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

Imke Driemel & Maria Kouneli

April 25, 2022

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.

---

[1]Outside of Africa, a similar phenomenon has been reported for the Trans-New Guinean language *Teiwa* (Sauerland et al. 2020).

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 and the matrix subject (Diercks & Rao 2019, Diercks et al. 2020), illustrated in (1).[2,3]

(1) a. âː-ŋgén **àː-lé** ∅-rú-è Kíbêːt.
        1SG-know 1SG-C 3-sleep-IPFV Kibeet.NOM
        'I know that Kibeet is sleeping.'
     b. îː-ŋgén **ìː-lé** ∅-rú-è Kíbêːt.
        2SG-know 2SG-C 3-sleep-IPFV Kibeet.NOM
        'You know that Kibeet is sleeping.'
     c. í-ŋgèn Kíplàŋgàt **kò-lé** ∅-rú-è Kíbêːt.
        3-know Kiplangat.NOM 3-C 3-sleep-IPFV Kibeet.NOM
        'Kiplangat knows that Kibeet is sleeping.'

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, Major & Torrence 2020, Major 2021, Major et al. 2022 for verbal analyses of such complementizers). Furthermore, we show that 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

---

[2]Kipsigis is the major variety of Kalenjin, a cluster of dialects of the Southern Nilotic branch of Nilo-Saharan, and it is spoken by about 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 elicitations in 2020-2021 with five native speakers (male, age range: 22-32) living in Nairobi. The five speakers consulted for questions about C-Agree all grew up in monolingual Kipsigis regions (two speakers in Narok County and three speakers in Bomet County). All five of them are also proficient in English and Swahili, the official languages of Kenya.

[3]Glossing abbreviations follow the Leipzig Glossing Rules with the addition of C = complementizer, IT = itive, and VENT = ventive. Tone is transcribed whenever possible, but some transcriptions are incomplete because of sound difficulties in Skype elicitations. Additionally, the tone on *le* is always transcribed as H, but it should be noted that it sometimes becomes low when it is followed by a word that starts with a H tone. The details of this sandhi phenomenon are currently not well-understood.

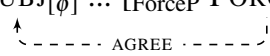eventuality-based models of attitude and speech reports (e.g. Kratzer 2013b, Elliott 2016, 2017, Moulton 2019).

The remainder of the paper is structured as follows: in Section 2, we provide a brief overview of previous theories of upwards-oriented complementizer agreement, before presenting the Kipsigis pattern in Section 3. We then develop our analysis in two steps: in 4 we argue that the Kipsigis "complementizer" is the lexical verb 'say' which agrees with a local subject, while in 5 we provide a semantic analysis. In Section 6, we conclude.
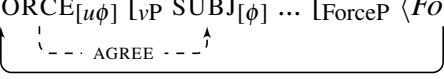
2   Previous theories of upwards-oriented C agreement

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). These two types of analyses are illustrated in (2) and (3), respectively.

(2)   *Upward Agree account*
$[_{vP}$ SUBJ$_{[\phi]}$ ... $[_{ForceP}$ FORCE$_{[u\phi]}$ ... $[_{FinP}$ ... $[_{TP}$ SUBJ ... $]]]]$
AGREE

(3)   *Downward Agree account*
$[_{vP}$ FORCE$_{[u\phi]}$ $[_{vP}$ SUBJ$_{[\phi]}$ ... $[_{ForceP}$ $\langle Force \rangle$ ... $[_{FinP}$ ... $[_{TP}$ SUBJ ... $]]]]]$
AGREE

Turning now to the nature of the goal, whereas Upward Agree accounts make the subject uniformly the target, Downward Agree approaches differ in terms of the agreement goal. For Lubukusu, Carstens (2016) proposes a direct Agree approach between the moved complementizer and the matrix subject, while 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. Thus, the complementizer moves to the matrix clause to check anaphoric ϕ-features. This idea is inspired by the indirect Agree analysis 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; this analysis has also been applied to Ibibio by Duncan & Torrence (2017), while a similar idea is proposed by Baker (2008) for Kinande.

Finally, some of these accounts also address the question of why the complementizer cannot probe downward from its base position into the embedded clause. Carstens

(2016) and McFadden & Sundaresan (2021) refer to the presence of a phase boundary to solve this problem, while Baker (2008) makes use of agreement parameter settings.

3 C agreement 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 and the typologically rare marked nominative system. The Kipsigis complementizer consists of the root of the lexical verb *le* 'say' and a person/number agreement prefix, glossed transparently in (4).

