Two Types of 'say'-Complementation in Kipsigis

Madeline Bossi

1. Introduction

In many languages, what appears to be a verb meaning 'say' (1a) can also be used alongside another verb in an attitude report (1b). This pattern is shown here for Kipsigis *le* 'say' (Kalenjin; Kenya).[1]

(1) a. Koo-∅-**le** Cheeroono kɔɔ-∅-al Kibeet tɛɛta.
PST2-3-LE C. PST2-3-buy K. cow
'Cherono said that Kibet bought a cow.'

b. Aɑ-ngen *({ko-**le** / ɑɑ-**le**}) kɔɔ-∅-al Kibeet tɛɛta.
1SG-know 3.SBJV-LE 1SG.SBJV-LE PST2-3-buy K. cow
'I know that Kibet bought a cow.'

A common analysis of this type of 'say'-complementation is that it reflects a diachronic link between 'say' verbs and complementizers (e.g. Heine & Kuteva 2002); that is, *le* in (1b) is only historically connected to the speech verb in (1a). Yet recent work questions this analysis, both across languages and within Kipsigis. For instance, Major (2021, 2023) analyzes 'say'-complementation in Uyghur (Turkic) and Avatime (Kwa) as involving the verb 'say' in a converbial clause. In a similar vein, for Kipsigis, Driemel & Kouneli (2022) argue that the morpheme *le* is always a speech verb—with verbal syntax and semantics across all of its uses.

However, in this paper, I show that systematic differences across constructions with *le* in Kipsigis make the uniformly verbal analysis in Driemel & Kouneli (2022) untenable. In particular, *le* only shows verbal syntactic behaviors when it transparently means 'say' and lacks these verbal behaviors in the absence of speech semantics. As a result, I propose that *le* represents two distinct forms. In some cases, it's the verb 'say', but in others, it's a complementizer devoid of speech semantics and only linked to the verb diachronically. This conclusion suggests a typology in which 'say'-based complementizers are verbs in some languages, complementizers in others, or even a combination thereof within a single language (see also Yue 2023).

The remainder of this paper is structured as follows. §2 provides key background on Kipsigis. In §3, I show that Kipsigis *le* is sometimes a speech verb and offer an analysis of these sorts of constructions. However, in §4, I show that *le* is not *always* a speech verb and provide an analysis of this other instance of *le*. With these facts and analyses established, §5 considers the typological implications of this work.

2. Background on Kipsigis

Kipsigis is a Nilo-Saharan language of the Kalenjin subgroup spoken in Western Kenya by 1.9 million people (Eberhard et al. 2021). All data come from my fieldwork on the language unless otherwise cited. Kipsigis is a verb-initial language with postverbal word order flexibility determined by information

* University of California, Berkeley. Kongoi mising' to Linus Kipkoech, Kiplangat Yegon, Chepkemoi Ronoh, and Lydia Bett for sharing their language with me. Thank you also to Amy Rose Deal, Emily Drummond, Peter Jenks, Maria Kouneli, and Line Mikkelsen. Financial support came from a UC Berkeley Oswalt Grant and an NSF DLI-DDRI grant. Field materials are archived with the California Language Archive: http://dx.doi.org/doi:10.7297/X2D79918.

[1] Abbreviations: 1 = 1st person, 2 = 2nd person, 3 = 3rd person, APPL = applicative, IPFV = imperfective, LP = local person, NEG = negation, PL = plural PST1 = recent past tense, PST2 = default past tense, PST3 = distant past tense, SBJV = subjunctive, SG = singular, REFL = reflexive, REL = relativizer.

 Proceedings of the 41st West Coast Conference on Formal Linguistics, ed. Nikolas Webster et al., 87-96. Somerville, MA: Cascadilla Proceedings Project.

structure (Bossi & Diercks 2019). Kipsigis verbs inflect with information about the tense, aspect, polarity, mood, subject, object, and valence of the verb, as illustrated in (2).

(2) Kii-mɑ-ɑɑ-koo-n-iin kɪtabʊʊt.
PST3-NEG-1SG-give-APPL-2SG.O book
'I didn't give you a book (long ago).'

Subject marking on verbs depends on the mood and phonological shape of the verb (Kouneli 2022). The indicative subject markers are listed in Table 1, where the difference between long vs. short vowels and ∅ vs. *ɪ-* in the third person is determined by the verb.

