

## Expletive agreement, evidentiality, and modality in Logooli \*

John Gluckman  
UCLA

Margit Bowler  
UCLA

**Abstract** We discuss and analyze two subject agreement markers in Logooli (Bantu, Kenya). We show that *e-* (class 9 subject agreement) and *ga-* (class 6 subject agreement) give rise to a variety of apparently evidential or modal meanings when they occur in constructions translated with “expletive” subjects. We propose a treatment of the Logooli data following Matthewson, Rullmann & Davis’s (2007) and Rullmann, Matthewson, & Davis’s (2008) choice function analysis of modality and evidentiality in St’át’imcets, and extend their original proposal to account for novel data in Logooli. We show that these two morphemes occur only with verbs that introduce modal bases, and propose that they differ from one another in the size of the subset of possible worlds that their associated choice functions select from the modal base. This in turn results in different interpretations based on the size of the subset of worlds that they select, the speaker’s ordering source, and the modal base provided by the verb.

**Keywords:** Bantu, choice functions, evidentiality, modality

Logooli (Bantu, Kenya) has two subject agreement markers, *e-* and *ga-*, that productively mark noun class 9 and noun class 6 subject agreement, respectively:<sup>1</sup>

- (1) a. i-ŋoombe e-futi-ki  
9-cow 9-inflate-AC  
‘The cow inflated’<sup>2</sup>

---

\* We would like to thank our wonderful Logooli consultant, Mwabeni Indire, for generously sharing his time and his language with us. Thanks also to Yael Sharvit, Maayan Abenina-Adar, and audience members at SALT 26, ACAL 47, UCLA Semantics Tea, and the UCLA American Indian Seminar. We particularly thank Mike Diercks, Mary Paster, and Meredith Landman at Pomona College for their generosity and expertise in all things Logooli. This research was funded in part by NSF GRFP grant DGE-1144087, and a Lenart Graduate Travel Fellowship. All errors are our own.

<sup>1</sup> Logooli (also called Luragooli, Maragoli, and Lulogooli, among other names) is a Bantu language in the Luhia subfamily. It is spoken in western Kenya and Tanzania by around 600,000 people (Lewis, Simons & Fennig 2016). Logooli exhibits a range of phenomena that are typically associated with Bantu languages, including two tones and a complex tense/aspect system. We ignore both for the purpose of this paper. See Leung (1991) and Samuels & Paster (2015) for discussions of Logooli tone.

<sup>2</sup> The following abbreviations are used in this paper:

- b. ami-ino ga-gw-ε  
6-tooth 6-fall-FV  
'The teeth fell'

However, these agreement markers also occur in response to prompts with expletive subjects. This occurs despite the fact that the utterances do not include either class 6 or class 9 subjects:

- (2) a. e-ror-ek-a ndee Sira a-gw-ε  
9-look-AC-FV that 1Sira 1-fall-FV  
'It looks like Sira fell'
- b. ga-ror-ek-a ndee Sira a-gw-ε  
6-look-AC-FV that 1Sira 1-fall-FV  
'It looks like Sira fell'

When *e-* and *ga-* occur in constructions like (2), the morphemes contribute additional meanings to the utterance. These meanings vary in part based on the verb that the agreement markers combine with; they can include marking direct versus indirect speaker perception of the evidence for the embedded proposition *P*, widespread versus restricted knowledge of *P*, and modal force, among other things. That is, even though *e-* and *ga-* might be referred to as marking an “expletive” subject, they clearly contribute something semantically non-trivial.<sup>3</sup>

In this paper, we provide an analysis of *e-* and *ga-* in the spirit of Matthewson et al.’s (2007) and Rullmann et al.’s (2008) treatment of modality and evidentiality in St’át’imcets (henceforth, RMD). Following RMD, we propose that *e-* and *ga-* are associated with choice functions that operate over the best possible worlds in the modal base as supplied by the verb and restricted by the contextually supplied ordering source. *e-* and *ga-* differ in the size of the subset that this choice function selects; conceptually speaking, *ga-* picks out a larger subset of worlds in the modal base than *e-*.

In the following sections, we show how this accounts for the various interpretations of *e-* and *ga-*, ranging from evidential to modal. This project also represents a

I-17 : class markers	COP : copula	PASS : passive	REC : reciprocal
SG/PL : singular/plural	FUT : future	POSS : possessive	REFL : reflexive
AC : anticausative	FV : final vowel	PROG : progressive	
CAUS : causative	NEG : negative	PRT : particle	

Logooli has approximately 17 noun classes. We use numbers to gloss these classes, following Bantuist convention.

<sup>3</sup> Syntactic theories typically treat English expletive subjects (also called “pleonastic subjects”) as semantically vacuous, although it has also been argued for English that expletive subjects may contribute evidential-like semantics (e.g., Grimm (2006), among others).

first pass at examining evidentiality in Bantu, a topic that is not yet well described or understood.

## 1 Logooli data

In the following sections, we report the interpretations of *e-* and *ga-* in combination with various semantic classes of verbs. In [section 2](#), we give our proposal and walk through a Logooli example according to this analysis. In [section 3](#), we discuss the predictions of our analysis.

### 1.1 Indirect (*e-*) versus direct (*ga-*) perception

In combination with verbs of perception like *kufana* ‘to seem’ and *kuroreka* ‘to appear,’ the use of *e-* conveys that the speaker has indirectly perceived evidence for the truth of the embedded proposition. Conversely, *ga-* conveys that the speaker has directly perceived evidence for the truth of the proposition:<sup>4</sup>

- (3) Context: It’s flu season, and Imali didn’t come to school. The speaker says:
- a. e-fan-a kuresa Imali a-saal-a  
9-seem-FV like 1Imali 1-be.sick-FV  
‘It seems like Imali is sick’
  - b. # ga-fan-a kuresa Imali a-saal-a  
6-seem-FV like 1Imali 1-be.sick-FV  
‘It seems like Imali is sick’
- (4) Context: The speaker sees Imali coughing and sneezing. They say:
- a. ? e-fan-a kuresa Imali a-saal-a  
9-seem-FV like 1Imali 1-be.sick-FV  
‘It seems like Imali is sick’
  - b. ga-fan-a kuresa Imali a-saal-a  
6-seem-FV like 1Imali 1-be.sick-FV  
‘It seems like Imali is sick’  
Speaker’s comment: “(4b) is only appropriate if you’re looking at Imali.”