(4) a. âː-ŋgén   àː-lé   ∅-rú-è   Kíbêːt.
         1SG-know 1SG-LE 3-sleep-IPFV Kibeet.NOM
         'I know that Kibeet is sleeping.'
     b. Kà-ó-mwá   òː-lé   ∅-rú-è   Kíbêːt.
         PST-2PL-say 2PL-LE 3-sleep-IPFV Kibeet.NOM
         'You(pl) said that Kibeet is sleeping.'
     c. Kíː-ŋgèn   kèː-lé   ∅-rú-è   Kíbêːt.
         IMP-know IMP-LE 3-sleep-IPFV Kibeet.NOM
         'It is known that Kibeet is sleeping.' (impersonal)[4]

Based on work with two native speakers, Diercks & Rao (2019) report an additional non-agreeing 'say'-based complementizer for Kipsigis, illustrated in (5).[5]

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

The five native speakers that we consulted all found the non-agreeing complementizer in sentences like (5) 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.[6] Table 1 gives the paradigm for the agreement prefixes on *le*. The prefixes are identical to the agreement prefixes of lexical verbs in (a sub-conjugation of) the subjunctive, which we discuss in detail in Section 4.1.
Diercks & Rao (2019) argue that the Kipsigis complementizer can only agree with the matrix subject. We do indeed find upwards-oriented agreement with the matrix subject

---

[4]The impersonal construction in Kipsigis is syntactically active. Morphologically, it is expressed by combining a first-person plural subject agreement prefix with 3rd person tonal melody.

[5]Our [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).

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

	SG	**PL**
1	àː-	kèː-
2	ìː-	òː-
3	kò-	
imp	kèː-	

Table 1: Agreement prefixes on *le* (=subjunctive subject prefixes)

with verbs from a variety of lexical classes (e.g. *jaːn* 'to believe', *mwa* 'to say', *ruaːtit* 'to dream', *taːm* 'to falsely accuse', *nɛrɛːtʃ* 'to be angry (about)').

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 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 (6) and an applied object in (7).

(6) Kà-∅-kás Kíplàŋgàt kobun **íɲê:** kò-lé/ **ìː-lé** kà-∅-tʃóːr Kíbêːt
    PST-3-hear Kiplangat.NOM from   2SG 3-LE/ 2SG-LE PST-3-steal Kibeet.NOM
    rabɪːnɪk.
    money
    'Kiplangat heard from you that Kibeet stole the money.'

(7) Koː-a-mwɑi-teː-tʃi **Tʃèbêːt** ɛːn tʊːjɛːt àː-lé/**kò-lé** kòː-∅-tʃóːr
    PST-1SG-say-IT-APPL Cheebeet at meeting 1SG-LE/3-LE PST-3-steal
    Kíbêːt rabɪːnɪk.
    Kibeet.NOM money
    'At the meeting, I said on Cheebeet's behalf that Kibeet stole the money.'

Another example of agreement with non-subject DPs can be seen in (8). The verb *wuːt* '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.[7] In this case, *le* agrees with the experiencer-indirect object, and not with the grammatical subject.

(8) Kɑ-∅-wuːt-u-**an**   **ɑː-le** kòː-∅-kɛ́r Kíbêːt   kurgeːt.
    PST-3-forget-VENT-1SG 1SG-LE PST-3-close Kibeet.NOM door
    'I forgot that Kibeet closed the door.'

Furthermore, impersonal agreement on the complementizer (see (4-c) above) is also available for a wide range of fully inflected lexical verbs in the matrix clause, in which case a hearsay or rumour interpretation arises; this is illustrated in (9) below.

(9) Kà-∅-kás Kíplàŋgàt   kèː-lé kà-∅-tʃóːr Kíbêːt   rabɪːnɪk.
    PST-3-hear Kiplangat.NOM IMP-LE PST-3-steal Kibeet.NOM money
    'Kiplangat heard (a rumour) that Kibeet stole the money.'

---

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

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.

(10) ko-ɑ-mwaa-un   ɑ-lɛ-ndʒin   ko-∅-ɪt   tuɣa amut
      PST-1SG-tell-2SG.OBJ   1SG-C-2SG.OBJ   PST-3-arrive cows yesterday
      'I DID tell you (sg) that the cows arrived yesterday.' (Diercks & 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 (6) and (7)) and suffixal object agreement for indirect objects of speech verbs.

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, Grimshaw 2015, Moulton 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 2020, 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 (11). Crucially, the "complementizer" is ungrammatical in this case.

(11) **kà-∅-lé** Kíbêːt   (***kò-lé**) ∅-rú-è   làːkwɛ̀ːt.
      PST-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 (11). 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

6

to this position (see (3) 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. In order to understand the importance of this distinction, a short detour into Kipsigis verbal inflection is needed. All verbs in the language inflect for tense, aspect, and mood, and previous literature has identified three moods: indicative, subjunctive, and imperative (Toweett 1979, Rottland 1982, Creider & Creider 1989). Setting the imperative aside, the main difference between the indicative and the subjunctive is that the former is used in root clauses, while the latter in subordinate clauses; the language lacks infinitives of the European type.[8] 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 (12) below, the verb *ru* 'sleep' has a short-voweled subject agreement prefix in its indicative (matrix) form in (12-a), but a long-voweled prefix in its subjunctive (embedded) form in (12-b).[9] For 3rd person subjects, the prefix is ∅ in most cells of the paradigm, while it is always *ko-* in the subjunctive.[10]

(12)   a.   Kì:-| ó |-rú.
           PST-2PL-sleep(IND)
           'You(pl) slept.'
      b.   Ó-mátʃ-é | ò: |-rú.
           2PL-want-IPFV 2PL-sleep.SBJV
           'You(pl) want to sleep.'