	SG	PL
1	a(a)-	kɪ(ɪ)-
2	ɪ(ɪ)-	ɔ(ɔ)-
3	∅/ɪ-	

Table 1: Indicative subject markers in Kipsigis

On the other hand, the subjunctive subject markers are listed in Table 2, where the difference between *kɔ-* vs. *kɔɔ-* in the third person is likewise determined by the verb.

	SG	PL
1	aa-	kɛɛ-
2	ɪɪ-	ɔɔ-
3	kɔ(ɔ)-	

Table 2: Subjunctive subject markers in Kipsigis

The indicative subject markers in Table 1 appear in all non-embedded contexts and in some finite embedded contexts. The subjunctive subject markers in Table 2 appear in a range of embedded contexts, including conditional antecedents, temporal clauses, and some complement clauses.

Finally, Kipsigis object markers are listed in Table 3. Unlike subject marking, these forms do not differ across verbs.

	SG	PL
1	-an	-ɛɛch
2	-ɪn	-aak
3	∅	

Table 3: Object markers in Kipsigis

3. *Le* is sometimes a speech verb

With this background established, here, I show that Kipsigis *le* behaves syntactically and semantically like a speech verb in some contexts—namely, in matrix clauses and in converbial clauses. These patterns motivate the analysis in Driemel & Kouneli (2022), in which *le* is a speech verb across all of its uses.

3.1. Le in matrix clauses

Kipsigis *le* can be used as an matrix attitude verb meaning 'say' (3). In these contexts, there is no overt complementizer and the embedded verb surfaces in the indicative mood, as evidenced by the indicative subject marker and the availability of tense marking on *sich* 'give birth' in (3), which is impossible in the subjunctive.

(3) Kɑ-∅-**le** Cheeroono koo-∅-sich Cheebeet laakwɛɛt.
PST1-3-LE C. PST2-3-give.birth C. child
'Cherono said (recently) that Chebet had a baby.'

Furthermore, in matrix contexts, *le* can bear the full range of verbal morphology including tense, aspect, subject marking, applicatives, and object marking, as seen in (4).

(4) **Kɑ-ɑ-leen-j-iin** ∅-nyoon-e Cheebeet.
PST1-1SG-LE-APPL-1SG.O 3-come.SG-IPFV C.
'I told you (recently) that Chebet is coming.'

In light of these behaviors, it is reasonable to treat *le* as a speech verb in matrix clauses.

3.2. Le in converbial clauses

Kipsigis also has converbial clauses, which modify the action described in the matrix clause. Verbs in converbial clauses bear the subjunctive subject markers found in Table 2 and can host a range of other types of verbal morphology including aspect, applicatives, and object marking (5).

(5) a. Ka-∅-ɪt Cheeroono ndoonyo [**ko**-ng'ɑlɑɑl-**e** ɛɛn sɑutiit ne mii parak].
PST1-arrive C. market 3.SBJV-speak-IPFV in voice REL.SG COP loud
'Cherono arrived at the market (recently), speaking in a loud voice.'

b. Koo-∅-keer-ɑn Cheebeet [**ɑɑ**-koo-**n-iin** kɪtabʊʊt].
PST2-3-see-1SG.O C. 1SG.SBJV-give-APPL-2SG.O book
'Chebet saw me giving you a book.'

Just like any other verb, *le* can also appear in these converbial clauses (6). As with the converbial clause in (5b), *le* can bear a range of verbal morphology in these constructions, including aspect, subjunctive subject marking, applicatives, reflexives, and object marking.

(6) a. Kɑ-ɑ-mwɑ-e [**ɑɑ-leelen** ka-∅-chɔɔr Kibeet rabɪɪnɪk].
PST1-1SG-tell-IPFV 1SG.SBJV-LE.IPFV PST1-3-steal K. money
'I was saying (recently) that Kibet stole the money.'
(Driemel & Kouneli 2022:ex. 16)

b. Koo-∅-chɑɑm-chi-kɛɛ Kibeet [**ko-leen-chi-kɛɛ** ng'ɑɑm].
PST2-3-whisper-APPL-REFL K. 3.SBJV-LE-APPL-REFL clever
'Kibet whispered to himself that he's clever.'
(Driemel & Kouneli 2022:ex. 20)

c. Kii-i-mwɑ-w-ɑɑn [**ii-leen-j-ɑɑn** kɪɪ-∅-al Kibeet tɛɛta].
PST3-2SG-tell-APPL-1SG.O 2SG.SBJV-LE-APPL-1SG PST3-3-buy K. cow
'You told me (long ago) that Kibet bought a cow.'

