Natural Language & Linguistic Theory
https://doi.org/10.1007/s11049-024-09622-8

C-Agree is local subject-verb agreement in Kipsigis

Imke Driemel[1] · **Maria Kouneli**[2]

Received: 9 April 2022 / Accepted: 27 February 2024

Abstract
Upwards-oriented complementizer agreement raises questions about the directionality and locality of agreement. Based on novel data from original fieldwork, we argue that what has been described as an agreeing 'say'-based complementizer in Kipsigis (Diercks and Rao 2019; Diercks et al. 2020) is the lexical verb 'say,' and what looks like C-Agree is in fact agreement between this verb and its locally introduced (often covert) subject. Our analysis highlights that 'say'-based complementizers might be of category V, not C, in more languages than previously thought (Koopman 1984; Major 2021), which means that some instances of what has been described as C-Agree may instantiate standard verbal agreement. Furthermore, we provide a semantic analysis of 'say'-based complementation in Kipsigis along the lines of contentful eventualities (Hacquard 2006; Kratzer 2013a).

Keywords Complementizer agreement · Complementation · 'Say'-based complementizers · Attitude semantics · Kipsigis · Nilotic

1 Introduction

A number of African languages have been reported to display upwards-oriented complementizer agreement, where the embedded C head agrees with the matrix subject: see for example Baker (2008) on Kinande, Idiatov (2010) on Mande languages, Diercks (2013) on Lubukusu, Duncan and Torrence (2017) on Ibibio, Nformi (2017) on Limbum, and Letsholo and Safir (2019) on Ikalanga.[1] This is different from

[1] Outside of Africa, a similar phenomenon has been reported for the Trans–New Guinean language Teiwa (Sauerland et al. 2020) and for Arapesh, spoken in Papua New Guinea (Baker 2008).

✉ M. Kouneli
maria.kouneli@rutgers.edu

I. Driemel
imke.driemel@york.ac.uk

1 Department of Language and Linguistic Science, University of York, YO10 5DD York, UK

2 Department of Linguistics, Rutgers University, 18 Seminary Pl, New Brunswick, NJ 08901, USA

Published online: 21 October 2024

the well-studied pattern of downwards-oriented complementizer agreement in Germanic, where in embedded clauses, a C head can show covariance with the ϕ-features of the embedded subject (Shlonsky 1994; Zwart 1997; Carstens 2003; van Koppen 2005, 2012; Fuß 2008, 2014; Haegeman and van Koppen 2012). While the Germanic pattern does not pose serious problems for standard approaches to agreement using Downward Agree (e.g. Chomsky 2000, 2001), upwards-oriented complementizer agreement raises a number of questions about the directionality and locality of Agree, with some studies arguing that Upward Agree (Bjorkman and Zeijlstra 2019; among others) is necessary for the analysis of the pattern (e.g. Nformi 2017; Letsholo and Safir 2019; McFadden and Sundaresan 2021).

Despite the theoretical significance of the phenomenon, however, both the properties of upwards-oriented C-Agree in individual languages and the extent of cross-linguistic variation are poorly understood, primarily because most known examples come from understudied languages. In this paper, we begin to fill this gap by carefully investigating the phenomenon in Kipsigis, a Nilotic language spoken in Kenya that has been reported to display an upwards-oriented agreement pattern between an embedded 'say'-based complementizer—glossed neutrally in this paper as LE—and the matrix subject (Diercks and Rao 2019; Diercks et al. 2020):[2, 3]

(1) a. ûː-ŋgén **ùː-lé** ∅-rú-è Kíbêːt.
1SG-know 1SG-LE 3-sleep-IPFV Kibeet.NOM
'I know that Kibeet is sleeping.'

b. îː-ŋgén **ìː-lè** ∅-rú-è Kíbêːt.
2SG-know 2SG-LE 3-sleep-IPFV Kibeet.NOM
'You know that Kibeet is sleeping.'

c. í-ŋgèn Kíplàŋgàt **kò-lé** ∅-rú-è Kíbêːt.
3-know Kiplangat.NOM 3-LE 3-sleep-IPFV Kibeet.NOM
'Kiplangat knows that Kibeet is sleeping.'

Kipsigis is also the only documented case where the C-like element can show additional (optional) cross-referencing with the matrix object, in the form of a suffix:

(2) Kù-∅-tʃúːm-ú-ún Tʃéːbêːt **kò-lèːn-tʃ(i)-ùn**
PST.CURR-3-whisper-APPL-1SG Cheebeet.NOM 3-LE-APPL-1SG
kà-∅-tʃɔ́ːr Kíplàŋgàt ràbíːník.
PST.CURR-steal Kiplangat.NOM money

[2] Kipsigis is the major variety of Kalenjin, a cluster of dialects of the Southern Nilotic branch of Nilo-Saharan, and it is spoken by about two million people in western Kenya (Eberhard et al. 2020). Unless indicated otherwise, data in this paper come from the authors' fieldwork. The authors had a series of Skype and Zoom elicitations in 2020–2024 with seven native speakers (male, age range: 22–32) living in Nairobi, while some data come from the second author's fieldwork conducted in Nairobi and Kilifi over four trips to Kenya between 2017 and 2022. The speakers who were consulted on questions about C-Agree all grew up in monolingual Kipsigis regions (two speakers in Narok County and five speakers in Bomet County). All of them are also proficient in English and Swahili, the official languages of Kenya.

[3] Glossing abbreviations follow the Leipzig Glossing Rules with the addition of C = complementizer, IMPRS = impersonal, IT = itive, PST.CURR = current past, PST.DIST = distant past, PST.REC = recent past, SBJVI = subjunctive Type I, SBJVII = subjunctive Type II, and VENT = ventive. Tone is transcribed whenever possible, but some transcriptions are incomplete because of sound difficulties in Skype elicitations.

'Cheebeet whispered to me that Kiplangat stole the money.'

Based on novel data from original fieldwork, we argue that what has been described as an (agreeing) 'say'-based complementizer in Kipsigis is in fact the lexical verb 'say,' not a complementizer (see also Koopman and Sportiche 1989; Özyıldız et al. 2018; Major 2021; Major and Torrence 2021; Major et al. 2022 for verbal analyses of such complementizers). Furthermore, we show that prefixal agreement is not always with the matrix subject (contra Diercks and Rao 2019); the pattern is best characterized as agreement with the source of information. We therefore present an analysis according to which what looks like C-Agree in (1) is an instance of agreement between the lexical verb 'say' and its locally introduced (often covert) subject. Downward Agree can straightforwardly account for instances of subject-verb agreement, and our analysis thus solves the locality and directionality problems posed by the (apparent) upwards-oriented nature of C-Agree. We also provide a semantic analysis in which the verbal category of the "complementizer" is reflected in its semantics, building on recent eventuality-based models of attitude and speech reports (e.g. Kratzer 2013b; Elliott 2016, 2017; Moulton 2019).

The remainder of the paper is structured as follows. In Sect. 2, we provide an overview of previous theories of upwards-oriented complementizer agreement; in Sect. 3, we present some background on complementation in Kipsigis and provide a description of the pattern of upwards-oriented complementizer agreement in the language. We then develop our analysis in three steps: in Sect. 4 we argue that the Kipsigis "complementizer" is the lexical verb 'say,' which agrees with a local subject; in Sect. 5 we motivate the syntactic structure that we assume for complementation; and in Sect. 6 we provide a semantic analysis, which will account for some of the distributional restrictions we find with *le* and clausal embedding predicates more generally. In Sect. 7, we conclude.

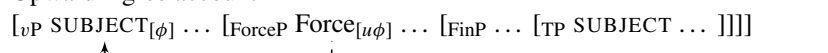
2 Previous theories of upwards-oriented C agreement

Since the theoretical analysis of upwards-oriented complementizer agreement in Lubukusu by Diercks (2013), there has been a growing body of literature on the implications of this pattern of C agreement for theories of Agree (e.g. Carstens 2016; Diercks et al. 2020; McFadden and Sundaresan 2021). There are two questions that are regularly discussed within the literature on upwards-oriented C-Agree: first, the direction of Agree and, second, the nature of the goal. We address each question in turn.

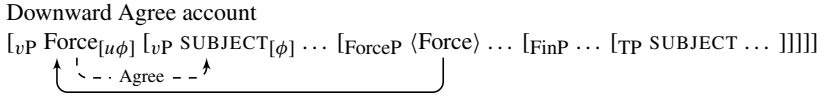
While a number of accounts implement upwards-oriented agreement directly via Upward Agree between the embedded C head and the matrix subject (Nformi 2017; Letsholo and Safir 2019; McFadden and Sundaresan 2021), other approaches maintain a Downward Agree analysis with an additional (covert) movement step of the embedded complementizer prior to Agree (Carstens 2016; Diercks and Rao 2019; Diercks et al. 2020). The two types of analyses are illustrated in (3) and (4), respectively. Reasons for the lack of downward probing of C into the embedded clause include the position of the complementizer with respect to a phase boundary (Carstens

2016; McFadden and Sundaresan 2021) and cross-linguistically determined parameter settings (Baker 2008).

(3) Upward Agree account



(4) Downward Agree account



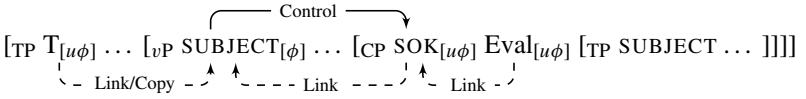
For Lubukusu, Carstens (2016) proposes an approach involving Downward Agree between the moved complementizer and the matrix subject: the Force head carries $u\phi$-features and moves into the matrix clause, where it adjoins to the vP, from which position it can undergo Agree with the ϕ-features of the matrix subject. Diercks and Rao (2019) and Diercks et al. (2020), in their analyses of Kipsigis and Lubukusu, adopt similar mechanics as Carstens (2016), but make the additional assumption that movement of the complementizer is triggered by anaphoricity requirements instead of a phase boundary. Thus, the complementizer moves to the matrix clause to check anaphoric ϕ-features. Diercks et al. (2020, 378) argue against Carstens's account based on the fact that Lubukusu allows raising to object, past the agreeing Force head, which is incompatible with the assumption that Force introduces a phase boundary. Crucially, the raised object can trigger object marking, which indicates that such raising constructions are A-movement. This excludes an analysis where the object only moves to the specifier of ForceP (Bruening 2002). Hence the raising facts point to the absence of a phase boundary, while Carstens's account of C-Agree relies on the presence of a phase boundary.

As for Upward Agree accounts, they generally cannot capture the observation that it is often the matrix subject that is solely targeted for C-Agree. Languages like Lubukusu and Ibibio always show agreement of the C head with the matrix subject even in the presence of a matrix object, though see Nformi (2017) for intervention effects that are triggered in such cases for at least certain matrix verbs in Limbum. As Diercks and Rao (2019) point out, Kipsigis constitutes a notable exception since the C-like element can show additional (optional) cross-referencing with the matrix object in the form of a suffix. However, Diercks and Rao still reject the Upward Agree analysis, based on the observation that the suffixal marker shows properties of a clitic instead of an agreement marker. Since clitic doubling is clause-bound, they argue that the C head must move into the matrix clause to act as a host for the clitic. Note, however, that this movement takes place covertly most of the time.

Another way to derive upwards-oriented agreement with C is to posit a silent element in the specifier of CP, which then acts as an intermediary between the ϕ-probe on C and the antecedent in the matrix clause. This indirect Agree analysis was originally put forth by Diercks (2013) for Lubukusu, where the complementizer first agrees with an anaphor in its specifier via Spec-Head Agree and this anaphor is subsequently bound by the matrix subject; see also Duncan and Torrence (2017) for a similar C-Agree analysis in Ibibio, as well as Gluckman (2022) on Nyala East. Given

that anaphors are often subject-oriented, this type of analysis provides a straightforward explanation for why the C head always cross-references the subject, at least in C-Agree languages like Lubukusu, Ibibio, Kinande, and Ikalanga. The suffixal agreement pattern in Kipsigis, however, is not captured by this analysis since, across languages, object-oriented anaphors are not found to our knowledge. Baker (2022) discusses the Kipsigis pattern as one of the motivations for analyzing the silent element in Spec,CP as a close relative of PRO rather than an anaphor and as being in turn in a control relation with the antecedent in the matrix clause. After all, we do find subject control predicates as well as object control predicates across languages. Baker takes agreement with matrix T as the decisive factor for the C-Agree patterns found across languages. He argues that what determines the antecedent for C-Agree is the ability to agree with matrix T, termed the T/Agree Condition. Evidence comes from constructions involving a thematic subject that nevertheless does not enter an Agree relation with matrix T, such as *by*-phrases of passives and causees in morphological causative constructions. In such environments, the thematic subject never triggers C-Agree in Lubukusu, Ibibio, Kinande, and Ikalanga. In Baker (2022), the dependency between embedded C and matrix T is derived by splitting the Agree mechanism into Agree-Link and Agree-Copy, where the former creates a pointer from probe to goal and only the latter copies ϕ-features (cf. Arregi and Nevins 2012; Marušič et al. 2015). Baker proposes that certain heads, such as T, undergo Agree-Link and Agree-Copy, whereas other heads, like C, only undergo Agree-Link. Once one head agrees via Agree-Copy, all heads in the chain created previously via Agree-Link copy the same ϕ-features. We see an illustration of C-Agree with the matrix subject in (5). The embedded C head, dubbed Eval, introduces a PRO-like DP (SOK = seat of knowledge) in its specifier. SOK is in an obligatory control relation (Landau 2013) with the nearest argument that best matches the theta role given to it by Eval; in most cases this is the matrix subject. First, Eval creates an Agree-Link to SOK. As a consequence of the control relation, SOK creates an Agree-Link with the controller in a second step. At this point, no ϕ-features have been copied. This only happens in the last step, when matrix T undergoes Agree-Link and Agree-Copy with the controller of SOK. Since ϕ-feature copying to all heads in the Agree chain including Eval can only be triggered by Agree-Copy between the controller and matrix T, we find C-Agree only with antecedents that show agreement with matrix T (T/Agree Condition).

(5) Agree-Link and Agree-Copy account



Note that Baker's (2022) account, as it has been presented so far, does not extend to the additional cross-referencing of the matrix object via a suffix on the complementizer that we find in Kipsigis. Baker proposes that in Kipsigis, Eval can introduce both SOK and OOK (= object of knowledge). The latter is then controlled by the matrix object. As for the obligatory presence of prefixal subject agreement with suffixal object agreement on the C-like element in Kipsigis, Baker draws an analogy to the 'say,' which can similarly only introduce a goal if it has also introduced an agent.

This is particularly interesting with respect to our own analysis, as we will claim that the C-like element simply constitutes the verb 'say' in Kipsigis.

The pattern of complementizer agreement in Kipsigis poses a challenge for all existing accounts. Even with the addition of OOK in Baker's account, it is unclear why there is no true object agreement in Kipsigis matrix clauses, as would be expected based on the T/Agree Condition (Baker 2022, 50). Recall that Diercks and Rao (2019) identify object markers as clitics instead of agreement markers, the result of (optional) clitic doubling. Hence, the suffixal cross-referencing on the complementizer violates Baker's T/Agree Condition. Additionally, we will show that even prefixal agreement on the C-like element can cross-reference the matrix object, in the absence of cross-referencing of the matrix subject. More generally, we will provide several arguments in favor of analyzing the C-like element as a verbal category, questioning C raising accounts that have been proposed for Kipsigis (Diercks and Rao 2019; Diercks et al. 2020).

3 Complementation in Kipsigis

In this section, we describe the pattern of C agreement in Kipsigis, based on previous descriptions as well as our own fieldwork. Before proceeding to details, we note that the language is *pro*-drop, with a VSO unmarked order (Bossi and Diercks 2019) and the typologically rare marked nominative system (Kouneli 2019; Kouneli and Nie 2021).

We start by describing complementation strategies in the language more generally, for which a short detour into mood inflection is needed. All verbs in Kipsigis inflect for tense, aspect, and mood, and previous literature has identified three moods: indicative, subjunctive, and imperative (Toweett 1979; Rottland 1982; Creider and Creider 1989). The language lacks infinitives of the European type.[4] Morphologically, the subjunctive differs from the indicative in the vowel length of the subject agreement prefix and in the tonal melody of the stem (see Toweett 1979 for detailed conjugation paradigms). Thus, we see that in (6) below, the verb *ru* 'sleep' has a short-voweled subject agreement prefix in its indicative (matrix) form in (6a) but a long-voweled prefix in its subjunctive (embedded) form in (6b).[5] For third person subjects, the prefix is ∅- in most cells of the paradigm, while it is always *ko(:)-* in the subjunctive.[6]

[4]While various tense and aspect distinctions are made in the indicative, only two aspect forms are distinguished in the subjunctive: the perfective and imperfective. It is also worth noting that Toweett (1979) and Rottland (1982) call this inflection of the verb the *governed verb form* and the *abhängige Verbform* (dependent verb form), respectively. We adopt the term "subjunctive" used in the description of Nandi and Kipsigis inflection in Creider and Creider (1989).

[5]In the remainder of the paper, subjunctive inflection will always be indicated in the glosses, while indicative will be left unglossed.

[6]The exact shape of the subject agreement prefix, as well as the tonal melody of the stem, varies not only by mood but also by the tense-aspect combination of the verb; it also depends on which conjugation class a given verb belongs to, Class I or II (not to be confused with Type I and Type II subjunctives, introduced below). The examples given in this section (including *le* itself) belong to Class I. The interested reader is referred to Toweett (1979), Rottland (1982), Creider and Creider (1989), and Kouneli (2022) for a complete description and sample conjugation paradigms.

(6) a. Kù-í-rù.
PST.CURR-2SG-sleep(IND)
'You slept.'

b. Í-mátʃ-é [ì:-rù].
2SG-want-IPFV 2SG-sleep.SBJV
'You want to sleep.'

The specific syntactic and semantic environments in which the subjunctive is used will be discussed shortly (see also the Appendix), but we note here that in the first person singular, there is a morphological distinction between two types of subjunctive in the (unmarked) perfective. An example can be seen in (7), where the lexical verb *ru* has a long vowel in its 1SG agreement prefix when it appears as the main verb in the second conjunct of a coordination (an environment that requires subjunctive in Kipsigis) but a short vowel when embedded under a volitional predicate (compare to the long vowel in (6b) above). There is no such morphological difference for other person-number combinations or for 1SG in the imperfective, where the vowel is always long. We will be calling the former type of subjunctive Type I and the latter subjunctive Type II, glossed henceforth as SBJVI and SBJVII. We assume that there is syncretism between the two types in all cells of the paradigm except 1SG in the perfective.