Our consultant noted that it’s not completely infelicitous to use (4a) in the context in (4). This is because it is also possible that Imali’s sneezing and coughing is due

<sup>4</sup> In this paper, we ignore the various “detransitivizing” verbal suffixes that can co-occur with *e-* and *ga-*. These can include (exclusively) the anticausative *-Vk*, the passive *-w*, and the reciprocal *-an* (in combination with the anticausative). The contributions of these suffixes in the expletive constructions are not yet well understood (cf. Seidl & Dimitriadis 2003).

to allergies, rather than to sickness. Since even direct perception is still compatible with speaker doubt, the use of *e-* in this context is still available. What is crucial, however, is that the direct perception in the context in (4) enables the speaker to use *ga-* and make the strongest claim possible, unlike in (3).

As a first pass, the data in (3-4) might seem to support an evidential analysis of *e-* and *ga-*. However, data in the following provides evidence against this treatment. We further address why a purely evidential analysis cannot account for the Logooli data in subsection 3.2.

## 1.2 Restricted (*e-*) versus general (*ga-*) knowledge

In combination with attitude report verbs like *kumanyeka* ‘to be known,’ *kusoverwa* ‘to be believed,’ and *kuvoleka* ‘to be said,’ *e-* conveys that a proposition is “privileged” or “restricted” knowledge. Conversely, *ga-* conveys “general” knowledge. Our consultant sometimes indicates this by glossing his Logooli utterances with “It is not widely {known / believed / said} that *P*” and “It is widely {known / believed / said} that *P*” for *e-* and *ga-*, respectively:

- (5) a. *ga-many-ek-a ndee Kurt Cobain y-i-isuŋga*  
 6-know-AC-FV that Kurt Cobain 1-REFL-kill  
 ‘It is (well) known that Kurt Cobain killed himself’
- b. *e-many-ek-an-i ndee Kurt Cobain y-aremban-a na m-kari*  
 9-known-AC-REC-FV that Kurt Cobain 1-argue-FV PRT 1-wife  
 w-eve  
 1-POSS  
 ‘It is (not well) known that Kurt Cobain argued with his wife’

The speaker need not have witnessed either the death of Kurt Cobain, or Kurt Cobain arguing with his wife, to felicitously utter (5a) or (5b). Instead, the speaker is merely relating the information that he knows, not relating how he acquired the information. This contrasts with the use of *e-* and *ga-* in combination with the perception verbs in subsection 1.1.

## 1.3 Less affectedness (*e-*) versus more affectedness (*ga-*)

In combination with emotive factive predicates like *kurereriza* ‘to be sad,’ *kufuniza* ‘to be surprised,’ and *kugenia* ‘to be odd’/‘to be strange,’ *e-* expresses that the speaker is less affected in terms of the relative emotion, while *ga-* expresses that the speaker is extremely affected in terms of the relative emotion:<sup>5</sup>

<sup>5</sup> These verbs form a natural class only when they occur with an expletive subject; *John is sad* reports John’s emotional state, whereas *John is odd* reports the speaker’s attitude towards John.

(6) Context: Maina is a huge Lakers fan. If the Lakers lose a game, he can say:

- a. # e-verer-iz-a ndee Lakers va-goot-w-i  
 9-be.sad-CAUS-FV that 2Lakers 2-defeat-PASS-FV  
 ‘It’s sad that the Lakers lost’<sup>6</sup>
- b. ga-verer-iz-a ndee Lakers va-goot-w-i  
 6-be.sad-CAUS-FV that 2Lakers 2-defeat-PASS-FV  
 ‘It’s sad that the Lakers lost’

(7) Context: Sira is a casual Lakers fan. If the Lakers lose a game, he can say:

- a. e-verer-iz-a ndee Lakers va-goot-w-i  
 9-be.sad-CAUS-FV that 2Lakers 2-defeat-PASS-FV  
 ‘It’s sad that the Lakers lost’
- b. # ga-verer-iz-a ndee Lakers va-goot-w-i  
 6-be.sad-CAUS-FV that 2Lakers 2-defeat-PASS-FV  
 ‘It’s sad that the Lakers lost’

Note that the contrast between (6) and (7) has nothing to do with whether the speaker has (in)direct evidence for the proposition, or whether the information is widely known or not. This usage of *e-* and *ga-* to express “speaker affectedness” is distinct from the two previous uses.

#### 1.4 Weaker modal force (*e-*) and stronger modal force (*ga-*)

In combination with modal verbs, *e-* signals weak modal force, whereas *ga-* signals strong modal force.<sup>7</sup> As shown in the following examples, our consultant sometimes glosses his Logooli utterances with “should” (for *e-*) and “must” (for *ga-*):

(8) Context: A school-age child is skipping school. He runs into another child skipping school, who tells him:

- a. e-dukan-a ndee u-zi-ε m-skolu m-soom-e  
 9-arrive-FV that 2sg-go-FV in-school PART-study-FV  
 Consultant’s gloss: ‘You should go to school.’  
 Literally: ‘It’s required that you go to school and study’

<sup>6</sup> Literally: ‘It’s saddening that the Lakers were defeated.’