The fact that *le* patterns with other Kipsigis verbs in appearing in converbial clauses and bearing a range of verbal morphology here also speaks to *le*'s status as a speech verb in these constructions.

Further evidence that *le* is a speech verb here comes from two sources. First, in a converbial clause, *le* is compatible with verbal modifiers like adverbs (7a) and PPs (7b).

(7) a. Koo-i-mwɑ-w-ɑɑn [ii-**leen**-j-ɑɑn **ɑmut** koo-∅-sich Chɛɛpkɔɛch
PST-2SG-tell-APPL-1SG.O 2SG.SBJV-LE-APPL-1SG.O yesterday PST2-3-birth C.
laakwɛɛt].
child
'You told me yesterday that Chepkoech had a baby.'

b. Koo-i-mwɑ-w-ɑɑn [ii-**leen**-j-ɑɑn **ɛɛn arageɛneɛɛt** kɔɔ-∅-al Kibeet tɛɛta].
PST2-2SG-tell-APPL-1SG.O 2SG.SBJV-LE-APPL-1SG.O in surprise PST2-3-buy K. cow
'You told me with surprise that Kibet bought a cow.'

This pattern is expected if *le* is a verb because verbs should, in principle, allow this type of modification. Second, as noted in Diercks & Rao (2019), when it co-occurs with a matrix speech verb, *le* in a converbial clause generates a VERUM focus interpretation of the matrix predicate. Sentences like (8) are most natural when "someone is challenging [the speaker]" or questioning them (Diercks & Rao 2019:386).[2]

(8) You and I were talking about the cows yesterday, and I told you that the cows slept. Today, I talk with you again, and you say "I didn't know that the cows slept yesterday. You never told me!" I counter your comment with:
Koo-ɑ-mwɑ-w-iin [ɑɑ-**leen**-j-iin koo-Ø-ruu-jɑ tuugɑ].
PST2-1SG-tell-APPL-2SG.O 1SG.SBJV-LE-APPL-2SG.O PST-3-sleep-PL cows
'I DID tell you that the cows slept.'
(Diercks & Rao 2019:ex. 39b)

Following Kerr & van der Wal (2022)—who link predicate doubling to VERUM focus across Bantu—I suggest that VERUM focus in sentences like (8) indicates the doubling of speech verbs: namely, *mwa* 'tell' and *le* 'say'. In this way, the VERUM focus effect seen in (8) lends support to the claim that *le* is sometimes a speech verb. Together, these syntactic and semantic patterns suggest that *le* is a speech verb in the contexts described here.

3.3. Analysis of verbal le

To capture the behaviors seen in §3.1-3.2, I follow Driemel & Kouneli (2022) in treating *le* as a speech verb with the denotation in (9) or an equivalent.

(9) $[\![\text{le}]\!]^{w,g} = \lambda e_v.say(e)$, or equivalent

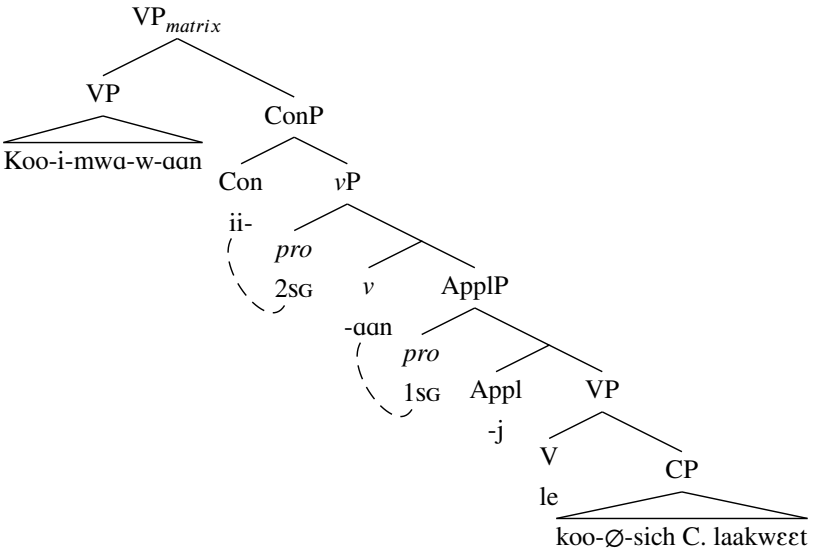
However, I diverge from them in analyzing sentences like (6) - (8) as involving converbial clauses, rather than complementation structures. This approach follows Major (2021, 2023), who argues for a converbial analysis of 'say'-complementation in Uyghur (Turkic) and Avatime (Kwa)

On my account, a sentence like (10a) has the converbial structure in (10b). Verbal inflection arises via standard agreement processes in Kipsigis; in particular, subject marking results from agreement with Con, while object marking results from agreement with *v*, which is the canonical object marking process in the language. Subjunctive subject markers surface—rather than indicative ones—because subjunctive is the default mood in contexts without finite T.

