word
    ‘word/news’

(9) (a-)ma-khuwa
    6AUG-6NC-word
    ‘words/news’

Certain nouns in Nyala East may also appear with the four EM markers. I refer to this as *nominal-EM*. Nominal-EM appears outside of an “inner” augment—which is obligatory in the presence of nominal-EM—and may appear with an additional augment—which is not obligatory.

(10) a. *(e-)vi-e-khuwa*
    $8$AUG-$8$EM-5NC-word
    ‘word/news’

    c. *(e-)li-e-khuwa*
    $5$AUG-$5$EM-5NC-word
    ‘word/news’

    b. *(a-)ka-e-khuwa*
    $6$AUG-$6$EM-5NC-word
    ‘word/news’

    d. ? *e-e-khuwa*
    $9$EM-5NC-word
    ‘word/news’

(11) a. *(e-)vi-a-ma-khuwa*
    $8$AUG-$8$EM-6AUG-6NC-word
    ‘words/news’

    b. *(a-)ka-a-ma-khuwa*
    $6$AUG-$6$EM-6AUG-6NC-word
    ‘words/news’
Morpho-syntactically, nouns with nominal-EM act like they are in the noun class of the EM. They trigger concordial agreement on adjectives, subject agreement, and object marking, all reflecting the noun class of the EM.

(12) Without EM → Subject/adjectival agreement reflects noun class.

\[\text{e-} \text{mbaha } \text{y-al-i } \text{e-layi} \]
\[\text{9NC-story } \text{9SM-COP-FV 9AGR-good} \]
\[\text{The story was good.}\]

(13) With EM → Subject/adjectival agreement reflects EM.

a. \[\text{e-vi-} \text{e-} \text{mbaha } \text{vi-al-i } \text{vi-layi} \]
\[\text{8AUG-8EM-9NC-story 8SM-COP-FV 8AGR-good} \]
\[\text{The story (that everyone heard) was good.}\]

b. \[\text{a-ka-} \text{e-} \text{mbaha } \text{ka-al-i } \text{ma-layi} \]
\[\text{6AUG-6EM-9story 6SM-COP-FV 6AGR-good} \]
\[\text{The story (that a few people heard) was good.}\]

c. \[\text{e-} \text{li-} \text{e-} \text{mbaha } \text{li-al-i } \text{e-layi} \]
\[\text{5AUG-5EM-9NC-story 5SM-COP-FV 5AGR-good} \]
\[\text{The story (that a specific person heard) was good.}\]

d. \[\text{e-} \text{e-} \text{mbaha } \text{y-al-i } \text{e-layi} \]
\[\text{9EM-9NC-story 9SM-COP-FV 9AGR-good} \]
\[\text{The story (that someone heard) was good.}\]

However, referentially, nouns with EM reflect the “inner” or “true” noun class of the noun. For instance, Class 8 strictly consists of (count) plurals (evitau ‘books’), but does not refer to a plurality when class 8 is used as nominal-EM.

(14) a. \[\text{e-vi-ou-rembo } \text{vi-li } \text{e-ua} \]
\[\text{8AUG-8AGR-14NC-beauty 8SM-COP 5NC-flower} \]
\[\text{(Everyone thinks that) Beauty is a flower.}\]

b. \[\text{*e-vi-tau } \text{vi-lala} \]
\[\text{8AUG-8NC-book 8AGR-one} \]
\[\text{*one books}\]

c. \[\text{e-vi-o-mu-pango } \text{vi-lala} \]
\[\text{8AUG-AGR-3AUG-3NC-plan 8AGR-one} \]
\[\text{one plan (that everyone came up with)}\]

Like verbal- and clausal-EM, nominal-EM reflects an “epistemic scale.” (And again, Class 9 EM is degraded on most nouns.)