<sup>7</sup> By “modal verbs,” we mean verbs that introduce modal bases ordered according to some contextually supplied ordering source. The verb in (8)-(9), *kudukana*, is morphologically complex; it consists of *kuduka* ‘to arrive’ together with the reciprocal suffix *-an*. Interestingly, we have not found any Logooli verbs that have solely modal semantics. Instead, Logooli modal verbs have both modal and non-modal uses. In addition to *kudukana*, these verbs include *kunyala* ‘to manage’/‘to be able,’ and *kwenyeka* ‘to be wanted’/‘to be expected’/‘should.’

- b. # ga-dukan-a ndee u-zi-ε m-solu m-soom-e  
6-arrive-FV that 2sg-go-FV in-school PRT-study-FV  
Consultant’s gloss: ‘You must go to school.’  
Literally: ‘It’s required that you go to school and study’
- (9) Context: A school-age child is skipping school. He runs into a police officer, who tells him:
  - a. # e-dukan-a ndee u-zi-ε m-skolu m-soom-e  
9-arrive-FV that 2sg-go-FV in-school PART-study-FV  
Consultant’s gloss: ‘You should go to school.’  
Literally: ‘It’s required that you go to school and study’
  - b. ga-dukan-a ndee u-zi-ε m-solu m-soom-e  
6-arrive-FV that 2sg-go-FV in-school PRT-study-FV  
Consultant’s gloss: ‘You must go to school.’  
Literally: ‘It’s required that you go to school and study’

Unlike the data in the previous sections, the use of *e-* and *ga-* in (8-9) does not supply any information about directness or indirectness of evidence, generality of knowledge, or speaker affectedness. Rather, the data in (8-9) strongly suggests that *e-* and *ga-* interact with the modal base.

### 1.5 Summary

We summarize the contributions of *e-* and *ga-* with respect to the different classes of predicates in Table 1 below. Given this range of meanings, our task now is to give a compositional semantics for this data.

	<b>Perception verbs</b>	<b>Attitude report verbs</b>	<b>Emotive factives</b>	<b>Modals</b>
<i>e-</i>	indirect evidence	restricted knowledge	less affected	weak
<i>ga-</i>	direct evidence	general knowledge	more affected	strong

**Table 1** Summary of the contributions of *e-* and *ga-*.

## 2 Proposal: Choice functions operating over modal bases

Our proposal is in the spirit of RMD’s treatment of modality and evidentiality in St’át’imcets. We propose that *e-* and *ga-* introduce choice functions, and follow

RMD in assuming that choice functions can take a set as an argument and return some subset of that set.<sup>8</sup> In the case of *e-* and *ga-*, the choice function  $f$  associated with these morphemes operates over the set of the best possible worlds as determined by the ordering source.  $f$  takes this set of worlds as its argument and returns a subset of these worlds. That is, for any set of possible worlds  $A$ ,  $f(A) \subseteq A$ .

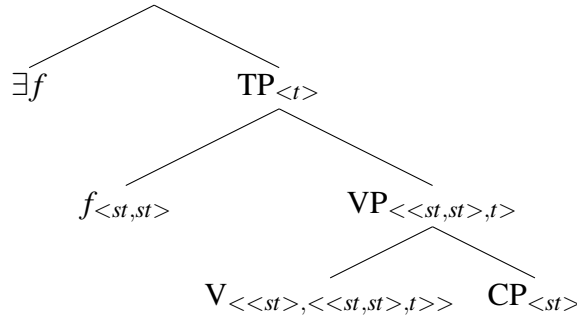
Given this assumption, the expletives in Logooli must combine with a lexical item that contributes a modal base for  $f$  to operate over.<sup>9</sup> This is supported by the set of Logooli verbs that *e-* and *ga-* can combine with; these verbs are typically treated as being intensional. This set of verbs includes *kumanyeka* ‘to be known,’ *kuvoleka* ‘to be said,’ *kufana* ‘to seem like,’ and so on.

Following this proposal, we give a basic denotation for a Logooli modal verb:<sup>10</sup>

$$(10) \quad \llbracket \text{MODAL VERB} \rrbracket^{w, OS, MB, ah} = \lambda P_{st} \lambda f_{st, st}: OS \text{ is appropriate for } MB \text{ and } f \text{ is a choice function. } \forall w' [w' \in f(\text{BEST}_{OS(ah, w)}(\text{MB}(ah, w))) \rightarrow P(w') = 1]$$

In this denotation, Logooli modal verbs combine with an embedded proposition  $P$  and a choice function  $f$ .  $P$  is true in all of the worlds picked by the choice function from the set of best possible worlds determined by the speaker’s modal base and ordering source. In (11), existential closure is provided at the top of the tree.

(11)



We crucially propose that *ga-* and *e-* differ in the size of the set of possible worlds that they select. The stronger of the two morphemes, *ga-*, selects a non-proper and non-empty subset of the best possible worlds in the modal base: that is, it returns the same set of best possible worlds. This effectively amounts to universal quantification. When a speaker uses *ga-*, they make a strong claim; they assert that the embedded proposition  $P$  is true in all the best possible worlds in the modal base.

<sup>8</sup> Choice functions were initially proposed by Reinhart (1997) to take a set and return an element from that set. This was used to account for the interpretation of indefinite DPs.

<sup>9</sup> We assume a traditional definition of a modal base as a set of worlds accessible from the actual world ordered according to some contextually supplied ordering source (following Kratzer 1991, Hacquard 2011, among many others).

<sup>10</sup> We abbreviate “ordering source” as *OS*, “modal base” as *MB*, and “attitude holder” as *ah*.

Conversely, when a speaker uses *e-*, they make a weaker claim. We propose that *e-* simply selects a non-empty subset of the best possible worlds in the modal base. In contrast to *ga-*, *e-* amounts to existential quantification. That is, *e-* asserts that there is some world in the modal base in which *P* is true.