The specific syntactic and semantic environments in which the subjunctive is used will be discussed in more detail in section 5.4 (see also the Appendix), but we note here that in the first person singular, there is a morphological distinction between two types of subjunctive. An example can be seen in (13), where the lexical verb *ru* has a long vowel in its 1SG agreement prefix when it appears as the main verb in the second conjunct of a coordination (an environment that requires subjunctive in Kipsigis), but a short vowel when embedded under a volitional predicate (compare to the long vowel in (12-b) above). There is no such morphological difference for other person-number combinations. We will be calling the former type of subjunctive Type I, and the latter subjunctive Type II.

---

[8] While various tense and aspect distinctions are made in the indicative, only two aspect forms are distinguished in the subjunctive: the perfective and imperfective. It is also worth noting that Toweett (1979) and Rottland (1982) call this inflection of the verb *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).

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

[10] The exact shape of the subject agreement prefix, as well as the tonal melody of the stem, varies not only by mood, but also by the tense-aspect combination of the verb; it also depends on which conjugation class (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 (2021) for a complete description and sample conjugation paradigms.

(13) a. Kɑ-∅-puːtʃ Tʃéːbêːt kɑːt (ak) ⟨aː⟩-tʃɑp tʃaːɪːk.
  PST-3-sweep Cheebeet house and 1SG-make.SBJVI tea
  'Cheebeet cleaned the house and I made tea.'

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

We observe in (14) 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 (12). We use here 1SG forms in order to illustrate the two types of subjunctive (which are morphologically indistinguishable in other cells of the paradigm). In (14-a), we see a matrix use of *le*, in which case there is indicative inflection. In (14-b), *le* is used to introduce an embedded clause, and it has Type I subjunctive. Finally, in (14-c), "matrix" *le* is embedded under a volitional predicate and shows up with Type II subjunctive.

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

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

 c. á-mátʃ-é ⟨à⟩-lé kìː-∅-tʃɔ́ːr Kíbêːt rabɪːnɪk.
  1SG-want-ipfv 1SG-LE.SBJVII PST-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: it is inflected in the indicative when used in the root clause, but in the subjunctive when it is embedded under a matrix verb (i.e. in verbal complementation).

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 (15), the imperfective form of *le* 'say' has the form *leːlen*, which exhibits irregular stem allomorphy.[11] 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.

(15) **Leːlen** lòɣɔ́jwèːk kɔ̀ː-∅-tʃɔ́ːr Kíbêːt rabɪːnɪk.
  LE.IPFV(IND) news.NOM PST-3-steal Kibeet.NOM money
  'The news say that Kibeet stole the money'

In (15), *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

---

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

the subjunctive only make a perfective vs. imperfective distinction). We see in (16) that when the matrix verb is inflected in the past imperfective, *le* can be appear in either its perfective or imperfective form.[12]

(16) Kɑ-ɑ-mwɑ-e   àː-lé/ɑː-leːlen   kà-∅-tʃɔ́ːr Kíbêːt   rabɩːnɩk.
      PST-1SG-say-IPFV 1SG-LE/1SG-LE.IPFV PST-3-steal Kibeet.NOM money
      'I was saying that Kibeet stole the money.'

The third argument in favor of a verbal analysis of *le* 'say' comes from a reevaluation of the suffixal agreement data presented in (10), which are repeated below as (17). Diercks & Rao (2019) give a list of *le* forms with object agreement, shown in Table 2.

(17) ko-ɑ-mwaa-un   ɑ-lɛ-ndʒin   ko-∅-ɩt   tuɣa amut
      PST-1SG-tell-2SG.OBJ 1SG-C-2SG.OBJ PST-3-arrive cows yesterday
      'I DID tell you (sg) that the cows arrived yesterday.' (Diercks & Rao 2019: 371)

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

Table 2: Suffixal agreement (Diercks & Rao 2019: 381)

Looking at Table 2, 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 2 being decomposable into an allomorph of *le – leːn –*, followed by the applicative 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 2. The decomposition of the suffixal forms is given in Table 3, with surface phonological forms in brackets.

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

Table 3: Suffixal agreement decomposed into APPL and object clitics

The morphemes making up the forms in Table 3 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

---

[12]Only imperfective - the morphologically marked aspect in Kipsigis - will be indicated in the glosses.

thematic roles (e.g. recipient, beneficiary).[13,14] An example is given in (18).

(18)   a.   Kà-∅-tʃáp   Kíbêːt   kímɲéːt.
           PST-3-make Kibeet.NOM ugali
           'Kibeet made ugali (type of food).'
       b.   Kɑ-∅-tʃɑp-**tʃi**   Kíbêːt   Tʃèːbêːt   kímɲéːt.
           PST-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 4, built with data from Toweett (1979: p.209).[15]

	SG	**PL**
1	-an	-ɛːtʃ
2	-in	-aːk
3	∅	

Table 4: Object clitics

The last piece of the reanalysis is the claim that the verb *le* has an allomorph *leːn*. We have already seen in (15) that *le* shows a type of allomorphy involving vowel lengthening and the consonant [n] in other cells of the paradigm as well (in that case, the non-past imperfective), while Zwarts (2004) reports two similar allomorphs for the cognate word in the Kalenjin dialect Endo.