(10) a. Koo-i-mwɑ-w-ɑɑn [ii-**leen**-j-ɑɑn koo-Ø-sich Chɛɛpkɔɛch laakwɛɛt].
PST-2SG-tell-APPL-1SG.O 2SG.SBJV-LE-APPL-1SG.O PST2-3-birth C. child
'You told me that Chepkoech had a baby.'

[2] I have updated the IPA transcription and segmentation of the examples from Diercks & Rao (2019) based on my replication of their data and my understanding of the language. I cite Diercks & Rao (2019), since they are the original data source, but this sentence does not look identical to the one in their paper.

b.



Further evidence for this converbial analysis comes from extraction patterns seen with this type of *le* clause. Because ConP is adjoined in a structure like (10b), it is predicted to be an island for syntactic movement. In line with this prediction, arguments cannot extract from within a converbial clause with *le*, shown in (11) with topic Ā-movement (Bossi 2023:Chap. 2 §2.5.5.2). Here, the topicalized constituent *Chɛɛpkɔɛch* 'Chepkoech' originates within ConP and, as a result, cannot move to clause-initial position with *ko* 'TOP'.

(11) * **Chɛɛpkɔɛch** ko koo-i-mwɑ-w-ɑɑn [ii-**leen**-j-ɑɑn koo-Ø-sich ____
C. TOP PST2-2SG-tell-APPL-1SG.O 2SG.SBJV-LE-APPL-1SG.O PST2-3-birth
laakwɛɛt].
child
Intended: 'You told me that Chepkoech had a baby.'

This same restriction is found in converbial clauses with other verbs, like *koo* 'give' in (12). Here, as in (11), the topicalized constituent *laakwɛɛt* 'child' cannot extract from within ConP to clause-initial position with *ko* 'TOP'.

(12) * **Laakwɛɛt** ko koo-keer-ɑn Cheebeet [ɑɑ-koo-chiin-i ____ kɪtabʊʊt].
child TOP PST2-3-see-1SG.O C. 1SG.SBJV-give-APPL.IPFV-LP book
Intended: 'Chebet saw me giving the child a book.'

The parallelism between (11) and (12) shows that movement out of adjoined converbial clauses is generally disallowed in Kipsigis and that clauses headed by *le* like the one in (11) behave as expected for a converbial clause. In this way, the data presented in this section suggest that *le* is a speech verb, at least in some contexts; however, in the follow section, I turn to another set of contexts where *le* appears obligatorily, but does not display the behaviors expected of a speech verb.

4. *Le* isn't always a speech verb

In addition to the contexts in §3, *le* also appears obligatorily in complementation structures with a range of matrix verbs. For instance, the attitude verbs *pwaat* 'think', *ngen* 'know', *ruaatit* 'dream', and many others require use of *le* (13).

(13) a. I-pwɑɑt-e Cheeroono *(ko-**le**) kɔɔ-Ø-al Kibeet tɛɛta.
3-think-IPFV C. 3.SBJV-LE PST2-3-buy K. cow
'Cherono thinks that Kibet bought a cow.'

b. I-ngen Cheebet *(ko-**le**) ∅-ru-e Cheeruyot.
3-know C. 3.SBJV-LE 3-sleep-IPFV C.
'Chebet knows that Cheruyot is sleeping.'

c. Koo-ɑɑ-ruɑɑtit *({ko-**le** / ɑɑ-**le**}) koo-tyen Kibeet.
PST2-1SG-dream 3.SBJV-LE 1SG.SBJV-LE PST2-sing/dance K.
'I dreamed that Kibet sang/danced.'

In this section, I describe the syntax and semantics of *le* in the types of structures seen in (13). I conclude that, although *le* is a speech verb in the contexts in §3, there are cases like those in (13) where *le* is demonstrably *not* a speech verb. In particular, I show that, in these contexts, *le*'s verbal syntax is restricted, and it doesn't contribute speech semantics. This empirical picture leads me to analyze *le* in sentences like (13) as a semantically bleached complementizer devoid of verbal syntax and semantics.