Given this difference in strength, we propose that the choice functions *e-* and *ga-* form the scale  $\langle e-, ga- \rangle$ . Since *e-* is weaker than *ga-*, if *f* in (11) is occupied by *e-*, the utterance is strengthened through scalar implicature to the meaning “*e-* and not *ga-*.” (We remain agnostic as to which mechanism should be used to derive this scalar implicature.) This strengthened meaning is in accordance with our Logooli consultant’s observation that it is generally infelicitous to use *e-* in contexts in which *ga-* is licensed.<sup>11</sup> On the other hand, since there is no stronger scalar alternative to *ga-*, utterances containing *ga-* in *f* cannot be further strengthened. It is infelicitous to use *ga-* in contexts in which *e-* is licensed simply because it makes too strong of an assertion.

## 2.1 Walkthrough of an example according to our analysis

The following sentences give examples of *e-* and *ga-* in combination with *kuholeka* ‘to hear.’ In these examples, the stronger morpheme, *ga-*, is used to indicate that the speaker has direct evidence for the embedded proposition. Conversely, the weaker morpheme, *e-*, is used to indicate indirect evidence.

(12) Context #1: The speaker heard a loud party happening next door to their apartment; that is, they have direct evidence for the embedded proposition. They say:

- a. ga-hol-ek-a kuresa vu-geni vu-are vu-rahe  
6-hear-AC-FV like 15-party 15-COP 15-good  
‘It sounds like the party was fun’
- b. # e-hol-ek-a kuresa vu-geni vu-are vu-rahe  
9-hear-AC-FV like 15-party 15-COP 15-good  
‘It sounds like the party was fun’

In (12), the speaker’s modal base contains all the worlds that are compatible with what they have heard. This includes things like loud music, singing, shouting, and so on. These worlds within their modal base are then ordered according to their ordering source, which includes facts like “Parties with loud music are fun,” “Parties with singing are fun,” and so on. In the context in (12), the embedded proposition *The party was fun* is true in all of the best worlds in the speaker’s modal base, as

<sup>11</sup> One noteworthy exception to this is (4), as previously mentioned.



determined by these facts. The speaker therefore uses *ga-* to indicate that *P* is true in all of these worlds.<sup>12</sup>

- (13) Context #2: The speaker’s friend tells them that a party he attended was fun. However, the speaker did not attend or overhear the party themselves; they only have indirect evidence for the embedded proposition. They say:
- a. # *ga-hol-ek-a* *kuresa vu-geni vu-are vu-rahe*  
6-hear-AC-FV like 15-party 15-COP 15-good  
‘It sounds like the party was fun’
  - b. *e-hol-ek-a* *kuresa vu-geni vu-are vu-rahe*  
9-hear-AC-FV like 15-party 15-COP 15-good  
‘It sounds like the party was fun’

In (13), the speaker’s modal base contains all the worlds that are compatible with what their friend has told them; that is, that the party was fun. However, the facts within the speaker’s ordering source that determine the best possible worlds differs from in (12). In this case, the speaker’s ordering source includes facts like “The speaker’s friend enjoys crowded parties, but the speaker does not,” “People generally think that parties with loud music are fun, but the speaker doesn’t like loud music,” “Secondhand information is generally unreliable,” and so on. Given this context and ordering source, the best worlds in the speaker’s modal base includes worlds in which the embedded proposition (*The party was fun*) is false. The speaker therefore uses *e-* to indicate that the set of best possible worlds in their modal base also includes worlds in which *P* is false.

The interpretation of *e-* and *ga-* as marking indirect versus direct evidence in (12-13) falls out from a combination of the speaker’s ordering source and the contents of their modal base, as determined by the verb *kuholeka* ‘to hear.’ This is a verb of perception, meaning that the best possible worlds in the modal base should reflect the perceptual evidence that the speaker receives from the actual world. When *e-* and *ga-* are used in combination with such a verb of perception, their choice reflects whether the perceptual evidence in their set of best possible worlds supports the truth of *P*, vis-a-vis their ordering source. We propose that a similar analysis holds for attitude report verbs, emotive factives, and modals.

### 3 Predictions of the analysis

In the following sections, we review some of the predictions that our analysis makes, and show that they are supported by Logooli data.

---

<sup>12</sup> We assume that speakers do not use *e-* in this context because they are obligated to make the strongest claim possible, given their evidence. This follows from Grice’s maxim of quantity (Grice 1975).

### 3.1 Modifying the ordering source

The choice of *e-* versus *ga-* hinges on the contents of the best possible worlds in the speaker's modal base. Since this set of possible worlds is determined by the speaker's ordering source, altering the set of facts within their ordering source can alter this set of best possible worlds. This in turn can affect the speaker's choice of *e-* versus *ga-*.

This is shown in the following example, in which two speakers must differ in their choice of *e-* versus *ga-* due to having different ordering sources. The difference in their ordering sources stems from their background knowledge, and not from any difference in perceptual input:

- (14) Context: Imali and Maina are watching Roger Federer play in a tennis match. Imali is a huge tennis fan and knows all the rules and statistics of the game. However, Maina is only vaguely familiar with the rules, and otherwise knows nothing about tennis. Roger Federer just served an ace.
- a. ga-ror-ek-a kuresia Federer a-kin-i vurahe karono  
6-look-AC-FV like 1Federer 1-play-FV well today  
'It looks like Federer is playing well today'  
✓ if Imali says this, # if Maina says this
  - b. e-ror-ek-a kuresia Federer a-kin-i vurahe karono  
9-look-AC-FV like 1Federer 1-play-FV well today  
'It looks like Federer is playing well today'  
# if Imali says this, ✓ if Maina says this

Imali, the knowledgeable speaker, can only felicitously use *ga-* in this context. This is because her ordering source contains the actual rules of tennis. She can therefore accurately judge whether Federer is playing well or poorly, according to these rules. Given her visual evidence, she uses *ga-* to make the strong claim that he is playing well. That is, in all of the best possible worlds in Imali's modal base, as determined by her ordering source, Roger Federer is playing well.

