

C-Agree is local subject-verb agreement in Kipsigis*

Imke Driemel & Maria Kouneli

August 21, 2023

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 & 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, and 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).

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 & Torrence (2017) on *Ibibio*, Nformi (2017) on *Limbum*, Letsholo & Safir (2019) on *Ikalanga*.¹ This is different from 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 & van Koppen 2012). While the Germanic pattern does not pose serious problems for standard approaches to agreement using Downward Agree (e.g.

*We are grateful to Donald Kibet, 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 Mike Diercks, Doreen Georgi, Travis Major, Deniz Özyıldız, Malte Zimmermann, and the audiences at the University of Potsdam, NYU, Universität Leipzig, the *Brussels Conference on Generative Linguistics 13*, WCCFL 39, GLOW 44, and DGfS 43 for useful feedback.

¹Outside 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).

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 & Zeijlstra 2019 a.o.) is necessary for the analysis of the pattern (e.g. Nformi 2017, Letsholo & Safir 2019, McFadden & 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 which has been reported to display an upwards-oriented agreement pattern between an embedded ‘say’-based complementizer, glossed transparently as LE, and the matrix subject (Diercks & Rao 2019, Diercks et al. 2020), illustrated in (1).^{2,3}

- (1) a. \hat{a} :ŋgén **â**:lé \emptyset -rú-è Kíbê:t.
 1SG-know 1SG-LE 3-sleep-IPFV Kibeet.NOM
 ‘I know that Kibeet is sleeping.’
 b. \hat{i} :ŋgén **i**:lè \emptyset -rú-è Kíbê:t.
 2SG-know 2SG-LE 3-sleep-IPFV Kibeet.NOM
 ‘You know that Kibeet is sleeping.’
 c. í-ŋgèn Kíplàngàt **kò**-lé \emptyset -rú-è Kíbê:t.
 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, as shown in (2).

- (2) Kà- \emptyset -tʃú:m-ú-án Tʃé:bê:t **kò-lè:n-tʃ(i)**-**àn**
 PST.CURR-3-whisper-APPL-1SG Chebeet.NOM 3-LE-APPL-1SG
 kà- \emptyset -tʃú:r Kíbê:t ràbí:ník.
 PST.CURR-steal Kibeet.NOM money
 ‘Chebeet whispered to me that Kibeet 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’, and not a complementizer (see also Koopman & Sportiche 1989, Özyıldız et al. 2018,

²Kipsigis is the major variety of Kalenjin, a cluster of dialects of the Southern Nilotic branch of Nilo-Saharan, and it is spoken by about 2 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 2020-2023 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 consulted for 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.

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

Major & Torrence 2021, Major 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 & Rao 2019), with the pattern being 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).

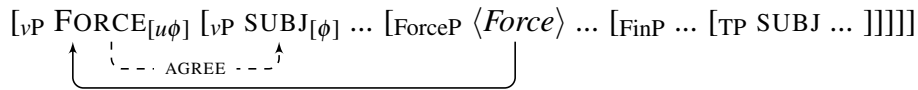
The remainder of the paper is structured as follows: in Section 2, we provide an overview of previous theories of upwards-oriented complementizer agreement; in Section 3, we present some background on complementation in Kipsigis and we provide a description of the pattern of upwards-oriented complementizer agreement in the language. We then develop our analysis in three steps: in Section 4 we argue that the Kipsigis “complementizer” is the lexical verb ‘say’ which agrees with a local subject, in Section 5 we motivate the syntactic structure that we assume for complementation, and in Section 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 Section 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 & 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 & Safir 2019, McFadden & 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 & 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 & Sundaresan 2021) and cross-linguistically determined parameter settings (Baker 2008).

- (3) *Upward Agree account*
- $$\begin{array}{c}
 [{}_{\nu P} \text{SUBJ}_{[\phi]} \dots [{}_{\text{ForceP}} \text{FORCE}_{[u\phi]} \dots [{}_{\text{FinP}} \dots [{}_{\text{TP}} \text{SUBJ} \dots]]] \\
 \uparrow \text{----- AGREE -----} \downarrow
 \end{array}$$

(4) *Downward Agree account*

For Lubukusu, Carstens (2016) proposes a Downward Agree approach between the moved complementizer and the matrix subject, where the FORCE head carries $u\phi$ -features and moves into the matrix clause, where it adjoins to the νP from which it can undergo Agree with the ϕ -features of the matrix subject. Diercks & Rao (2019) and Diercks et al. (2020), in their analyses of Kipsigis and Lubukusu, adopt similar mechanics to 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 Carsten’s account based on the fact that Lubukusu allows for 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 data point to the absence of a phase boundary, while Carsten’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 which 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 which are triggered for at least certain matrix verbs in Limbum in such cases. As Diercks & 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 & Rao (2019) 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 times.

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, which is subsequently bound by the matrix subject; see also Duncan & Torrence (2017) for a similar C-Agree analysis in Ibibio. 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 Kisigis, however, is not captured by this analysis since object-oriented anaphors are generally not found across languages. Baker (2022) discusses the Kisigis pattern as one of the motivations to analyze the silent element in Spec,CP as a close relative to PRO rather than an anaphor, which is in turn in a control relation with

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 the matrix subject. More generally, we will provide several arguments in favor analyzing the C-like element as a verbal category, questioning C-raising accounts which have been proposed for Kipsigis (Diercks & 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 & Diercks 2019) and the typologically rare marked nominative system (Kouneli 2019, Kouneli & 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 & 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 (6-a), but a long-voweled prefix in its subjunctive (embedded) form in (6-b).⁵ For 3rd person subjects, the prefix is \emptyset in most cells of the paradigm, while it is always *ko(:)-* in the subjunctive.⁶

- (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 1st 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

⁴While 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 *governed verb form* and *abhängige Verbform* (dependent verb form), respectively. We adopt the term ‘subjunctive’ used in the description of Nandi and Kipsigis inflection in Creider & Creider (1989).

⁵In the remainder of the paper, subjunctive inflection will always be indicated in the glosses, while indicative will be left un glossed.

⁶The 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 (Class I or II) a given verb belongs to. The examples given in this section (including *le* itself) belong to Class I. The interested reader is referred to Toweett (1979), Rottland (1982), Creider & Creider (1989), and Kouneli (2022) for a complete description and sample conjugation paradigms.

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 (6-b) 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. Kà-∅-pú:tʃ Tʃé:bê:t ká:t (ák) [à:-]tʃáp tʃà:ík.
 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 subjunctive Type II, with subjunctive Type I being restricted to coordination contexts (as in (7-a)), temporal adjunct clauses, and conditionals (see Appendix for details and examples). Thus, we find various verbs - more prominently, volitional predicates - that always select for a subjunctive Type II complement. In this case, there is no complementizer present, as already seen in (6-b). This is the only complementation strategy for those verbs.

A second class of verbs always select for clausal complements where the verb is inflected in the indicative. The more prominent verbs in this class are factive verbs, with two examples seen in (8). In (8-a), we see that *sír* ‘to pass’ is inflected for indicative mood when embedded under the adjectival attitude *pájpáj* ‘to be happy’. Similarly, in (8-b) *ɲa:l* ‘to lie’ is inflected for indicative under *nere:tʃ* ‘to be angry’. For verbs that select for indicative complements, the presence of the complementizer is obligatory.

- (8) a. À:-pájpáj [à:-lé kò:-∅-sír Kíplàngàt].
 1SG-happy 1SG-LE PST.REC-3-pass Kiplangat.NOM
 ‘I’m happy that Kiplangat passed (the exams).’
- b. α-nere:tʃ-i [α:-le kò:-∅-ɲa:l-an Kíbê:t].
 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 *tʃám* ‘to whisper’. Concretely, we see *tʃáp* ‘to make’ inflected for indicative in (9-a), but for subjunctive II in (9-b).

- (9) a. Kà-∅-tʃám Kíbê:t [kò:-lé kà-∅-tʃáp kímpé:t].
 PST.CURR-3-whisper Kibeet.NOM 3-LE PST.CURR-3-make ugali
 ‘Kibeet whispered that he made ugali.’

- b. Kà-∅-tʃá:m-ú-án Kíbê:t [à-tʃáp
 PST.CURR-3-whisper-VENT-1SG Kibeet.NOM 1SG-make.SBJVII
 kímpé:t].
 ugali
 ‘Kibeet whispered to me to make ugali.’

To summarize, in a similar fashion to 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 vs. 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: ✓ is used whenever an indicative/subjunctive complement is possible, ✗ whenever an indicative/subjunctive complement is impossible, while a ✗[?] indicates that we have never encountered uses of those predicates with complements of a given mood, but do not have actual ungrammatical examples at hand. Nevertheless, the data we do have indicate that those predicates with a ✗[?] are likely to either prohibit or at least strongly disprefer complements of that mood. The table is organized in three blocks: in the first block, we present those verbs that can select for either mood (for those verbs, we also include a column with the difference in meaning depending on the mood of the complement); 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.

Predicate	Indicative (+le)	Subjunctive (Type II)	Meaning difference (IND vs. SBJV)
<i>mwa</i> ‘say’	✓	✓	assertive/reportative vs. directive
<i>tʃam</i> ‘whisper’	✓	✓	assertive/reportative vs. directive
<i>sír</i> ‘write’	✓	✓	assertive/reportative vs. directive
<i>tép</i> ‘ask’	✓	✓	assertive/reportative vs. directive
<i>maɲ</i> ‘expect’	✓	✓	no clear difference
<i>kas</i> ‘hear’	✓	✓	hear that... vs. hear (someone) Xing
<i>keɾ</i> ‘see’	✓	✓	see that... vs. see (someone) Xing
<i>ɲen</i> ‘know’	✓	✓	know that... vs. know how to...
<i>pwaɪ</i> ‘think/remember’	✓	✓	think/remember that... vs. remember (someone) Xing
<i>ɲon</i> ‘complain’	✓	✗ [?]	
<i>jan</i> ‘believe’	✓	✗ [?]	
<i>tam</i> ‘falsely accuse’	✓	✗ [?]	
<i>rua:tit</i> ‘dream’	✓	✗ [?]	
<i>po:r</i> ‘show’	✓	✗ [?]	
<i>naj</i> ‘realize’	✓	✗ [?]	
<i>ra:gin</i> ‘worry’	✓	✗ [?]	
<i>ne:ɾ:ɬ</i> ‘be angry’	✓	✗ [?]	
<i>pajpaj</i> ‘happy’	✓	✗ [?]	
<i>matʃ</i> ‘want’	✗	✓	
<i>jaɲ</i> ‘make/do’	✗	✓	
<i>mje</i> ‘good’	✗ [?]	✓	
<i>kara:ran</i> ‘good/beautiful’	✗ [?]	✓	
<i>ja</i> ‘bad’	✗ [?]	✓	

Table 1: Mood selection

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 & Rao 2019).⁷ We now turn to the core properties of the complementizer - the focus of our paper.