Further evidence for the presence of an applicative suffix on the complementizer comes from reflexives and reciprocals. Kipsigis has a verbal suffix *-kɛː* used to form reflexives and reciprocals, illustrated in (19) below.[16]

(19)   Ki-keːr-e-**kɛː**.
       1PL-look-IPFV-REFL
       'We are looking at ourselves/at each other.'

The suffix *-kɛː* can appear after the applicative *-tʃi*, in which case it takes scope over the applicative. With (at least) verbs of speech, when the applied argument position is occupied by *-kɛː*, suffixal agreement on *le* can include both the applicative and the reflexive/reciprocal suffix, as shown in (20).

(20)   Koː-∅-tʃɑːm-**tʃi-kɛː**   Kíbêːt   ko-leːn-**tʃi-kɛː**   ŋɑːm.
       PST-3-whisper-APPL-REFL Kibeet.NOM 3SG-LE-APPL-REFL.SBJVI clever

---

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

[14]The applicative *-tʃi* has an allomorph *-ʲi* when attached to verbs ending in an alveolar obstruent. It also has the allomorph *-u* for 1/2 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).

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

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

'Kibeet whispered to himself that he's clever.'

Summarizing, morphological 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, and it can host applicative and reflexive/reciprocal morphology (even when used in complementation). Before closing this section, however, 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 (21-a), but not when *le* is used in complementation, irrespective of whether there is matrix negation present, as shown in (21-b) and (21-c). Diercks et al. (2020) argue that the ungrammaticality of negation in complementation uses indicates that *le* is a complementizer, and not a verb.

(21)   a.   Mɑ-ɑ-le   ∅-rú-è   là:kwɛ̀:t.
             NEG-1SG-LE(IND) 3-sleep-IPFV child.NOM
             'I didn't say that the child is sleeping.'
       b.   Ma-a-mwa   (*mɑ-)ɑ:-le   ∅-rú-è   là:kwɛ̀:t.
             NEG-1SG-say   NEG-1SG-LE.SBJVI 3-sleep-IPFV child.NOM
             'I didn't say that the child is sleeping.'
       c.   Ka-a-mwa   (*mɑ-)ɑ:-le   ∅-rú-è   là:kwɛ̀:t.
             PST-1SG-say   NEG-1SG-LE.SBJVI 3-sleep-IPFV child.NOM
             'I didn't say that the child is sleeping.'

However, what (21) 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, evidence in favor of this view comes from data like (22) below. What we see in (22) is a lexical verb (inflected in the subjunctive) embedded under a matrix predicate, similar to the make-up of complementation structures with *le*. Interestingly, we observe in this case the same pattern as in (21) with respect to negation: the negative prefix *ma-* is ungrammatical when attached to the embedded verb, as shown in (22-b)-(22-c).[17] Thus, we see that there is a class of subjunctives in the language that does not tolerate negation. Whatever the reason for this might be, what data like (22) show is that unavailability of negation in complementation uses of *le* does not constitute an argument against its analysis as a verb.[18]

---

[17]The examples in (22) 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 (22) 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; as we show in the next two sections, 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.

[18]It 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".

(22)   a. â:-ŋgén   à-pír   pè:k.
            1SG-know 1SG-hit.SBJVII water
            'I know how to swim (*lit*: to hit water).'

       b. Mɑ-ɑ(:)-ŋgen (*mɑ)-à-pír   pè:k.
            1SG-know   NEG-1SG-hit.SBJVII water
            'I don't know how (not) to swim.'

       c. *â:-ŋgén   mɑ-à-pír   pè: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 (21-b)-(21-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) acknowledge that *le* behaves like a verb. In our analysis, on the other hand, *le* is uniformly a verb, and differences in behavior between matrix and complementation uses arise from differences in the syntactic position of *le* (matrix verb vs. embedded under another verb, reflected in its mood inflection), a view that is supported by data like (22).

    Finally, the analysis of *le* as a verb makes the prediction that it should in principle be compatible with adverbial modification. As can be see in (23), this prediction is borne out: the adverb *mu:tjɑ* 'slowly' appears after (the imperfective form of) *le*, which is the expected position if the adverb modifies *le*, but not if it modifies the matrix verb.

(23)   [Kɑ-∅-mwɑ-e   Kibe:t   [ko-le:len   **mu:tjɑ** [ka-∅-tʃɔ:r
         PST-3-say-IPFV Kibeet.NOM 3-LE.IPFV.SBJVI slowly PST-3-steal
         Kıplaŋgat   rabı:nık]]].
         Kiplangat.NOM money
         'Kibeet was saying slowly that Kiplangat stole the money.'

4.2 Agreement with the source of information

In section 3, we showed that *le* in Kipsigis does not always agree with the matrix subject. Rather, agreement with other DPs in the matrix clause is possible if those DPs act as the source of the information reported in the embedded clause (recall (6) and (7)). In this section we provide two further arguments in favor of the claim that agreement is sensitive to the source of information and four arguments in favor of treating the local subject of *le* 'say' as a pronoun that establishes co-reference with a matrix or discourse antecedent.