Maina, the less knowledgeable speaker, can only felicitously use *e-*. This falls out from the fact that his ordering source does not contain the actual rules of tennis.<sup>13</sup> Since he does not know what makes a tennis player "good" in the actual world, there are worlds in his modal base in which Roger Federer is not playing well according to these actual rules. As a result, he uses *e-* to make a weaker claim about Federer's performance.

<sup>13</sup> We also assume in this context that Maina is aware of the fact that he is not knowledgeable about the actual rules of tennis. That is, we assume that Maina is aware of his own ignorance. If Maina were to mistakenly believe that he is an expert on tennis, he could then felicitously utter *ga-* in the context in (14).

### 3.2 Embedding and shifting

Some evidentials have been argued to operate at the speech act level as illocutionary operators (Faller 2002; Murray 2010). These illocutionary evidentials typically resist embedding. We find, however, that Logooli *e-* and *ga-* can both embed and shift. This is compatible with our treatment of *e-* and *ga-* as choice functions over sets of possible worlds, as opposed to speech act level illocutionary operators:

(15) Context: Sira has not seen Imali, but has heard through a secondhand source that she is sick. However, the speaker has seen Imali, and they think that she looks sick. Sira tells the speaker that according to this secondhand source, it appears that Imali is sick. (That is: the speaker has direct evidence for *P*, and the subject has indirect evidence for *P*.) The speaker can report:

- a. Sira a-ganagan-a ndee e-ror-ek-a ndee Imali ne mu-rwaye  
1Sira 1-think-FV that 9-look-AC-FV that 1Imali COP 1-sick  
'Sira thinks that it appears that Imali is sick'
- b. # Sira a-ganagan-a ndee ga-ror-ek-a ndee Imali ne mu-rwaye  
1Sira 1-think-FV that 6-look-AC-FV that 1Imali COP 1-sick  
'Sira thinks that it appears that Imali is sick'

In (15), the speaker can felicitously utter *e-* despite the fact that they have direct evidence for *P* (i.e., evidence that would typically license the use of *ga-*). This use of *e-* has therefore been shifted; it is interpreted relative to the modal base of the matrix subject, rather than of the speaker.<sup>14</sup>

Conversely, in (16) below, the speaker can felicitously utter *ga-* despite the fact that they do not have direct evidence for *P*. This use of *ga-* reflects the modal base of the matrix subject, rather than the speaker:

(16) Context: Sira has seen Imali, and tells the speaker that he thinks that Imali looks sick. The speaker has not seen Imali. (That is: the speaker has indirect evidence for *P*, and the subject has direct evidence for *P*.) The speaker can report:

- a. # Sira a-ganagan-a ndee e-ror-ek-a ndee Imali ne mu-rwaye  
1Sira 1-think-FV that 9-look-AC-FV that 1Imali COP 1-sick  
'Sira thinks that it appears that Imali is sick'

<sup>14</sup> Our consultant noted that *ga-* could be felicitous in (15) only if the speaker considers the source of Sira's information to be unreliable. The use of *e-* here reflects that Sira, not the speaker, does not consider the source to be reliable.

- b. Sira a-ganagan-a ndee ga-ror-ek-a ndee Imali ne mu-rwaye  
 1Sira 1-think-FV that 6-look-AC-FV that 1Imali COP 1-sick  
 ‘Sira thinks that it appears that Imali is sick’

Therefore, when *e-* or *ga-* is embedded under an attitude verb like *kuganagana* ‘to think,’ their use reflects the knowledge state of the local attitude holder, and not of the speaker. That is, *e-* and *ga-* can shift. We note that this parallels data on English modals, which also shift when embedded:

- (17) Howard thinks that Sarah {must / should / might} go to London, but I disagree.

### 3.3 Unavailability of *e-* and *ga-* in combination with non-modal verbs

In (11), we proposed that the choice function associated with *e-* and *ga-* is of type  $\langle st, st \rangle$ . Given this semantic type, the choice function must combine with a verb that supplies a modal base. *e-* and *ga-* therefore should be unable to combine with verbs that do not supply a modal base.

This prediction is borne out by the Logooli data. The agreement morphemes cannot grammatically combine with any non-modal verbs. For instance, *e-* and *ga-* cannot combine with weather verbs, as shown in (18) below. This differs from the English expletive subject *it*, which can grammatically co-occur with non-modal verbs, as shown in (19):

- |      |    |   |    |   |
|------|----|---|----|---|
| (18) | a. | riova ri-val-a<br>5sun 5-shine-FV<br>Prompt: ‘It is sunny’<br>Literally: ‘The sun is shining’ | b. | * {e- / ga-} val-a<br>9- / 6- shine-FV<br>Intended: ‘It is sunny’ |
|------|----|---|----|---|

- (19) It is raining outside.

We note that the data in (18) does not preclude the existence of a genuine expletive morpheme in Logooli. However, we have not yet found any such item, and neither *e-* nor *ga-* can serve this function.

### 3.4 Non-speaker orientation

In subsection 1.2, we described data in which the speaker uses *e-* or *ga-* to signal whether the embedded proposition *P* is “privileged” knowledge or “general” knowledge, respectively. We propose that the classification of *P* as “privileged” or “general” knowledge depends on the group of individuals that the speaker “associates” themselves with, or speaks on behalf of (cf. Moltmann (2006, 2012)). That

is, when a speaker reports that *P* is widely known, they are expressing that *P* is widely known among the group of individuals that they speak on behalf of.