(7) a. Kù-∅-púːtʃ Tʃéːbêːt káːt (ák) àː-tʃáp tʃàːíːk.
PST.CURR-3-sweep Cheebeet house and 1SG-make.SBJVI tea
'Cheebeet swept the house and I made tea.'

b. á-mátʃ-é [à-rú].
1SG-want-IPFV 1SG-sleep.SBJVII
'I want to sleep.'

As has already been mentioned, Kipsigis lacks infinitives, which means that the subjunctive is widely used in complementation contexts. The subjunctive used in complementation is Type II, with subjunctive Type I being restricted to coordination contexts (as in (7a)), temporal adjunct clauses, and conditionals (see Appendix for details and examples). Thus, we find various verbs—most prominently, volitional predicates—that always select for a subjunctive Type II complement. In this case, there is no complementizer present, as already seen in (6b). This is the only complementation strategy for these verbs.

A second class of verbs always selects for a clausal complement where the verb is inflected in the indicative. The most prominent verbs in this class are factive verbs, with two examples seen in (8). In (8a), we see that *síːr* 'to pass' is inflected for indicative mood when embedded under the adjectival attitude *pájpáj* 'to be happy.' Similarly, in (8b) *ŋaːl* 'to lie' is inflected for indicative under *nereːtʃ* 'to be angry.' For verbs that select for indicative complements, the presence of the complementizer is obligatory.

(8) a. Àː-pájpáj [ùː-lé kòː-∅-síːr Kíplàŋgàt].
1SG-happy 1SG-LE PST.REC-3-pass Kiplangat.NOM
'I'm happy that Kiplangat passed (the exams).'

b. ɑ-nereːtʃ-i [ɑː-le koː-∅-ŋɑːl-ɑn Kíbêːt].
1SG-angry-IPFV 1SG-LE PST.REC-3-lie-1SG Kibeet.NOM
'I'm angry that Kibeet lied to me.'

Finally, many verbs can select for either a subjunctive complement or an indicative complement, with interpretive differences. For example, when communication verbs select for indicative, the reading of the complement clause is an assertive/reportative one, while a directive meaning arises if subjunctive II is used. This is illustrated in (9) for the verb *tʃaːm* 'to whisper.' We see *tʃáp* 'to make' inflected for indicative in (9a) but for subjunctive II in (9b).

(9) a. Kà-∅-tʃáːm Kíbêːt [kò-lé kà-∅-tʃáp kímɲéːt].
PST.CURR-3-whisper Kibeet.NOM 3-LE PST.CURR-3-make ugali
'Kibeet whispered that he made ugali.'

b. Kù-∅-tʃáːm-ú-án Kíbêːt [à-tʃáp kímɲéːt].
PST.CURR-3-whisper-VENT-1SG Kibeet.NOM 1SG-make.SBJVII ugali
'Kibeet whispered to me to make ugali.'

To summarize, in a similar fashion as mood selection in European languages, lexical verbs in Kipsigis are divided into those that only select for subjunctive complements, those that only select for indicative complements, and those that can select either. While a complete investigation of the lexical semantics of the verbs that select for subjunctive versus indicative is beyond the scope of the paper, Table 1 provides a list of all predicates that we have tested so far. The table is to be read as follows: ✓ means that a complement of the given mood (indicative or subjunctive) is possible, and ✗ means that it is impossible. ✗? indicates that we have never encountered a use of the given predicate with a complement of the given mood, but do not have actual ungrammatical examples at hand. Nevertheless, the data we do have indicate that those predicates with a ✗? for one of the moods are likely to either prohibit or at least strongly disprefer complements of that mood. The table is organized into three blocks: in the first block, we present the verbs that can select for either mood (for these verbs, we also include a column explaining the difference in meaning); in the second block, we present the verbs that predominantly appear with indicative complements; in the third block, we present the verbs that select for subjunctive complements.

As was already mentioned, the complementizer is required whenever there is an indicative clausal complement. In other words, the complementizer is never optional (see also Diercks and Rao 2019).[7] We now turn to the core properties of the complementizer—the focus of our paper.

[7]For completeness, we note that there are, to our knowledge, two verbs that select for an indicative complement where the use of the complementizer is prohibited: *le* 'to say' and *par* 'to think (with negative bias).' The former provides evidence for the verbal analysis of this "complementizer" and will be discussed in Sect. 4.1, while the latter is discussed in detail in Bossi (2023a).

We also note that, modulo matrix uses of *le* that will be discussed in Sect. 4, we are not aware of any uses of *le* outside of complementation of the type discussed here (i.e. matrix predicate followed by indicative complement).

Table 1 Mood selection

Predicate	Indicative (+*le*)	Subjunctive (Type II)	Meaning difference (IND vs. SBJV)
mwa 'say'	✓	✓	Assertive/reportative vs. directive
tʃaːm 'whisper'	✓	✓	Assertive/reportative vs. directive
síːr 'write'	✓	✓	Assertive/reportative vs. directive
tɛp 'ask'	✓	✓	Assertive/reportative vs. directive
maŋ 'expect'	✓	✓	No clear difference
kas 'hear'	✓	✓	Hear that … vs. hear (someone) Xing
keːr 'see'	✓	✓	See that … vs. see (someone) Xing
ŋen 'know'	✓	✓	Know that … vs. know how to …
pwaːɪ 'think/remember'	✓	✓	Think/remember that … vs. remember (someone) Xing
ɲoːn 'complain'	✓	✗?	
jaːn 'believe'	✓	✗?	
taːm 'falsely accuse'	✓	✗?	
ruaːtit 'dream'	✓	✗?	
poːr 'show'	✓	✗?	
naj 'realize'	✓	✗?	
raːgin 'worry'	✓	✗?	
nɛrɛːtʃ 'be angry'	✓	✗?	
pajpaj 'happy'	✓	✗?	
matʃ 'want'	✗	✓	
jaj 'make/do'	✗	✓	
mje 'good'	✗?	✓	
karaːran 'good/beautiful'	✗?	✓	
ja 'bad'	✗?	✓	

The Kipsigis complementizer consists of the root of the lexical verb *le* 'say' and a person/number agreement prefix:

(10) a. ûː-ŋgén ùː-lé ∅-rú-è Kíbêːt.
1SG-know 1SG-LE 3-sleep-IPFV Kibeet.NOM
'I know that Kibeet is sleeping.'

b. Kà-ɔ́-mwá òː-lè ∅-rú-è Kíbêːt.
PST.CURR-2PL-say 2PL-LE 3-sleep-IPFV Kibeet.NOM
'You (PL) said that Kibeet is sleeping.'

c. Kíː-ŋgèn kèː-lé ∅-rú-è Kíbêːt.
IMPRS-know IMPRS-LE 3-sleep-IPFV Kibeet.NOM
'It is known that Kibeet is sleeping' (impersonal).[8]

[8]The impersonal construction in Kipsigis is syntactically active. Morphologically, it is expressed by combining a first person plural subject agreement prefix with third person tonal melody.

Table 2 Agreement prefixes on *le* (= subjunctive subject prefixes)

	SG	PL
1	à:-	kè:-
2	ì:-	ò:-
3	kò-	
IMPRS	kè:-	

Based on work with two native speakers, Diercks and Rao (2019) report an additional, nonagreeing 'say'-based complementizer for Kipsigis, illustrated in (11).[9]

(11) ɑ-ŋgɛn *(ɑ-le/kɔlɛ) ko-∅-ruuja tuɣa amut.
1SG-know 1SG-C/that PST-3-sleep cows yesterday
'I know (that) the cows slept yesterday.'
(Diercks and Rao 2019, 372)

The native speakers that we consulted all found the nonagreeing complementizer in sentences like (11) ungrammatical. We therefore conclude that our speakers only have an agreeing complementizer. It is possible that there is speaker variation, with the nonagreeing complementizer reported by Diercks and Rao only available in the grammar of a subset of speakers.[10] Table 2 gives the paradigm for the agreement prefixes on *le*. The prefixes are identical to the agreement prefixes of lexical verbs in subjunctive Type I, a fact that we discuss in detail in Sect. 4.1.

Diercks and Rao (2019) argue that the Kipsigis complementizer can only agree with the matrix subject. It is clear from our data, however, that the complementizer may agree with nonsubject DPs in the matrix clause, a possibility that is not fully explored in Diercks and Rao (2019). Whenever matrix objects can qualify as the source of information reported in the embedded clause, agreement with *le* becomes an option, as shown for a PP object in (12) and an applied object in (13).[11]

[9] Our [ATR] and vowel length transcriptions sometimes differ from those in Diercks and Rao (2019). Their transcriptions possibly contain some typos, since they display mismatches in the [ATR] values of vowels within a single word, which is prohibited in Kipsigis due to the language's dominant [ATR] vowel harmony system (Hall et al. 1974; Halle and Vergnaud 1981; Baković 2000; Nevins 2010). In this paper, we have maintained the original transcriptions and glosses for examples from Diercks and Rao (2019).

[10] Mike Diercks (p.c.) informs us that the speakers that they worked with came from Nakuru and Kericho, while our speakers all come from Bomet and Narok (these are all counties in western Kenya). It is therefore possible that there is dialectal variation.

[11] In (12), we assume that the form *kòlé* reflects agreement with the third person subject. An anonymous reviewer asks whether in this case *kòlé* could instead be the nonagreeing complementizer reported in Diercks and Rao (2019). This is in principle a possibility, since the third person agreeing form and the nonagreeing form are claimed to be morphologically identical (Diercks and Rao, however, do not provide tonal transcriptions, so this claim cannot be fully evaluated). Nevertheless, our speakers always reject the use of *kole* in contexts without possible third person targets, unlike the speakers consulted by Diercks and Rao. This is why we conclude that our speakers do not have the nonagreeing form in their grammar, as we discussed above.

(12) Kà-∅-kás Kíplàŋgàt kòbún **ín̂ɛ̂ː** kò-lé/**ɪ́ː-lè**
PST.CURR-3-hear Kiplangat.NOM from 2SG 3-LE/2SG-LE
kà-∅-tʃɔ́ːr Kíbêːt ràbíːník.
PST.CURR-3-steal Kibeet.NOM money
'Kiplangat heard from you that Kibeet stole the money.'

(13) Kòː-ɑ́-mwɑ̀j-têː-tʃí **Tʃèbêːt** éːn tòːjéːt ɑ̀ː-lé/**kò-lé**
PST.REC-1SG-say-IT-APPL Cheebeet at meeting 1SG-LE/3-LE
kòː-∅-tʃɔ́ːr Kíbêːt ràbíːník.
PST.REC-3-steal Kibeet.NOM money
'At the meeting, I said on Cheebeet's behalf that Kibeet stole the money.'

Another example of agreement with nonsubject DPs can be seen in (14). The verb *wuːt* 'to forget' appears in a syntactic frame in which the grammatical subject is invariably third person, and the experiencer is expressed as an indirect object introduced by the applicative.[12] In this case, *le* agrees with the experiencer–indirect object, not with the grammatical subject.

(14) Kɑ̀-∅-wúːt-ú-**ɑ́n** **ɑ̀ː-lé** kɔ̀ː-∅-kɛ́r Kíbêːt
PST.CURR-3-forget-VENT-1SG 1SG-LE PST.REC-3-close Kibeet.NOM
kúrgéːt.
door
'I forgot that Kibeet closed the door.'

Furthermore, impersonal agreement on the complementizer (see (10c) above) is also available for a wide range of fully inflected lexical verbs in the matrix clause, in which case a hearsay or rumor interpretation arises; this is illustrated in (15) below.

(15) Kà-∅-kás Kíplàŋgàt kèː-lé kà-∅-tʃɔ́ːr
PST.CURR-3-hear Kiplangat.NOM IMPRS-LE PST.CURR-3-steal
Kíbêːt ràbíːník.
Kibeet.NOM money
'Kiplangat heard (a rumor) that Kibeet stole the money.'

Diercks and Rao (2019) additionally report a pattern of what they call *object agreement*, where the complementizer (optionally) agrees with the indirect object of the matrix verb, in addition to agreement with the subject. In this case, the prefix on the complementizer tracks the ϕ-features of the subject, while the suffix tracks the ϕ-features of the object:

(16) Ko-ɑ-mwaa-un ɑ-lɛ-ndʒin ko-∅-ɪt tuɣa amut.
PST-1SG-tell-2SG.OBJ 1SG-C-2SG.OBJ PST-3-arrive cows yesterday
'I DID tell you (SG) that the cows arrived yesterday.'
(Diercks and Rao 2019, 371)

[12] This type of syntax for the verb 'forget' is attested in other languages as well (e.g. it is one of the possible case frames for *olvidarse* 'to forget' in Spanish; Rivero 2004).

We henceforth term this pattern "suffixal agreement" since our data reveal two types of object agreement: prefixal object agreement for objects that act as the source of information (as in (12) and (13)) and suffixal object agreement for indirect objects of communication verbs. Diercks and Rao report that suffixal agreement is associated with a verum focus interpretation, reflected in their translations of such examples. However, we have not reliably replicated this finding for all of our speakers, and it will thus not play a role in our analysis.

4 C agreement is verbal agreement

We first argue in Sect. 4.1 that what has been described as a 'say'-based complementizer in Kipsigis is, in fact, the lexical verb 'say'; in other words, it is of category V, not C. In Sect. 4.2, we present novel data from the language showing that the ϕ-features on *le* track the source of the information reported in the embedded clause, not necessarily the matrix subject (contra Diercks and Rao 2019).

4.1 *Le* is a verb

Even though 'say'-based complementizers have been linked to verbal properties before (e.g. Lord 1976; Güldemann 2008; Grimshaw 2015; Özyıldız et al. 2018; Halpert 2019; Letsholo and Safir 2019; Moulton 2019; Bondarenko 2020b; Demirok et al. 2020), analyses of these complementizers as elements of category V, not C, have not always been pursued in the literature (though see Koopman 1984; Koopman and Sportiche 1989; Kinyalolo 1993; Knyazev 2016; Major 2021; Major and Torrence 2021; Major et al. 2022 for exceptions). We provide here four main arguments in favor of analyzing the Kipsigis complementizer as a lexical verb 'say': it can be used as a matrix verb, it inflects for mood and aspect, it can host applicative and reflexive verbal morphology even when used in complementation, and it can be modified by adverbs.

We begin with the observation that *le* 'say' can act as a matrix verb, as shown in (17). Crucially, the "complementizer" is ungrammatical in this case.

(17) **Kù-∅-lé** Kíbêːt (***kò-lé**) ∅-rú-è là:kwèːt.
PST.CURR-3-LE Kibeet.NOM 3-LE 3-sleep-IPFV child.NOM
'Kibeet said that the child is sleeping.'

The VSO word order of the language makes it clear that *le* occupies the position of the lexical verb in (17). Matrix uses of *le* are also reported in Diercks and Rao (2019), but Diercks et al. (2020) take this as evidence in favor of an analysis in which the C head (*le*) overtly raises to the matrix clause. More specifically, they argue that a silent speech verb occupies the matrix verb position and that *le*, which is base-generated in C, moves to this position (see (4) in Sect. 2 for details on this type of analysis for upwards-oriented C-Agree). Such an analysis, however, faces certain challenges

once additional data about the morphology of *le* in matrix versus complementation uses are considered.

We turn, then, to our second argument: inflection for mood and aspect. The first observation is that *le* 'say' is inflected in the subjunctive mood when used as a "complementizer" but in the indicative when used in matrix clauses. We observe in (19) that the inflection of *le* 'say' in matrix versus complementation contexts shows the same contrast between indicative and two types of subjunctive that we see in lexical verbs like *ru* 'sleep' in (18). We use here perfective 1SG forms in order to illustrate the two types of subjunctive (which are morphologically indistinguishable in other cells of the paradigm). In (19a), we see a matrix use of *le*, in which case there is indicative inflection (cf. (18a)). In (19b), *le* is used to introduce an embedded clause, and it has Type I subjunctive, otherwise seen in—among other contexts—conditionals and coordination (cf. (18b)). Finally, in (19c), "matrix" *le* is embedded under a volitional predicate and shows up with Type II subjunctive (cf. (18c)). From now on, we always gloss subjunctive inflection on *le*.

(18) a. Kìː-á-rú.
PST.DIST-1SG-sleep(IND)
'I slept.'

b. iŋgot àː-rú
if 1SG-sleep.SBJVI
'if I sleep'

c. á-mátʃ-é à-rú.
1SG-want-IPFV 1SG-sleep.SBJVII
'I want to sleep.'

(19) a. Kìː-á-lé kìː-∅-tʃɔ́ːr Kíbêːt ràbíːník.
PST.DIST-1SG-LE(IND) PST.DIST-3-steal Kibeet.NOM money
'I said that Kibeet stole the money.'

b. Kìː-á-mwá àː-lé kìː-∅-tʃɔ́ːr Kíbêːt
PST.DIST-1SG-say 1SG-LE.SBJVI PST.DIST-3-steal Kibeet.NOM
ràbíːník.
money
'I said that Kibeet stole the money.'

c. á-mátʃ-é à-lé kìː-∅-tʃɔ́ːr Kíbêːt ràbíːník.
1SG-want-IPFV 1SG-LE.SBJVII PST.DIST-3-steal Kibeet.NOM money
'I want to say that Kibeet stole the money.'

In a C raising account (Diercks and Rao 2019; Diercks et al. 2020), it is an accident that the complementizer is inflected in the subjunctive. The mood inflection follows naturally, however, if *le* is a verb.

The C raising account also faces problems when it comes to matrix uses of *le* in the imperfective (so far, we have mostly seen perfective examples). As can be seen in (20), the imperfective form of *le* 'say' is *leːlen*, which exhibits irregular stem

Table 3 Suffixal agreement (Diercks and Rao 2019, Table 3)

	SG	PL
1	-lɛ-ndʒ-ɑn	-lɛ-ndʒ-ɛtʃ
2	-lɛ-ndʒ-in	-lɛ-ndʒ-ɔːɣ
3	-lɛ-ndʒ-i	

allomorphy.[13] In the verbal analysis pursued here, *le* is a lexical verb and is thus predicted to inflect for aspect. In a C raising account, on the other hand, *le* is a C head that raises into a matrix verb position. It is unlikely, however, that an element of category C would show irregular stem allomorphy conditioned by aspect.

(20) **Lèːlén** lɔ̀ɣójwɛ̀ːk kɔ̀ː-∅-tʃóːr Kíbêːt ràbíːníːk.
LE.IPFV(IND) news.NOM PST.REC-3-steal Kibeet.NOM money
'The news says that Kibeet stole the money.'