    First, agreement on *le* is subject to an animacy restriction, as shown by the contrast in (24). In both (24-a) and (24-b) there are two possible antecedents for agreement on *le*: the 1SG (animate) subject and a 3rd person source of information, introduced by the applicative *-e:n*. The source of information is animate in (24-a), but inanimate in (24-b), and what we observe is ungrammaticality of agreement with the source DP in the latter case. Interestingly, one of our consultants made the following comment: "*kole* is bad here [in (24-b)] because the door cannot talk and *kole* is for living things". This is in line with our arguments in favor of *le* being the lexical verb 'say'. [19]

---

[19]For some speakers, agreement with DPs denoting the source of information is not only sensitive to

(24) a. Ka-a-kas-ɛːn **Alice** àː-lé/**kò-lé** ka-kɔ-∅-ɪt
PST-1SG-hear-APPL Alice 1SG-LE/3-LE.SBJVI PST-PRF-3-arrive
làːgôːk.
children.NOM
'I heard from Alice that the children have arrived.'

b. Ka-a-kas-ɛːn **kurgeːt** àː-lé/***kò-lé** ka-kɔ-∅-ɪt
PST-1SG-hear-APPL door 1SG-LE/3-LE.SBJVI PST-PRF-3-arrive
làːgôːk.
children.NOM
'I heard from the door that the children have arrived.'

Second, *le* can agree with benefactive arguments introduced by the applicative *-tʃi*, but only if they can act as a the source of information. Thus, we see that agreement is possible in (7), repeated here as (25), but not in (26), where the benefactive argument of the predicate *kas* 'hear' cannot be construed as a source.

(25) Koː-ɑ-mwɑi-teː-tʃi **Tʃèbêːt** ɛːn tʊːjeːt àː-lé/**kò-lé** kɔ̀-∅-tʃɔ́ːr
PST-1SG-say-IT-APPL Cheebeet at meeting 1SG-LE/3-LE.SBJVI PST-3-steal
Kíbêːt rabɪːnɪk.
Kibeet.NOM money
'At the meeting, I said on Cheebeet's behalf that Kibeet stole the money.'

(26) *Kɑ-ɑ-kɑs-ʲi **Kìbêːt kò-lé** ∅-jàːtʃ-é kò-wá Nairobi.
PST-1SG-hear-APPL Kibeet 3-LE.SBJVI 3-must-IPFV 3-go(SBJV) 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 (27).[20]

(27) Ka-ɪ-kas [PP kobun **Kìplàŋgàt**] **kò-lé/** ìː-lé kà-∅-tʃɔ́ːr
PST-2SG-hear from Kiplangat 3-LE/ 2SG-LE.SBJVI PST-3-steal

Kíbêːt   rabɪːnɪk.
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 *kèː-lé* indicating a rumour interpretation, as we saw in (9), which receives a natural explanation if the pronominal subject co-refers with an impersonal antecedent in the discourse. 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 (28) for 3rd person and (29) for 2nd person.[21]

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

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

(29)   Context: *We are having an argument about who stole the money. **You** have presented convincing arguments that it is Kibeet who stole the money, and I say to **you**:*

ɑ-jɑːn-i   **ìː-lé**   kà-∅-tʃɔ́ːr Kíbêːt   rabɪːnɪk.
1SG-believe-IPFV 2SG-LE.SBJVI PST-3-steal Kibeet.NOM money
'I believe you that Kibeet stole the money.'

Since the contexts given in (27)-(29) do not ensure the necessary locality relations, we exclude an analysis involving a direct binding relation between the local subject and its antecedent. This leaves open the possibility for an account in which the subject is bound indirectly by a covert binder which itself is coreferent with the matrix antecedent. Such analyses are for example prominently pursued within the literature on long-distance reflexives (Anand & Hsieh 2005, Anand 2006, Charnavel 2020) and logophoric pronoun systems (Koopman & Sportiche 1989, Safir 2004, Speas 2004, Anand 2006). There is reason to doubt an anaphoric status of the subject. Since it serves as the goal for agreement with *le*, we would expect *Anaphor Agreement Effects* (Rizzi 1989, Woolford 1999, Sundaresan 2016, Murugesan 2020). As was shown in (19), reflexivization in Kipsigis takes place through the ϕ-invariant verbal suffix *-kɛː*, a strategy which is in complementary distribution with cliticization in non-anaphoric contexts, shown in (30).

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

Since this effect arguably qualifies as a case of *anaphoric* agreement (Woolford 1999: 264) and is absent with prefixal agreement on *lé*, we conclude that the subject does not instantiate an anaphor. Instead we propose that agreement takes place between *le* and a

---

[21]There is variation in our consultants' judgments regarding these examples. Three speakers find (29), but not (28), acceptable, while one speaker shows the opposite pattern accepting (28), but not (29).

covert pronoun introduced by *le*. This *pro* is coindexed with the matrix/discourse antecedent via the assignment function, thereby avoiding the need for c-command by the antecedent.

A final argument in favor of the *pro* analysis comes from the fact that the subject of *le* can be overtly realized under certain discourse conditions, shown in (31).

(31)   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:*

       Ka-a-mwa   à:-lé   **anɛ:**   kà-∅-tʃɔ́:r Kíbê:t
       PST-1SG-say 1SG-LE.SBJVI 1SG.NOM PST-3-steal Kibeet.NOM
       rabi:nik.
       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 analysis in the next section.