4.1. Restricted verbal syntax

In this section, I outline two restrictions on *le*'s syntactic behavior that pose challenges for the uniformly verbal analysis of *le* in Driemel & Kouneli (2022). First, the extensive verbal inflection seen on *le* in §3 is impossible when it co-occurs with matrix verbs that don't involve speech. Even if the matrix verb bears imperfective aspect, this same inflection is impossible on *le* when the matrix verb isn't a speech predicate, as with *keer* 'see' in (14). Here, *le* can only bear prefixal marking that is default 3rd person (*ko-* '3.SBJV') or that tracks the matrix subject (*aa-* '1SG.SBJV').

(14) I notice that you're dancing, but I don't say anything.
A-keer-e {ɑɑ-**le** / ko-**le** / *ɑɑ-**leelen** / *ko-**leelen**} i-tyen-i.
1SG-see-IPFV 1SG.SBJV-LE 3.SBJV-LE 1SG.SBJV-LE.IPFV 3.SBJV-LE.IPFV 2SG-dance-IPFV
'I see that you're dancing.'

In a similar vein, even if the matrix verb bears object marking, this same inflection is impossible on *le* when the matrix verb isn't a speech predicate, as with *par* 'show' in (15). Here, as in (14), *le* can only bear prefixal marking that is default 3rd person (*ko-* '3.SBJV') or that tracks the matrix subject (*aa-* '1SG.SBJV).

(15) We're hiding under the bed in a game of hide-and-seek. We can't speak—otherwise, the seeker will find us—but I point to a pair of shoes under the bed.
Kɑ-ɑ-pɑr-u-iin {ɑɑ-**le** / ko-**le** / *ɑɑ-**leen**-j-iin /
PST1-1SG-show-APPL-2SG.O 1SG.SBJV-LE 3.SBJV-LE 1SG.SBJV-LE-APPL-2SG.O
*ko-**leen**-j-iin} koo-mii kwoosiek kɪtanda arɪt.
1SG.SBJV-LE-APPL-2SG.O PST2-COP shoes bed under
'I showed you that there were shoes under the bed.'

These restrictions on *le*'s aspect and object marking are unexpected if *le* is uniformly a verb, given that verbs should be equally able to inflect in these ways, especially if the matrix verb bears this same type of inflection. However, if *le* is only a verb in some of its uses, this type of asymmetry is expected.

Second, in contrast to the pattern in §3, *le* can't take verbal modifiers when it co-occurs with matrix verbs that don't involve speech. Adverbs (16a) and PPs (16b) can't modify *le* in sentences with matrix verbs like *ngen* 'know'.

(16) a. * Aɑ-ngen ɑɑ-**le** **kɔmyɛ** kɔɔ-∅-al Kibeet kɪtabʊʊt.
1SG-know 1SG.SBJV-LE well PST2-3-buy K. book
Intended: 'I know well that Kibet bought a book.'

b. * Koo-ɑ-nɑi ɑɑ-**le** **ɛɛn arageɛnɛɛt** kɔɔ-∅-chɔɔr Kibeet pɪasɪnɪɪk.
PST2-1SG-know 1SG.SBJV-LE in surprise PST2-3-steal K. potatoes
Intended: 'I knew with surprise that Kibet stole the potatoes.'

The ungrammaticality of the sentences in (16) is due to the placement of the modifiers and not their semantics, since *ngen* 'know' is otherwise compatible with these modifiers, as seen in (17).

(17) a. Aɑ-ngen **kɔmyɛ** ɑɑ-**le** kɔɔ-∅-al Kibeet kɪtabʊʊt.
1SG-know well 1SG.SBJV-LE PST2-3-buy K. book
'I know well that Kibet bought a book.'

b. Koo-ɑ-nɑi **ɛɛn aragɛɛnɛɛt** ɑɑ-**le** kɔɔ-∅-chɔɔr Kibeet pɪasɪnɪɪk.
PST2-1SG-know in surprise 1SG.SBJV-LE PST2-3-steal K. potatoes
'I knew with surprise that Kibet stole the potatoes.'

As previously, this asymmetry is unexpected if *le* is a verb across all of its uses, since it should uniformly allow modification with adverbs and PPs. On the other hand, if *le* is only a verb in some of its uses, this type of asymmetry is expected.