In (20), *le* is in the matrix verb position. What is more striking, however, is that *le* can inflect for aspect even when used in complementation contexts (as a reminder, verbs in the subjunctive only make a perfective vs. imperfective distinction). We see in (21) that when the matrix verb is inflected in the past imperfective, *le* can appear in either its perfective or imperfective form.[14]

(21) Ká-ɑ-mwá-é àː-lé/àː-lèːlén
PST.CURR-1SG-say-IPFV 1SG-LE.SBJVI/1SG-LE.IPFV.SBJVI
kà-∅-tʃóːr Kíbêːt ràbíːníːk.
PST.CURR-3-steal Kibeet.NOM money
'I was saying that Kibeet stole the money.'

The third argument in favor of a verbal analysis of *le* 'say' comes from a reevaluation of the suffixal agreement data presented in (16), which are repeated below as (22). Diercks and Rao (2019) give a table of *le* forms with object agreement, reproduced as Table 3.

(22) Ko-ɑ-mwaa-un ɑ-lɛ-ndʒin ko-Ø-ɪt tuɣa amut.
PST-1SG-tell-2SG.OBJ 1SG-C-2SG.OBJ PST-3-arrive cows yesterday
'I DID tell you (SG) that the cows arrived yesterday.'
(Diercks and Rao 2019, 371)

Looking at Table 3, we observe that all forms share not only *le* but also a [ndʒ] consonant sequence. This indicates the possibility (acknowledged by Diercks and Rao 2019 themselves) that there is a hidden morpheme present between *le* and the person/number suffixal agreement. We argue here that this is indeed the case, with the forms reported in Table 3 being decomposable into an allomorph of *leːn*—followed by the applicative suffix *-tʃi*, followed by the regular object clitics in the language.

[13]The imperfective is usually expressed via a suffix, whose exact form is determined by a number of factors, including TAM and conjugation class. We again refer the interested reader to Toweett (1979), Rottland (1982), Creider and Creider (1989), and Kouneli (2022) for details on Kipsigis conjugation.

[14]Only imperfective—the morphologically marked aspect in Kipsigis—will be indicated in the glosses.

Table 4 Suffixal agreement decomposed into APPL and object clitics

	SG	PL
1	-leːn-tʃi-ɑn [leːndʒɑːn]	-leːn-tʃi-eːtʃ [leːndʒeːtʃ]
2	-leːn-tʃi-in [leːndʒiːn]	-leːn-tʃi-ɑːk [leːndʒɑːk]
3	-leːn-tʃi [leːndʒi]	

Table 5 Object clitics

	SG	PL
1	-an	-ɛːtʃ
2	-in	-aːk
3	∅	

Regular phonological processes (e.g. voicing of obstruents after nasals and vowel coalescence rules; Kouneli 2019, Chap. 2) give the surface forms that we see in Table 3. The decomposition of the suffixal forms is given in Table 4, with surface phonological forms in brackets.

The morphemes making up the forms in Table 4 are independently attested. The suffix *-tʃi* is the most common applicative morpheme in the language (Toweett 1979; Rottland 1982; Creider and Creider 1989), used to introduce applied arguments with a variety of thematic roles (e.g. recipient, beneficiary).[15, 16] An example is given in (23).

(23) a. Kà-∅-tʃáp Kíbêːt kímɲéːt.
PST.CURR-3-make Kibeet.NOM ugali
'Kibeet made ugali (type of food).'

b. Kà-∅-tʃáp-**tʃí** Kíbêːt Tʃèːbêːt kímɲéːt.
PST.CURR-3-make-APPL Kibeet.NOM Cheebeet ugali
'Kibeet made ugali for Cheebeet/on behalf of Cheebeet.'

The object clitics that we have postulated in Table 4 are simply the regular object clitics in the language, summarized in Table 5, which was constructed with data from Toweett (1979, 209).[17] The last piece of the reanalysis is the claim that the verb *le* has an allomorph *leːn*. In Kalenjin languages, there are at least 10 CV verbs that have a CVː(*n*/*l*/*r*) allomorph when followed by other morphemes, with *le* being such a verb (e.g. Zwarts 2004, 116 reports the allomorphs *le* and *leːl* for the cognate word in Endo-Marakwet). To see that this is not unique to *le* in Kipsigis, consider the

[15] There is another applicative suffix *-eːn*, which is mostly used for instruments and sources (Toweett 1979; Rottland 1982).

[16] The applicative *-tʃi* has an allomorph *-^{j}i* when attached to verbs ending in an alveolar obstruent. It also has the allomorph *-u* for first/second person applied arguments for most (but not all) lexical verbs. This has been analyzed as a specialized use of the ventive suffix *-u* in Kalenjin/Southern Nilotic languages (Rottland 1982; Creider and Creider 1989; Zwarts 2004; Mietzner 2009).

[17] The clitics take the [ATR] value of the stem. Additionally, the vowel of 1SG and 2SG clitics is lengthened in the presence of a local person subject (not indicated in the table; Toweett 1979; Creider and Creider 1989).

allomorphy displayed by the verb *ɲo* 'to come' in (24): the root has the form *ɲo* when in stem-final position in (24a) but the form *ɲoːn* when followed by the applicative in (24b). The surface phonological form in (24b) is the same as that found on *le* with 1SG suffixal agreement (see Table 4), further strengthening the point that those forms include an applicative suffix followed by an object clitic.

(24) a. Kù-∅-ɲò káːt.
PST.CURR-3.IND-come home
'He/she came home.'
b. ɲoːn-tʃi-ɑn. [ɲóːndʒáːn]
come-APPL-1SG
'Come to me!'

Further evidence for the presence of an applicative suffix on the complementizer comes from reflexives and reciprocals. Kipsigis has a verbal suffix *-kɛː* used to form reflexives and reciprocals, illustrated in (25) below.[18]

(25) Kí-kéːr-è-**kêː**.
1PL-look-IPFV-REFL
'We are looking at ourselves/at each other.'

The suffix *-kɛː* can appear after the applicative *-tʃi*, in which case it takes scope over the applicative. With (at least) communication verbs, when the applied argument position is occupied by *-kɛː*, suffixal agreement on *le* can include both the applicative and the reflexive/reciprocal suffix, as shown in (26).

(26) Kóː-∅-tʃùːm-**tʃì̀-kêː** Kíbêːt kò-lèːn-**tʃì̀-kêː**
PST.REC-3-whisper-APPL-REFL Kibeet.NOM 3-LE-APPL-REFL.SBJVI
ŋûːm.
clever
'Kibeet whispered to himself that he's clever.'

Finally, our fourth argument: if *le* is indeed a verb, it is predicted that it should in principle be compatible with adverbial modification. As can be seen in (27), this prediction is borne out: the adverb *muːtja* 'slowly' can appear after (the imperfective form of) *le*, which is the expected position if the adverb modifies *le* but not if it modifies the matrix verb. It is not clear at this point whether there are semantic differences depending on the position of the adverb (following the matrix predicate vs. following *le*).

(27) [Kù-∅-mwá-é Kíbêːt [kò-lèːlén **mùːtjà**
PST.CURR-3-say-IPFV Kibeet.NOM 3-LE.IPFV.SBJVI slowly
[kà-∅-tʃóːr Kíplàŋgàt ràbíːník]]].
PST.CURR-3-steal Kiplangat.NOM money
'Kibeet was saying slowly that Kiplangat stole the money.'

[18]This suffix is unique in being outside of the [ATR] harmony domain of the verb.

Summarizing, data that were not explored in Diercks and Rao (2019) and Diercks et al. (2020) strongly support the analysis of *le* as a verb: it inflects for mood and aspect, it can host applicative and reflexive/reciprocal morphology (even when used in complementation), and it can be modified by adverbs.

The analysis of *le* as a verb implies that clauses introduced by *le* (in a descriptive sense) will differ from European CPs in at least some of their distributional properties. This is borne out. For example, when a *le* clause is placed in subject position, all of our speakers provide a translation that involves the verb 'say,' as shown in (28); they insist that such a sentence cannot mean 'That Kibeet stole the money is bad.' When asked to translate that English sentence into Kipsigis, they give the paraphrase in (29), which involves two separate clauses (and no copy of *le*).

(28) Já **[kè:-lé/kò-lé/à:-lé** kà-∅-tʃó:r Kíbê:t
bad IMPRS-LE.SBJVI/3-LE.SBJVI/1SG-LE.SBJVI PST-3-steal Kibeet.NOM
ràbí:ník].
money
#'That Kibeet stole the money is bad.'
'It is bad for people/him/her/me to say that Kibeet stole the money.'

(29) Kà-∅-tʃó:r Kíbê:t ràbí:ník. Já.
PST.CURR-3-steal Kibeet.NOM money bad
'Kibeet stole the money. (This) is bad.'

Similarly, *le* is incompatible with specificational uses of CPs, as shown in (30).[19]

(30) Context: We are organizing an event (for which we need money), but Kibeet stole the money and so we cannot organize it. Cheebeet (who doesn't know that Kibeet stole the money) comes in and asks: Why didn't the event take place? What was the problem?
Ta:bʊ:t ko (*ko-le) kɔ̀:-∅-tʃó:r Kíbê:t ràbí:ník.
problem TOP 3-LE.SBJVI PST.REC-3-steal Kibeet.NOM money
'The problem was that Kibeet stole the money.'

Before closing this section, it is worth examining a negation-related argument that Diercks et al. (2020) provide against a verbal analysis. The negative morpheme *ma-* can attach to *le* when it is used as a matrix verb, as in (31a), but not when *le* is used in complementation, irrespective of whether there is matrix negation present, as shown in (31b) and (31c). Diercks et al. (2020) argue that the ungrammaticality of negation in complementation uses indicates that *le* is a complementizer, not a verb.

(31) a. Má-ɑ-lè ∅-rú-è là:kwɛ̀:t.
NEG-1SG-LE(IND) 3-sleep-IPFV child.NOM
'I didn't say that the child is sleeping.'
b. Má-a-mwá (*mɑ-)ù:-lé ∅-rú-è là:kwɛ̀:t.
NEG-1SG-say NEG-1SG-LE.SBJVI 3-sleep-IPFV child.NOM
'I didn't say that the child is sleeping.'

[19] In this example, the noun *ta:bʊ:t* 'problem' appears in a preverbal topic position (this position is further discussed in Sect. 5.1).

c. *Kà-a-mwá mɑ-à:-lé ∅-rú-è là:kwɛ̀:t.
PST.CURR-1SG-say NEG-1SG-LE.SBJVI 3-sleep-IPFV child.NOM
Intended: 'I didn't say that the child is sleeping.'

However, what (31) shows is an asymmetry between matrix and complementation uses of *le* with respect to the availability of negation. While this is something that needs to be explained, the data do not suggest that the explanation lies in the verbal versus complementizer status of *le*. While we do not have a concrete explanation at this point (though see Fn. 41 for a suggestion), evidence against Diercks et al.'s argument comes from data like (32) below. What we see in (32) is a lexical verb embedded under a matrix predicate, and interestingly, we observe in this case the same pattern as in (31) with respect to negation: the negative prefix *ma-* is ungrammatical when attached to the embedded verb, as shown in (32b) and (32c).[20] Thus, we see that there are embedded verbs in the language that do not tolerate negation. Whatever the reason for this might be, what data like (32) show is that unavailability of negation in complementation uses of *le* does not constitute an argument against its analysis as a verb.[21]

(32) a. ɑ̂:-ŋgén à-pír pé:k.
1SG-know 1SG-hit.SBJVII water
'I know how to swim' (lit. 'to hit water').
b. Mɑ̂-ɑ(:)-ŋgén (*mɑ)-à-pír pé:k.
1SG-know NEG-1SG-hit.SBJVII water
'I don't know how (not) to swim.'
c. *ɑ̂:-ŋgén mɑ-à-pír pé:k.
1SG-know NEG-1SG-hit.SBJVII water
Intended: 'I don't know how to swim' or 'I know how not to swim.'

Furthermore, if the explanation for the ungrammaticality of negation in (31b) and (31c) is the C status of *le*, as argued by Diercks et al. (2020), then it is not clear why negation is possible in matrix uses, where Diercks et al. (2020) acknowledge that *le* behaves like a verb. In our analysis, on the other hand, *le* is uniformly a verb, and differences in behavior between matrix and complementation uses arise from differences in the syntactic position of *le* (matrix verb vs. embedded under another verb), a view that is supported by data like (32).

[20]The examples in (32) are reminiscent of control clauses with subjunctives in Greek and other Balkan languages (e.g. Iatridou 1988; Terzi 1992; Varlokosta 1993; Krapova 2001; Landau 2004; Roussou 2009). Preliminary data suggest that we find control in (32) too, but a more detailed investigation is needed to confirm the behavior of such structures in Kipsigis. The question that arises, however, is whether complementation with *le* might involve control, especially since Baker (2022) has recently advocated for an analysis of complementizer agreement that involves control, as discussed in Sect. 2. We will show in the next sections, however, that there is evidence for the presence of a structural subject of *le* that behaves like *pro* (not PRO), which argues against control as the right analysis (irrespective of whether one adopts a predicational or propositional analysis). It is left as a question for further research, though, why *le* structures and control clauses pattern alike with respect to negation.

[21]It is also worth noting that negation is impossible with restructuring infinitives (Wurmbrand 2001; among others) in European languages despite their verbal status; this is another argument in favor of dissociating the (un)availability of negation from the lexical category of the "complementizer."

4.2 Agreement with the source of information

In Sect. 3, we showed that *le* in Kipsigis does not always agree with the matrix subject. Rather, agreement with other DPs in the matrix clause is possible, if those DPs act as the source of the information reported in the embedded clause (recall (12) and (13)). In this section we provide two further arguments in favor of the claim that agreement is sensitive to the source of information and four arguments in favor of treating the local subject of *le* 'say' as a pronoun that establishes coreference with a matrix or discourse antecedent.

First, agreement on *le* is usually subject to an animacy restriction, as shown by the contrast in (33).[22] In both (33a) and (33b) there are two possible antecedents for agreement on *le*: the 1SG (animate) subject and a third person source of information, introduced by the applicative *-eːn*. The source of information is animate in (33a), but inanimate in (33b), and what we observe is ungrammaticality of agreement with the source DP in the latter case. Interestingly, one of our consultants made the following comment: "*Kole* is bad here [in (33b)] because the door cannot talk and *kole* is for living things." This is in line with our arguments in favor of *le* being the lexical verb 'say.'[23]

(33) a. Ká-a-kás-éːn **Alice** àː-lé/**kò-lé**
PST.CURR-1SG-hear-APPL Alice 1SG-LE/3-LE.SBJVI
ká-kò-∅-ít lùːgôːk.
PST.CURR-PRF-3-arrive children.NOM
'I heard from Alice that the children have arrived.'

b. Ká-a-kás-éːn **kúrgéːt** àː-lé/***kò-lé**
PST.CURR-1SG-hear-APPL door 1SG-LE/3-LE.SBJVI
ká-kò-∅-ít lùːgôːk.
PST.CURR-PRF-3-arrive children.NOM
'I heard from the door that the children have arrived.'

Second, *le* can agree with benefactive arguments introduced by the applicative *-tʃi*, but only if they can act as the source of information. Thus, we see that agreement is possible in (13), repeated here as (34), but not in (35), where the benefactive argument of the predicate *kas* 'hear' cannot be construed as a source.

(34) Kòː-á-mwàj-têː-tʃí **Tʃèbêːt** éːn tòːjéːt
PST.REC-1SG-say-IT-APPL Cheebeet at meeting
àː-lé/**kò-lé** kòː-∅-tʃóːr Kíbêːt ràbíːník.
1SG-LE.SBJVI/3-LE.SBJVI PST.REC-3-steal Kibeet.NOM money
'At the meeting, I said on Cheebeet's behalf that Kibeet stole the money.'

[22] The only exceptions to this generalization that we are aware of are inanimate nouns of the repository of information type (e.g. book, radio, news). See Anand et al. (2019), among others, for discussion on the ability of those nouns to act as subjects of speech act predicates.

[23] For some speakers, agreement with DPs denoting the source of information is sensitive not only to animacy but also to how reliable the source is judged to be by the speaker (Culy 1994; Speas 2004). For example, in a context where Alice in (33a) is known to be an unreliable person (e.g. someone who lies often), one consultant reports that third person agreement on *le* is no longer possible. Thanks to Deniz Özyıldız for creating the "unreliable Alice" context.

(35) *Kɑ-ɑ-kɑs-ʲi **Kìbêːt kò-lé** ∅-jàːtʃ-é kò-wá
PST.CURR-1SG-hear-APPL Kibeet 3-LE.SBJVI 3-must-IPFV 3-go(SBJV)
Nairobi.
Nairobi
Intended: 'I heard on Kibeet's behalf that one should go to Nairobi.'

The data presented so far show that ϕ-features encoded on *le* result from agreement with the source of information, which does not always coincide with the matrix subject (contra Diercks and Rao 2019). Since we analyze *le* as a lexical verb, the analysis that suggests itself is one in which *le* agrees with a locally merged subject. Since the subject is covert, however, further investigation is needed regarding its status and the cause for coindexation with a matrix antecedent. In the following, we will thus explore whether the local subject is an anaphor or a pronoun and, in the latter case, whether the relation is established via coreference or binding.

A binding relation is brought into question by the fact that c-command is not necessary for agreement. The verb *le* can agree with the source even if the source is embedded in a PP:[24]

(36) Ká-ɪ-kàs [PP kòbùn **Kìplàŋgàt**] **kò-lé**/ìː-lè
PST.CURR-2SG-hear from Kiplangat 3-LE/2SG-LE.SBJVI
kà-∅-tʃɔ́ːr Kíbêːt ràbíːník.
PST.CURR-3-steal Kibeet.NOM money
'You heard from Kiplangat that Kibeet stole the money.'

Further support for this claim comes from *le* clauses that take the impersonal form *kèː-lé* indicating a rumor interpretation, as we saw in (15), which receives a natural explanation if the pronominal subject corefers with an impersonal antecedent in the discourse.[25] Here, we provide two additional scenarios in which *le* can agree with an antecedent that is (saliently) present only in the preceding discourse, not in the matrix clause: (37) is a third person example, and (38) is a second person example.[26]

(37) Context: You are an investigative journalist and you have one **informant.** No one knows your **informant** but the people you talk to (including your editor) know you only get your information from **him.** So, you go to your editor and say:

[24]In (36), as well as (71a) later in the paper, the transcription for 2SG past forms shows the underlying representation of the morphemes, but a regular phonological process of vowel coalescence (Kouneli 2019, Chap. 2) applies to *ka-* and *ɪ-*, resulting in the form *kɛː-* on the surface. Similarly, we give underlying forms in (58), where vowel coalescence applies between the ventive and 2SG object clitic.