5   Analysis

In this section, we lay out the main components of the analysis in 5.2 and 5.3, that is the syntax and semantics of *le* complementation and the Type I subjunctive, including an introduction to an eventuality-based framework of attitude predicates in 5.1. Furthermore, we provide some first insights into an analysis of the Type II subjunctive in Kipsigis in 5.4.

5.1   Background on 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 (32-a). 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, see (32-b).

(32)   *Hintikkan vs. Davidsonian semantics*
    a. $[\![\text{ believe }]\!]^{w,g} = \lambda p \lambda x. \forall w' \in \text{DOX}_{x,w} : p(w')$
    b. $[\![\text{ believe }]\!]^{w,g} = \lambda p \lambda x \lambda e. believe(e) \wedge \text{EXP}(e,x) \wedge \forall w' \in \text{CONT}_{e,w} : p(w')$

As was shown in the previous sections, the Kipsigis *le* morpheme can inflect for aspect and come with adverbial modification, as can the matrix predicate, shown in (33). Thus,

we adopt an eventuality-based framework for attitude predicates in Kipsigis, for the matrix verb as well as the morpheme *le*.

(33) a. [Kɑ-ɑ-tʃɑːm-e   **muːtjɑ** [aː-leːlen   [ka-∅-tʃɔːr
          PST-1SG-whisper-IPFV slowly   1SG-LE.IPFV.SBJVI PST-3-steal
          Kíbêːt   rabɪːnɪk]]].
          Kibeet.NOM money
          'I was whispering slowly that Kibeet stole the money.'
      b. [Kɑ-ɑ-tʃɑːm-e   [aː-leːlen   **muːtjɑ** [ka-∅-tʃɔːr
          PST-1SG-whisper-IPFV 1SG-LE.IPFV.SBJVI slowly   PST-3-steal
          Kíbêːt   rabɪːnɪk]]].
          Kibeet.NOM money
          'I was whispering slowly that Kibeet stole the money.'

Given the addition of the eventuality argument, others have pursued a full thematic separation of the eventuality argument and the content function, where the latter is introduced as a silent modal, a mood particle, or a complementizer in the left periphery of the embedded clause (Kratzer 2006, 2016, Moulton 2009, 2019, Bogal-Allbritten 2015, Grano 2016, Özyıldız et al. 2018, Demirok et al. 2020).[22] This move shifts the semantic action from the attitude predicate to the complementizer and/or mood of the embedded clause, compare (32-b) to (34-a) and (34-b). Attitude predicate and embedded clause combine via predicate modification in most approaches. Another consequence of the Neo-Davidsonian approach is that the attitude holder, for *believe* the experiencer, will be introduced by a separate head.

(34) *Neo-Davidsonian semantics*
      a. ⟦ *believe* ⟧^{w,g} = $\lambda e.believe(e)$
      b. ⟦ MOOD/COMP/MODAL ⟧^{w,g} = $\lambda p \lambda e. \forall w' \in \text{CONT}_{e,w} : p(w')$

We will adopt this separation, as it allows us to model semantic differences of the interpretation of the attitude predicate coming from the complement clause, which will become relevant in section 5.4.

5.2 *le*-clauses as sets of contentful eventualities

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. We model this as in (36), where the embedded proposition (an indicative CP) is a sister to the verb *le* 'say', which itself is part of a subjunctive TP embedded under the matrix predicate. In this section, we focus on the semantics of these structures, and we further discuss the syntactic choices in 5.3.

   We propose that embedded clauses headed by agreeing forms of *le* constitute sets of