Taken together, these syntactic patterns indicate that *le* isn't always a verb; instead, they suggest that it's a verb in some cases (§3) but not in others. Furthermore, in cases where *le* isn't a verb—like the ones described in this section—inflection on *le* is restricted to only prefixal marking.

4.2. No speech semantics

Building on the syntactic patterns described in §4.1, here I show that *le* doesn't always contribute speech semantics either. First, *le* is obligatory in attitude reports that don't semantically involve speech. This pattern can be seen in (18) where speech is explicitly denied and in (19) where the attitude holder is an inanimate object incapable of speech.

(18) Koo-ɑɑ-ngen *({ko-**le** / ɑɑ-**le**}) koo-miit-een Kibeet, lɑkini mɑ-ɑ-mwɑ kiy.
PST2-1SG-know 3.SBJV-LE 1SG.SBJV-LE PST2-COP-in K. but NEG-1SG-say thing
'I knew that Kibet was there, but I didn't say anything.'

(19) Koo-∅-keer kamɛra *(ko-**le**) koo-∅-it Kibeet.
PST-3-see camera 3.SBJV-LE PST-3-arrive K.
'The camera saw that Kibet arrived.'

If *le* uniformly contributes speech semantics, the data in (18) - (19) are unexpected, since *le* should presumably introduce speech here too.

There are two possible counter-arguments to this generalization—neither of which hold up in Kipsigis. First, one might claim that *le* is a light verb that only contributes speech semantics in combination with other functional structure like *v* (Grimshaw 2015, Major 2021). As a result, *le* would only lack speech semantics in contexts without *v*. However, this view is at odds with the one in Driemel & Kouneli (2022), who claim that prefixal marking on *le* is due to agreement with *v*. On their view, even if *le* is a light speech verb, it should always contribute speech semantics, since it necessarily combines with *v* when it has prefixal marking. Second, one might claim that *le* contributes internal speech. However, this must not be the case, since inanimate objects like the camera in (19) are presumably incapable of internal speech and, more crucially, *le* cannot generally be used to convey internal speech, as evidenced by its infelicity in (20). Here, *le* is only appropriate if Chebet speaks about the weather aloud. To convey internal speech, speakers use *pwaat* 'think' alongside *le* (21).

(20) Describing a picture of Chebet with a thought bubble showing sunny weather outside.
Koo-∅-**le** Cheebet koo-∅-lɑɑl-e asɪsta ɑmut.
PST2-3-LE C. PST2-3-shine-IPFV sun yesterday
'Chebet said that the sun was shining yesterday.'

(21) Koo-i-pwɑɑt-e Cheebet ko-**le** koo-∅-lɑɑl-e asɪsta ɑmut.
PST2-3-think-IPFV C. 3.SBJV-LE PST2-3-shine-IPFV sun yesterday
'Chebet thought that the sun was shining yesterday.'

A second piece of evidence that *le* doesn't always contribute speech semantics comes from the fact that complementation structures with *le* can be factive. Even when negated, *ngen* 'know' plus *le* entails the truth of the embedded clause, rendering the continuation in (22b) unnatural.

(22) a. Koo-mɑ-ii-ngen ii-**le** koo-∅-tun Kibeet Cheeroono...
PST2-NEG-2SG-know 2SG.SBJV-LE PST-3-marry K. C.
'You didn't know that Kibet married Cherono...'

b. # Lɑkini mɑ-∅-tun Kibeet Cheeroono.
but NEG-3-marry K. C.
'but Kibet didn't marry Cherono.'

This pattern of factivity is surprising if *le* contributes speech semantics, since speech contexts aren't generally factive across languages. In fact, as noted by Major (2021, 2023), verbal 'say'-complementation is typically incompatible with factivity. On the other hand, if *le* doesn't contribute speech semantics in (22a), the observed pattern of factivity is entirely expected.

In this way, the semantic patterns described here suggest that *le* doesn't always contribute speech semantics; instead, they suggest that it does in some cases (§3) but not in others. This semantic conclusion goes hand-in-hand with the syntactic conclusion from §4.1 in suggesting that *le* isn't always a speech verb.