[25]We acknowledge that a pronominal analysis is not the only way to analyze impersonals in Kipsigis. Alternatively, existential closure could provide an appropriate paraphrase for a rumor interpretation, along the lines of *I heard that someone says* … , as suggested by a reviewer. A detailed investigation of the impersonal outside of complementation contexts would be needed to decide which one of the derivations of the *kèː-lé* form is more likely to derive a rumor interpretation. Also note the parallelism to evidential marking on the 'say'-based complementizer in Nyala East (Gluckman 2022, Sect. 3.3).

[26]There is variation in our consultants' judgments regarding these examples. Three speakers find (38) but not (37) acceptable, while one speaker shows the opposite pattern, accepting (37) but not (38).

Ka-a-kas **kò-lé** kà-∅-tʃɔ́ːr Kíbêːt ràbíːník.
PST.CURR-1SG-hear 3-LE.SBJVI PST.CURR-3-steal Kibeet.NOM money
'I heard that Kibeet stole the money.'

(38) Context: We are having an argument about who stole the money. **You** have presented convincing arguments that it is Kibeet who stole the money, and I say to **you**:
ɑ̂ː-jɑ́ːn-í **ìː-lè** kà-∅-tʃɔ́ːr Kíbêːt ràbíːník.
1SG-believe-IPFV 2SG-LE.SBJVI PST.CURR-3-steal Kibeet.NOM money
'I believe you that Kibeet stole the money.'

Since the contexts given in (36)–(38) do not ensure the necessary locality relations, we exclude an analysis involving a direct binding relation between the local subject and its antecedent. This leaves open the possibility of an account in which the subject is bound indirectly by a covert binder that itself is coreferent with the matrix antecedent. Such analyses are for example prominently pursued within the literature on long distance reflexives (Anand and Hsieh 2005; Anand 2006; Charnavel 2020) and logophoric pronoun systems (Koopman and Sportiche 1989; Safir 2004; Speas 2004; Anand 2006).

There is reason to doubt an anaphoric status for the subject. Since it serves as the goal for agreement with *le*, we would expect anaphor agreement effects (Rizzi 1989; Woolford 1999; Sundaresan 2016; Murugesan 2022). As was shown in (25), reflexivization in Kipsigis takes place through the ϕ-invariant verbal suffix *-kɛː*, a strategy that is in complementary distribution with cliticization in nonanaphoric contexts, as shown in (39).

(39) Kɑ-ɑ-keːr-(*ɑn)-kɛː. / Kɑ-ɑ-keːr-kɛː-(*an).
PST.CURR-1SG-see-1SG.OBJ-REFL PST.CURR-1SG.OBJ-see-REFL-1SG
'I saw myself.'

Since this effect arguably qualifies as a case of anaphoric agreement (Woolford 1999, 264) and is absent with prefixal agreement on *lé*, we conclude that the subject does not instantiate an anaphor. Instead we propose that agreement takes place between *le* and a covert pronoun introduced by *le*. This *pro* is coindexed with the matrix/discourse antecedent via the assignment function, thereby avoiding the need for c-command by the antecedent. More details will be given in the next section.

A final argument in favor of the *pro* analysis comes from the fact that the subject of *le* can be overtly realized under certain discourse conditions:[27]

(40) Context: We are having a conversation, and I keep saying that Kibeet stole the money, but you don't want to believe me. So finally, I say:
Kà-a-mwá ɑ̀ː-lé **anɛː** kà-∅-tʃɔ́ːr
PST.CURR-1SG-say 1SG-LE.SBJVI 1SG.NOM PST.CURR-3-steal
Kíbêːt ràbíːník.
Kibeet.NOM money
'I said that Kibeet stole the money.'

[27]The exact conditions that license the overt realization of the subject of *le* are not clear. We have only been able to elicit overt subjects when the matrix predicate is a speech verb.

Having argued for the presence of a local subject and the status of *le* as a verb, we turn to the syntactic analysis in the next section.

5 Syntactic analysis

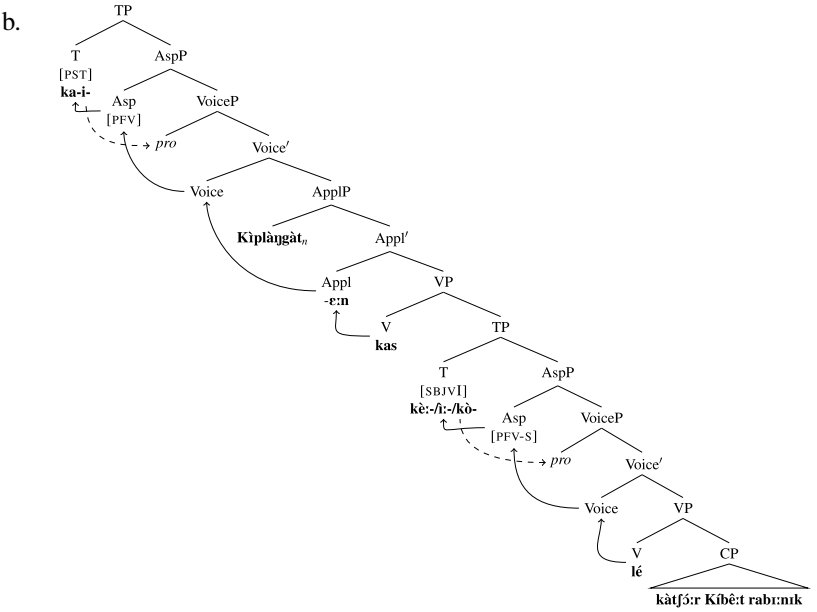
In our analysis, *le* is a verb, and its agreement morphology reflects agreement with a *pro* subject. Thus, any embedded clause that appears with *le* in reality consists of two clauses: the clause headed by *le* itself and the embedded proposition. For a full syntactic analysis, there are two questions that remain to be answered: (i) what is the size of those clauses, and (ii) how do they combine with each other and with the matrix predicate? Our answer to the first question is that the clause containing *le* is a TP, while the embedded proposition is a CP. We give arguments for this analytical choice in Sect. 5.1. Regarding the second question, we argue that all embedding makes use of a complementation structure, where the TP containing *le* is a sister to the matrix verb and the indicative CP complement is a sister to *le*. Arguments for this type of structure are presented in Sect. 5.2.

Our complete analysis, including the different agreement options, is presented in (41b), which is the structure for the sentence in (41a). We choose a verb of perception, since such verbs most naturally allow both an agent and a source of information, making agreement possibilities on *le* more transparent. We assume that Voice introduces the external argument of *le* (Kratzer 1996), while the source argument of the matrix predicate *kas* 'hear' enters the derivation via a high Appl head (Marantz 1993; Pylkkänen 2008).[28] To account for the verb initiality of Kipsigis, we assume that V moves via Voice and Asp to T (or a higher projection; see Bossi and Diercks 2019), shown by the arrows in (41b). The dashed arrows indicate Agree between T and the subject, respectively. Of special interest is the subjunctive T head probing for the ϕ-features of the agent of the saying event—a free pronoun serving as a goal for Downward Agree. Prefixal agreement on *le* follows straightforwardly, as the ϕ-features of *pro* vary with its denotation. The form *iːle* is chosen if *pro* points to the addressee of the utterance, whereas *kòlé* appears if *pro* is coindexed with *Kɪ̀plàŋgàt*, which is the source argument from the matrix clause. Another option is the impersonal form *kèːlé*, which leads to a rumor interpretation: recall (15). In this case, *pro* corefers with an impersonal antecedent in the discourse.

(41) a. Ká-ɪ-kás-ɛ́ːn Kɪ̀plàŋgàt kèː-lé/ìː-lè/kò-lé
PST.CURR-2SG-hear-APPL Kiplangat IMPRS-LE/2SG-LE/3-LE.SBJVI
kà-∅-tʃóːr Kíbêːt ràbíːník.
PST.CURR-3-steal Kibeet.NOM money
'You heard from Kiplangat that Kibeet stole the money.'

[28] In Pylkkänen (2008), source arguments are introduced by a low applicative. We choose here a high applicative for presentation purposes, but this is not crucial for the analysis. Further work on the behavior of arguments introduced by *-eːn*, which is also used to introduce instruments, is needed to determine whether it should be best analyzed as a high or low applicative in Kipsigis.

b.



5.1 The size of clausal complements

In the syntactic structure we provide, *le* clauses are TP sisters to the matrix verb, while *le* itself takes a CP complement: see (41b). In this section, we provide empirical arguments in favor of this choice for the size of the clauses involved.

Starting with the indicative CP complement of *le*, we discuss data showing that *le* can generally introduce a CP. Kipsigis has a topicalization strategy where a DP topic moves to the left periphery and is followed by the overt topic marker *ko* (Driemel and Kouneli 2022); this is illustrated in (42). Following previous work on Nilotic (van Urk 2015), we assume that the preverbal topic position in Kipsigis is Spec,CP.

(42) Kìbêːt **kó** kà-∅-ám kímŋéːt.
Kibeet TOP PST.CURR-3-eat ugali
'Kibeet ate ugali.'

As shown in (43), *le* can introduce clauses with an overt topic marker, indicating that the embedded clause is a CP.

(43) âː-ŋgén àː-lé [Kìbêːt **kó** kà-∅-tʃɔ́ːr ràbíːník].
1SG-know 1SG-LE.SBJVI Kibeet TOP PST.CURR-3-steal money
'I know that Kibeet stole the money.'

A similar argument can be made on the basis of embedded questions. Kipsigis is generally *wh*-in-situ, as shown in (44), which displays the standard VSO order.[29]

[29]However, the language has extensive scrambling, with focused elements showing a preference for the immediately postverbal position (Bossi and Diercks 2019). Since *wh*-words are inherently focused, they often scramble to that position.

(44) Kà-∅-tʃɔ́ːr **ŋàː** ràbíːníːk?
PST.CURR-3-steal who.NOM money
'Who stole the money?'

We see in (45) that embedded *wh*-questions are introduced by *le*. Under the standard assumption that interrogative clauses are CPs, these data show that *le* can take a CP complement.

(45) Mâ-ɑ(ː)-ŋgén àː-lé [kà-∅-tʃɔ́ːr **ŋàː** ràbíːníːk].
NEG-1SG-know 1SG-LE.SBJVI PST.CURR-3-steal who.NOM money
'I don't know who stole the money.'

Summarizing, *le* can combine with clauses that are clearly CPs, and its behavior in these cases is identical to its behavior with indicative complements (e.g. it displays the same morphology and agreement possibilities). We therefore conclude that the most straightforward assumption for the category of the indicative complement in (41b) is a CP.

Moving on to the category of the *le* clause itself, it is clear that it contains at least a VoiceP and an AspP: as has been extensively argued, *le* has a thematic subject (which can even be overt: see (40)), and it can inflect for aspect. Nevertheless, *le* clauses also display certain properties that point towards a reduced clausal structure. First, we never see an overt complementizer co-occurring with *le* in complementation structures. Second, we see in (46) that the subject of *le* cannot be topicalized (in contrast to the subject of the embedded proposition, as shown in (42) above).

(46) *Kà-a-mwá [ánɛ̂ː kó àː-lé [kà-∅-tʃɔ́ːr
PST.CURR-1SG-say 1SG TOP 1SG-LE.SBJVI PST.CURR-3-steal
Kíbêːt ràbíːníːk]].
Kibeet.NOM money
Intended: 'I said that Kibeet stole the money.'

Given these properties, we conclude that there is no evidence for the presence of a C layer, and we follow previous work according to which (at least some) subjunctives are TPs (e.g. Alexiadou et al. 2012; Pietraszko 2017, 2020).[30, 31] According to the same work, however, the languages under investigation (Greek, Romanian, and Ndebele) distinguish between CP and TP subjunctives. As was already mentioned in Sect. 3, Kipsigis does distinguish between two types of subjunctive, which are morphologically different in perfective 1SG only (see Appendix for fur-

[30] *Le* does not show tense distinctions, and it is incompatible with negation, as discussed in Sect. 4. In some theories, these properties could be explained if the T layer is also absent (e.g. Wurmbrand 2001). Nevertheless, the presence of subject agreement on the verb (which is standardly associated with T) points towards the presence of T in Kipsigis *le* clauses.

[31] In their grammar of the related dialect Nandi, Creider and Creider (1989) claim that full clauses cannot be coordinated in the language, and we have already seen in (7a) that the verb of the second conjunct in what could be clausal coordination must inflect for subjunctive Type I (not indicative) in Kipsigis. An investigation of clausal coordination is beyond the scope of this paper, but if CP coordination is disallowed in Kipsigis, examples such as (7a) could be interpreted as evidence for the lack of a C layer in subjunctive Type I.

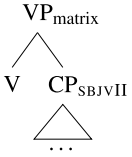
ther discussion). It is thus possible that while subjunctive Type I is a TP subjunctive, subjunctive Type II is a CP subjunctive. Type II subjunctives are used (without an overt complementizer) as clausal complements to a variety of lexical verbs, as was discussed in Sect. 3, which is consistent with their analysis as CP complements.

While we leave a complete investigation of the subjunctive as a topic for further research, we provide here a sketch of a morphological analysis for the relationship between subjunctives Type I and Type II. As a reminder, the two subjunctives are syncretic in all cells of the paradigm except for perfective 1SG. In most morphological frameworks, this would be analyzed in terms of a shared feature. Such an analysis, however, is not trivial if one assumes that the difference between the two types of subjunctive lies in the size of the clause. A possible solution is an analysis along the lines of Pietraszko (2017), where subjunctive morphology does not spell out mood features, but is rather determined positionally. More specifically, subjunctive morphology could be spelling out agreement features on a deficient T head (as a reminder, there are no tense distinctions in the subjunctive in Kipsigis). The two subjunctives would then not be sharing any specific mood feature, but rather they would have in common the presence of a deficient T head. In subjunctive Type II—but not Type I—there is a C head present above T, and one could posit an allomorphy rule where a deficient 1SG T head is spelled out differently in the context of C.

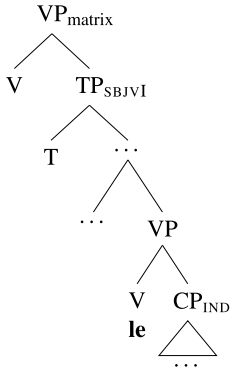
5.2 Complementation versus adjunction

The most common assumption in the syntactic literature on complementation is that a CP headed by the complementizer (e.g. *that* in English) is merged as a sister to the matrix verb. In Kipsigis, however, the element mediating the relationship between the matrix verb and the embedded proposition is the verb *le*, not C. Thus the embedded indicative CP is a sister to *le*, not the main verb. This has consequences for the analysis of mood selection in the language, which was discussed in Sect. 3: we argued that predicates in Kipsigis can select for indicative complements, subjunctive (Type II) complements, or both. We are now in a position to revise this description. Predicates can select for subjunctive complements, as shown in (47), but the nature of indicative "selection" is indirect: under our analysis, verbs select for a *le* clause that then introduces the indicative complement, as in (48). What this means is that *le* is possibly the only lexical verb in Kipsigis that can directly merge with an indicative CP (see Major 2021 for a similar claim for the verb 'say' in Uyghur). In Sect. 6.1, we provide a semantic explanation for this fact.

(47) SBJVII mood selection



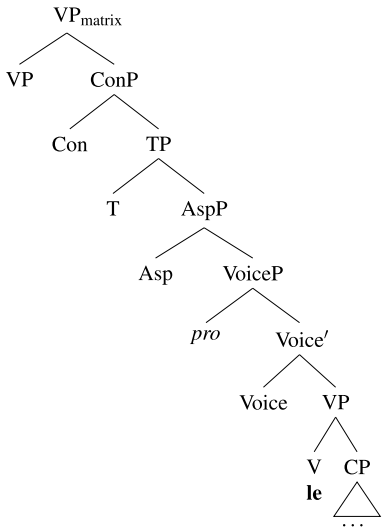
(48) Indicative mood selection



What is less clear is the nature of the relationship between the clause containing *le* and the matrix predicate in (48). While we have argued for a complementation structure, where the *le* clause is a sister to the verb of the matrix clause, there are two plausible alternatives to this view: the *le* clause could combine with the matrix predicate via either adjunction or coordination. These alternatives are motivated by existing analyses of 'say' complementation in other languages (e.g. Major 2021) and by the presence of subjunctive Type I on *le*, which is also seen in adjunct clauses and in coordination structures in Kipsigis. In this section, we discuss, and eventually reject, these two alternatives.

Starting with adjunction, Major (2021) argues for Uyghur that 'say' is part of a converbial clause that is adjoined to the matrix predicate (see also Major et al. 2022 on Lubukusu). This is illustrated in (49). In our structure, we have included a TP since T is necessary to account for agreement in Kipsigis, but converbial clauses are usually more truncated in other languages.

(49)



The structure in (49) is partly motivated by the presence of converbial morphology on the Uyghur verb 'say' when used as a complementizer. The same morphology is used in the language for converbial adjuncts that add a manner modification to a VP (e.g. the equivalent of sentences like 'I entered the house **running**'). At first glance, we find a parallel in Kipsigis: as Bossi (2023b) argues, subjunctive can be used for this type of adjunct clause in Kipsigis as well, illustrated here with the example in (50).

(50) Kɑ-ɑ-we káː **[ɑ̀ː-lɑbɑt-*(i)]**.
PST.CURR-1SG-go.1SG house 1SG-run-IPFV.SBJV
'I entered the house running.'

Nevertheless, we find three important differences between adjunct clauses like (50) and *le* clauses. First, imperfective morphology is obligatory in converbial clauses like (50), indicating that aspect plays an important role in the syntactic and semantic makeup of these clauses. *Le*, on the other hand, is primarily used in the perfective; even though imperfective is possible in some cases (see (21)), its use is restricted, and we have not encountered any context where it is obligatory. Also note that subjunctive Type I and Type II are syncretic in the imperfective (even for 1SG), which means that it is impossible to deternine the type of subjunctive that converbs appear in. It is therefore not even clear that the verb in converbial clauses in Kipsigis has exactly the same morphology as *le*.

Second, while converbial clauses like the one in (50) can be used to answer a 'how' question, *le* clauses can never do so. This is illustrated by the contrast in (51) and (52) below, which indicates that *le* clauses differ from converbial clauses in not acting as manner modifiers.[32]

(51) A: Kɑ-∅-weːndi-tɑ ɑno kaː Tʃéːbêːt?
PST.CURR-3-go-IT how home Cheebeet.NOM
'How did Cheebeet enter the house?'
B: Ko-lɑbɑt-i.
3-run-IPFV.SBJV
'Running.'

(52) A: Ka-∅-tʃaːm-ta ɑno Tʃéːbêːt?
PST.CURR-3-whisper-IT how Cheebeet.NOM
'How did Cheebeet whisper?' (With the intended meaning 'What did Cheebeet whisper?')
B: #Kò-lé kà-∅-tʃɔ́ːr Kíbêːt ràbíːník.
3-LE.SBJVI PST.CURR-3-steal Kibeet.NOM money
'(Saying) that Kibeet stole the money.'