---

[22]Another line of approach takes the content function to be encoded via an additional thematic role (Elliott 2016, 2017, Bondarenko 2020, Portner & Rubinstein 2020).
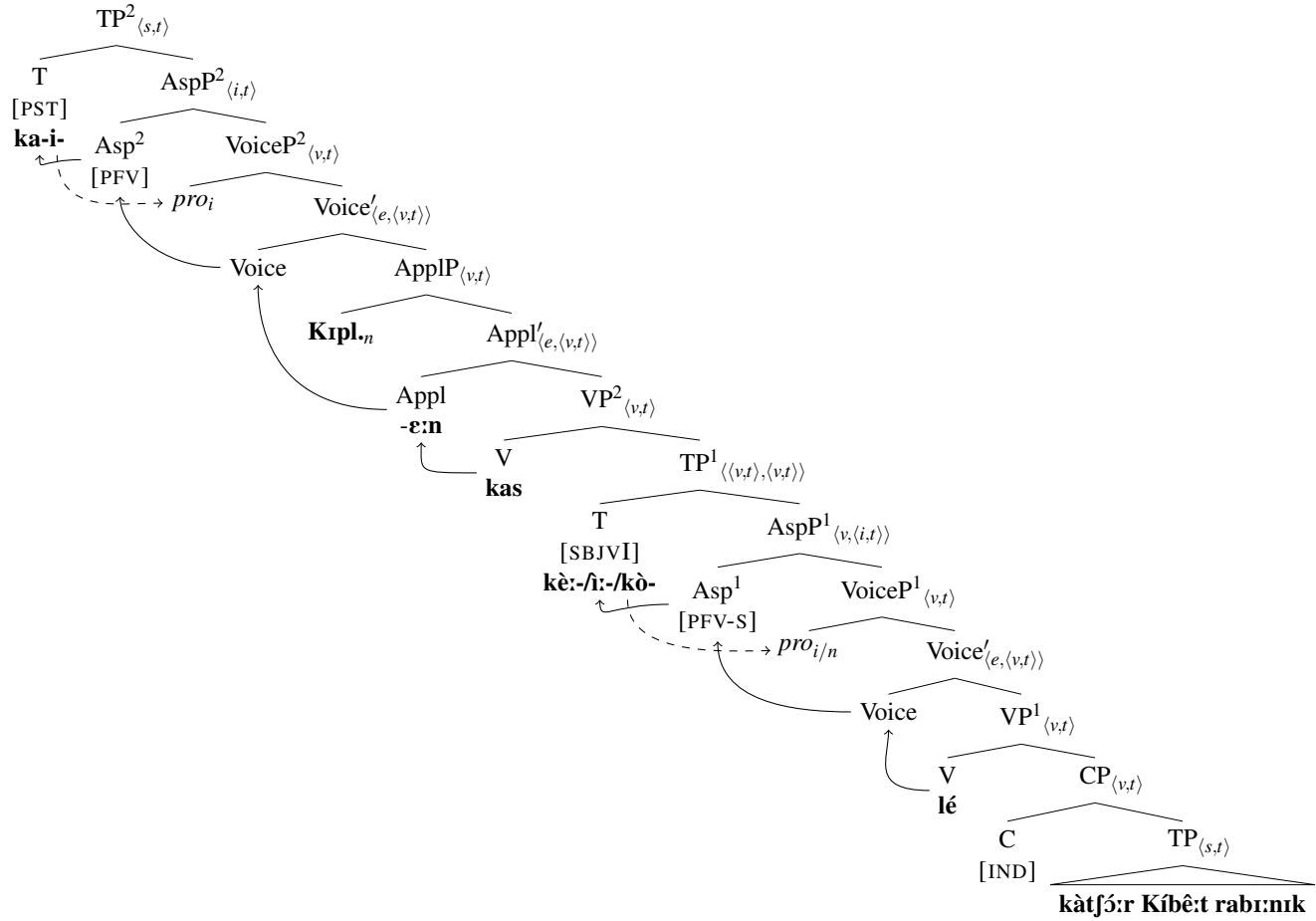
contentful saying events, where the verbal nature of *le* 'say' is reflected in its semantics. In order to do so, we adopt an eventuality-based framework where the relation between the attitude holder and the proposition is mediated by *contentful eventualities*. Let us introduce the main analysis by illustrating our proposal with the structure in (36), based on the example in (35). 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.

(35) Ka-i-kas-ɛːn   Kìplàŋgàt kèː-lé/ ìː-lé/ kò-lé   kà-∅-tʃɔ́ːr
      PST-2SG-hear-APPL Kiplangat IMP-LE/ 2SG-LE/ 3-LE.SBJVI PST-3-steal
      Kíbêːt   rabɪːnɪk.
      Kibeet.NOM money
      'You heard from Kiplangat that Kibeet stole the money.'

We assume that Voice introduces the external argument of *le* (Kratzer 1996), while the source argument of the matrix predicate *kas* 'hear' enters the derivation via a high Appl head (Marantz 1993, Pylkkänen 2008).[23] 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 (36). The dashed arrows indicate Agree between T and the subject, respectively. Of special interest is the subjunctive T head probing for the ϕ-features of the agent of the saying event – a free pronoun serving as a goal for Downward Agree. Prefixal agreement on *le* follows straightforwardly, as the ϕ-features of *pro* vary with its denotation. The form *iːle* is chosen if *pro* points to the addressee of the utterance, whereas *kòlé* appears if *pro* is co-indexed with *Kìplàŋgàt*, that is the source argument from the matrix clause. Another option is the impersonal form *kèːlé* which leads to a rumour interpretation, recall (9). In this case, *pro* co-refers with an impersonal antecedent in the discourse.

---

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

(36)   Analysis of (35):

$\text{TP}^2_{\langle s,t\rangle}$

T [PST] **ka-i-**

$\text{AspP}^2_{\langle i,t\rangle}$

Asp^2 [PFV]

$\text{VoiceP}^2_{\langle v,t\rangle}$

pro_i

Voice

$\text{Voice}'_{\langle e,\langle v,t\rangle\rangle}$

KIpl.$_n$

$\text{ApplP}_{\langle v,t\rangle}$

Appl **-ɛ:n**

$\text{Appl}'_{\langle e,\langle v,t\rangle\rangle}$

V **kas**

$\text{VP}^2_{\langle v,t\rangle}$

T [SBJVI] **kè:-/ì:-/kò-**

$\text{TP}^1_{\langle\langle v,t\rangle,\langle v,t\rangle\rangle}$

Asp^1 [PFV-S]

$\text{AspP}^1_{\langle v,\langle i,t\rangle\rangle}$

$pro_{i/n}$

Voice

$\text{VoiceP}^1_{\langle v,t\rangle}$

V **lé**

$\text{Voice}'_{\langle e,\langle v,t\rangle\rangle}$

C [IND]