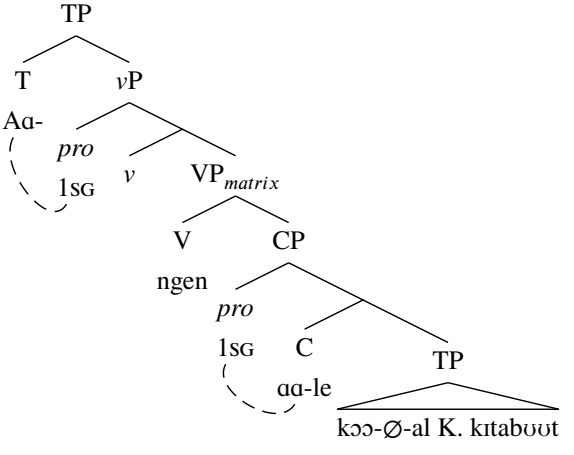
4.3. Analysis of complementizer le

To capture the behaviors seen in §4.1-4.2, I propose that, when *le* doesn't show the properties of a speech verb, it is a semantically-bleached complementizer with the denotation in (23) or an equivalent. This version of *le* is only linked to verbal *le* diachronically. In this way, my account differs from the uniformly verbal approach in Driemel & Kouneli (2022).

(23) $[\![\text{le}]\!]^{w,g} = \lambda e_v.\forall w' \in \text{CONT}(e) : p(w')$, or equivalent

On my view, complementizer *le* involves standard complementation syntax. A sentence like (24a) has the structure in (24b). Importantly, complementizer *le* is restricted to prefixal marking; the limited nature of true complementizer agreement in Kipsigis means that this pattern can be captured via existing accounts of upward-oriented complementizer agreement without actually requiring Agree to probe upwards (e.g. Diercks 2013, Carstens 2016). For instance, I take an "Indirect Agree" approach in (24b), in which *le* Agrees locally with a null *pro* in its specifier that is co-referential with a matrix argument (Diercks 2013). As previously, complementizer *le* bears subjunctive subject markers—rather than indicative ones—because subjunctive marking is the default in contexts without finite T.

(24) a. Aɑ-ngen ɑɑ-**le** kɔɔ-∅-al Kibeet kɪtabʊʊt.
1SG-know 1SG.SBJV-LE PST2-3-buy K. book
'I know that Kibet bought a book.'

b.


Further evidence for this complementation analysis comes from extraction patterns seen with this type of *le* clause, which differ from those seen in §3.3 with *le* in a converbial clause. Unlike converbial *le*, complementizer *le* heads an embedded clause; as a result, extraction should be possible out of the embedded clause. In line with this prediction, arguments can extract from within clauses with complementizer

le, seen in (25) with topic Ā-movement. Unlike with converbial *le*, here the topicalized constituent *Kibeet* 'Kibet' can move to clause initial position with *ko* 'TOP' from within the embedded clause.

(25) **Kibeet** ko ii-ngen ii-**le** ka-∅-chɔɔr ____ pıasıınık.
K. TOP 2SG-know 2SG.SBJV-LE PST1-3-steal potatoes
'You know that Kibet stole potatoes.'

The different extraction patterns seen with converbial *le* in (11) versus complementizer *le* in (25) pose real challenges for the uniformly verbal approach in Driemel & Kouneli (2022). If *le* is always a speech verb in an embedding structure—as proposed in Driemel & Kouneli (2022)—then extraction should always be permitted out of a clause headed by *le*. On the other hand, if *le* is always a speech verb in a converbial structure, then extraction should never be permitted out of a clause headed by *le*. Neither of these sets of predictions are upheld; instead, the data reveal an asymmetry in whether or not extraction is permitted, which correlates with other syntactic and semantic behaviors of *le*. The only way to capture this asymmetry—alongside the other syntactic and semantic behaviors discussed here—is through the type of split analysis that I propose.

5. Implications

In this paper, I argue that Kipsigis *le* has two distinct forms: one speech verb and one complementizer. These different forms of *le* have different semantics and appear in different syntactic structures. In particular, in some contexts, *le* is a speech verb, while in others, it's a semantically bleached complementizer that is only linked to verbal *le* diachronically.