Crucially, Kipsigis differs in this respect from languages like Uyghur, where 'how' questions can target clauses headed by the 'say'-based complementizer, as shown in

[32] Kipsigis 'how' questions require the presence of the itive suffix on the verb, for reasons that are currently not understood (see also Bossi and Diercks 2019).

(53).[33] Major (2024) takes these data as evidence for a converbial analysis of 'say' clauses in that language.

(53) A: Mahinur néme/kim*(-ni) qandaq oyla-y-du?
Mahinur what/who-ACC how think-NONPST-3
'What does Mahinur think about what/who?'
B: Mahinur Tursun-ni ket-t-i de-p oyla-y-du.
Mahinur Tursun-ACC leave-PST-3 say-CNV think-NONPST-3
'Mahinur thinks something, saying that Tursun left.'
(Adapted from Major 2024, 24–25)

The third difference lies in extraction possibilities. As was discussed in the previous section, Kipsigis has a topic position in the left periphery. Topicalization exhibits island effects, which indicates that movement is involved. Illustrative examples are given in (54) and (55) below.

(54) Complex NP island
*__Kibeːt__$_i$ ko kɑ-∅-somɑn Tʃébêːt [kìtàbóːt ne
Kibeet TOP PST.CURR-3-read Cheebeet.NOM book REL.SG
kiː-∅-sir-e __ìnéːndèt__$_i$/__$_i$].
PST.DIST-3-write-IPFV 3SG.NOM
Intended: 'Kibeet, Cheebeet read the book that he wrote.'
(Driemel and Kouneli 2022, 14)

(55) Adjunct island
*__Kibeːt__$_i$ ko ka-kɪ-sɪːndan-ɛːtʃ [ɑmun mɑ-∅-ɲoː __$_i$].
Kibeet TOP PST.CURR-1PL-win-1PL(IMPRS) because NEG-3-come
Intended: 'Kibeet, they beat us (at the race) because he didn't come.'
(Driemel and Kouneli 2022, 14)

Topicalization out of converbial clauses as in (50) is impossible, as shown in (56). This is consistent with their status as adjuncts, which generally behave as islands in the language (see e.g. (55)). For what follows, we also provide the contexts used during elicitation for extraction because speakers often reject fronting to the topic position if not presented with an appropriate pragmatic context (in this case, that of a contrastive topic—see Driemel and Kouneli 2022 for details).

(56) Context: Multiple people enter the house holding different things. What was everyone holding? Who was holding the flower?

[33] Readers might wonder at this point how questions involving attitude predicates are formed in the language. The *wh*-word *neː* 'what' is used, with agreeing *le* being optionally present, as shown in (i). It is not clear yet whether there are semantic differences associated with the presence versus absence of *le*. The presence of *le* in questions, however, is another argument in favor of its verbal analysis; if it were a complementizer, its presence would be quite surprising.

(i) Îː-ŋgén (iː-lè) nêː?
2SG-know.IPFV 2SG-LE.SBJVI what
'What do you know?'

***Maʊwaːt**$_i$ kó kɑ-ɑ-we káː [ɑː-nɑm-e __$_i$].
flower TOP PST.CURR-1SG-go.1SG house 1SG-hold-IPFV.SBJV
Intended: 'The flower, I entered the house holding (it).'

Extraction out of *le* clauses, on the other hand, is always possible, as illustrated in (57)–(59) below for a variety of matrix predicates: a communication verb in (57), a communication verb with a goal argument in (58), and a doxastic predicate in (59).[34]

(57) Context: We are at an event with multiple people attending and multiple dishes available. Who ate what? What did Kibeet eat? (What did Cheebeet eat?)
Kìbêːt$_i$ kó ka-a-mwa [ɑː-le ka-∅-am __$_i$
Kibeet TOP PST.CURR-1SG-say 1SG-LE.SBJVI PST.CURR-3-eat
kimɲeːt].
ugali
'Kibeet, I said that he ate ugali.'

(58) Context: We are at an event with multiple people attending and multiple dishes available. Who ate what? Who ate ugali? (Who ate meat?)
Kímɲéːt$_i$ kó kɑ-ɑ-mwɑ-u-in [a-leːn-tʃi-in
ugali TOP PST.CURR-1SG-say-VENT-2SG 1SG-LE-APPL-2SG.SBJVI
ka-∅-am Kíbêːt __$_i$].
PST.CURR-3-eat Kibeet.NOM
'Ugali, I told you that Kibeet ate it.'

(59) Context: We are at an event with multiple people attending and multiple dishes available. Who ate what? What did Cheebeet eat? (What did Kibeet eat?)
Tʃèbêːt$_i$ kó âː-jáːn-íˊ [àː-lé ka-∅-am __$_i$
Cheebeet TOP 1SG-believe-IPFV 1SG-LE.SBJVI PST.CURR-3-eat
peːndɑ].
meat
'Cheebeet, I believe that she ate meat.'

Data like (57)–(59) indicate that *le* clauses do not have the same structure as converbial clauses in the language, and they point away from an adjunct analysis along the lines of (49).[35]

[34] Bossi (2023b) reports that extraction is impossible (only) when the matrix predicate is a speech verb; the only example she provides is the equivalent of (58), that is, an example where a goal argument is also present. Our speakers, however, confidently judged such examples as grammatical. While we cannot explain this discrepancy from Bossi's findings, we note the following: (i) in her handout, Bossi is not explicit about the pragmatic context used to elicit those data, and speakers often reject topic fronting if the context is not salient, and (ii) her speakers also have a nonagreeing form of *le* (absent in our speakers' grammar), indicating the possibility of dialectal differences.

[35] It is well-known that adjunct islands can sometimes be violated in English and other languages, provided that the events denoted by the matrix predicate and the predicate inside the adjunct meet certain semantic conditions (roughly, when the main event and the adjoined event form an integrated semantic unit; Truswell 2007; among others). One could therefore argue that extraction out of *le* clauses is possible because the saying event forms an integrated semantic unit with the matrix predicate. Note, however, that the English

The extraction data above also provide evidence against the other alternative analysis of *le* clauses: an analysis according to which these clauses are adjoined to the matrix predicate via coordination. Such an analysis might appear attractive in light of data like (7a), repeated here as (60), where subjunctive Type I is used for the inflection of the second conjunct and the conjunction marker *ak* can be omitted.

(60) Kù-∅-púːtʃ Tʃéːbêːt káːt (ák) àː-tʃáp tʃàːíːk.
PST.CURR-3-sweep Cheebeet house and 1SG-make.SBJVI tea
'Cheebeet swept the house, and I made tea.'

However, extraction out of the second conjunct of examples like (60) is impossible irrespective of whether the conjunction marker is present or not, as shown in (61). This is not surprising, since movement in such cases would violate the Coordinate Structure Constraint, which is otherwise active in the language.

(61) Context: Different people were assigned tasks of making different beverages for the guests. Who made what? What did Cheebeet make?
***Tʃèːbêːt**$_i$ kó kɑ-ɑ-puːtʃ káːt [(ák) kò-tʃap __$_i$
Cheebeet TOP PST.CURR-1SG-sweep house and 3-make.SBJVI tʃàːíːk].
tea
Intended: 'Cheebeet, I swept the house, and she made tea.'

Summing up, *le* clauses behave differently from both converbial (adjunct) clauses and the second conjunct in coordination structures. Given these differences, we conclude that a structure in which *le* clauses are sisters to the matrix predicate can best account for the extraction data that we observe. As for the use of subjunctive (Type I) on *le*, we note that there is no reason why such morphology would be incompatible with a complementation structure. According to the morphological analysis sketched in the previous section, subjunctive morphology does not spell out specific mood features, nor does it indicate a particular way in which a clause is merged in the syntax. Rather, it signals the presence of a deficient T head.

While subjunctive Type I morphology on *le* is compatible with our analysis, what might seem problematic is that verbs (= matrix predicates) can select for a TP just in case this TP contains *le*. In other words, in our analysis, *le* clauses can be complements of V, but it is not clear that TP clauses can generally act as verbal complements in the language.[36] We acknowledge that this is a potential weakness of our analysis, but we have shown in this section that there is overwhelming evidence against adjunction of the *le* clause. This means that complementation is the correct structure. While we will unfortunately have to leave the analysis of the selection mechanism as a topic for further research, we speculate that a shift from adjunction to complementation

equivalent of the ungrammatical Kipsigis sentence in (56) is grammatical for many speakers, indicating that Kipsigis disallows extraction out of adjuncts irrespective of the semantics of the events involved. We thank Malte Zimmermann for bringing this to our attention.

[36]We note, however, that we cannot be entirely sure that clauses of size TP are never sisters to verbs in other contexts; complementation in Kipsigis is severely understudied, and it is possible that the relevant constructions have not been discovered yet.

might be a possible grammaticalization path of the verb 'say' into a complementizer. For example, Balusu (2020) claims that in Stage I 'say' clauses are adjuncts, in Stage II they are serial verbs, and in Stage III they are truly embedded (while still being verbs). Such a grammaticalization path would fit well with what has been reported in the recent literature on 'say'-based complementation, where both adjunction (e.g. for Uyghur) and serialization (e.g. for Avatime) have been proposed (e.g. Major 2021). Kipsigis would, thus, constitute an example of a language in the subordination stage. Finally, Burukina (2023) has recently analyzed the 'say'-based complementizer in Mari (Uralic) as a C head that can introduce an argument in its specifier. This C is at least historically related to converbs, and could thus represent an example of the final stage of the grammaticalization path, where 'say' is a C head.

We close this section by pointing out that analyzing *le* clauses in terms of subordination can potentially explain another prominent difference between Kipsigis and other languages with 'say'-based complementation: while the presence of the verb 'say' usually implies *absence* of factivity (see e.g. Major 2024 on Uyghur), agreeing *le* is readily used with factive complements in Kipsigis. For example, we have already seen that agreeing *le* is used with predicates like *ŋgeɲ* 'to know' (1) or *pajpaj* 'to be happy (that)' (8). Examples (62) and (63) below confirm that these predicates are factive in Kipsigis: according to our speakers, a continuation that negates the content of the embedded clause leads to a contradiction. This contrasts with nonfactive verbs like *pwaːt* 'to think,' where the continuation is felicitous, as expected (64). Furthermore, emotive factives (of the type that we see in (63)) uniformly behave as factive predicates even in those African languages that otherwise seem to lack (or have a very small set of) factive predicates of the European type (Ken Safir, p.c.).[37]

(62) **âː-ŋgén** **àː-lé** kòː-∅-síːr Kíplàŋgàt testi,
1SG-know.IPFV 1SG-LE.SBJVI PST.REC-3-pass Kiplangat.NOM test
#kɔbaːtɛːn maː-∅-siːr.
but NEG-3-pass
'I know that Kiplangat passed the test, #but he didn't pass.'

(63) **À:-pájpáj** **àː-lé** kòː-∅-síːr Kíplàŋgàt **#kɔbaːtɛːn**
1SG-happy 1SG-LE PST.REC-3-pass Kiplangat.NOM but
maː-∅-siːr.
NEG-3-pass
'I'm happy that Kiplangat passed (the exams), #but he didn't pass.'

(64) **Kɑ-ɑ(ː)-pwɑːt-i** **àː-lé** kɔ̀-∅-tʃɔ́ːr Kíbêːt
PST.CURR-1SG-think-IPFV 1SG-LE.SBJVI PST.REC-3-steal Kibeet.NOM
ràbíːník, **kɔbaːtɛːn maː-∅-tʃɔːr** (Kíbêːt ràbíːník).
money but NEG-3-steal Kibeet.NOM money
'I thought that Kibeet stole the money, but he didn't (steal the money).'

In contrast, Uyghur exhibits factivity alternations associated with the distribution of 'say': as shown in (65a), the verb *bil* 'to know' has a factive interpretation when the

[37] On a related note, Baker (2022) reports that emotive factives are generally incompatible with the agreeing complementizer in African languages, with the exception of Kipsigis.

complement clause is participial (without 'say'), but not when the complement clause is introduced by the converbial form of 'say' in (65b).

(65) a. Mahinur Tursun-ning ket-ken-**lik-i-ni**
Mahinur Tursun-GEN leave-PTCP-COMP-3POSS-ACC
bil-i-du, **#biraq u ket-mi-d-i.**
know-NONPST-3 but he leave-NEG-PST-3
'Mahinur knows that Tursun left, #but he didn't leave.'

b. Mahinur Tursun-(ni) ket-t-i **de-p** **bil-i**-du, **biraq u**
Mahinur Tursun-ACC leave-PST-3 say-CVB know-NONPST-3 but he
ket-mi-d-i.
leave-NEG-PST-3
'Mahinur knows (something), saying that Tursun left, but he didn't leave.'
(Major 2024, 26, emphasis in the original)

Building on work by Bochnak and Hanink (2021), Major (2024) argues that the lack of factivity in (65b) is due to the adjunct status of the converbial clause headed by 'say.' The fact that no such factivity alternations arise in Kipsigis follows straightforwardly from our analysis, since *le* clauses are complements and not adjuncts.

6 Semantic analysis

After an introduction to the eventuality-based framework of attitude predicates in Sect. 6.1, we lay out the main components of the semantic analysis in Sect. 6.2, namely the semantics of *le* complementation and the Type I subjunctive, with a focus on prefixal agreement. An extension to suffixal agreement on *le* will be made in Sect. 6.3.

6.1 *Le* clauses as sets of contentful eventualities

A classic Hintikkan semantics treats attitude predicates as quantifiers over worlds, determined by the attitude verb and the attitude holder. The verb *believe*, for example, quantifies over worlds compatible with the subject's doxastic alternatives:

(66) Hintikkan semantics
$[\![believe]\!]^{w,g} = \lambda p \lambda x . \forall w' \in \text{DOX}_{x,w} : p(w')$

A shortcoming of this analysis is that attitude predicates are not analyzed as full-fledged verbs, which come with aspect morphology and/or adverbial modification. Hence, recent proposals in this domain have argued for the addition of an eventuality argument to attitude predicates, as a way of combining Davidsonian event semantics (Davidson 1967) with Hintikkan attitude semantics (Hintikka 1969). In order to make this connection, certain eventualities must be claimed to have propositional content. Following Hacquard (2006, 2010) and Anand and Hacquard (2008), we can define a CONT(ENT) function from eventualities to sets of possible worlds compatible with those eventualities. The denotation for *believe* under such an approach is given in (67).

(67) Davidsonian semantics
$[\![believe]\!]^{w,g} = \lambda p \lambda x \lambda e.believe(e) \wedge \text{EXP}(e, x) \wedge \forall w' \in \text{CONT}_{e,w} : p(w')$

As was shown in the previous sections, the Kipsigis *le* morpheme can inflect for aspect and come with adverbial modification, as can the matrix predicate, as shown in (68). Thus, we adopt an eventuality-based framework for attitude predicates in Kipsigis, along the lines of (67), for communication verbs, attitude verbs, and crucially also the morpheme *le*.

(68) a. [Ká-ɑ-tʃáːm-é **mùːtjà** [àː-lèːlén
PST.CURR-1SG-whisper-IPFV slowly 1SG-LE.IPFV.SBJVI
[kà-∅-tʃóːr Kíbêːt ràbíːník]]].
PST.CURR-3-steal Kibeet.NOM money
'I was whispering slowly that Kibeet stole the money.'

b. [Ká-ɑ-tʃáːm-é [àː-lèːlén **mùːtjà**
PST.CURR-1SG-whisper-IPFV 1SG-LE.IPFV.SBJVI slowly
[kà-∅-tʃóːr Kíbêːt ràbíːník]]].
PST.CURR-3-steal Kibeet.NOM money
'I was whispering slowly that Kibeet stole the money.'

Before we come to the semantic decomposition of a concrete example, let us briefly address another extension of the Davidsonian account, which we will not adopt. Given the addition of the eventuality argument, some works have pursued a full thematic separation of the eventuality argument and the content function, where the latter is introduced as a silent modal, a mood particle, or a complementizer in the left periphery of the embedded clause (Kratzer 2006, 2016; Moulton 2009, 2019; Bogal-Allbritten 2015; Grano 2016; Özyıldız et al. 2018; Demirok et al. 2020).[38] This move shifts the semantic action from the attitude predicate to the complementizer/mood of the embedded clause. We will not adopt this separation since the selectional restrictions in the Kipsigis complementation patterns lead us to believe that the content function is a crucial component of the lexical entry for *le*. In fact, we propose that *le* is the only verb that can encode the content function, as it seems to be the only verb that can embed an indicative clause, as was shown in Table 1 and argued in Sect. 5.2. More concretely, we propose the lexical entry in (69a) for *le*. The communication/attitude verbs *le* combines with, however, simply denote eventualities, shown exemplarily in (69b)–(69d) for the verbs *mwa* 'to say,' *jaːn* 'to believe,' and *kas* 'to hear.'

(69) a. $[\![le]\!]^{w,g} = \lambda p_{\langle s,t \rangle} \lambda e_v.say(e) \wedge \forall w' \in \text{CONT}(e)\text{: } p(w')$
b. $[\![mwa]\!]^{w,g} = \lambda e_v.say(e)$
c. $[\![jaːn]\!]^{w,g} = \lambda e_v.believe(e)$
d. $[\![kas]\!]^{w,g} = \lambda e_v.hear(e)$

Under the assumption that CONT cannot be introduced by a silent complementizer or mood operator in Kipsigis, all communication/attitude verbs other than *le* necessarily

[38] Another line of approach takes the content function to be encoded via an additional thematic role (Elliott 2016, 2017; Portner and Rubinstein 2020).

combine with *le* to introduce indicative complement clauses, as they do not encode CONT themselves. This explains why *le* is obligatory with indicative complementation, as was shown in Sect. 3. An immediate prediction of the lexical entry in (69a) is that *le* should not be able to take a nominal argument as a complement. This prediction is borne out, as is shown with the content nouns in (70a). Note that content nouns are possible with verbs that do not have this built-in restriction, which is shown for *mwa* in (70b).[39]

(70) a. *Kɑ-∅-le(ːn-tʃi-ɑn) Kíbêːt saːɛːt/lɔɣɔjwɛːk/ɑtiːndɑːɲɑːt.
PST-3-LE-APPL-1SG Kibeet.NOM prayer/news/story
Intended: 'Kibeet told (me) a prayer/the news/a story.'

b. Kɑ-∅-mwɑː-(ʊ-ɑn) Kíbêːt saːɛːt/lɔɣɔjwɛːk/ɑtiːndɑːɲɑːt.
PST-3-say-VENT-1SG Kibeet.NOM prayer/news/story
'Kibeet told (me) a prayer/the news/a story.'