The Kipsigis complementizer consists of the root of the lexical verb *le* ‘say’ and a person/number agreement prefix, glossed transparently in (10).

- (10) a. â:-ŋgén ù:-lé Ø-rú-è Kíbê:t.
 1SG-know 1SG-LE 3-sleep-IPFV Kibeet.NOM
 ‘I know that Kibeet is sleeping.’
 b. Kà-ó-mwá ò:-lè Ø-rú-è Kíbê:t.
 PST.CURR-2PL-say 2PL-LE 3-sleep-IPFV Kibeet.NOM
 ‘You(pl) said that Kibeet is sleeping.’
 c. Kí:-ŋgèn kè:-lé Ø-rú-è Kíbê:t.
 IMP-know IMP-LE 3-sleep-IPFV Kibeet.NOM
 ‘It is known that Kibeet is sleeping.’ (impersonal)⁸

Based on work with two native speakers, Diercks & Rao (2019) report an additional non-agreeing ‘say’-based complementizer for Kipsigis, illustrated in (11).⁹

- (11) α-ŋgɛn *(α-le/kɔɛ) ko-Ø-ruuja tuya amut
 1SG-know 1SG-C/that PST-3-sleep cows yesterday
 ‘I know (that) the cows slept yesterday.’ (Diercks & Rao 2019: 372)

The native speakers that we consulted all found the non-agreeing 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 non-agreeing complementizer reported by Diercks & Rao (2019) to only be 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 which we discuss in detail in Section 4.1.

Diercks & 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 non-subject DPs in the matrix clause, a possibility that is not fully explored in

⁷For completeness, we note that there are, to our knowledge, two verbs which 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 Section 4.1, while the latter is discussed in detail in Bossi (2023a).

We also note that, modulo matrix uses of *le* which will be discussed in Section 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).

⁸The impersonal construction in Kipsigis is syntactically active. Morphologically, it is expressed by combining a 1st plural subject agreement prefix with 3rd person tonal melody.

⁹Our [ATR] and vowel length transcriptions sometimes differ from those in Diercks & 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 & Vergnaud 1981, Baković 2000, Nevins 2010). In this paper, we have maintained the original transcriptions and glosses for examples from Diercks & Rao (2019).

¹⁰Mike 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.

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

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

Diercks & Rao (2019). Whenever matrix objects can qualify as the source of information reported in the embedded clause, agreement with *le* becomes an option, shown here for a PP object in (12) and an applied object in (13).¹¹

- (12) Kà-∅-kás Kíplàngàt kòbún íjê: kò-lé/ ì:-lè
 PST.CURR-3-hear Kiplangat.NOM from 2SG 3-LE/ 2SG-LE
 kà-∅-tʃó:r Kíbê:t ràbí:ník.
 PST.CURR-3-steal Kibeet.NOM money
 ‘Kiplangat heard from you that Kibeet stole the money.’

- (13) Kò:-á-mwàj-tê:-tʃí Tjèbê:t é:n tò:jé:t à:-lé/kò-lé
 PST.REC-1SG-say-IT-APPL Cheebet at meeting 1SG-LE/3-LE
 kò:-∅-tʃó:r Kíbê:t ràbí:ník.
 PST.REC-3-steal Kibeet.NOM money
 ‘At the meeting, I said on Cheebet’s behalf that Kibeet stole the money.’

Another example of agreement with non-subject DPs can be seen in (14). The verb *wú:t* ‘to forget’ appears in a syntactic frame in which the grammatical subject is invariably 3rd 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, and not with the grammatical subject.

- (14) Kà-∅-wú:t-ú-ún à:-lé kò:-∅-kér Kíbê:t
 PST.CURR-3-forget-VENT-1SG 1SG-LE PST.REC-3-close Kibeet.NOM
 kúrgé:t.
 door
 ‘I forgot that Kibeet closed the door.’

Furthermore, impersonal agreement on the complementizer (see (10-c) above) is also available for a wide range of fully inflected lexical verbs in the matrix clause, in which

¹¹In (12), we assume that the form *kòlé* reflects agreement with the 3rd person subject. An anonymous reviewer asks whether in this case *kòlé* could instead be the non-agreeing complementizer reported in Diercks & Rao (2019). This is in principle a possibility, since the 3rd person agreeing form and the non-agreeing form are claimed to be morphologically identical (Diercks & Rao 2019, however, do not provide tonal transcriptions, so this claim cannot be fully evaluated). Nevertheless, our speakers always reject the use of *kòlé* in contexts without possible 3rd person targets, unlike the speakers consulted by Diercks & Rao (2019). This is why we conclude that our speakers do not have the non-agreeing form in their grammar, as was discussed earlier.

¹²This 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).

case a hearsay or rumour interpretation arises; this is illustrated in (15) below.

- (15) Kà-∅-kás Kíplàngàt kè:-lé kà-∅-tʃó:r Kíbê:t
 PST.CURR-3-hear Kiplangat.NOM IMP-LE PST.CURR-3-steal Kibeet.NOM
 ràbí:ník.
 money
 ‘Kiplangat heard (a rumour) that Kibeet stole the money.’

Diercks & 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-α-mwaa-un α-lɛ-ndʒin ko-∅-it tuya 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 & Rao 2019: 371)

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 & Rao (2019) 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

In this section, we first argue in 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, and not C. In 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, and not necessarily the matrix subject (*contra* Diercks & 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, Moulton 2019, Halpert 2019, Letsholo & Safir 2019, Bondarenko 2020), analyses of these complementizers as elements of category V, and not C, have been sporadic in the literature (e.g. Koopman 1984, Koopman & Sportiche 1989, Kinyalolo 1993, Knyazev 2016, Özyıldız et al. 2018, Demirok et al. 2020, Major & Torrence 2021, Major 2021, Major et al. 2022). 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 be modified by adverbs, and it can host applicative and reflexive verbal morphology even when used in complementation.

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) **kà-∅-lé** Kíbê:t (***kò-lé**) ∅-rú-è là:kwè:t.
 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 & 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 *le* (which is base-generated in C) moves to this position (see (4) in Section 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 vs. complementation uses are considered.

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 vs. 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 (19-a), we see a matrix use of *le*, in which case there is indicative inflection (cf. (18-a)). In (19-b), *le* is used to introduce an embedded clause, and it has Type I subjunctive, otherwise seen in - among other contexts - conditionals and coordination (cf. (18-b)). Finally, in (19-c), “matrix” *le* is embedded under a volitional predicate and shows up with Type II subjunctive (cf. (18-c)). From now on, we always gloss subjunctive inflection on *le*.

- (18) a. Kì:-[á]-rú.
 PST.DIST-1SG-sleep(IND)
 ‘I slept’
 b. Iṅgot [à]-rú...
 if 1SG-sleep.SBJVI
 ‘If I sleep...’
 c. á-mátj-é [á]-rú.
 1SG-want-IPFV 1SG-sleep.SBJVII
 ‘I want to sleep.’
- (19) a. Kì:-[á]-lé kì:-∅-tjó:r Kíbê:t ràbí:ník.
 PST.DIST-1SG-LE(IND) PST.DIST-3-steal Kibeet.NOM money
 ‘I said that Kibeet stole the money.’
 b. Kì:-á-mwá [à]-lé kì:-∅-tjó:r Kíbê:t
 PST.DIST-1SG-say 1SG-LE.SBJVI PST.DIST-3-steal Kibeet.NOM
 ràbí:ník.
 money
 ‘I said that Kibeet stole the money.’

- c. á-mátʃ-é [à]-lé kì:-Ø-tʃó:r Kíbê:t ràbí:ník.
 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 & 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’ has the form *le:lén*, which exhibits irregular stem 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) **Lè:lén** lòjójwè:k kò:-Ø-tʃó:r Kíbê:t ràbí:ník.
 LE.IPFV(IND) news.NOM PST.REC-3-steal Kibeet.NOM money
 ‘The news say 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-α-mwá-é à:-lé/à:-lè:lén kà-Ø-tʃó:r
 PST.CURR-1SG-say-IPFV.SBJVI 1SG-LE/1SG-LE.IPFV PST.CURR-3-steal
 Kíbê:t ràbí:ník.
 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 & Rao (2019) give a list of *le* forms with object agreement, shown in Table 3.

- (22) ko-α-mwaa-un α-lɛ-ndʒin ko-Ø-it tuya 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 & 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 & 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 *le* – *le.n* –, followed by the applicative

¹³The 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 & Creider (1989), Kouneli (2022) for details on Kipsigis conjugation.

¹⁴Only imperfective - the morphologically marked aspect in Kipsigis - will be indicated in the glosses.

	SG	PL
1	-lɛ-ndʒ-an	-lɛ-ndʒ-ɛtʃ
2	-lɛ-ndʒ-in	-lɛ-ndʒ-ɔ:ɣ
3	-lɛ-ndʒ-i	

Table 3: Suffixal agreement (Diercks & Rao 2019: 381)

suffix *-tʃi*, followed by the regular object clitics in the language. Regular phonological processes (e.g. voicing of obstruents after nasals and vowel coalescence rules; Kouneli 2019: Chapter 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.