This group can vary contextually. We therefore predict that in the right context, the speaker can use *ga-* to report “widely known” information that is only “widely known” in certain communities, or even that the speaker themselves might not believe. This prediction is supported by the following example:

- (20) Context: A scientist visits a village where everyone believes that the world is flat. Although the scientist himself doesn’t believe that the world is flat, and many people outside the village do not think that the world is flat, he can felicitously state:
- ga-ganagan-aj-w-a    yeeno ndee ri-lova    ri-a    bameka  
 9-think-PROG-PASS-FV here    that 5-world 5-COP flat  
 ‘It’s (widely) thought here that the world is flat’

In this context, the scientist reports what people in the village generally think, not what he thinks, or what other people outside of the village think. That is, the use of *ga-* in (20) reflects the modal bases and ordering sources of the people in the village, not of the scientist.

We take this data as further evidence against an analysis of *e-* and *ga-* as evidentials. The data in (20) shows that these morphemes can be non-speaker-oriented; however, evidentials are typically strongly speaker-oriented, and resist shifting (cf. the data in subsection 3.2) (Aikhenvald 2004; Sauerland & Schenner 2007)).

### 3.5 Scalar alternatives and continuations

In 2, we proposed that *e-* and *ga-* form a scale in which *ga-* is a stronger scalar alternative to *e-*:  $\langle e-, ga- \rangle$ . As a result, we predict that these morphemes should pattern like other scalar items with respect to the availability of continuations.

A weaker scalar item can be strengthened to a stronger scalar item through an overt continuation introduced by *in fact*. This continuation presupposes the truth of the prejacent, and asserts an additional proposition, which may strengthen the prejacent. However, an assertion containing a strong scalar item cannot be followed by an *in fact* continuation with a weaker scalar item; the weaker scalar item cannot be true if *in fact* presupposes that the stronger item is also true. We show this data in (21) below with the English quantifiers  $\langle some, all \rangle$ :

- (21) a. ✓ Vince ate some of the cookies... in fact, he ate all of the cookies.  
 b. # Vince ate all of the cookies... in fact, he ate some of the cookies.

We observe similar behavior with respect to the Logooli expletives. An utterance containing *e-* can be felicitously followed by a continuation with *ga-*, as in (22a);

however, an utterance with *ga-* cannot felicitously be followed by a continuation with *e-*, as in (22b):

- (22) a. e-ror-ek-a ndee Sira a-ze-e... naandio ga-ror-ek-a ndee Sira  
 9-look-AC-FV that 1Sira 1-go-FV in.fact 6-look-AC-FV that 1Sira  
 a-ze-e  
 1-go-FV  
 ‘It looks like Sira left... in fact, it (really) looks like Sira left.’
- b. # ga-ror-ek-a ndee Sira a-ze-e... naandio e-ror-ek-a ndee Sira  
 6-look-AC-FV that 1Sira 1-go-FV in.fact 9-look-AC-FV that 1Sira  
 a-ze-e  
 1-go-FV  
 # ‘It (really) looks like Sira left... in fact, it looks like Sira left.’

This pattern is expected if *e-* is a weaker scalar item than *ga-*.

### 3.6 Preliminary prediction: Challengeability

Like RMD, we predict that the Logooli morphemes form part of the propositional content of an utterance. We predict therefore that the contribution of these morphemes can be overtly challenged in the discourse. In doing so, a discourse participant challenges the basis by which the other speaker has restricted the set of best possible worlds in which *P* is true; that is, the other speaker’s choice function.

We find that this prediction is borne out by our data, suggesting that *e-* and *ga-* are included within the propositional content of an utterance. However, we additionally find that only *e-* assertions can be challenged in this way. A Logooli speaker can use *ga-* to challenge an *e-* assertion, but cannot use *e-* to challenge a *ga-* assertion. We show this contrast in (23-24) below:

- (23) a. Context: Sira, Kageha, and Maina are eating dinner. Sira got up and left his plate on the table. Some time has passed, and he hasn’t yet returned. Noticing Sira’s abandoned plate of food on the table, Kageha says:  
 e-ror-ek-a ndee Sira a-zi-e  
 9-look-AC-FV that 1Sira 1-go-FV  
 ‘It (kind of) looks like Sira left.’
- b. Maina also notices that Sira’s coat, hat, and car are gone. He responds:  
 yago ne agirigare daave. ga-ror-ek-a ndee Sira a-zi-e  
 10that COP 10truth NEG 6-look-AC-FV that 1Sira 1-go-FV  
 ‘That’s not true. It (really) looks like Sira left.’

Here, Maina contradicts the strength of Kageha's assertion containing *e-*. In the context in (23), this is understood as challenging the strength of Kageha's evidence for *P*. Maina finds that there is enough evidence to use *ga-* to make the strong claim that Sira has left, as opposed to the weaker claim made with *e-* in Kageha's initial assertion.

Conversely, an utterance containing *ga-* cannot be challenged with an utterance containing *e-*:

(24) Context: Sira's coat is not on the coat rack. Kageha thinks this means that Sira has left, because Sira always wears his coat when he leaves the house. Kageha says:

- a. ga-ror-ek-a ndee Sira a-zi-e  
6-look-AC-FV that 1Sira 1-go-FV  
'It (really) looks like Sira left.'

- b. However, Maina knows that sometimes Sira likes to wear his coat inside the house. Maina responds:

# yago ne agirigare daave. e-ror-ek-a ndee Sira a-zi-e  
 10that COP 10truth NEG 9-look-AC-FV that 1Sira 1-go-FV

‘That’s not true. It (kind of) looks like Sira left.’<sup>15</sup>

We propose that this pattern falls out from our assumption of the scalar alternatives in 2 and the scalar implicature which pragmatically strengthens assertions including *e-*. We assume that the Logooli utterance *yago ne agirigare daave* ‘that’s not true’ targets this pragmatic strengthening. By uttering ‘that’s not true’ in (23b), Maina challenges the pragmatic strengthening of the assertion that Kageha makes with *e-* in (23a). This can then be followed by an assertion containing *ga-*, which is logically compatible with (unstrengthened) *e-*.