In this section, we introduced the framework of contentful eventualities and utilized it to explain the uniqueness of *le* in being able to combine with other matrix predicates to introduce indicative complement clauses, as one of the two main complementation strategies in Kipsigis. We derived the special status of *le* via the distribution of the content function. The next section will provide the fully fleshed-out analysis of *le* complementation in Kipsigis.

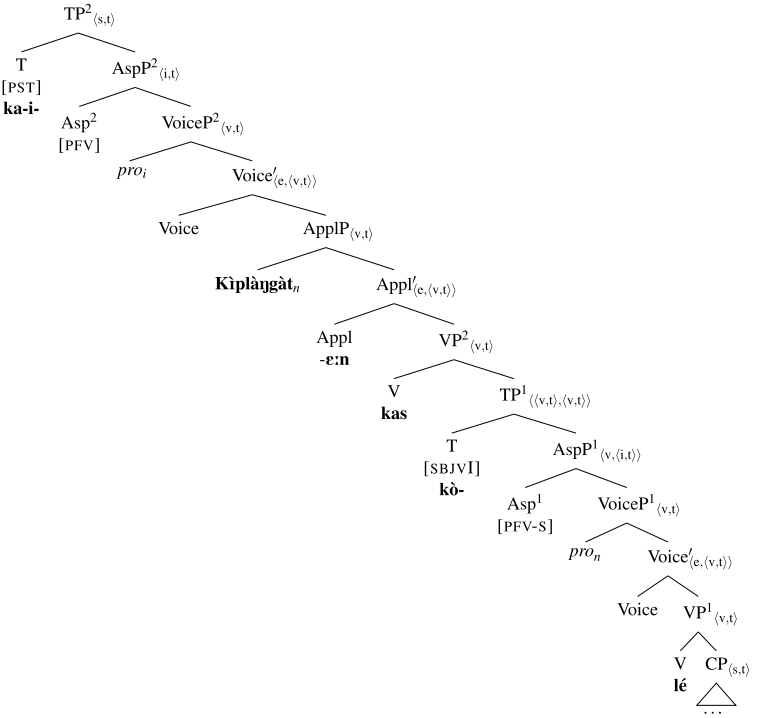
6.2 The semantic components of *le* complementation

Let us now turn to a concrete example. To reiterate, we propose that embedded clauses headed by agreeing forms of *le* constitute sets of contentful saying events, where the verbal nature of *le* 'say' is reflected in its semantics. In order to implement this, we adopt an eventuality-based framework where the relation between the attitude holder and the proposition is mediated by contentful eventualities. We will illustrate our proposed semantics based on the example with a reception verb that we used in Sect. 5, namely (41a), repeated in (71a) with added bracketing; this bracketing reflects the syntactic choices made for the underlying structure, which we argued for in Sect. 5. The tree in (71b) is based on (41b), now enriched with semantic types. We first focus on the *kòlé* derivation, where *pro* is coindexed with the applied argument *Kɪ̀plàŋgàt*.

(71) a. [$_{TP^2}$ Ká-ɪ-kás-ɛ́ːn Kɪ̀plàŋgàt
PST.CURR-2SG-hear-APPL Kiplangat
[$_{TP^1}$ kèː-lé/ìː-lè/kò-lé [$_{CP}$ kà-∅-tʃóːr
IMPRS-LE/2SG-LE/3-LE.SBJVI PST.CURR-3-steal
Kíbêːt ràbíːník]]].
Kibeet.NOM money
'You heard from Kiplangat that Kibeet stole the money.'

[39] While the entry in (69b) does not exclude a combination with content nouns, it does not per se allow this combination either. One way to derive (70b) is by adopting a theta head that introduces individual type arguments, since its existence has independently been argued for in the context of hyperraising in 'say'-based complementation languages in general (Bondarenko 2020b) as well as in Kipsigis in particular (Driemel and Kouneli 2021).

b.



Since *le* is not a complementizer but a verbal category, it introduces a saying eventuality: see the repeated entry in (72a). Following Hacquard (2006), we assume that speech and attitude eventualities have propositional content, that is, that they define sets of possible worlds. The *le* morpheme in Kipsigis is unique in that it not only denotes a saying eventuality but also introduces the content function CONT (see also Kratzer 2006; Moulton 2009). The content function takes eventualities and outputs sets of worlds compatible with the worlds accessible in those eventualities. As for (71a), these would be worlds in which Kibeet steals the money. Thus, *le* combined with the embedded CP results in a set of saying events the content of which is that Kibeet steals the money (72b). This analysis ensures that the agreement morpheme on *le* will always track the source of the information of the embedded clause, as the verb comes with its own Voice layer that introduces the agent of the saying event, where Voice combines with VP via Event Identification (Kratzer 1996), as is shown in (72c). Thus, attitude holder and proposition are connected indirectly via the attitude eventuality. The analysis crucially also predicts that *le* can be modified by an adverb, as was shown in (27)/(68b).

(72) a. $[\![le]\!]^{w,g} = \lambda p_{\langle s,t\rangle} \lambda e_v . say(e) \wedge \forall w' \in \text{CONT}(e)\text{: } p(w')$

b. $[\![\text{VP}^1]\!]^{w,g} = \lambda e_v . say(e) \wedge \forall w' \in \text{CONT}(e)$: Kibeet steals the money in w'

c. $[\![\text{VoiceP}^1]\!]^{w,g} = \lambda e_v.say(e) \wedge \text{AG}(e) = g(n) \wedge \forall w' \in \text{CONT}(e)$: Kibeet steals the money in w'

Given that the content function is part of the lexical entry of *le*, we derive the special status of *le* clauses in the language. As was pointed out in Sect. 5.2, *le* is the only verb that can introduce an indicative clause. Hence, verbs other than *le* have to combine with *le* to take indicative complements. If it is true that propositional content is necessarily introduced by CONT and that *le* is the only lexeme encoding CONT, we predict that no other matrix verbs can take indicative clauses as complements directly, without also making use of *le*. At the same time, *le* can act as the matrix verb on its own since it additionally encodes a saying event. In the remainder of this section, we lay out how *le* clauses combine with matrix predicates and how our semantics extends from speech reports to attitude reports.

The next two points concern the aspectual information and the implementation of the subjunctive on *le* in (71a). Given that *le* can show aspect morphology, as was shown in (21), we include an AspP layer in (71b). In general, subjunctive Type I connects clauses, as it is also used more widely in coordinate clauses: recall example (7a). With respect to *le* complementation, subjunctive Type I expresses a causal relation between the event introduced by *le* and the event introduced by the matrix predicate. In order to integrate this CAUSE relation, we have to consider the analysis of aspect. Traditionally, aspect is assumed to existentially close off the eventuality argument and introduce a time argument; denotations are given in (73) for perfective and imperfective aspect. In unembedded scenarios, that is, in indicative clauses, the standard account can be adopted. Hence, the denotation in (73a) can be directly taken to be encoded by Asp^2 in (71b).[40]

(73) Aspect (cf. Kratzer 1998; Paslawska and von Stechow 2003)

a. $[\![\text{PFV}]\!] = \lambda P_{\langle \text{v,t} \rangle} \lambda t.\exists e[\tau(e) \subseteq t \wedge P(e)]$ Asp^2 in (71b)

b. $[\![\text{IPFV}]\!] = \lambda P_{\langle \text{v,t} \rangle} \lambda t.\exists e[t \subseteq \tau(e) \wedge P(e)]$

Following Parsons (1990) and Thomason (2014), we take CAUSE to be a relation between eventualities. So in order to let SBJVI encode a causal relation between the matrix event and the saying event, SBJVI has to be able to access the event argument of *le*. This is not provided by the lexical entries in (73). Grano (2021) proposes to define variants of the aspectual morphemes in such cases: see (74), where the eventuality argument is passed up rather than existentially closed off as in (73). We will adopt this idea and take (74a) to be encoded by Asp^1, resulting in the denotation in (75) for AspP^1.

(74) Aspect under subjunctive (Grano 2021)

a. $[\![\text{PFV-S}]\!] = \lambda P_{\langle \text{v,t} \rangle} \lambda e.\lambda t[\tau(e) \subseteq t \wedge P(e)]$ Asp^1 in (71b)

b. $[\![\text{IPFV-S}]\!] = \lambda P_{\langle \text{v,t} \rangle} \lambda e.\lambda t[t \subseteq \tau(e) \wedge P(e)]$

[40] The aspect denotations make use of τ, which, applied to an event, produces the event time (Krifka 1998). The difference between perfective and imperfective is that for the former, the run time of the event is included in the reference time, whereas for the latter, the reference time is included in the run time of the event.

(75) $[\![\text{AspP}^1]\!]^{w,g} = \lambda e \lambda t[\tau(e) \subseteq t \wedge say(e) \wedge \text{AG}(e) = g(n) \wedge \forall w' \in \text{CONT}(e)$: Kibeet steals the money in $w']$

The entry for subjunctive Type I is provided in (76a); it combines the saying events in (75) with the hearing events in (76b), the result of which is the denotation of VP^2, shown in (76c). The lexical entry in (76a) is inspired by Özyıldız et al. (2018), who provide a similar entry for a gerundive affix serving a similar linking function in Turkish complementation.[41]

(76) a. $[\![\text{SBJVI}]\!]^{w,g} = \lambda P_{\langle \text{v}, \langle i,t \rangle \rangle} \lambda Q_{\langle \text{v,t} \rangle} \lambda e''.\exists e' \exists t[\text{CAUSE}(e', e'') \wedge P(e')(t) \wedge Q(e'')]$

b. $[\![kas]\!]^{w,g} = \lambda e_v.hear(e)$

c. $[\![\text{TP}^1]\!]^{w,g}([\![kas]\!]^{w,g}) = [\![\text{VP}^2]\!]^{w,g}$
$= \lambda e''.\exists e' \exists t[\text{CAUSE}(e', e'') \wedge \tau(e') \subseteq t \wedge say(e') \wedge \text{AG}(e') = g(n)$
$\wedge \forall w' \in \text{CONT}(e')$: Kibeet steals the money in $w' \wedge hear(e'')]$

Finally, both the experiencer and the source of the hearing event are added via Event Identification in the matrix clause, resulting in the denotation in (77).

(77) $[\![\text{VoiceP}^2]\!]^{w,g} = \lambda e''.\exists e' \exists t[\text{CAUSE}(e', e'') \wedge \tau(e') \subseteq t \wedge say(e') \wedge \text{AG}(e') = g(n) \wedge \forall w' \in \text{CONT}(e')$: Kibeet steals the money in $w' \wedge hear(e'') \wedge \text{SOURCE}(e'') = kiplangat_n \wedge \text{EXP}(e'') = g(i)]$, defined iff $g(i)$ is addressee[42]

The CAUSE relation is bidirectional, where direction is resolved by context. In (77), the agent of the saying event corefers with the source of the hearing event, indicated by 3SG agreement on *le* (recall that we provide the *kòlé* derivation above). In this case, the CAUSE relation can only be interpreted in a way such that the saying event causes the hearing event to take place. In other words, Kiplangat being the agent of the saying event causes the addressee to enter a hearing event with Kiplangat as the source.

(78) You heard from Kiplangat$_i$ [*pro*$_i$ *kò-lé* [Kibeet stole the money]]
$\rightsquigarrow say(e')$ causes $hear(e'')$ to take place

The reverse relation, however, holds in case the agent of the saying event corefers with the subject of the matrix predicate, that is, if *le* inflects for 2SG; in other words,

[41] The way SBJVI connects the *le* clause with the matrix predicate suggests a promising explanation why negation is unable to appear on embedded *le*; recall the examples in (31). Since negation is standardly taken to be a propositional operator but event semantics introduces sets of events, existential closure of events is often suggested as the bridge between the event domain and the propositional domain (see Penka 2010; Winter and Zwarts 2011 for discussion). In this context, note that the entire *le* clause operates in the event domain; existential closure of the saying event e' in (76c) only comes in via application of SBJVI. At the same time, SBJVI takes the matrix verb as an argument. Hence, there is no position for negation to take scope in between the matrix verb and *le*. We believe that this is the root of the problem of why negation is incompatible with embedded *le* clauses more generally in Kipsigis. The incompatibility with negation is not expected when *le* acts as the matrix predicate, since matrix predicates are tensed and thus allow negation to apply to a propositional argument.

[42] ϕ-features on pronouns denote partial identity functions of type ⟨e, e⟩ (Sauerland 2003, 2008; Heim 2008); for free pronouns the relevant assignment is given by the utterance context.

the *ì:lé* derivation of (71a). In this case, the hearing event causes the saying event to take place, as the addressee is the agent of both the hearing and saying events. This interpretation can be understood more abstractly as representing the addressee's own interpretation of Kiplangat's words.

(79) You$_i$ heard from Kiplangat [*pro*$_i$ *ì:-lé* [Kibeet stole the money]]
$\rightsquigarrow$ $hear(e'')$ causes $say(e')$ to take place

The two readings are also attested for a 'say'-based complementizer combining with reception verbs in Turkish, where Özyıldız et al. (2018) identify the reading in (78) as a speech report and the reading in (79) as an attitude report. The speech report reading represents a saying event that causes the addressee to hear such a speech event. The attitude report reading, however, represents a hearing event that causes an abstract saying event, or as Özyıldız et al. (2018, 302) put it: "the internal event of mentally representing/interpreting" the hearing event. In a sense, the saying event happens internally as the agent conducts an internal monologue caused by the event introduced by the matrix predicate.

Our data support the discussion in Kratzer (2013a), Grano (2016), and Major (2021) in that SAY complementation in Kipsigis is not limited to speech event interpretations but readily allows attitude readings. The eventuality introduced by *le* can encode either a speech event or a mental state, where the latter specifically can occur under nonspeech matrix verbs, that is, in situations that do not involve speaking. Özyıldız et al. (2018) describe the attitude reading as a "mental utterance," and Demirok et al. (2020) call the reading "inner speech," while Major (2021) classifies this reading as stative SAY.[43] Interestingly, for perception verbs such as in (71a), where both the speech report reading and the attitude report reading are possible, one of our consultants consistently mentions a commitment effect for *le* clauses showing agreement with the matrix subject on behalf of the subject's referent, that is, under the attitude report reading. The Kipsigis complementation pattern, thus, aligns with the previous literature, as attitude readings have been reported for covert SAY in English as well as overt SAY in Turkish.

The analysis presented in this section can be extended to a variety of communication and attitude verbs shown to combine with *le* clauses throughout this paper; some are presented in (80). Examples of verbs that appear with *le* clauses in our field notes include *jaːn* 'to believe,' *bwaːt* 'to think/remember,' *taːm* 'to (falsely) accuse,' *ɲoːn* 'to complain,' *naj* 'to realize,' *ruaːtit* 'to dream,' *raːgin* 'to worry,' *pajpaj* 'to be happy,' abd *nɛrɛːtʃ* 'to be angry' (recall Table 1).

(80) a. Kà-∅-tʃáːm Kíbêːt kò-lé kà-∅-tʃáp kímɲéːt.
PST.CURR-3-whisper Kibeet.NOM 3-LE.SBJVI PST.CURR-3-make ugali
'Kibeet whispered that he made ugali.'

[43] Major (2021) makes a distinction between eventive and stative SAY in languages with SAY complementation, but the obligatory presence of agreement on *le* in Kipsigis suggests an eventive SAY syntax in Major's typology. In other words, the syntax of *le* always corresponds to eventive SAY, but its semantics corresponds to either eventive or stative SAY. We therefore chose not to pursue this line of analysis.

b. ɑ̂ː-jáːn-í ɑ̀ː-lé mógôl ŋwɔ̀ŋ.
1SG-believe-IPFV 1SG-LE.SBJVI round earth.NOM
'I believe that the Earth is round.'

c. Àː-pájpáj ɑ̀ː-lé kòː-∅-síːr Kíplàŋgàt.
1SG-happy 1SG-LE.SBJVI PST.REC-3-pass Kiplangat.NOM
'I'm happy that Kiplangat passed (the exams).'

As for communication verbs like *tʃáːm* 'to whisper' (80a) and the doxastic verb *jaːn* 'to believe' (80b), it is reasonable to assume that the eventuality introduced by the matrix verb causes a saying event to take place, either as a speech event or as a mental event. In this sense, the causal direction matches the one in (79). So for example in (80b), entering a believing eventuality leads the speaker to entertain the mental utterance that the Earth is round. On the other hand, the sentence in (80a) conveys the meaning that Kibeet enters a whispering event that causes the speech event whose content is that he made *ugali*. For other verbs, however, the opposite causal direction seems to be suitable: for example, for fiction predicates like *ruaːtit* 'to dream' and emotive factives like *nɛrɛːtʃ* 'to be angry' and *pajpaj* 'to be happy.' The latter is seen in (80c). In these cases, it is plausible to assume that the mental utterance causes the attitude holder to enter a dream event or a state of happiness/sadness.[44]

We speculate that factivity more generally arising with predicates like 'to be happy' in (80c) is a consequence of the directional relation indicated in (78). Recall from examples (62) and (63) in Sect. 5.2 that the matrix verbs in Kipsigis for 'to know' and 'to be happy' trigger factive inferences. These effects follow from the assumption that in such cases the mental 'say' event always causes the state/event denoted by the matrix predicate to take place. Since causation is often assumed to imply causal precedence, our view of factivity aligns with recent accounts for complementation patterns in Buryat (Bondarenko 2020a) and Bangla (Banerjee 2023) in which factivity is analyzed as a preexistence condition rather than a requirement of the complement clause to be true. Note that Bondarenko (2020a) and Banerjee (2023) also pursue a Davidsonian analysis of complementation where attitude verbs simply denote eventualities but their content is introduced by different means. We believe this is not a coincidence. Bochnak and Hanink (2021, 999) in their work on the complementation patterns in Washo point out that, within an eventuality-based account, there is no obvious way to encode a factivity presupposition in the traditional sense since, for example, a matrix verb like 'know' does not serve as the linking element anymore between the attitude holder's doxastic alternatives and the proposition denoted by the complement clause. Hence, there is not one lexical element that could presuppose that the proposition is true in all of the attitude holder's doxastic alternatives. Instead of encoding the factivity trigger lexically, Davidsonian-based accounts

[44] One question we leave open for now is how the semantics of SBJVI extends to the semantics of SBJVII. Recall from Table 1 and the minimal pair in (9) that a set of verbs select for either SBJVI+*le* or SBJVII, with a meaning difference often aligning with a reportative versus directive flavor (*Kibeet whispered that he made* ugali vs. *Kibeet whispered to me to make* ugali). We tentatively proposed that SBJVII mood selection involves an embedded CP, as was shown in (47). On the meaning side, this additional structure presumably corresponds to a more complex semantics, for example, involving an INTEND component and a self-referential *CAUSE relation along the lines of a recent proposal made for intention-based predicates in Grano (2022). We leave the details of this proposal to future work.

rather derive factivity compositionally via the way factive verbs combine with their complement clauses, which is often motivated by the observation that the clauses occur with some form of nominalization (Özyıldız 2017; Bondarenko 2020a; Bochnak and Hanink 2021). The complementation patterns in Kipsigis then pose a challenge, as we do not (at least so far) observe an interaction between a factive reading of the matrix verb and the form of the complement clause. For this reason, we think it is promising to view factivity in Kipsigis as a consequence of the CAUSE semantics introduced by SBJVI—specifically, as the result of the causal direction in which a mental utterance causes the state/event denoted by the matrix predicates to take place, that is, a state of knowing in (62) or a state of happiness in (63). Since the CAUSE relation is bidirectional and a factivity inference is only triggered in one direction, we expect the *le* clauses to come with the same morphosyntax across factive and nonfactive predicates. While this view naturally accounts for the observation that the matrix verb *pwaːt* can mean 'think' or 'remember' when combining with a *le* clause (Table 1), it does raise the question why we do not observe such factivity alternations elsewhere in the language. We want to stress, however, that we have only investigated this topic with a handful of predicates so far. Thus, we do not exclude the possibility that SBJVII could play a role for factivity with at least some of the predicates (see e.g. the perception verbs 'hear' and 'see' in Table 1) or that other properties such as extraction pattern in the predicted ways. We leave this topic to future research.