	SG	PL
1	-le:n-tʃi-an (le:ndʒa:n)	-le:n-tʃi-ɛ:tʃ(le:ndʒɛ:tʃ)
2	-le:n-tʃi-in (le:ndʒi:n)	-le:n-tʃi-a:k (le:ndʒa:k)
3	-le:n-tʃi (le:ndʒi)	

Table 4: Suffixal agreement decomposed into APPL and object clitics

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 & 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 Kíbê:t kímpé:t.
 PST.CURR-3-make Kibeet.NOM ugali
 ‘Kibeet made ugali (type of food).’
- b. Kû-∅-tʃáp-tʃí Kíbê:t Tʃê:bê:t kímpé:t.
 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 above are simply the regular object clitics in the language, summarized in Table 5, built with data from Toweett (1979: p.209).¹⁷ The last piece of the reanalysis is the claim that the verb *le* has an allomorph *le:n*. In Kalenjin languages, there are at least ten CV verbs that have a CV:(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 *le:l* for the cognate word in Endo-Marakwet). To see that this is not unique to *le* in Kipsigis, consider the allomorphy displayed by the verb *jo* ‘to

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

¹⁶The applicative *-tʃi* has an allomorph *-ʃi* when attached to verbs ending in an alveolar obstruent. It also has the allomorph *-u* for 1st/2nd 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 & Creider 1989, Zwarts 2004, Mietzner 2009).

¹⁷The 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 & Creider 1989).

	SG	PL
1	-an	-ɛ:tʃ
2	-in	-a:k
3	∅	

Table 5: Object clitics

come’ in (24): the root has the form *ɲo* when in stem-final position in (24-a), but the form *ɲo:n* when followed by the applicative in (24-b). The surface phonological form in (24-b) 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. Kà-∅-ɲò ká:t.
PST.CURR-3.IND-come home
‘He/she came home.’
- b. ɲo:n-tʃi-an (ɲó:ndʒá:n)
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 *-kɛ:* used to form reflexives and reciprocals, illustrated in (25) below.¹⁸

- (25) Kí-ké:r-è-kê:
1PL-look-IPFV-REFL
‘We are looking at ourselves/at each other.’

The suffix *-kɛ:* 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 *-kɛ:*, suffixal agreement on *le* can include both the applicative and the reflexive/reciprocal suffix, as shown in (26).

- (26) Kó:-∅-tʃù:m-tʃi-kê: Kíbê:t kò-lè:n-tʃi-kê: ɲâ:m.
PST.REC-3-whisper-APPL-REFL Kibeet.NOM 3-LE-APPL-REFL.SBJVI clever
‘Kibeet whispered to himself that he’s clever.’

If *le* is indeed a verb, it is also predicted that it should in principle be compatible with adverbial modification. As can be seen in (27), this prediction is borne out: the adverb *mutja* ‘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*).

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

- (27) [Kà-∅-mwá-é Kíbê:t [kò-lè:lén mù:tjà
 PST.CURR-3-say-IPFV Kibeet.NOM 3-LE.IPFV.SBJVI slowly
 [kà-∅-tʃó:r Kíplàngàt ràbí:ník]].
 PST.CURR-3-steal Kiplangat.NOM money
 ‘Kibeet was saying slowly that Kiplangat stole the money.’

Summarizing, data that were not explored in [Diercks & 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 which 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 this 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 IMP-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 comes in (who doesn’t know that Kibeet stole the money) and asks: Why didn’t the event take place? What was the problem?*
 Ta:bʊ:t ko (*ko-le) kò-∅-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. More specifically, the negative morpheme *ma-* can attach to *le* when it is used as a matrix verb, as in (31-a), but not when *le* is used in complementation, irrespective of whether there is matrix negation present, as shown in (31-b) and (31-c). [Diercks et al. \(2020\)](#) argue that the ungrammaticality of negation in complementation uses indicates that *le* is a complementizer, and not a verb.

¹⁹In this example, the noun *ta:bʊ:t* ‘problem’ appears in a pre-verbal topic position (this position is further discussed in Section 5.1).

- (31) a. Mǎ-ǎ-lè Ø-rú-è là:kwè:t.
 NEG-1SG-LE(IND) 3-sleep-IPFV child.NOM
 ‘I didn’t say that the child is sleeping.’
- b. Mǎ-a-mwá (*mǎ-)à:l-é Ø-rú-è là:kwè:t.
 NEG-1SG-say NEG-1SG-LE.SBJVI 3-sleep-IPFV child.NOM
 ‘I didn’t say that the child is sleeping.’
- c. Kà-a-mwá (*mǎ-)à:l-é Ø-rú-è là:kwè:t.
 PST.CURR-1SG-say NEG-1SG-LE.SBJVI 3-sleep-IPFV child.NOM
 ‘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 vs. complementizer status of *le*. While we do not have a concrete explanation at this point (though see footnote 36 for a suggestion), evidence in favor of this view comes from data like (32) below. What we see in (32) is a lexical verb embedded under a matrix predicate. 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 (32-b)-(32-c).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. â:-ɲgén à-pír pé:k.
 1SG-know 1SG-hit.SBJVII water
 ‘I know how to swim (*lit*: to hit water).’
- b. Mâ-ǎ(:)-ɲgén (*mǎ)-à-pír pé:k.
 1SG-know NEG-1SG-hit.SBJVII water
 ‘I don’t know how (not) to swim.’
- c. *â:-ɲgén mǎ-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 (31-b)-(31-c) were 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) acknowledged that *le*

²⁰The 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 Section 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* (and 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.

²¹It is also worth noting that negation is impossible with restructuring infinitives (Wurmbrand 2001 a.o.) 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”.

- (34) Kò:-á-mwàj-tê:-tʃí Tʃèbê:t é:n tò:jé:t à:-lé/kò-lé
 PST.REC-1SG-say-IT-APPL Cheebet at meeting 1SG-LE/3-LE.SBJVI
 kò:-∅-tʃó:r Kíbê:t ràbí:ník.
 PST.REC-3-steal Kibeet.NOM money
 ‘At the meeting, I said on Cheebet’s behalf that Kibeet stole the money.’
- (35) *Ká-a-kas-ji Kíbê:t kò-lé ∅-jà:tʃ-é kò-wá
 PST.CURR-1SG-hear-APPL Kibeet 3-LE.SBJVI 3-must-IPFV 3-go(SBJV)
 Nairobi.
 Nairobi
 ‘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 & 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 considering its status and the cause for co-indexation 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 co-reference or binding. A binding relation is questioned by the fact that c-command is not necessary for agreement. The verb *lé* can agree with the source even if the source is embedded in a PP, see (36).²⁴

- (36) Ká-ɪ-kàs [PP kòbùn Kìplàngàt] kò-lé/ ì:-lè
 PST.CURR-2SG-hear from Kiplangat 3-LE/ 2SG-LE.SBJVI
 kà-∅-tʃó:r Kíbê:t ràbí:ník.
 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 which take the impersonal form *kè:-lé* indicating a rumour interpretation, as we saw in (15), which receives a natural explanation if the pronominal subject co-refers with an impersonal antecedent in the discourse.²⁵ Here, we provide two additional scenarios in which *le* can agree with an antecedent that is only (saliently) present in the preceding discourse and not in the matrix clause, see (37) for 3rd person and (38) for 2nd person.²⁶

²⁴In (36), as well as (64-a) 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: Chapter 2) applies to *ka-* and *ɪ-*, resulting in the form *kɛ:-* on the surface. Similarly, we give underlying forms in (75), where vowel coalescence applies between the ventive and 2SG object clitic.

²⁵We 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 rumour interpretation, i.e., 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 *kè:-lé* form is more likely to derive a rumour interpretation.

²⁶There 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).

- (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 à:lé anɛ: kà-Ø-tʃó:r Kíbê:t
 PST.CURR-1SG-say 1SG-LE.SBJVI 1SG.NOM PST.CURR-3-steal Kibeet.NOM
 ràbí:ník.
 money
 'I said that Kibeet stole the money.'

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

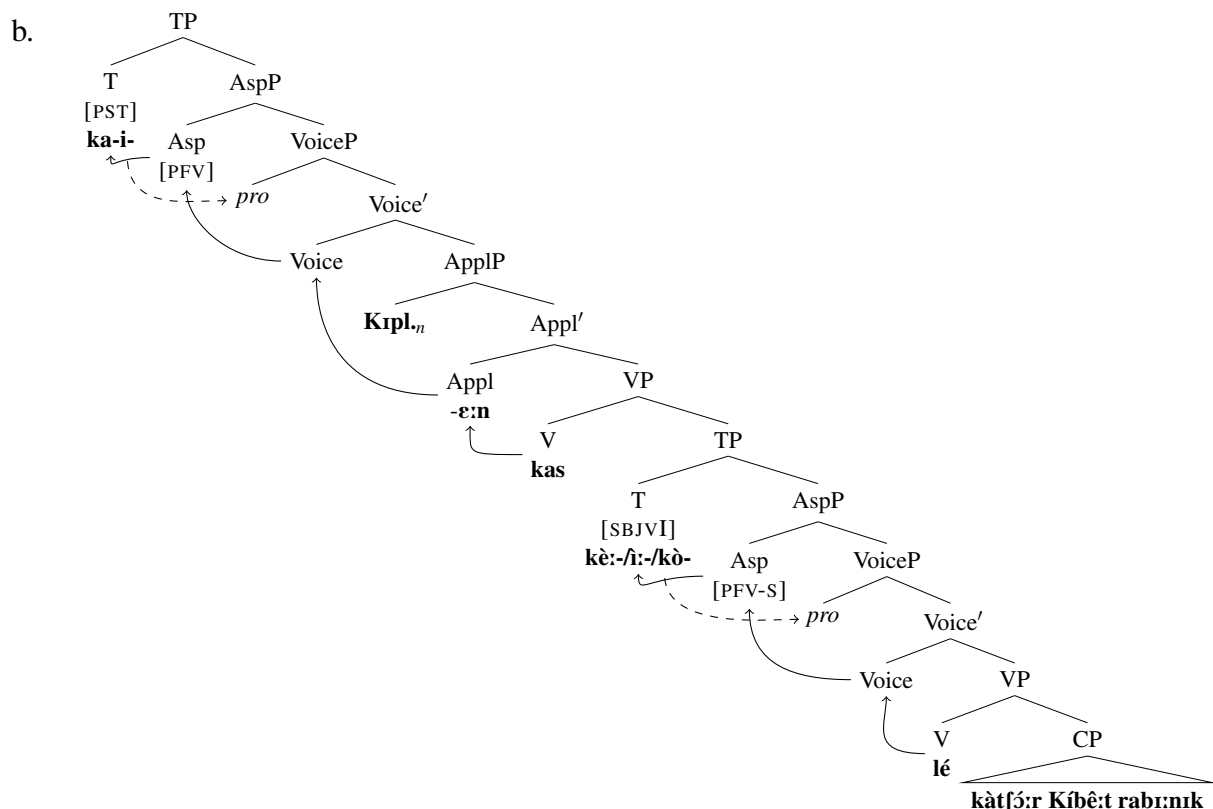
5 Syntactic analysis

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 these clauses 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 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 5.2.