- evi-esuvira : ‘a belief according to everyone/based on a lot of things’
- aka-esuvira : ‘a belief according to a few people/based on a few things’
- eli-esuvira : ‘a belief according to someone specific/based on a specific thing’
- ?y-esuvira : ‘a belief according to someone/based on something’
<table>
<thead>
<tr>
<th>Class</th>
<th>Gloss</th>
<th>(e)vi-</th>
<th>(a)ka-</th>
<th>(e)li-</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ousiiva</td>
<td>sadness</td>
<td>eviousiiva</td>
<td>akousiiva</td>
<td>eliousiiva</td>
<td>yousiiva</td>
</tr>
<tr>
<td>oulembo</td>
<td>beauty</td>
<td>evioulembo</td>
<td>akoulembo</td>
<td>elioulembo</td>
<td></td>
</tr>
<tr>
<td>outinyu</td>
<td>difficulty</td>
<td>evioutinyu</td>
<td>akoutinyu</td>
<td>elioutinyu</td>
<td>youtinyu</td>
</tr>
<tr>
<td>oung ali</td>
<td>truth</td>
<td>evioung'ali</td>
<td>akoung'ali</td>
<td>elioung'ali</td>
<td>young'ali</td>
</tr>
<tr>
<td>embaha</td>
<td>story/stories</td>
<td>eviembaha</td>
<td>akeembaha</td>
<td>eliembaha</td>
<td>yembaha</td>
</tr>
<tr>
<td>omupango</td>
<td>plan</td>
<td>eviomupango</td>
<td>akomupango</td>
<td>eliomupango</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Epistemic marking on nouns

The nouns that appear with nominal-EM form a natural class: they are the nouns that are evaluated relative to an individual’s beliefs/judgements. Nominal-EM is ungrammatical on any other noun. Empirically, nominal-EM can appear on any noun that is either propositional (i.e., has propositional content) or subjective. A non-exhaustive list is provided in Table 1.

Unlike with verbal-EM, nominal-EM does not require a (selected) CP. This is most evident with subjective nouns like ourembo ‘beauty,’ which does not subcategorize for a CP.

5. Interim summary

Summarizing the data above, the first observation is that across the three categories, EM clearly makes the same semantic contribution. In all three cases, it indicates a source vis-a-vis the relevant category it appears on. I moreover note that the morphological reflex of the EM across its three uses is revealing: it is ostensibly agreement morphology. Finally, it is important to observe that verbal-EM is exceptional in that it has a co-occurrence restrictions: verbal-EM is only possible in the presence of a selected CP. (Cf Iatridou & Embick 1997 for a similar observation in Greek)
c. * [ Ingwe ya-vichir-e ] na
9SM-win-FV and EM-be.surprised-CAUS-FV
[Intended: ‘Ingwe won and it surprised him/her.’]

Not only does verbal-EM require a selected CP, the CP must be syntactically local. An intervening
NP bearing $\phi$-features is an intervener for verbal-EM.

\begin{align*}
&\{\text{vi-} \\
&\text{ka-} \\
&\text{li-} \\
&e-\}
\end{align*}

\begin{align*}
\text{ali mbo } & \text{ Ingwe ya-khil-il-e} \\
\text{EM-COP COMP } & \text{9SM-win-ASP-FV}
\end{align*}

‘It’s sure that Ingwe won.’ (Lit: ‘It’s that Ingwe won.’)

\begin{align*}
&\{\text{*vi-} \\
&\text{*ka-} \\
&\text{*li-} \\
&\text{*e-}\}
\end{align*}

\begin{align*}
\text{al-i ou-veyi mbo } & \text{ Ingwe ya-khil-il-e} \\
\text{EM-COP-FV 14NC-lie COMP } & \text{9SM-win-ASP-FV}
\end{align*}

[Intended: ‘It was a lie that Ingwe won.’]

\begin{align*}
\text{w-ali } & \text{ ou-veyi mbo } \text{ Ingwe ya-khil-il-e} \\
\text{14SM-COP 14NC-lie COMP } & \text{9SM-win-ASP-FV}
\end{align*}

‘It was a lie that Ingwe won.’

Thus verbal-EM is clearly distinguished from clausal- and nominal-EM, which have no such
requirement on an associated selected constituent.

In the analysis below I make sense of these facts by analyzing all epistemic marking as a reflex of
a functional projection (D, C, T) agreeing with a perspectival operator. The analysis accounts for the
disparity between nouns and clauses on the one hand and verbs on the other by positing perspectival
operators only in two domains, DPs and CPs. Verbal-EM thus must derive from an inherently non-local
agreement relationship, through agreement with one of those operators. Since verbal-EM always appears
in the presence of a syntactically local CP, I conclude that the source of verbal-EM is a perspectival
operator in the embedded CP (Gluckman, to appear).