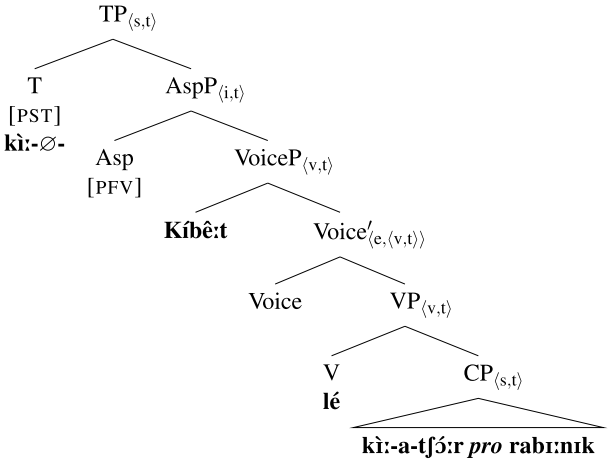
One might wonder at this point if the CAUSE semantics is too weak to capture the restrictions on SAY complementation in Kipsigis. Given the bidirectionality, for example, do we now make the false prediction that *kas* 'hear' and *le* can occur in the opposite order in (71a)? The same question arises for the data in (80). Such word orders are in principle not predicted by our analysis. Patterns of the form [SUBJECT *le mwa/kas/tʃáːm* CLAUSE] cannot be generated given the special status of *le*. Recall from (69) that it is only *le* that introduces the content function and selects for a proposition. Hence, only *le* is able to introduce indicative semantics. As such *le* cannot serve as the second argument of SBJVI (76a), which has to be a predicate of events, thereby ruling out the flipped word orders where *le* precedes the matrix predicate.

Finally, let us briefly address *le* in matrix position. As shown in Sect. 4.1, *le* is able to act as the matrix predicate on its own without the requirement to combine with a *le* clause to introduce indicative complement clauses. We provide another example in (81a), with the underlying syntactic structure in (81b) and the semantic contribution of VoiceP in (81c). Note that the denotation of VoiceP is derived with a lexical entry for *le* that is identical to the one we proposed in (72a). In prose, (81c) denotes a set of saying events whose agent is Kibeet, and in all worlds compatible with the content of such events, the speaker steals the money. Hence, the derivation of (81a) is straightforward under the current account. Since we provide a syntax and semantics of *le* along the lines of a lexical verb, we predict that *le* naturally occurs as such in environments where it acts as the sole embedding predicate introducing the speech event. Also note that *le* is inflected for indicative mood, thereby predicting the absence of CAUSE semantics that is normally associated with the occurrence of subjunctive.

(81) a. [$_{TP}$ Kìː-∅-lé Kíbêːt [$_{CP}$ kìː-á-tʃóːr ràbíːník]].
PST-3-LE(IND) Kibeet.NOM PST-1SG-steal(IND) money

'Kibeet said that I stole the money.'

b.



c. $[\![\text{matrix VoiceP in (81a)}]\!]^{w,g}$
$= \lambda e_v.say(e) \wedge \text{AG}(e) = kibeet \wedge \forall w' \in \text{CONT}(e)$: $g(i)$ steals the money in w'; defined iff $g(i)$ is speaker

This section presented the main components of the semantic analysis, implementing the analysis of prefixal agreement on *le* and deriving its special status in Kipsigis complementation, as pertains to the combinatorial possibilities with other perception, communication, and attitude predicates.

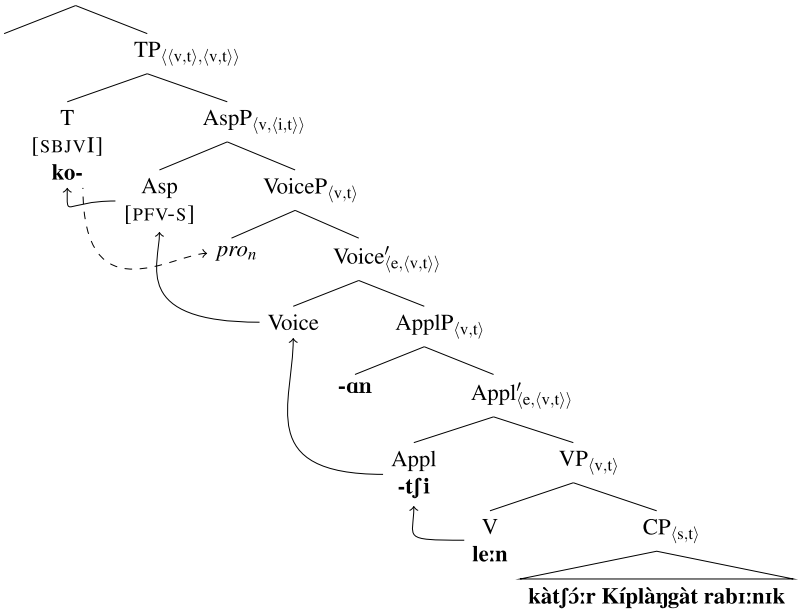
6.3 Extension to suffixal agreement

Let us now turn to suffixal agreement, which was discussed in Sect. 4.1. We repeat example (2) from the introduction in (82), which shows that *le* not only shows prefixal agreement with the matrix subject but also an object clitic introduced by APPL.

(82) Kù-∅-tʃáːm-ú-án Tʃéːbêːt kò-lèːn-tʃ(i)-àn
PST.CURR-3-whisper-APPL-1SG Cheebeet.NOM 3-LE-APPL-1SG.SBJVI
kà-∅-tʃɔ́ːr Kíplàŋgàt ràbíːník.
PST.CURR-steal Kiplangat.NOM money
'Cheebeet whispered to me that Kiplangat stole the money.'

The occurrence of suffixal agreement is predicted under an account that treats *le* as a verb. In such cases, *le* introduces an applied argument in addition to a subject, as shown in the partial derivation in (83).

(83)



We provide the denotation of matrix VoiceP in (84). Suffixal agreement, decomposed into APPL and a 1SG object clitic, introduces a goal argument for the embedded saying event, matching the goal argument of the matrix whispering event. As in (80a), the subject of *le* corefers with the matrix subject, and the sentence receives a reading in which the whispering event causes the saying event to take place.

(84) $[\![\text{matrix VoiceP in (83)}]\!]^{w,g}$
$= \lambda e''.\exists e'\exists t[\text{CAUSE}(e', e'') \wedge \tau(e') \subseteq t \wedge say(e') \wedge \text{AG}(e') = g(n) \wedge$
$\text{GOAL}(e') = g(i) \wedge \forall w' \in \text{CONT}(e')$: Kiplangat steals the money in
$w' \wedge whisper(e'') \wedge \text{GOAL}(e'') = g(i) \wedge \text{AG}(e'') = cheebeet_n]$; defined
iff $g(i)$ is speaker

Supportive evidence for our analysis comes from the fact that for some matrix verbs some speakers allow applied arguments on *le* only, without the need for an applied object in the matrix clause. Examples of such verbs are *ɲo:ɲ* 'complain' and *si:r* 'write':

(85) Kò:-á-ɲó:ɲ ù:-lé:n-**tʃí** Kìbê:t kò:-jû:tʃ-è:n
PST.REC-1SG-complain 1SG-LE-APPL.SBJVI Kibeet PST.REC-bad-PL
ùmìtwú:gík.
food.NOM
'I complained to Kibeet that the food was bad.'

(86) Kò:-á-sí:r ù:-lé:n-**tʃí** Tʃè:bê:t à-tʃèlɛ́wànì.
PST.REC-1SG-write 1SG-LE-APPL.SBJVI Cheebeet 1SG-be.late
'I wrote to Cheebeet that I will be late.'

This section concludes our syntactic and semantic account of *le* as a clausal embedder under attitude predicates.

7 Conclusion

In this paper, we have argued that the Kipsigis "complementizer" is in fact a verb and that C-Agree is verbal agreement with a locally introduced subject, which is in most cases a covert pronoun. Our analysis resolves the problems for locality and directionality of Agree posed by the upwards-oriented C agreement pattern. Whereas previous analyses have argued for the presence of a C head or a hybrid status of 'say'-based complementizers functioning as an element of category V or C depending on context, we assign the "complementizer" *le* in Kipsigis the category V throughout all complementation occurrences in the language. If this line of approach is feasible in more languages with 'say'-based complementation (e.g. Koopman 1984; Koopman and Sportiche 1989; Major 2021), it could indicate that some instances of what has been described as C-Agree may instantiate standard verbal agreement instead. This is significant because many reported cases of upwards-oriented complementizer agreement involve 'say'-based complementizers, not noun-y complementizers of the Indo-European type.[45] This observation has broader implications for theories of agreement, since it calls into question the existence of genuine agreement between an element of category C and a matrix subject. Similarly, for the Germanic C agreement pattern, alternative analyses not employing C-Agree have been proposed, arguing for allomorphy (Weisser 2019) or clitic doubling (van Alem 2023a,b) instead.

While we present a clear-cut case, in Kipsigis, of a syntactic reanalysis of a 'say'-based complementizer as a lexical verb, other languages with 'say'-based complementation might not warrant such an analysis. Specifically in languages where the morpheme in question seems to be polyfunctional beyond the contexts of speech verbs and complementation, neither C nor V are suitable to capture the distribution. Several examples can be found in Güldemann's (2008) discussion of quotative verbs found in African languages. For instance, a recent case study by Kiemtoré (2023) of 'say'-based complementation in the West African language Jula reveals several more functions, including similative, desiderative, and naming constructions, which arguably require a broader syntactic category.

Finally, our analysis provides support for recent accounts of complementation phenomena within a Neo-Davidsonian framework (Hacquard 2006, 2010; Kratzer 2006, 2013a; Grano 2016, 2022). Such a framework makes it possible for the syntactic analysis of *le* as a verb to be reflected in the semantic composition. More importantly, it lets us derive the unique status of *le* with a denotation that combines the semantics of an eventuality with the content function. This lexical entry is distinct from all other communication and attitude predicates, which simply denote an eventuality, thus requiring them to combine with *le* to be able to embed a proposition. The fact that *le*

[45] According to a survey of Bantu complementizers in Gluckman (2023), this type of agreement is also possible for complementizers derived from manner deictics and pronouns; complementizers derived from demonstratives (as in English) or the copula, on the other hand, never exhibit agreement. Investigating the category of manner deictic and pronominal complementizers is an interesting topic for further research.

takes over the function of a lexical verb and a clausal embedder might be the key to deriving the trademark characteristic of 'say'-based "complementizers" more generally, as the denotation allows them to occur naturally either as the main predicate in sentences with clausal complementation or as the embedding predicate in combination with another matrix verb.

Appendix: Subjunctive

Various uses of the two types of subjunctive have been mentioned throughout the paper. In this appendix, we briefly summarize all environments known to us where Type I and II subjunctive are licensed in Kipsigis. Beyond its use in verbal complementation (see discussion in Sect. 3), Type II subjunctive is also used for purpose clauses and after modals.

(87) Purpose clauses (Toweett 1979, 199)

a. (à)sí à-pîr
so 1SG-hit.SBJVII
'so that I hit'

b. (à)sí à:-kát
so 1SG-greet.SBJVII
'so that I greet'

(88) Modals (Toweett 1979, 225)

a. ∅-ɲɑ:l-u à-pîr …
3-must-IPFV 1SG-hit.SBJVII
'I must hit …'

b. Mje à-pîr …
good 1SG-hit.SBJVII
'It is good that I hit …'

As for Type I subjunctive (the form that *le* has when used in complementation), its use is more restricted. As shown in (89), it is the form of the verb used in conditional clauses. It is also found with some temporal adjunct clauses, of the type illustrated in (90).

(89) iŋgot à:-rú
if 1SG-sleep.SBJVI
'if I sleep'

(90) Kò:-∅-tʃáp-é kímɲé:t Tʃé:bê:t ko:n ɑ:-ɲo:
PST.REC-3-make-IPFV ugali Cheebeet.NOM when 1SG-come.SBJVI
ká:t.
house
'Cheebeet was making ugali when I entered the house.'

The other prominent use of Type I subjunctive is in coordination (which was discussed extensively in Sect. 5.2). As shown in (91), if two clauses are coordinated

with the marker *ak* 'and' in Kipsigis, the second conjunct must be in subjunctive Type I. Interestingly, the coordinator *ak* is optional.

(91) Kà-∅-púːtʃ Tʃéːbêːt káːt (ák) àː-tʃáp tʃàːíːk.
PST.CURR-3-sweep Cheebeet house and 1SG-make.SBJVI tea
'Cheebeet swept the house and I made tea.'

Looking at the environments in which the two types of subjunctive are used, it seems that subjunctive Type II has many similarities to the subjunctive of European languages, especially those Balkan languages that lack infinitives. Subjunctive Type I, on the other hand, does not have a clear parallel, except perhaps for consecutive tenses of some East African languages (we thank an anonymous reviewer for making this connection). We leave a complete investigation of verbal mood in Kipsigis as a topic for further research.

Acknowledgements We are grateful to Enock Kirui, Wesley Kirui, Hillary Mosonik, Victor Mutai, Vincent Mutai, Philemon Ronoh, and Nathan Rotich for their valuable work as linguistic consultants. We'd also like to thank Mark Baker, Doreen Georgi, Travis Major, Ken Safir, Malte Zimmermann, three anonymous reviewers, and audiences at Universität Potsdam, NYU, Universität Leipzig, Humboldt-Universität zu Berlin, USC, Rutgers University, UQAM, BCGL 13, the DGfS 43 workshop "On the nouniness of propositional arguments," WCCFL 39, and GLOW 44 for useful feedback on earlier stages of this work.

Funding This research was partially funded by the DFG Research Training Group GRK 2011 Interaction of Grammatical Building Blocks (IGRA) and DFG grant 439622645 ("Layers of morphosyntactic number in Eastern Sudanic").

Declarations

Competing Interests There are no conflicts of interest for the authors.

References

Alexiadou, Artemis, Elena Anagnostopoulou, Gianina Iordachioaia, and Mihaela Marchis. 2012. In support of Long Distance Agree. In *Local modelling of non-local dependencies in syntax*, eds. Tibor Kiss and Gereon Müller, 55–81. Berlin: de Gruyter.

Anand, Pranav. 2006. De de se. PhD diss., Massachusetts Institute of Technology.

Anand, Pranav, and Valentine Hacquard. 2008. Epistemics with attitude. In *Proceedings of SALT XVIII*, eds. Tova Friedman and Satoshi Ito, 37–54. Ithaca, NY: Cornell University.

Anand, Pranav, and Feng-fan Hsieh. 2005. Long-distance reflexives in perspective. In *Proceedings of the 24th West Coast Conference on Formal Linguistics*, eds. John Alderete, Chung-hye Han, and Alexei Kochetov, 43–51. Somerville: Cascadilla Proceedings Project.

Anand, Pranav, Jane Grimshaw, and Valentine Hacquard. 2019. Speech act predicates and their subjects. In *Tokens of meaning: A festschrift for Lauri Kartunnen*. Chicago: University of Chicago Press.

Arregi, Karlos, and Andrew Nevins. 2012. *Morphotactics: Basque auxiliaries and the structure of spellout*. Dordrecht: Springer.

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

Baker, Mark. 2022. Upward complementizer agreement. Chapter 2 of book-length ms. "Complementizers relating to noun phrases: Rare constructions within a theory of universal grammar," Rutgers University. https://sites.rutgers.edu/mark-baker/research/.

Baković, Eric. 2000. Harmony, dominance, and control. PhD diss., Rutgers University.

Balusu, Rahul. 2020. The quotative complementizer says "I'm too baroque for *that*." In *Proceedings of 8th Formal Approaches to South Asian Languages*, eds. Bhamati Dash and Gurmeet Kaur, 1–12.

Banerjee, Arka. 2023. Factive complements are not always unique entities: A case study with Bangla 'remember.' In *Proceedings of WoLLIC 29*, eds. Helle Hvid Hansen, Andre Scedrov, and Ruy J. G. B. de Queiroz, 83–100. Cham, Switzerland: Springer.

Bjorkman, Bronwyn M., and Hedde Zeijlstra. 2019. Checking up on (ϕ-)agree. *Linguistic Inquiry* 50: 527–569.

Bochnak, M. Ryan, and Emily A. Hanink. 2021. Clausal embedding in Washo: Complementation vs. modification. *Natural Language and Linguistic Theory* 40: 979–1022. https://doi.org/10.1007/s11049-021-09532-z.

Bogal-Allbritten, Elizabeth. 2015. Decomposing attitudes: The view from Navajo. Paper presented at the 89th meeting of the Linguistic Society of America.

Bondarenko, Tatiana. 2020a. Factivity from pre-existence: Evidence from Barguzin Buryat. *Glossa: A Journal of General Linguistics* 5: 1–35. https://doi.org/10.5334/gjgl.1196.

Bondarenko, Tatiana. 2020b. Hyperraising and Logical Form: Evidence from Buryat. Poster presented at the 43rd Generative Linguistics in the Old World conference. http://tbond.scripts.mit.edu/tb/wp-content/uploads/2020/03/hyperraising_handout.pdf.

Bossi, Madeline. 2023a. Epistemic modality across syntactic categories in Kipsigis. PhD diss., University of California, Berkeley.