Our complete analysis, including the different agreement options, is presented in (41-b), which is the structure for the sentence in (41-a). 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, Pykkänen 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 & Diercks (2019)), shown by the arrows in (41-b). 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 *i:le* is chosen if *pro* points to the addressee of the utterance, whereas *kòlé* appears if *pro* is co-indexed with *Kìplàngàt*, that is the source argument from the matrix clause. Another option is the impersonal form *kè:lé* which leads to a rumour interpretation, recall (15). In this case, *pro* co-refers with an impersonal antecedent in the discourse.

²⁸In Pykkänen (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 *-en*, 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.

- (41) a. Ká-i-kás-é:n Kiplàngàt kè:-lé/ ì:-lè/ kò-lé
 PST.CURR-2SG-hear-APPL Kiplangat IMP-LE/ 2SG-LE/ 3-LE.SBJVI
 kà-∅-tʃó:r Kíbê:t ràbí:nf̩k.
 PST.CURR-3-steal Kibeet.NOM money
 ‘You heard from Kiplangat that Kibeet stole the money.’



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 (41). 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 & Kouneli 2022), illustrated in (42). Following previous work on Nilotic (van Urk 2015), we assume that the pre-verbal topic position in Kipsigis is SpecCP.

- (42) Kíbê:t kó kà-∅-ám kímjé:t.
 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) â:-ŋgén à:-lé [Kìbê:t kó kà-∅-tʃó:r ràbí:ník].
 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

- (44) Kà-∅-tʃó:r ñà: ràbí:ník?
 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-a(:)-ŋgén à:-lé [kà-∅-tʃó:r ñà: ràbí:ník].
 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 (41-b) 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, shown in (42) above).

- (46) *Kà-a-mwá [ánê: kó à:-lé [kà-∅-tʃó:r
 PST.CURR-1SG-say 1SG TOP 1SG-LE.SBJVI PST.CURR-3-steal
 Kìbê:t ràbí:ník]].
 Kibeet.NOM money
 ‘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

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

³⁰*le* does not show tense distinctions and it is incompatible with negation, as discussed in Section 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.

³¹In their grammar of the related dialect Nandi, Creider & Creider (1989) claim that full clauses cannot

work, however, the languages under investigation (Greek, Romanian, and Ndebele) distinguish between CP and TP subjunctives. As was already mentioned in Section 3, Kipsigis does distinguish between two types of subjunctive, which are morphologically different in perfective 1SG only (see Appendix for further 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 Section 3, which is consistent with their analysis as CP complements.

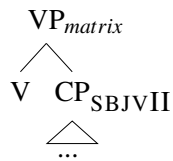
While we leave a complete investigation of the subjunctive as a topic for further research, we provide here a sketch of a morphological analysis for the relationship between subjunctive Type I and II. As a reminder, the two subjunctives are syncretic in all cells of the paradigm except for perfective 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 vs. adjunction

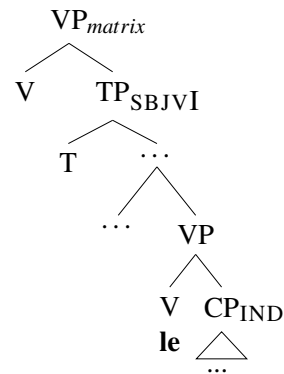
The most common assumption in the syntactic literature on complementation is that a CP headed by the complementizer (e.g. *that* in English) is merged as a sister to the matrix verb. In Kipsigis, however, the element mediating the relationship between the matrix verb and the embedded proposition is the verb *le*, and 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 Section 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 which 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 Section 6.1, we provide a semantic explanation for this fact.

be coordinated in the language, and we have already seen in (7-a) that the verb of the second conjunct in what could be clausal coordination must inflect for subjunctive Type I (and 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 (7-a) could be interpreted as evidence for the lack of a C layer in subjunctive Type I.

(47) SBJVII mood selection:



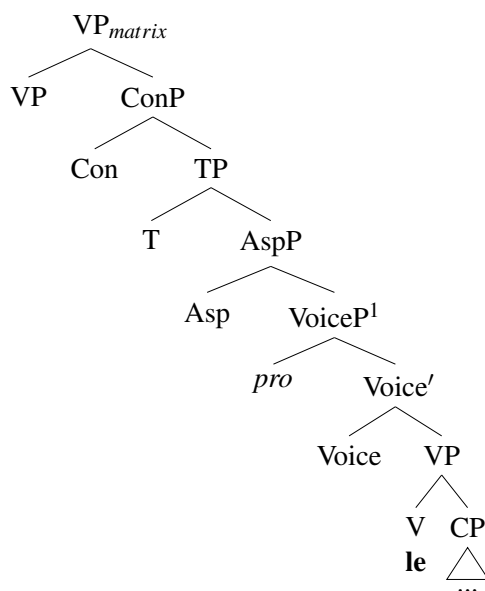
(48) Indicative mood selection:



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 which 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 clauses in Kipsigis as well, illustrated here with the example in (50).

- (50) Ká-a-we ká: [á:-labat-*(i)].
 PST.CURR-1SG-go.1SG house 1SG-run-IPFV.SBJV
 ‘I entered the house running.’

Nevertheless, we find two 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 make-up 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 Types I and II are syncretic in the imperfective (even for 1SG), and so it is not possible to determine whether converbs employ the same type of subjunctive that we find on *le*.

The second and more important 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 (51)–(52) below.

- (51) *Complex NP island*
 ***Kibet**_i [ko] ka-Ø-soman Tjébê:t [kìtábú:t ne
 Kibeet TOP PST.CURR-3-read Cheebeet.NOM book REL.SG
 ki:-Ø-sir-e ìné:ndèt_i/__ i].
 PST.DIST-3-write-IPFV 3SG.NOM
 Intended: ‘Kibeet, Cheebeet read the book that he wrote.’
 ([Driemel & Kouneli 2022](#): p.14)

- (52) *Adjunct island*
 ***Kibet**_i [ko] ka-ki-si:ndan-ε:tf [amun ma-Ø-jo: __ i].
 Kibeet TOP PST.CURR-1PL-win-1PL(IMP) because NEG-3-come
 Intended: ‘Kibeet, they beat us (at the race) because he didn’t come.’
 ([Driemel & Kouneli 2022](#): p.14)

Topicalization out of converbial clauses as in (50) is impossible, as shown in (53). This is consistent with their status as adjuncts, which generally behave as islands in the language (see, for example (52)). 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 & Kouneli 2022](#) for details).

- (53) Context: *Multiple people enter the house holding different things. What was everyone holding? Who was holding the flower?*
 ***Mauwat**_i kó ka-a-we ká: [a:-nam-e __ i].
 flower TOP PST.CURR-1SG-go.1SG house 1SG-hold-IPFV.SBJV
 ‘The flower, I entered the house holding (it).’

Extraction out of *le*-clauses, on the other hand, is always possible, illustrated in (54)–(56) below for a variety of matrix predicates: a communication verb in (54), a communication verb with a goal argument in (55) and a doxastic predicate in (56).³²

- (54) 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?)*
Kíbê:t_i kó ka-a-mwa [a:-le ka-∅-am ___ i
 Kibeet TOP PST.CURR-1SG-say 1SG-LE.SBJVI PST.CURR-3-eat
 kimpe:t].
 ugali
 ‘Kibeet, I said that he ate ugali.’
- (55) Context: *We are at an event with multiple people attending and multiple dishes available. Who ate what? Who ate ugali? (Who ate meat?)*
Kímpé:t_i kó ka-a-mwa-u-in [a-le:n-tʃi-in
 ugali TOP PST.CURR-1SG-say-VENT-2SG 1SG-LE-APPL-2SG.SBJVI
 ka-∅-am Kíbê:t ___ i].
 PST.CURR-3-eat Kibeet.NOM
 ‘Ugali, I told you that Kibeet ate it.’
- (56) 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?)*
Tfêbê:t_i kó â:-já:n-í [â:-lé ka-∅-am ___ i
 Cheebeet TOP 1SG-believe-IPFV 1SG-LE.SBJVI PST.CURR-3-eat
 pe:nda].
 meat
 ‘Cheebeet, I believe that she ate meat.’

Data like (54)–(56) indicate that *le*-clauses do not have the same structure as converbial clauses in the language, and they point against an adjunct analysis along the lines of (49).

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 (7-a), repeated here as (57), where subjunctive type I is used for the inflection of the second conjunct and the conjunction marker *ak* can be omitted.

- (57) Kâ-∅-púrtʃ Tfêbê:t ká:t (ák) â:-tʃáp tʃá:rík.
 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 (57) is impossible

³²Bossi (2023b) reports that extraction is impossible (only) when the matrix predicate is a speech verb; the only example she provides is the equivalent of (55), i.e. 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 non-agreeing form of *le* (absent in our speakers’ grammar), indicating the possibility of dialectal differences.

irrespective of whether the conjunction marker is present or not, as shown in (58). This is not surprising, since movement in such cases would violate the Coordinate Structure Constraint, which is otherwise active in the language.

- (58) Context: *Different people were assigned tasks of making different beverages for the guests. Who made what? What did Cheebet make?*
 *Tfè:bè:t_i kó kɑ-ɑ-pu:tʃ ká:t [(ák) kò-tʃap ____i tʃà:í:k].
 Cheebet TOP PST.CURR-1SG-sweep house and 3-make.SBJVI tea
 ‘Cheebet, 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, despite them sharing subjunctive Type I inflection. 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.