6. Analysis

Starting with clausal-EM, we adopt Diercks’s (2013) analysis of the agreeing complementizer in the
closely related Lubukusu. Diercks analyses the agreeing complementizer as agreement (formally, Agree
Chomsky 2000, 2001) with an “anaphoric” operator in spec-CP. The function of the operator is to fix the
source or perspective of the embedded clause.

\begin{align*}
&\text{Masika ka-mb-ol-ir-e} \quad \text{a-chi} \quad \text{Ingwe ya-vich-ir-e} \\
&\text{Masika 1SM-1SG.OM-say-APPL-FV 1SM-COMP 9Ingwe 9SM-win-ASP-FV}
\end{align*}

‘Masika told me that Ingwe won.’
Importantly, Op’s reference (and class features) is *syntactically* independent of its antecedent; it is established via semantic/pragmatic principles (Diercks, 2013; Sundaresan, 2018). Diercks demonstrates this independence by showing that Op and the higher subject do not exhibit intervention effects caused by an intervening DP: the 1st person indirect object in (18a) cannot trigger 1st person agreement on the complementizer.\(^8\) Thus, in (18b), Op is merged with Class 1 features; it does not acquire them through syntactic mechanisms. C agrees with Op and realizes Class 1 agreement.

We propose that clausal-EM is derived from C agreeing with an operator that is imbued with class features which index a contextually salient group of people or evidence:

\[(19)\] Where X ranges over the contextually salient people or evidence:

a. Op\(_{φ,8}\) : all of X
b. Op\(_{φ,6}\) : some of X
c. Op\(_{φ,5}\) : a specific member of X
d. Op\(_{φ,9}\) : a non-specific member of X.

\[(20)\] a. \{vi-\(\phi\), ka-\(\phi\), li-\(\phi\)\} chi Masika ka-chi-e Nairobi EM-COMP 1Masika 1SM-go-FV Nairobi

‘Masika went to Nairobi (according to . . . ).’

Turning to verbal-EM, as noted above, it is not possible to analyze verbal-EM in a parallel fashion—i.e., by positing a dedicated perspectival operator in the verbal (or temporal) domain. Verbal-EM is *syntactically* dependent on a local CP. It is thus distinct from the relationship illustrated in (18b) between subject agreement on \(T\) in the matrix clause and the perspectival operator. I therefore analyze verbal-EM as subject agreement (located on \(T\)) with the perspectival operator in the embedded clause.

\(^8\) Dierck’s analysis involves LF raising and checking/matching of the operator with the subject in spec-TP. See Carstens (2016) and Diercks et al. (2020) for an alternative analysis that relates the subject-orientation to head-movement of C to \(v\). In the analysis of clausal-EM below, the requirement that Op must “match” a higher subject is vacuously satisfied since there is no higher subject.
In support of this analysis, we note that verbal-EM and an embedded complementizer (with or without clausal-EM) are subject to intervention effects (see (17) above). In (22), only class 14 subject agreement is permitted on the copula; verbal-EM is not permitted.

(22) w-ali ou-vehi chi Ingwe ya-khil-il-e
14SM-COP 14NC-lie EM-COMP 9Ingwe 9AGR-win-ASP-FV
‘It’s a lie that Ingwe won.’

Because nominal-EM is not dependent on another constituent, I suggest that it works in a parallel fashion to clausal-EM: there is an “agreeing determiner” in Nyala East, whose function is to “orient” the perspective of any noun which can be oriented (i.e., subjective/propositional nouns). I assume that D is realized as the augment in Nyala East (Visser, 2008; Ndayiragije et al., 2012; Gambarage, 2019), which is null with V-initial noun classes.⁹

(23) a. vi-ka-li-e
  EM-see-NEUT-FV Masika ka-chi-e Nairobi
  1SM-go-FV Nairobi
‘(According to . . . ) it looks like Masika went to Nairobi.’

Op bears φ features, and so can be targeted for agreement by other heads, therefore nouns with nominal-EM trigger agreement and concord referencing the EM class features. Crucially, Op does not change the referential properties of the noun, and so nouns with EM reflect the reference of the “inner” noun clause.

7. Remaining issues and discussion

In sum, Nyala East suggests a semantic parallelism between the CP and DP domain—a parallelism that does not extend to the verbal domain. However, we also emphasize that the perspective in CPs and DPs in Nyala East still has distinct respective behaviors. Unlike the agreeing complementizer, D does not track anything like subject-orientation. I would suggest that this reflects a diachronic rather than synchronic fact: the agreeing complementizer was historically a verb (proto-Bantu *-ti ‘say’) and so was already able to reflect a non-local dependency. The agreeing determiner was never a verb, and so is not capable (yet) of a non-local dependency.

⁹ An alternative analysis that we ultimately must reject is treating nominal-EM as an instance of “reduced” associative/linker, otherwise available in colloquial speech. This strategy crucially changes the referential properties of the entire DP, and so does not explain why, e.g., eviouvehi refers to a lie, rather than ‘things of a lie,’ with an elided evinju ‘things.’
References