Bossi, Madeline. 2023b. Two types of 'say'-complementation in Kipsigis. Handout from talk given at the 41st West Coast Conference on Formal Linguistics, University of California, Santa Cruz.

Bossi, Madeline, and Michael Diercks. 2019. V1 in Kipsigis: Head movement and discourse-based scrambling. *Glossa: A Journal of General Linguistics* 4: 65.

Bruening, Benjamin. 2002. Raising to object and proper movement. Ms., University of Delaware.

Burukina, Irina. 2023. External merge in spec,CP: Complementizers projecting an argument. *Syntax* 26: 85–105.

Carstens, Vicki. 2003. Rethinking complementizer agreement: Agree with a case-checked goal. *Linguistic Inquiry* 34: 393–412.

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

Charnavel, Isabelle. 2020. Logophoricity and locality: A view from French anaphors. *Linguistic Inquiry* 51: 671–723.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, eds. Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge: MIT Press.

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

Creider, Chet, and Jane Tapsubei Creider. 1989. *A grammar of Nandi*. Hamburg: Helmut Buske.

Culy, Christopher. 1994. Aspects of logophoric marking. *Linguistics* 32: 1055–1094.

Davidson, Donald. 1967. The logical form of action sentences. In *The logic of decision and action*, ed. Nicholas Resher, 81–95. Pittsburgh: University of Pittsburgh Press.

Demirok, Ömer, Deniz Özyıldız, and Balkız Öztürk. 2020. Complementizers with attitude. In *Proceedings of the North East Linguistic Society 49*, eds. Maggie Baird and Jonathan Pesetsky. Amherst: Graduate Linguistics Student Association.

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

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

Diercks, Michael, Marjo van Koppen, and Michael Putnam. 2020. Agree probes down: Anaphoric feature valuation and phase reference. In *Agree to agree: Agreement in the minimalist programme*, eds. Peter W. Smith, Johannes Mursell, and Katharina Hartmann, 347–389. Berlin: Language Science Press.

Driemel, Imke, and Maria Kouneli. 2021. ⟨v, t⟩-type complementation in Kipsigis. Ms. In *Proceedings of the 39th West Coast Conference on Formal Linguistics*. To appear. https://ling.auf.net/lingbuzz/006045.

Driemel, Imke, and Maria Kouneli. 2022. A dedicated topic position in Kipsigis. In *Proceedings of TripleA 8: Fieldwork perspectives on the semantics of understudied languages of Africa, Asia, Australia, and Oceania*, ed. Henrison Hsieh, 1–20. Singapore: ScholarBank@NUS.

Duncan, Philip T., and Harold Torrence. 2017. Verbal complementizers and the Indirect Agree relation in Ibibio. Handout from talk given at the 91st meeting of the Linguistic Society of America.

Eberhard, David M., Gary F. Simons, and Charles D. Fennig, eds. 2020. *Ethnologue: Languages of the world*, 23rd edn. Dallas: SIL International.

Elliott, Patrick D. 2016. Explaining DPs vs. CPs without syntax. In *Proceedings of the 52nd annual meeting of the Chicago Linguistic Society*, 171–185. Chicago: Chicago Linguistic Society.

Elliott, Patrick D. 2017. Elements of clausal embedding. PhD diss., University College London. Revised January 2020.

Fuß, Eric. 2008. Multiple agreement and the representation of inflection in the C-domain. *Linguistische Berichte* 213: 78–108.

Fuß, Eric. 2014. Complementizer agreement (in Bavarian): Feature inheritance or feature insertion? In *Bavarian syntax: Contributions to the theory of syntax*, eds. Günther Grewendorf and Helmut Weiß, 51–82. Amsterdam: Benjamins.

Gluckman, John. 2022. Perspectival domains in nouns and clauses. *Natural Language and Linguistic Theory* 41: 159–205. https://doi.org/10.1007/s11049-022-09544-3.

Gluckman, John. 2023. Towards a typology of Bantu complementizers. Handout from talk given at the 54th Annual Conference on African Linguistics, University of Connecticut.

Grano, Thomas. 2016. Semantic consequences of syntactic subject licensing: Aspectual predicates and concealed modality. In *Proceedings of Sinn und Bedeutung 20*, eds. Nadine Bade, Polina Berezovskaya, and Anthea Schöller, 306–322. Tübingen: University of Tübingen.

Grano, Thomas. 2021. Intention reports and eventuality abstraction in a theory of mood choice. Slides for talk given at Semantics and Linguistic Theory 31. https://www.dropbox.com/s/yf9ggqvd6vmwq4k/Grano_poster_SALT31.pdf.

Grano, Thomas. 2022. Intention reports and eventuality abstraction in a theory of mood choice. Ms., Indiana University. https://www.dropbox.com/s/brxs5gpyvib9ltr/intend_mood_distr_f22.pdf.

Grimshaw, Jane. 2015. The light verbs *say* and SAY. In *Structures in the mind*, eds. Ida Toivonen, Piroska Csúri, and Emile van der Zee, 79–99. Cambridge: MIT Press.

Güldemann, Tom. 2008. *Quotative indexes in African languages: A synchronic and diachronic survey*. Berlin: de Gruyter Mouton.

Hacquard, Valentine. 2006. Aspects of modality. PhD diss., Massachusetts Institute of Technology.

Hacquard, Valentine. 2010. On the event relativity of modal auxiliaries. *Natural Language Semantics* 18: 79–114.

Haegeman, Liliane, and Marjo van Koppen. 2012. Complementizer agreement and the relation between C^0 and T^0. *Linguistic Inquiry* 43: 441–454.

Hall, Beatrice, R. M. R. Hall, Martin D. Pam, Amy Myers, Stephen A. Antell, and Godfrey K. Cherono. 1974. African vowel harmony systems from the vantage point of Kalenjin. *Afrika und Übersee* 57: 241–267.

Halle, Morris, and Jean-Roger Vergnaud. 1981. Harmony processes. In *Crossing the boundaries in linguistics: Studies presented to Manfred Bierwisch*, 1–22. Dordrecht: Springer.

Halpert, Claire. 2019. How to be an embedded clause: Say complementizers in Bantu. In *Proceedings of TripleA 5*, eds. M. Ryan Bochnak, Miriam Butt, Erlinde Meertens, and Mark-Matthias Zymla, 31–46. Tübingen: University of Tübingen.

Heim, Irene. 2008. Features on bound pronouns. In *Phi-Theory: Phi-features across modules and interfaces*, eds. Daniel Harbour, David Adger, and Susana Béjar, 35–56. Oxford: Oxford University Press.

Hintikka, Jaakko. 1969. Semantics for propositional attitudes. In *Philosophical logic*, eds. J. W. Davis, D. J. Hockney, and W. K. Wilson, 21–45. Dordrecht: Reidel.

Iatridou, Sabine. 1988. On nominative case assignment and a few related things. In *Papers on case and agreement II*, 175–196. Cambridge: MIT Press.

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

Kiemtoré, Alassane. 2023. Issues in Jula complementation: Structures, relations and matters of interpretation. PhD diss., Stuttgart University.
Kinyalolo, Kasangati Kikuni W. 1993. On some syntactic properties of ɖɔ. *Lingua* 91: 201–233.
Knyazev, Mikhail. 2016. Complementizers in Kalmyk. In *Complementizer semantics in European languages*, eds. Kasper Boye and Petar Kehayov, 655–679. Berlin: Mouton de Gruyter.
Koopman, Hilda. 1984. *The syntax of verbs*. Dordrecht: Foris.
Koopman, Hilda, and Dominique Sportiche. 1989. Pronouns, logical variables, and logophoricity in Abe. *Linguistic Inquiry* 20: 555–588.
Kouneli, Maria. 2019. The syntax of number and modification: An investigation of the Kipsigis DP. PhD diss., New York University.
Kouneli, Maria. 2022. Inflectional classes in Kipsigis. *Glossa: A Journal of General Linguistics* 47(1).
Kouneli, Maria, and Yining Nie. 2021. Across-the-board tonal polarity in Kipsigis: Implications for the morphology-phonology interface. *Language* 97(2): 111–138.
Krapova, Iliyana. 2001. Subjunctives in Bulgarian and Modern Greek. In *Comparative syntax of Balkan languages*, eds. María Luisa Rivero and Angela Ralli, 105–126. Oxford: Oxford University Press.
Kratzer, Angelika. 1996. Severing the external argument from its verb. In *Phrase structure and the lexicon*, eds. Johan Rooryck and Laurie Zaring, 109–137. Dordrecht: Kluwer.
Kratzer, Angelika. 1998. More structural analogies between pronouns and tenses. In *Proceedings of SALT VIII*, eds. Devon Strolovitch and Aaron Lawson, 92–110. Ithaca, NY: Cornell University.
Kratzer, Angelika. 2006. Decomposing attitude verbs. Handout from talk in honor of Anita Mittwoch on her 80th birthday.
Kratzer, Angelika. 2013a. Modality and the semantics of embedding. Slides for talk given at Amsterdam Colloquium, December 20.
Kratzer, Angelika. 2013b. Modality for the 21st century. In *L'interface langage-cognition/The language-cognition interface: Actes du 19e Congrès International des Linguistes Genève*, eds. Stephen R. Anderson, Jacques Moeschler, and Fabienne Reboul, 179–199. Librairie Droz.
Kratzer, Angelika. 2016. Evidential moods in attitude and speech reports. Slides for talk. https://works.bepress.com/angelika_kratzer/10/.
Krifka, Manfred. 1998. The origins of telicity. In *Events and grammar*, ed. Susan Rothstein, 197–235. Dordrecht: Kluwer.
Landau, Idan. 2004. The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory* 22: 811–877.
Landau, Idan. 2013. *Control in generative grammar*. Cambridge: Cambridge University Press.
Letsholo, Rose, and Ken Safir. 2019. Complement clause C-agreement beyond subject phi-agreement in Ikalanga. In *African linguistics across the disciplines: Selected papers from the 48th Annual Conference on African Linguistics*, eds. Samson Lotven, Silvina Bongiovanni, Phillip Weirich, Robert Botne, and Samuel Gyasi Obeng, 159–186. Berlin: Language Science Press.
Lord, Carol. 1976. Evidence for syntactic reanalysis: From verb to complementizer in Kwa. In *Papers from the parasession on diachronic cyntax (12th regional meeting), Chicago Linguistic Society*, eds. Sanford Steever, Carol Walker, and Salikoko Mufwene, 179–191. Chicago: University of Chicago.
Major, Travis. 2021. On the nature of "say" complementation. PhD diss., University of California, Los Angeles.
Major, Travis. 2024. Re-analyzing 'say' complementation: Implications for case theory and beyond. *Natural Language and Linguistic Theory* 42: 1125–1190. https://doi.org/10.1007/s11049-023-09594-1.
Major, Travis, and Harold Torrence. 2021. "Say"-chains, not "say"-complementation. In *Proceedings of the 38th West Coast Conference on Formal Linguistics*, eds. Rachel Soo, Una Y. Chow, and Sander Nederveen, 283–293. Somerville: Cascadilla Proceedings Project.
Major, Travis, Justine Sikuku, and Michael Diercks. 2022. "Say"-complementation: Reanalyzing agreeing complementizers as verbs. Ms., University of Southern California, Moi University, and Pomona College.
Marantz, Alec. 1993. Implications of asymmetries in double object constructions. In *Theoretical aspects of Bantu grammar*, ed. Sam A. Mchombo, 113–150. Stanford: CSLI Publications.
Marušič, Franc, Andrew Nevins, and Bill Badecker. 2015. The grammars of conjunction agreement in Slovenian. *Syntax* 18: 39–77.
McFadden, Thomas, and Sandhya Sundaresan. 2021. Unifying species of C-agreement. In *Proceedings of the fifty-first annual meeting of the North East Linguistic Society*, eds. Alessa Farinella and Angelica Hill, 91–100. Amherst: Graduate Linguistics Student Association.
Mietzner, Angelika. 2009. *Räumliche Orientierung in nilotischen Sprachen*. Vol. 24 of *Nilo-Saharan: Linguistic analyses and documentation*. Cologne: Rüdiger Köppe.

Moulton, Keir. 2009. Natural selection and the syntax of clausal complementation. PhD diss., University of Massachusetts Amherst.
Moulton, Keir. 2019. (Non)-complement clauses and in-situ saturation: Consequences for cross-clausal A-dependencies. Slides for talk given at Generative Linguistics in the Old World in Asia 12. https://glowlinguistics.org/asia12/workshop-talk-by-keir-moulton/.
Murugesan, Gurujegan. 2022. Deriving the anaphor–agreement effect and the violations of it. *Syntax* 25: 39–83.
Nevins, Andrew. 2010. *Locality in vowel harmony*. Cambridge: MIT Press.
Nformi, Jude. 2017. Complementizer agreement and intervention effects. Ms., Leipzig University.
Özyıldız, Deniz. 2017. Attitude reports with and without true belief. In *Proceedings of SALT 27*, eds. Dan Burgdorf, Jacob Collard, Sireemas Maspong, and Brynhildur Stefánsdóttir, 397–417.
Özyıldız, Deniz, Travis Major, and Emar Maier. 2018. Communicative reception reports as hearsay: Evidence from indexical shift in Turkish. In *Proceedings of the 36th West Coast Conference on Formal Linguistics*, eds. Richard Stockwell, Maura O'Leary, Zhongshi Xu, and Z. L. Zhou, 296–305. Somerville: Cascadilla Proceedings Project.
Parsons, Terence. 1990. *Events in the semantics of English*. Cambridge: MIT Press.
Paslawska, Alla, and Arnim von Stechow. 2003. Perfect readings in Russian. In *Perfect explorations*, eds. Artemis Alexiadou, Monika Rathert, and Arnim von Stechow, 307–362. Berlin: Mouton de Gruyter.
Penka, Doris. 2010. *Negative indefinites*. Oxford: Oxford University Press.
Pietraszko, Joanna. 2017. Inflectional dependencies: A study of complex verbal expressions in Ndebele. PhD diss., University of Chicago.
Pietraszko, Joanna. 2020. An argument for true c-selection in clausal complementation. Handout from talk given at the Brussels Conference on Generative Linguistics 13.
Portner, Paul, and Aynat Rubinstein. 2020. Desire, belief, and semantic composition: Variation in mood selection with desire predicates. *Natural Language Semantics* 28: 343–393.
Pylkkänen, Liina. 2008. *Introducing arguments*. Cambridge: MIT Press.
Rivero, María Luisa. 2004. Spanish quirky subjects, person restrictions, and the Person-Case Constraint. *Linguistic Inquiry* 35: 494–502.
Rizzi, Luigi. 1989. On the anaphor-agreement effect. *Rivista Di Linguistica* 2: 27–42.
Rottland, Franz. 1982. *Die Südnilotischen Sprachen: Beschreibung, Vergleichung und Rekonstruktion*. Vol. 7 of *Kölner Beiträge zur Afrikanistik*. Berlin: Dietrich Reimer.
Roussou, Anna. 2009. In the mood for control. *Lingua* 119: 1811–1836.
Safir, Ken. 2004. Person, context and perspective. *Rivista Di Linguistica* 16: 107–153.
Sauerland, Uli. 2003. A new semantics for number. In *Proceedings of SALT 13*, eds. Robert B. Young and Yuping Zhou, 258–275. Ithaca, NY: CLC Publications.
Sauerland, Uli. 2008. On the semantic markedness of phi-features. In *Phi-Theory: Phi-features across modules and interfaces*, eds. Daniel Harbour, David Adger, and Susana Béjar, 35–56. Oxford: Oxford University Press.
Sauerland, Uli, Bart Hollebrandse, and Frantisek Kratochvíl. 2020. When hypotaxis looks like parataxis: Embedding and complementizer agreement in Teiwa. *Glossa: A Journal of General Linguistics* 5: 89.
Shlonsky, Ur. 1994. Agreement in Comp. *The Linguistic Review* 11: 351–375.
Speas, Margaret. 2004. Evidentiality, logophoricity and the syntactic representation of pragmatic features. *Lingua* 114: 255–276.
Sundaresan, Sandhya. 2016. Anaphora vs. agreement: A new kind of Anaphor Agreement Effect in Tamil. In *The impact of pronominal form on interpretation*, eds. Patrick Grosz and Pritty Patel-Grosz, 77–105. Berlin: de Gruyter.
Terzi, Arhonto. 1992. PRO in finite clauses: A study of the inflectional heads of the Balkan languages. PhD diss., City University of New York.
Thomason, Richmond. 2014. Formal semantics for causal constructions. In *Causation in grammatical structures*, eds. Bridget Copley and Fabienne Martin, 58–75. Oxford: Oxford University Press.
Toweett, Taaitta. 1979. *A study of Kalenjin linguistics*. Nairobi: Kenya Literature Bureau.
Truswell, Robert. 2007. Extraction from adjuncts and the structure of events. *Lingua* 117: 1355–1377.
van Alem, Astrid. 2023a. Complementizer agreement is clitic doubling: Evidence from intervention effects in Frisian and Limburgian. Ms., Potsdam University.
van Alem, Astrid. 2023b. Life of Phi: Phi-features in West Germanic and the syntax-morphology interface. PhD diss., Leiden University.
van Koppen, Marjo. 2005. *One probe—two goals: Aspects of agreement in Dutch dialects*. Utrecht: Landelijke Onderzoekschool Taalwetenschap/National Research School of Linguistics.

van Koppen, Marjo. 2012. The distribution of phi-features in pronouns. *Natural Language and Linguistic Theory* 30: 135–177.
van Urk, Coppe. 2015. A uniform syntax for phrasal movement: A case study of Dinka Bor. PhD diss., Massachusetts Institute of Technology.
Varlokosta, Spyridoula. 1993. Control in Modern Greek. In *University of Maryland working papers in linguistics*, eds. Carol A. Mason, Susan M. Powers, and Cristina Schmitt, Vol. 1, 144–163.
Weisser, Philipp. 2019. Telling allomorphy from agreement. *Glossa: A Journal of General Linguistics* 4: 86.
Winter, Yoad, and Joost Zwarts. 2011. Event semantics and abstract categorial grammar. In *The mathematics of language: 12th biennial conference*, eds. Makoto Kanazawa, András Kornai, Marcus Kracht, and Hiroyuki Seki, 174–191. Berlin: Springer.
Woolford, Ellen. 1999. More on the anaphor agreement effect. *Linguistic Inquiry* 30: 257–287.
Wurmbrand, Susi. 2001. *Infinitives: Restructuring and clause structure*. Berlin: Mouton de Gruyter.
Zwart, Jan-Wouter. 1997. *Morphosyntax of verb movement: A minimalist approach to the syntax of Dutch*. Dordrecht: Kluwer.
Zwarts, Joost. 2004. *The phonology of Endo: A southern Nilotic language*. Munich: LINCOM.