6 Semantic analysis

After an introduction to the eventuality-based framework of attitude predicates in section 6.1, we lay out the main components of the semantic analysis in 6.2, that is 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 section 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, see (59).

- (59) *Hintikkan semantics*

$$\llbracket \textit{believe} \rrbracket^{w,g} = \lambda p \lambda x. \forall w' \in \text{DOX}_{x,w} : p(w')$$

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), Anand & Hacquard (2008), we can define a CONT(ENT) function from eventualities to sets of possible worlds compatible with that eventuality. The denotation for *believe* under such an approach is given in (60).

CONT themselves. This explain why *le* is obligatory with indicative complementation, as was shown in section 3. An immediate prediction of the lexical entry in (62-a) 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 (63-a). Note that content nouns are possible with verbs that do not have this built-in restriction, which is shown for *mwa* in (63-b).³⁴

- (63) a. *K α - \emptyset -le:(n-tfi-an) Kíbê:t sa:ɛ:t/ lɔɔjwɛ:k/ ati:nda:pa:t.
 PST-3-LE-APPL-1SG Kibeet.NOM prayer/ news/ story
 ‘Kibeet told (me) a prayer/the news/a story.’
 b. K α - \emptyset -mwa:(u-an) Kíbê:t sa:ɛ:t/ lɔɔjwɛ:k/ ati:nda:pa:t.
 PST-3-say-VENT-1SG Kibeet.NOM prayer/ news/ story
 ‘Kibeet told (me) a prayer/the news/a story.’

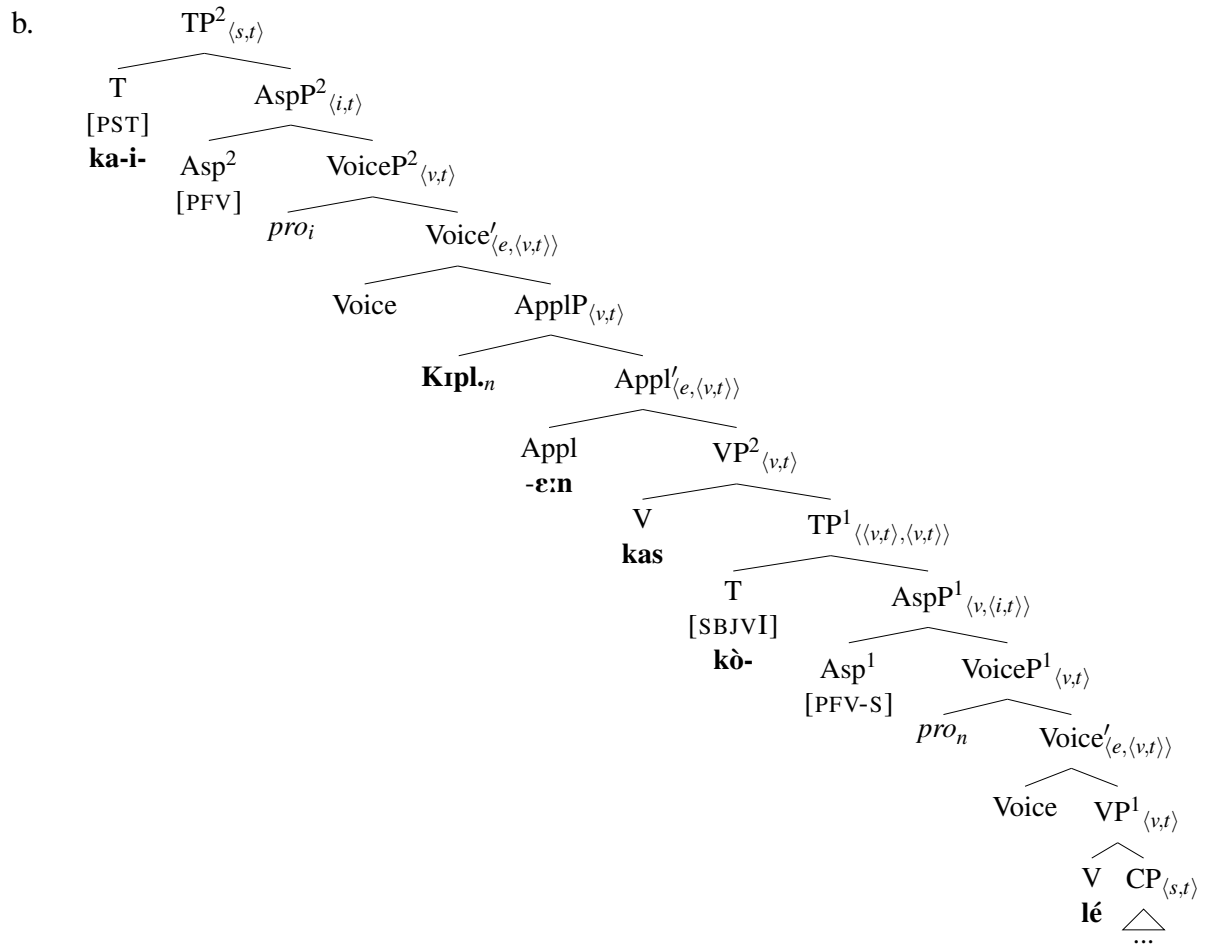
In this section, we introduced the framework of contentful eventualities and utilized it to explain the uniqueness of *le* to be able to combine with other matrix predicates to introduce indicative complement clauses as one of the two main complementation strategies in Kipsigis. We derived the special status 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 do so, 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 from section 5, repeated in (64-a). The bracketing in (64-a) reflects the syntactic choices made for the underlying structure, which we argued for in section 5. The tree in (64-b) is based on (41-b) from the previous section, now enriched with semantic types. We first focus on the *kòlé* derivation, where *pro* is co-indexed with the applied argument *Kìplàngàt*.

- (64) a. [_{TP}² Ká-I-kás-é:n Kìplàngàt [_{TP}¹ kè:r-lé/ ì:l-è/
 PST.CURR-2SG-hear-APPL Kiplangat IMP-LE/ 2SG-LE/
 kò-lé [_{CP} kà- \emptyset -tʃó:r Kíbê:t ràbí:ník]]
 3-LE.SBJVI PST.CURR-3-steal Kibeet.NOM money
 ‘You heard from Kiplangat that Kibeet stole the money.’

³⁴While the entry in (62-b) does not exclude a combination with content nouns, it does not per se allow for this combination either. One way to derive (63-b) is by adopting a theta-head that introduces individual-type arguments since its existence has independently been argued for in the context of hyper-raising in say-based complementation languages in general (Bondarenko 2020) as well as in Kipsigis in particular (Driemel & Kouneli 2021).



Since *le* is not a complementizer but a verbal category, it introduces a saying eventuality, see the repeated entry in (65-a). Following Hacquard (2006), we assume that speech and attitude eventualities have propositional content, that is 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 that eventuality. As for (64-a), 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 (65-b). 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 which introduces the agent of the saying event, where Voice combines with VP via Event Identification (Kratzer 1996), as is shown in (65-c). 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)/(61-b).

- (65)
- $\llbracket le \rrbracket^{w,g} = \lambda p_{\langle s,t \rangle} \lambda e_v. say(e) \wedge \forall w' \in \text{CONT}(e) : p(w')$
 - $\llbracket VP^1 \rrbracket^{w,g} = \lambda e_v. say(e) \wedge \forall w' \in \text{CONT}(e) : \text{Kibeet steals the money in } w'$
 - $\llbracket VoiceP^1 \rrbracket^{w,g} = \lambda e_v. say(e) \wedge \text{AG}(e) = g(n) \wedge \forall w' \in \text{CONT}(e) : \text{Kibeet}$

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 section 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 le is the only lexeme encoding CONT, we predict that no other matrix verbs can take indicative clauses directly as complements 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 (64-a). Given that le can show aspect morphology, as was shown in (21), we include an AspP layer in (64-b). In general, subjunctive Type I connects clauses, as it is also used more widely in coordinate clauses, recall example (7-a). 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 function, 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 (66) for perfective and imperfective aspect. In unembedded scenarios, that is in indicative clauses, the standard account can be adopted. Hence, the denotation in (66-a) can be directly taken to be encoded by Asp² in (64-b).³⁵

(66) *Aspect (cf. Kratzer 1998, Paslawska & von Stechow 2003)*

- a. $\llbracket \text{PFV} \rrbracket = \lambda P_{\langle v,t \rangle} \lambda t. \exists e [\tau(e) \subseteq t \wedge P(e)]$ Asp² in (64-b)
 b. $\llbracket \text{IPFV} \rrbracket = \lambda P_{\langle v,t \rangle} \lambda t. \exists e [t \subseteq \tau(e) \wedge P(e)]$

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 (66). Grano (2020, 2021) proposes to define variants of the aspectual morphemes in such cases, see (67), where the eventuality argument is passed up rather than existentially closed off, as in (66). We will adopt this idea and take (67-a) to be encoded by Asp¹, resulting in the denotation in (68) for AspP¹.