Conversely, utterances containing *ga-* cannot be pragmatically strengthened; *ga-* has no stronger scalar alternative. As a result, the use of *yago ne agirigare daave* ‘that’s not true’ is infelicitous in (24), since it cannot target any operation of pragmatic strengthening.<sup>16</sup>

### 3.7 Preliminary prediction: Raising

When the use of an expletive morpheme is precluded, the meaning is equivalent to that of *ga-*. That is, in the absence of a choice function, the default interpretation is simply that of universal quantification over the best possible worlds in the modal

15 Note also that it’s not necessary for either Kageha or Maina to have seen Sira in this context. Rather, the choice of *e-* or *ga-* is determined purely by what is known about Sira’s general behavior, coupled with the other visual evidence that the speakers have for *P*.

16 The utterance *yago ne agirigare daave* ‘that’s not true’ can also challenge non-scalar assertions:

- (i) a. i-mbwa ne i-siru  
 9-dog COP 9-brown  
 ‘The dog is brown.’  
 b. yago ne agirigare daave. i-mbwa ne i-mwamu  
 10that COP 10truth NEG 9-dog COP 9-black  
 ‘That’s not true. The dog is black.’

At present, it is an open question as to why this utterance is infelicitous in combination with unstrengthened scalar items. However, we note that this is similar to English data. (2b) below is infelicitous without the inclusion of the scalar particle *only*, or without focus intonation on the scalar item *some*:

- (ii) a. Leroy ate all of the cookies.  
 b. # That’s not true, he ate some of the cookies.



base. In Logooli, this arises in hyper-raising/copy-raising constructions.<sup>17</sup> In these constructions, an overt referential subject appears as the subject of the main verb, rather than either of the expletive agreement markers, as in (25a) below:

- (25) Context: You hear Imali coughing and sneezing.
- a. Imali a-hol-ek-a kuresia ne mu-rwaye  
 Imali 1-hear-AC-FV like COP 1-sick  
 ‘Imali sounds like she’s sick.’
- b. ga-hol-ek-a kuresia Imali ne mu-rwaye  
 6-hear-AC-FV like 1Imali COP 1-sick  
 ‘It sounds like Imali is sick.’
- c. # e-hol-ek-a kuresia Imali ne mu-rwaye  
 9-hear-AC-FV like 1Imali COP 1-sick  
 ‘It sounds like Imali is sick.’

Default universal quantification has been proposed for other constructions; for instance, conditionals are typically proposed to have default universal quantification over the antecedent (Kratzer 1991). It is also reflected in cross-linguistic data; RMD report that St’át’imcets speakers prefer a default universal strength interpretation of St’át’imcets modals. Furthermore, Gillon (2006) reports that Skwxwú7mesh speakers prefer a default identity function interpretation of the choice function introduced by Skwxwú7mesh determiners. Although we currently do not have an explanation for why the default interpretation of (25a) should be universal quantification, we note that this is in line with several other proposals.

Further support for this proposal comes from dialectal variation in Logooli. Mountjoy-Venning & Diercks (2016) report that for some Logooli speakers, hyper-raising and the *ga-* agreement marker can co-occur. They refer to these constructions as instances of “non-agreeing raising,” as in (26) below. In this construction, the matrix verb lacks the expected class 10 subject agreement *zi-*; instead, *ga-* occurs on the matrix verb:<sup>18</sup>

17 Hyper-raising is A-movement of the subject out of a finite embedded clause. This differs from copy-raising, which involves a pronominal subject in the embedded clause that is co-referential with the subject of the main verb (e.g., *Howard seemed like he enjoyed the concert*) (Diercks 2012; Halpert 2012; Carstens & Diercks 2012). Since Logooli, like other Bantu languages, is pro-drop, differentiating between the two constructions can be difficult.

18 Only *ga-* is licensed in this construction; *e-* cannot occur.

- (26)    *zi-ŋombe ga-ror-ek-a*    (nde) *zi-r-ii*  
           10-cow   6-look-AC-FV (that) 10-eat-PAST  
           ‘The cows seem to have eaten’  
           (Mountjoy-Venning & Diercks 2016, ex. (4); their translation)

Since the subject of (25a) is in fact the noun *ziŋombe* ‘cows’ (as shown by Mountjoy-Venning & Diercks 2016), we must account for how the  $\langle st, st \rangle$  argument of the modal verb is saturated. We assume that in such non-agreeing raising constructions, the modal verb combines with an additional phonologically null argument of type  $\langle st, st \rangle$ . This morpheme contributes the same semantics as *ga-*, i.e., universal quantification over the best possible worlds.<sup>19</sup>

#### 4 Conclusion and further puzzles

In this paper, we have argued that the Logooli morphemes *e-* and *ga-* introduce choice functions that combine with a universally quantified modal base. These choice functions select a subset of the best possible worlds of this modal base, and differ in the size of the subset they select. Logooli speakers use *e-* and *ga-* to make weak and strong assertions, respectively, about the embedded proposition *P*; their interpretation is dependent on contextual factors such as the information state of the speaker and the group of individuals that the speaker “associates” with. Our Logooli data can be accounted for with a similar analysis to RMD’s account for modality and evidentiality in St’át’imcets. This provides additional evidence for the proposal that modality and evidentiality are not distinct categories (argued by e.g., RMD, Palmer 1986, among others, contra e.g., de Haan 1999).

There is a great deal of puzzling data on *e-* and *ga-* that remains to be investigated. For instance, some Logooli verbs that introduce modal bases can occur only with one of the two agreement morphemes; the perception verbs *kufunya* ‘to smell’/‘to taste’ and *kuholeka* ‘to feel’ are only compatible with *e-*.<sup>20</sup> Conversely, the emotive factive *kugaasa* ‘to be perfect’ is only compatible with *ga-*. At present, we have no explanation as to why these restrictions occur.