This Kipsigis case study challenges the recent cross-linguistic hypothesis that 'say'-based complementizers might be uniformly analyzable as speech verbs. Much recent work highlights that 'say'-complementation can have verbal properties across a variety of languages (see e.g. Halpert 2019, Letsholo & Safir 2019, Major et al. 2022 on Bantu; Driemel & Kouneli 2022 on Nilo-Saharan; Major & Torrence 2021, Major 2021 on Kwa; Baker & Vinokurova 2010, Ozyildiz 2017, Predolac 2017, Major 2021, 2023 on Turkic; Bondarenko 2020 on Mongolic; Spadine 2020 on Simitic; a.o.). However, the Kipsigis facts suggest a typology in which 'say'-based complementizers are verbs in some languages, complementizers in others, or even a combination thereof within a single language, as seen in Kipsigis. In fact, this conclusion meshes nicely with recent work on a similar type of 'say'-complementation in Sakha (Turkic) by Yue (2023); according to Yue (2023), Sakha has both a 'say'-based complementizer and a converbial construction, which often look similar but can be disentangled via their case marking properties. While more work is needed to flesh out the typology of 'say'-complementation fully, this Kipsigis case study constitutes another step in that direction.

References

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

Bondarenko, Tatiana. 2020. Factivity from pre-existence: Evidence from Barguzin Buryat. *Glossa* 5(1). 1–35.

Bossi, Madeline. 2023. *Epistemic modality across syntactic categories in Kipsigis*. Berkeley, CA: UC Berkeley dissertation.

Bossi, Madeline & Michael Diercks. 2019. V1 in Kipsigis: Head movement and discourse-based scrambling. *Glossa: a journal of general linguistics* 4(1).

Carstens, Vicki. 2016. Delayed valuation: A reanalysis of goal features, "upward" complementizer agreement, and the mechanics of Case. *Syntax* 19(1). 1–42.

Diercks, Michael. 2013. Indirect agree in Lubukusu complementizer agreement. *Natural Language & Linguistic Theory* 31(2). 357–407.

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

Driemel, Imke & Maria Kouneli. 2022. C-agree is local subject-verb agreement in Kipsigis. Unpublished manuscript.

Eberhard, David M., Gary F. Simons & Charles D. Fennig. 2021. Kipsigis. In *Ethnologue: Languages of the world*, 24th. Dallas, TX: SIL International.

Grimshaw, Jane. 2015. The light verbs say and SAY. In *Structures in the mind: Essays on language, music, and cognition in honor of Ray Jackendo*, vol. 79. MIT Press.

Halpert, Claire. 2019. How to be an embedded clause: 'Say' complementizers in Bantu. In M. Ryan Bochnak, Miriam Butt, Erlinde Meertens & Mark-Matthias Zymla (eds.), *Proceedings of TripleA 5: Fieldwork perspectives on the semantics of African, Asian and Austronesian languages*. Universitätsbibliothek Tübingen, Publikationssystem.

Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Kerr, Elisabeth & Jenneke van der Wal. 2022. Indirect verum marking in 10 Bantu languages. Presentation at Bantu9.

Kouneli, Maria. 2022. Inflectional classes in Kipsigis. *Glossa* 47(1).

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

Major, Travis. 2021. *On the nature of "say" complementation*. UCLA dissertation.

Major, Travis. 2023. Re-analyzing "say" complementation: Implications for case theory and beyond. Unpublished manuscript.

Major, Travis, Justine Sikuku & Michael Diercks. 2022. "Say"-complementation: Re-analyzing agreeing complementizers as verbs. Unpublished manuscript.

Major, Travis & Harold Torrence. 2021. "Say"-chains, not "say"-complementation. In Rachel Soo, Una Y. Chow & Sander Nederveen (eds.), *Proceedings of the 38th West Coast Conference on Formal Linguistics*, 283–293. Somerville, MA: Cascadilla Proceedings Project.

Ozyildiz, Deniz. 2017. Attitude reports with and without true belief. In *Semantics and Linguistic Theory*, vol. 27, 397–417. Linguistic Society of America.

Predolac, Esra. 2017. *The syntax of sentential complementation in Turkish*. Ithaca, NY: Cornell University dissertation.

Spadine, Carolyn. 2020. *The structure of attitude reports: Representing context in grammar*. Cambridge, MA: MIT dissertation.

Yue, Christine Soh. 2023. Sakha "say" complementization: A Case-by-Agree approach. Presentation at NELS 53.

Proceedings of the 41st West Coast Conference on Formal Linguistics

edited by Nikolas Webster, Yağmur Kiper, Richard Wang, and Sichen Larry Lyu

Cascadilla Proceedings Project Somerville, MA 2025

Copyright information

ISBN 978-1-57473-483-6 hardback

Ordering information

Orders for the printed edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Bossi, Madeline. 2025. Two Types of 'say'-Complementation in Kipsigis. In *Proceedings of the 41st West Coast Conference on Formal Linguistics*, ed. Nikolas Webster et al., 87-96. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #3734.