(67) *Aspect under subjunctive (Grano 2020)*

- a. $\llbracket \text{PFV-S} \rrbracket = \lambda P_{\langle v,t \rangle} \lambda e. \lambda t [\tau(e) \subseteq t \wedge P(e)]$ Asp¹ in (64-b)
 b. $\llbracket \text{IPFV-S} \rrbracket = \lambda P_{\langle v,t \rangle} \lambda e. \lambda t [t \subseteq \tau(e) \wedge P(e)]$

³⁵The 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 runtime of the event is included in the reference time, whereas for the latter, the reference time is included in the runtime of the event.

$$(68) \quad \llbracket \text{AspP}^1 \rrbracket^{w,g} = \lambda e \lambda t [\tau(e) \subseteq t \wedge \text{say}(e) \wedge \text{AG}(e) = g(n) \wedge \forall w' \in \text{CONT}(e) : \text{Kibeet steals the money in } w']$$

The entry for subjunctive Type I is provided in (69-a) combining the saying events in (68) with the hearing events in (69-b), the result of which is the denotation of VP^2 , shown in (69-c). The lexical entry in (69-a) is inspired by Özyıldız et al. (2018), who provide a similar entry for a gerundive affix serving a similar linking function in Turkish complementation.³⁶

$$(69) \quad \begin{array}{l} \text{a. } \llbracket \text{SBJVI} \rrbracket^{w,g} = \lambda P_{\langle v, \langle i, t \rangle \rangle} \lambda Q_{\langle v, t \rangle} \lambda e'' . \exists e' \exists t [\text{CAUSE}(e', e'') \wedge P(e')(t) \wedge Q(e'')] \\ \text{b. } \llbracket \text{kas} \rrbracket^{w,g} = \lambda e_v . \text{hear}(e) \\ \text{c. } \llbracket \text{TP}^1 \rrbracket^{w,g} (\llbracket \text{kas} \rrbracket^{w,g}) = \llbracket \text{VP}^2 \rrbracket^{w,g} \\ \quad = \lambda e'' . \exists e' \exists t [\text{CAUSE}(e', e'') \wedge \tau(e') \subseteq t \wedge \text{say}(e') \wedge \text{AG}(e') = g(n) \wedge \\ \quad \forall w' \in \text{CONT}(e') : \text{Kibeet steals the money in } w' \wedge \text{hear}(e'')] \end{array}$$

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 (70).

$$(70) \quad \llbracket \text{VoiceP}^2 \rrbracket^{w,g} = \lambda e'' . \exists e' \exists t [\text{CAUSE}(e', e'') \wedge \tau(e') \subseteq t \wedge \text{say}(e') \wedge \text{AG}(e') = g(n) \wedge \forall w' \in \text{CONT}(e') : \text{Kibeet steals the money in } w' \wedge \text{hear}(e'') \wedge \text{SOURCE}(e'') = \text{kiplangat}_n \wedge \text{EXP}(e'') = g(i)], \text{ defined iff } g(i) \text{ is addressee}^{37}$$

The CAUSE function is bidirectional, where direction is resolved by context. In (70), the agent of the saying event co-refers 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 function 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.

$$(71) \quad \text{You heard from Kiplangat}_i \text{ [} \textit{pro}_i \textit{ kò-lé} \text{ [Kibeet stole the money]]} \\ \rightsquigarrow \text{say}(e') \text{ causes } \text{hear}(e'') \text{ to take place}$$

The reverse relation, however, holds in case the agent of the saying event co-refers with the subject of the matrix predicate, that is if *le* inflects for 2SG, or in other words the *ì:lé* derivation of (64-a). In this case, the hearing event causes the saying event to take place,

³⁶The way SBJVI connects the *le*-clause with the matrix predicate suggests a promising explanation to 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 & 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 (69-c) 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 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 they are tensed and thus allow for negation to apply to a propositional argument.

³⁷ ϕ -features on pronouns denote partial identity functions of type $\langle e, e \rangle$ (Sauerland 2003, 2008, Heim 2008); for free pronouns the relevant assignment is given by the utterance context.

as the addressee is both the agent of the hearing and the saying event. This interpretation can be understood more abstractly as representing the addressee’s own interpretation of Kiplangat’s words.

- (72) You_i heard from Kiplangat [*pro*_i ì:-lé [Kibeet stole the money]]
 ~> *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 (71) as a speech report and the reading in (72) as an attitude report. The speech report reading represents a saying-event which causes the addressee to hear such a speech event. The attitude report reading, however, represents a hearing event which 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 for attitude readings. The eventuality introduced by *le* can encode either a speech event or a mental state, where the latter specifically can occur under non-speech matrix verbs, i.e. in situations that do not involve speaking. Özyıldız et al. (2018) describe the attitude reading as a ‘mental utterance’, Demirok et al. (2020) called the reading ‘inner speech’, while Major (2021) classifies this reading as stative SAY.³⁸ Interestingly, for perception verbs such as in (64-a), where both the speech report 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, i.e. 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 of which are presented in (73). Examples of verbs that appear with *le*-clauses in our field-notes include *ja:m* ‘to believe’, *bwa:t* ‘to think/remember’, *ta:m* ‘to (falsely) accuse’, *ɲo:m* ‘to complain’, *naj* ‘to realize’, *rua:tit* ‘to dream’, *ra:gin* ‘to worry’, *pajpaj* ‘to be happy’, *neɾe:tf* ‘to be angry’ (recall Table 1).

- (73) a. Kà-∅-tʃá:m Kíbê:t kò-lé kà-∅-tʃáp
 PST.CURR-3-whisper Kibeet.NOM 3-LE.SBJVI PST.CURR-3-make
 kímpé:t.
 ugali
 ‘Kibeet whispered that he made ugali.’

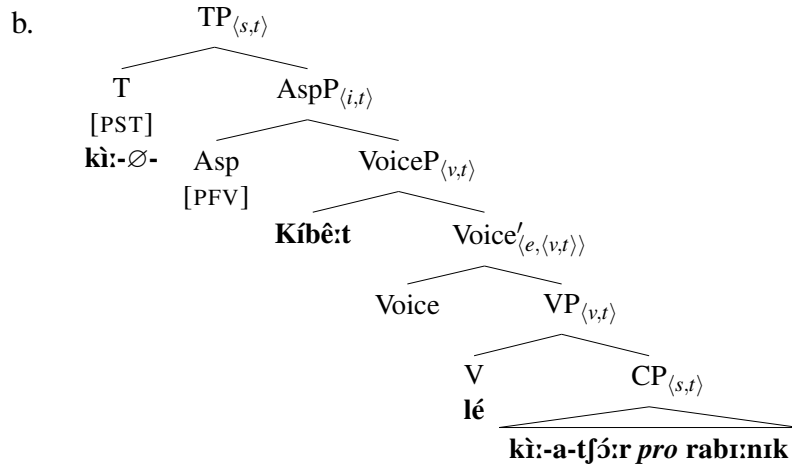
³⁸Major (2021) makes a distinction between eventive and stative SAY in languages with SAY-complementation, but the obligatory presence of agreement on *le* in Kipsigis suggests an eventive SAY syntax in Major’s typology. In other words, the syntax of *le* always corresponds to eventive SAY, but its semantics correspond to either eventive or stative SAY. We therefore chose not to pursue this line of analysis.

- b. â:-já:n-í ù:-lé mógôl ηwòŋ.
 1SG-believe-IPFV 1SG-LE.SBJVI round earth.NOM
 ‘I believe that the Earth is round.’
- c. À:-pájpáj ù:-lé kò:-Ø-sír Kíplàngàt.
 1SG-happy 1SG-LE.SBJVI PST.REC-3-pass Kiplangat.NOM
 ‘I’m happy that Kiplangat passed (the exams).’

As for communication verbs like *tʃám* ‘to whisper’ (73-a) and the doxastic verb *jam* ‘to believe’ (73-b), it is reasonable to assume that the eventuality introduced by the matrix verb causes a say-event to take place, either as a speech event or as a mental event. In this sense, the causal direction matches the one in (72). So for example in (73-b), entering a believe-eventuality leads the speaker to entertain the mental utterance that the Earth is round. On the other hand, the sentence in (73-a) conveys the meaning that Kibeet enters a whisper-event which causes the speech event whose content is that he made ugali. For other verbs, however, the opposite causal direction seems to be suitable, e.g., for fiction predicates like *ruatit* ‘to dream’ or emotive factives like *nerɛ:tf* ‘to be angry’ or *pajpaj* ‘to be happy’. The latter is repeated in (73-c). 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.

Finally, let us briefly address *le* in matrix position. As shown in section 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 (74-a), with the underlying syntactic structure in (74-b) and the semantic contribution of VoiceP in (74-c). Note that the denotation of VoiceP is derived with a lexical entry for *le* which is identical to the one we proposed in (65-a). In prose, (74-c) 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 (74-a) 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 which is normally associated with the occurrence of subjunctive.

- (74) a. [TP Kì:-Ø-lé Kíbê:t [CP kì:-á-tʃó:r ràbí:ník]]
 PST-3-LE(IND) Kibeet.NOM PST-1SG-steal(IND) money
 ‘Kibeet said that I stole the money.’



- c. $\llbracket \text{matrix VoiceP in (74-a)} \rrbracket^{w,g}$
 $= \lambda e_v. say(e) \wedge AG(e) = kibeet \wedge \forall w' \in \text{CONT}(e) : g(i)$ 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.

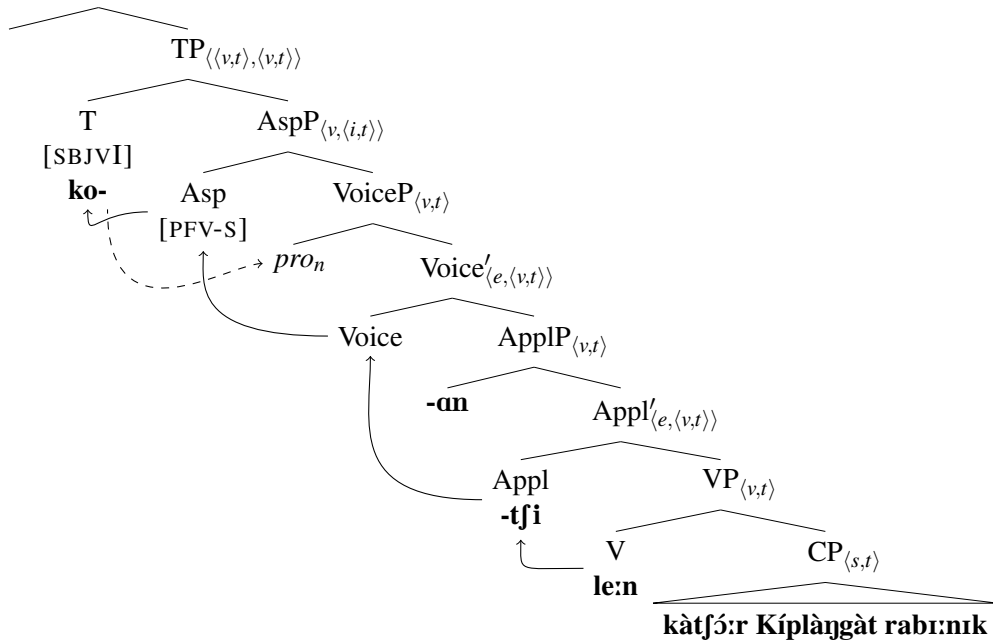
6.3 Extension to suffixal agreement

Let us now turn to suffixal agreement, which was discussed in section 4.1. We repeat example (2) from the introduction in (75), which shows that *le* not only shows prefixal agreement with the matrix subject but also an object clitic introduced by APPL.