Furthermore, preliminary fieldwork in Kenya suggests that there is a great deal of speaker variation with respect to the use and interpretation of these morphemes in Logooli. At present, it appears that this variation may be areal, although this hypothesis remains to be confirmed. In addition to establishing the facts regarding

<sup>19</sup> Alternatively, we could stipulate that Logooli modal verbs have two lexical entries: one that occurs with *e-* and *ga-*, and one that occurs with overt referential subjects in non-agreeing raising constructions and does not involve quantification over possible worlds. We have no strong objections to this latter analysis; both approaches are equally stipulative.

<sup>20</sup> The verb *kuholeka* ‘to feel’ is homophonous with the verb for ‘to hear.’

*e-* and *ga-* in Logooli, we are also interested in investigating similar expletive morphemes in other Luhia languages, and in Bantu more generally. We are particularly interested in establishing the modal facts for languages that have a greater number of these morphemes, such as Lubukusu and Wanga.

## References

- Aikhenvald, Alexandra. 2004. *Evidentiality*. Oxford: Oxford University Press.
- Carstens, Vicki & Michael Diercks. 2012. Parameterizing Case and Activity: Hyper-raising in Bantu. In *Proceedings of the 40th Annual Meeting of the Northeast Linguistic Society*, Amherst, MA: GLSA.
- Diercks, Michael. 2012. Parameterizing Case: Evidence from Bantu. *Syntax* 15(3). 253–286. doi:10.1111/j.1467-9612.2011.00165.x.
- Faller, Martina. 2002. *Semantics and Pragmatics of Evidentials in Cuzco Quechua*: Stanford University PhD dissertation.
- Gillon, Carrie. 2006. *The Semantics of Determiners: Domain restriction in Skwxwú7mesh*. Vancouver, Canada: University of British Columbia PhD dissertation. doi:10.14288/1.0092938.
- Grice, Paul. 1975. Logic and Conversation. In Peter Cole & Jerry L. Morgan (eds.), *Syntax and Semantics 3: Speech arts*, Elsevier.
- Grimm, Scott. 2006. An empirical view on raising to subject. *ZAS Papers in Linguistics* 25. 83–110.
- de Haan, Ferdinand. 1999. Evidentiality and Epistemic Modality: Setting boundaries. *Southwest Journal of Linguistics* 18(1). 83–101.
- Hacquard, Valentine. 2011. Modality. In Klaus von Stechow, Claudia Maienborn & Paul Portner (eds.), *Semantics: An International handbook of Natural Language Meaning* 2, 1484–1515. Mouton de Gruyter.
- Halpert, Claire. 2012. *Argument licensing and agreement in Zulu*. Cambridge, MA: Massachusetts Institute of Technology PhD dissertation.
- Kratzer, Angelika. 1991. Modality. In Armin von Stechow & Dieter Wunderlich (eds.), *Semantics: An International handbook of Contemporary Research*, 639–650. Mouton de Gruyter.
- Leung, Elizabeth Woon-Yee. 1991. *The Tonal Phonology of Llogoori: A study of Llogoori verbs*. Ithaca, NY: Cornell University PhD dissertation.
- Lewis, Paul, Gary Simons & Charles Fennig. 2016. *Ethnologue: Languages of the World*, nineteenth edition. Online version: <http://www.ethnologue.com>.
- Matthewson, Lisa, Hotze Rullmann & Henry Davis. 2007. Evidentials as epistemic modals: Evidence from St’at’imcets. *The Linguistic Variation Yearbook* 7. doi:10.1075/livy.7.07mat.

- Moltmann, Friederike. 2006. Generic *one*, arbitrary PRO and the first person. *Natural Language Semantics* 14(3). 257–281. doi:10.1007/s11050-006-9002-7.
- Moltmann, Friederike. 2012. Two kinds of first-person-oriented content. *Synthese* 184(2). 157–177. doi:10.1007/s11229-010-9730-6.
- Mountjoy-Venning, Cliff & Michael Diercks. 2016. Raising in Logoori. handout from Annual Conference on African Linguistics (ACAL) 47.
- Murray, Sarah. 2010. *Evidentiality and the Structure of Speech Acts*: Rutgers, the State University of New Jersey PhD dissertation.
- Palmer, Frank. 1986. *Mood and Modality*. Cambridge: Cambridge. doi:10.1017/CBO9781139167178.
- Reinhart, Tanya. 1997. Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy* 20(4). 335–397. doi:10.1023/A:1005349801431.
- Rullmann, Hotze, Lisa Matthewson, & Henry Davis. 2008. Modals as distributive indefinites. *Natural Language Semantics* 16(4). 317–357. doi:10.1007/s11050-008-9036-0.
- Samuels, Alex & Mary Paster. 2015. Verbal tone in Logoori. handout from Workshop on Luyia Bantu languages at Annual Conference on African Linguistics (ACAL) 46, University of Oregon.
- Sauerland, Uli & Mathias Schenner. 2007. Embedded evidentials in Bulgarian. In E. Puig-Waldmüller (ed.), *Proceedings of Sinn und Bedeutung 11*, Barcelona: Universitat Pompeu Fabra.
- Seidl, Amanda & Alexis Dimitriadis. 2003. Statives and reciprocal morphology in Swahili. *Typologie des langues d’Afrique et universaux de la grammaire 1* doi:10.1.1.529.202.

John Gluckman  
UCLA Department of Linguistics  
3125 Campbell Hall  
Los Angeles, CA 90095  
johnglu@ucla.edu

Margit Bowler  
UCLA Department of Linguistics  
3125 Campbell Hall  
Los Angeles, CA 90095  
margitbowler@gmail.com