- (75) Kà-Ø-tfá:m-ú-án Tfé:bê:t kò-lè:n-tf(i)-àn
PST.CURR-3-whisper-APPL-1SG Cheebet.NOM 3-LE-APPL-1SG.SBJVI
kà-Ø-tfó:r Kíbê:t ràbí:ník.
PST.CURR-steal Kibeet.NOM money
‘Cheebet whispered to me that Kibeet 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, shown for the partial derivation in (76).

(76)



We provide the denotation of matrix VoiceP in (77). 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 (73-a), the subject of *le* co-refers with the matrix subject, and the sentence receives a reading in which the whisper-event causes the saying-event to take place.

- (77) $\llbracket \text{matrix VoiceP in (76)} \rrbracket^{w,g}$
 $= \lambda e'' . \exists e' \exists t [\text{CAUSE}(e', e'') \wedge \tau(e') \subseteq t \wedge \text{say}(e') \wedge \text{AG}(e') = g(n) \wedge \text{GOAL}(e') = g(i) \wedge \forall w' \in \text{CONT}(e') : \text{Kiplangát steals the money in } w' \wedge \text{whisper}(e'') \wedge \text{GOAL}(e'') = g(i) \wedge \text{AG}(e'') = \text{Kibeet}_n]$, 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 *no:n* ‘complain’ and *si:r* ‘write’, shown in (78) and (79).

- (78) Kò:-á-*no:n* ù:-lé:n-*tjí* Kibê:t kò:-jâ:tj-*è:n*
 PST.REC-1SG-complain 1SG-LE-APPL.SBJVI Kibeet PST.REC-bad-PL
 ùmítwá:gík.
 food.NOM
 ‘I complained to Kibeet that the food was bad.’

- (79) Kò:-á-*si:r* ù:-lé:n-*tjí* Tjè:bê:t à-tjèlèwàni.
 PST.REC-1SG-write 1SG-LE-APPL.SBJVI Chebet 1SG-be.late
 ‘I wrote to Chebet 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 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 & 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 all reported cases of upwards-oriented complementizer agreement involve ‘say’-based complementizers, and not noun-y complementizers of the Indo-European type. 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 of a syntactic reanalysis of a ‘say’-based complementizers as a lexical verb in Kipsigis, 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 the discussion of quotative verbs found in African languages by [Güldemann \(2008\)](#). For instance, a recent case study by [Kientoré \(2023\)](#) of ‘say’-based complementation in the West-African language Jula reveals several more functions including similitive, 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, 2020](#), [Moulton 2019](#)). Such a framework allows us to let the syntactic analysis of *le* as a verb be reflected in the semantic composition. More importantly, it lets us derive the unique status of *le* by 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* 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.

References

- 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*: Potsdam University dissertation.
- Alexiadou, Artemis, Elena Anagnostopoulou, Gianina Iordachioaia & Mihaela Marchis. 2012. In support of Long Distance Agree. In Tibor Kiss & Gereon Müller (eds.), *Local modelling of non-local dependencies in syntax*, 55–81. Berlin: de Gruyter.
- Anand, Pranav. 2006. *De de se*: MIT dissertation.
- Anand, Pranav, Jane Grimshaw & Valentine Hacquard. 2019. Speech Act Predicates and their Subjects. In *Tokens of Meaning: A Festschrift for Lauri Karttunen*, UChicago Press.
- Anand, Pranav & Valentine Hacquard. 2008. Epistemics with Attitude. In T. Friedman & S. Ito (eds.), *Proceedings of SALT XVIII.*, 37–54. Ithaca, NY: Cornell University.
- Anand, Pranav & Feng-fan Hsieh. 2005. Long-distance reflexives in perspective. In John Alderete, Chung-hye Han & Alexei Kochetov (eds.), *Proceedings of the 24th West Coast Conference on Formal Linguistics*, 43–51. Somerville, MA: Cascadilla Proceedings Project.
- Arregi, Karlos & 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 Manuscript *Complementizers Relating to Noun Phrases: Rare Constructions within a Theory of Universal Grammar*, intended to be published in Language Science Press, Rutgers University, <https://sites.rutgers.edu/mark-baker/research/>.
- Baković, Eric. 2000. *Harmony, dominance, and control*: Rutgers University dissertation.
- Bjorkman, Bronwyn M. & Hedde Zeijlstra. 2019. Checking up on (ϕ -)Agree. *Linguistic Inquiry* 50. 527–569.
- Bogal-Allbritten, Elizabeth. 2015. Decomposing attitudes: The view from navajo. Paper presented at the 89th Meeting of the LSA.
- Bondarenko, Tatiana. 2020. Hyperraising and Logical Form: evidence from Buryat. GLOW 43 poster, tbond.scripts.mit.edu.
- Bossi, Madeline. 2023a. *Epistemic modality across syntactic categories in Kipsigis*: UC Berkeley dissertation.
- Bossi, Madeline. 2023b. Two types of ‘say’-complementation in kipsigis. Handout from talk given at WCCFL 41 at UCSC.
- Bossi, Madeline & 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.
- 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 R. Martin, D. Michaels & J. Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.
- Creider, Chet & Jane Tapsubei Creider. 1989. *A grammar of Nandi*. Hamburg: Helmut Buske Verlag Hamburg.
- Culy, Christopher. 1994. Aspects of logophoric marking. *Linguistics* 32. 1055–1094.
- Davidson, Donald. 1967. The logical form of action sentences. In Nicholas Rescher (ed.), *The Logic of Decision and Action*, 81–95. Pittsburgh: University of Pittsburgh Press.
- Demirok, Ömer, Deniz Özyıldız & Balkız Öztürk. 2020. Complementizers with attitude. In Maggie Baird & Jonathan Pesetsky (eds.), *Proceedings of the North East Linguistics Society 49*, Amherst: GLSA.
- Diercks, Michael. 2013. Indirect agree in Lubukusu complementizer agreement. *Natural Language & Linguistic Theory* 31. 357–407.
- Diercks, Michael, Marjo van Koppen & Michael Putnam. 2020. Agree Probes Down: Anaphoric Feature Valuation and Phase Reference. In Peter W Smith, Johannes Mursell & Katharina Hartmann (eds.), *Agree to Agree: Agreement in the Minimalist Programme*, 347–389. Berlin: Language Science Press.
- Diercks, Michael & Meghana Rao. 2019. Upward-oriented complementizer agreement with subjects and objects in Kipsigis. In Emily Clem, Peter Jenks & Hannah Sande (eds.), *Theory and description in African Linguistics: Selected papers from the 47th Annual Conference on African Linguistics*, 369–393. Berlin: Language Science Press.
- Driemel, Imke & Maria Kouneli. 2021. $\langle v, t \rangle$ -type complementation in Kipsigis. To appear in Proceedings of WCCFL 39, <https://ling.auf.net/lingbuzz/006045>.
- Driemel, Imke & Maria Kouneli. 2022. A dedicated topic position in Kipsigis. In Henrison Hsieh (ed.), *Proceedings of TripleA 8: Fieldwork Perspectives on the Semantics of Understudied Languages of Africa, Asia, Australia, and Oceania*, 1–20. Singapore: NUS ScholarBank.
- Duncan, Philip T. & Harold Torrence. 2017. Verbal complementizers and the Indirect Agree relation in Ibibio. Handout from talk given at the 91st Meeting of the LSA.
- Eberhard, David M., Gary F. Simons & Charles D. Fennig (eds.). 2020. *Ethnologue: Languages of the world*. Dallas, Texas: SIL International 23rd edn.
- Elliott, Patrick D. 2016. Explaining DPs vs. CPs without syntax. In *Proceedings of the 52nd annual meeting of the Chicago Linguistics Society*, 171–185. Chicago: Chicago Linguistic Society.
- Elliott, Patrick D. 2017. *Elements of clausal embedding*: University College London dissertation. Revised version, 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 Günther Grewendorf & Helmut Weiß (eds.), *Bavarian Syntax: Contributions to the theory of syntax*, vol. 220, 51–82. Amsterdam/Philadelphia: John Benjamins.
- Grano, Thomas. 2016. Semantic consequences of syntactic subject licensing: Aspectual predicates and concealed modality. In Nadine Bade, Polina Berezovskaya & Anthea Schöller (eds.), *Proceedings of of Sinn und Bedeutung 20*, 306–322. Tübingen: Universität Tübingen.
- Grano, Thomas. 2020. Intention reports and eventuality abstraction in a theory of mood choice. Manuscript, Indiana University [\[link\]](#).
- Grano, Thomas. 2021. Intention reports and eventuality abstraction in a theory of mood choice. Slides for SALT 31, [\[link\]](#).
- Grimshaw, Jane. 2015. The Light Verbs Say and SAY. In Ida Toivonen, Piroska Csúri & Emile van der Zee (eds.), *Structures in the Mind*, 79–99. Oxford: 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*: MIT dissertation.
- Hacquard, Valentine. 2010. On the event relativity of modal auxiliaries. *Natural Language Semantics* 18. 79–114.
- Haegeman, Liliane & Marjo van Koppen. 2012. Complementizer agreement and the relation between C^o and T^o . *Linguistic Inquiry* 43. 441–454.
- Hall, Beatrice, R.M.R. Hall, Martin D. Pam, Amy Myers, Stephen A. Antell & Godfrey K. Cheron. 1974. African Vowel Harmony Systems from the Vantage Point of Kalenjin. *Afrika und Übersee* 241-267. 241–267.
- Halle, Morris & 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 M Ryan Bochnak, Miriam Butt, Erlinde Meertens & Mark-Matthias Zymla (eds.), *Proceedings of triplea 5*, 31–46. Tübingen: University of Tübingen.
- Heim, Irene. 2008. Features on bound pronouns. In Daniel Harbour, David Adger & Susana Béjar (eds.), *Phi-Theory. Phi-features across Modules and Interfaces*, 35–56. Oxford: Oxford University Press.
- Hintikka, Jaakko. 1969. Semantics for propositional attitudes. In J. W. Davis, D. J. Hockney, & W. K. Wilson (eds.), *Philosophical Logic*, 21–45. Reidel: Dordrecht.
- Iatridou, Sabine. 1988. On Nominative Case Assignment and a Few Related Things. In *Papers on Case and Agreement II, MIT Working Papers in Linguistics 19*, 175–196. Cambridge, MA: MIT Press.
- Idiatov, Dmitry. 2010. Person-number agreement on clause linking markers in Mande. *Studies in Language* 34. 832–868.
- Kiemtoré, Alassane. 2023. *Issues in Jula complementation: Structures, relations and matters of interpretation*: Stuttgart University dissertation.
- Kinyalolo, Kasangati Kikuni W. 1993. On some syntactic properties of $\dot{q}\dot{\bar{a}}$. *Lingua* 91. 201–233.

- Knyazev, Mikhail. 2016. Complementizers in Kalmyk. In Kasper Boye & Petar Kehayov (eds.), *Complementizer semantics in European Languages*, 655–679. Berlin: Mouton de Gruyter.
- Koopman, Hilda. 1984. *The Syntax of Verbs*. Dordrecht: Foris.
- Koopman, Hilda & Dominique Sportiche. 1989. Pronouns, logical variables, and logophoricity in Abe. *Linguistic Inquiry* 20. 555–588.
- van Koppen, Marjo. 2005. *One Probe – Two Goals: Aspects of Agreement in Dutch Dialects*. Utrecht: LOT.
- van Koppen, Marjo. 2012. The distribution of phi-features in pronouns. *Natural Language & Linguistic Theory* 30. 135–177.
- Kouneli, Maria. 2019. *The syntax of number and modification: An investigation of the Kipsigis DP*: New York University dissertation.
- Kouneli, Maria. 2022. Inflectional classes in Kipsigis. *Glossa: A journal of general linguistics* 47(1).
- Kouneli, Maria & Yining Nie. 2021. Across-the-board tonal polarity in Kipsigis: Implications for the morphology-phonology interface. *Language* 97(2). e111–e138.
- Krapova, Iliyana. 2001. Subjunctives in Bulgarian and Modern Greek. In María Luisa Rivero & Angela Ralli (eds.), *Comparative Syntax of Balkan Languages*, 105–126. Oxford: Oxford University Press.
- Kratzer, Angelika. 1996. Severing the external argument from its verb. In Johan Rooryck & Laurie Zaring (eds.), *Phrase Structure and the Lexicon*, 109–137. Dordrecht: Kluwer.
- Kratzer, Angelika. 1998. More Structural Analogies Between Pronouns and Tenses. In Devon Strolovitch & Aaron Lawson (eds.), *Proceedings of SALT VIII.*, 92–110. Ithaca, NY: Cornell University.
- Kratzer, Angelika. 2006. Decomposing attitude verbs. Handout from a talk in honor of Anita Mittwoch on her 80th birthday.
- Kratzer, Angelika. 2013a. Modality and the semantics of embedding. Slides, Amsterdam Colloquium December 20, 2013.
- Kratzer, Angelika. 2013b. Modality for the 21st century. In Stephen R. Anderson, Jacques Moeschler & Fabienne Reboul (eds.), *L'interface Langage-Cognition/The Language-Cognition Interface: Actes du 19e Congrès International des Linguistes Genève*, 179–199. Librairie Droz.
- Kratzer, Angelika. 2016. Evidential moods in attitude and speech reports. Slides, available at https://works.bepress.com/angelika_kratzer/10/.
- Krifka, Manfred. 1998. The origins of telicity. In Susan Rothstein (ed.), *Events and Grammar*, 197–235. Kluwer: Dordrecht.
- Landau, Idan. 2004. The scale of finiteness and the calculus of control. *Natural Language & Linguistic Theory* 22. 811–877.
- Landau, Idan. 2013. *Control in Generative Grammar*. Cambridge: Cambridge University Press.
- Letsholo, Rose & Ken Safir. 2019. Complement clause C-agreement beyond subject phi-agreement in Ikalanga. In Samson Lotven, Silvina Bongiovanni, Phillip Weirich, Robert Botne & Samuel Gyasi Obeng (eds.), *African linguistics across the disciplines: Selected papers from the 48th Annual Conference on African Linguistics*,

- 159–186. Berlin: Language Science Press.
- Lord, Carol. 1976. Evidence for syntactic reanalysis: From verb to complementizer in kwa. In Sanford Steever, Carol Walker & Salikoko Mufwene (eds.), *Papers from the Parasession on Diachronic Syntax (12th Regional Meeting)*, Chicago Linguistics Society, 179–191. Chicago, Illinois: University of Chicago.
- Major, Travis. 2021. *On the nature of "say" complementation*. Los Angeles, CA: UCLA dissertation.
- Major, Travis, Justine Sikuku & Michael Diercks. 2022. “say”-complementation: reanalyzing agreeing complementizers as verbs. Ms., USC, Moi University and Pomona College.
- Major, Travis & Harold Torrence. 2021. “Say”-chains, not “say”-complementation. In Rachel Soo, Una Y. Chow & Sander Nederveen (eds.), *Proceedings of the 38th west coast conference on formal linguistics*, 283–293. Somerville, MA: Cascadilla Proceedings Project.
- Marantz, Alec. 1993. Implications of asymmetries in double object constructions. In Sam A. Mchombo (ed.), *Theoretical aspects of Bantu grammar*, 113–150. Stanford, CA: CSLI Publications.
- Marušič, Franc, Andrew Nevins & Bill Badecker. 2015. The grammars of conjunction agreement in slovenian. *Syntax* 18. 39–77.
- McFadden, Thomas & Sandhya Sundaresan. 2021. Unifying species of C-agreement. In Alessa Farinella & Angelica Hill (eds.), *Proceedings of the Fifty-First Annual Meeting of the North East Linguistic Society*, 91–100. Amherst, MA: GLSA.
- Mietzner, Angelika. 2009. *Räumliche Orientierung in nilotischen Sprachen*, vol. 24 Nilo-Saharan: Linguistic Analyses and Documentation. Köln: Rüdiger Köppe Verlag.
- Moulton, Keir. 2009. *Natural Selection and the Syntax of Clausal Complementation*: University of Massachusetts Amherst dissertation.
- Moulton, Keir. 2019. Decomposing attitude verbs. Slides, GLOW in Asia XII, Workshop on (Non-)Complementation, <http://sicogg.or.kr/GLOW-Asia-12-2019/>.
- 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, MA: MIT Press.
- Nformi, Jude. 2017. Complementizer agreement and intervention effects. Ms., Leipzig University.
- Özyıldız, Deniz, Travis Major & Emar Maier. 2018. Communicative reception reports as hearsay: Evidence from indexical shift in Turkish. In Richard Stockwell, Maura O’Leary, Zhongshi Xu & Z.L. Zhou (eds.), *Proceedings of the 36th West Coast Conference on Formal Linguistics*, 296–305. Somerville, MA: Cascadilla Proceedings Project.
- Parsons, Terence. 1990. *Events in the semantics of English*. Cambridge, MA: MIT Press.
- Paslawska, Alla & Arnim von Stechow. 2003. Perfect readings in Russian. In Artemis Alexiadou, Monika Rathert & Arnim von Stechow (eds.), *Perfect Explorations*, 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*: University of Chicago dissertation.
- 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 & 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, MA: 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 Kölner Beiträge zur Afrikanistik. Berlin: Dietrich Reimer Verlag.
- 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 R Young & Y Zhou (eds.), *Proceedings of SALT 13*, 258–275. Ithaca, NY: CLC Publications.
- Sauerland, Uli. 2008. On the semantic markedness of phi-features. In Daniel Harbour, David Adger & Susana Béjar (eds.), *Phi-Theory. Phi-features across Modules and Interfaces*, 35–56. Oxford: Oxford University Press.
- Sauerland, Uli, Bart Hollebrandse & 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 Patrick Grosz & Pritty Patel-Grosz (eds.), *The Impact of Pronominal Form on Interpretation*, 77–105. Berlin: de Gruyter.
- Terzi, Arhonto. 1992. *PRO in finite clauses: A study of the inflectional heads of the Balkan languages*: City University of New York dissertation.
- Thomason, Richmond. 2014. Formal semantics for causal constructions. In Bridget Copley & Fabienne Martin (eds.), *Causation in grammatical structures*, 58–75. Oxford: Oxford University Press.
- Toweett, Taaita. 1979. *A study of Kalenjin linguistics*. Nairobi: Kenya Literature Bureau.
- van Urk, Coppe. 2015. *A uniform syntax for phrasal movement: A case study of Dinka Bor*: MIT dissertation.
- Varlokosta, Spyridoula. 1993. Control in Modern Greek. In Carol A. Mason, Susan M. Powers & Cristina Schmitt (eds.), *University of Maryland Working Papers in Linguistics*, vol. 1, 144–163.
- Weisser, Philipp. 2019. Telling Allomorphy from Agreement. *Glossa* 4. 86.
- Winter, Yoad & Joost Zwarts. 2011. Event semantics and abstract categorial grammar. In Makoto Kanazawa, András Kornai, Marcus Kracht & Hiroyuki Seki (eds.), *Mol*

12. *Incs*, vol. 6878, 174–191. Berlin, Heidelberg: 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*. Kluwer: Dordrecht.
- Zwarts, Joost. 2004. *The Phonology of Endo: A Southern Nilotic Language*. München: LINCOM Studies in African Linguistics 59.

A 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 Section 3), Type II subjunctive is also used for purpose clauses and after modals.

(80) 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'

(81) Modals (Toweett 1979: 225)

- a. Ø-ɲa: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 (82), 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 (83).

(82) Iɲgot à:-rú...

- if 1SG-sleep.SBJVI
'if I sleep...'

(83) Kò:-Ø-tfáp-é kímjé:t Tfé:bê:t ko:n a:-ɲo: kát.
PST.REC-3-make-IPFV ugali Cheebeet.NOM when 1SG-come.SBJVI 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 Section 5.2). As shown in (84), 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.

(84) Kà-Ø-pú:tf Tfé:bê:t kát (ák) à:-tfáp tfà:í:k.
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 the language as a topic for further research.