$\text{VP}^1_{\langle v,t\rangle}$

$\text{CP}_{\langle v,t\rangle}$

$\text{TP}_{\langle s,t\rangle}$

kàtʃɔ́:r Kíbê:t rabɪ:nɪk

We will now present the key points of the semantic analysis. We choose to illustrate the *kòlé* derivation, where *pro* is co-indexed with the applied argument *Kɩplàŋgàt*. Since *le* is not a complementizer but a verbal category, it introduces a saying eventuality, see (37-a). Following Hacquard (2006), we assume that speech and attitude eventualities have propositional content, that is they define sets of possible worlds. With Kratzer (2006), we fully decompose attitude eventuality and propositional content. A content function CONT, introduced by C in (37-b), takes eventualities and outputs sets of worlds compatible with that eventuality in which Kibeet steals the money. Both (37-a) and (37-b) combine via Predicate Modification, resulting in a set of saying events the content of which is that Kibeet steals the money (37-c). This decomposition 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). 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 (23)/(33-b).

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

in w'

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

The next two points concern the aspectual information and the implementation of the subjunctive on *le* in (35). Given that *le* can show aspect morphology, as was shown in (16), we include an AspP layer in (36). In general, subjunctive Type I connects clauses, as it is also used more widely in coordinate clauses, recall example (13-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 (38) for perfective and imperfective aspect. In unembbeded scenarios, that is in indicative clauses, the standard account can be adopted. Hence, the denotation in (38-a) can be directly taken to be encoded by Asp^2 in (36).[24]

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

  a. $[\![\text{PFV}]\!] = \lambda P_{\langle v,t \rangle} \lambda t. \exists e[\tau(e) \subseteq t \wedge P(e)]$ Asp^2 in (36)

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

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 (38). Grano (2020, 2021) proposes to define variants of the aspectual morphemes in such cases, see (39), where the eventuality argument is passed up rather than existentially closed off, as in (38). We will adopt this idea and take (39-a) to be encoded by Asp^1, resulting in the denotation in (40) for AspP^1.

(39) Aspect under subjunctive (Grano 2020)

  a. $[\![\text{PFV-S}]\!] = \lambda P_{\langle v,t \rangle} \lambda e. \lambda t[\tau(e) \subseteq t \wedge P(e)]$ Asp^1 in (36)

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

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

The entry for subjunctive Type I is provided in (41-a) combining the saying events in (40) with the hearing events in (41-b), the result of which is the denotation of VP^2, shown in (41-c). The lexical entry in (41-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.

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

---

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

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

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

Finally, both the experiencer and the source of the hearing event are added via Event Identification in the matrix clause, resulting in the denotation in (42).

(42) $[\![\ \text{VoiceP}^2\ ]\!]^{w,g} = \lambda e''.\exists e'\exists t[\text{CAUSE}(e',e'') \wedge \tau(e') \subseteq t \wedge say(e') \wedge \text{AG}(e') = g(n) \wedge \forall w' \in \text{CONT}(e') : \text{Kibeet steals the money in } w' \wedge hear(e'') \wedge \text{SOURCE}(e'') = kiplangat_n \wedge \text{EXP}(e'') = g(i)],$
defined iff $g(i)$ is addressee[25]

The CAUSE function is bidirectional, where direction is resolved by context. In (42), 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. 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 *ìːlé* derivation of (35). In this case, the hearing event causes the saying event to take place, as the adressee 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. The two readings are also attested for Turkish where Özyıldız et al. (2018) identify the former reading as a speech report and the latter reading as an attitude report. The analysis presented for (35) can be extended to a variety of attitude verbs shown to combine with *le*-clauses throughout this paper, most of which will receive the more abstract attitude report reading. Examples of such verbs that appear with *le*-clauses in our fieldnotes include *jaːn* ' to believe', see (43), *bwaːt* 'to think/remember', *taːm* 'to (falsely) accuse', *ɲoːn* 'to complain', *naj* 'to realize', *ruaːtit* 'to dream', *raːgin* 'to worry', *pajpaj* 'to be happy', *nɛrɛːtʃ* 'to be angry'.

(43) αː-jαːn-i   αː-le   mʊ́gɔ̂l ŋwɔ̀ɲ.
     1SG-believe-IPFV 1SG-LE.SBJV I round earth.NOM
     'I believe that the Earth is round.'

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' or a 'thought event', whereas Major (2021) classifies this reading as stative SAY.[26] Interestingly, for perception verbs

---

[25]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.

[26]Major (2021) makes a distinction between eventive and stative SAY in languages with SAY-

such as in (35), where both the speech event and the attitude 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 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.

Let us now turn to suffixal agreement, which was discussed in section 4.1. We provide another example in (44) where *le* not only shows prefixal agreement with the matrix subject but also an object clitic introduced by APPL.

(44) Kɑ-∅-tʃɑːm-u-ɑn   Tʃéːbêːt   ko-leːn-tʃ(i)-ɑn   kà-∅-tʃɔ́ːr
      PST-3-whisper-APPL-1SG Cheebeet.NOM 3-LE-APPL-1SG.SBJVI PST-steal
      Kíbêːt   rabɪːnɪk.
      Kibeet.NOM money
      'Chebeet 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 (45).

(45)
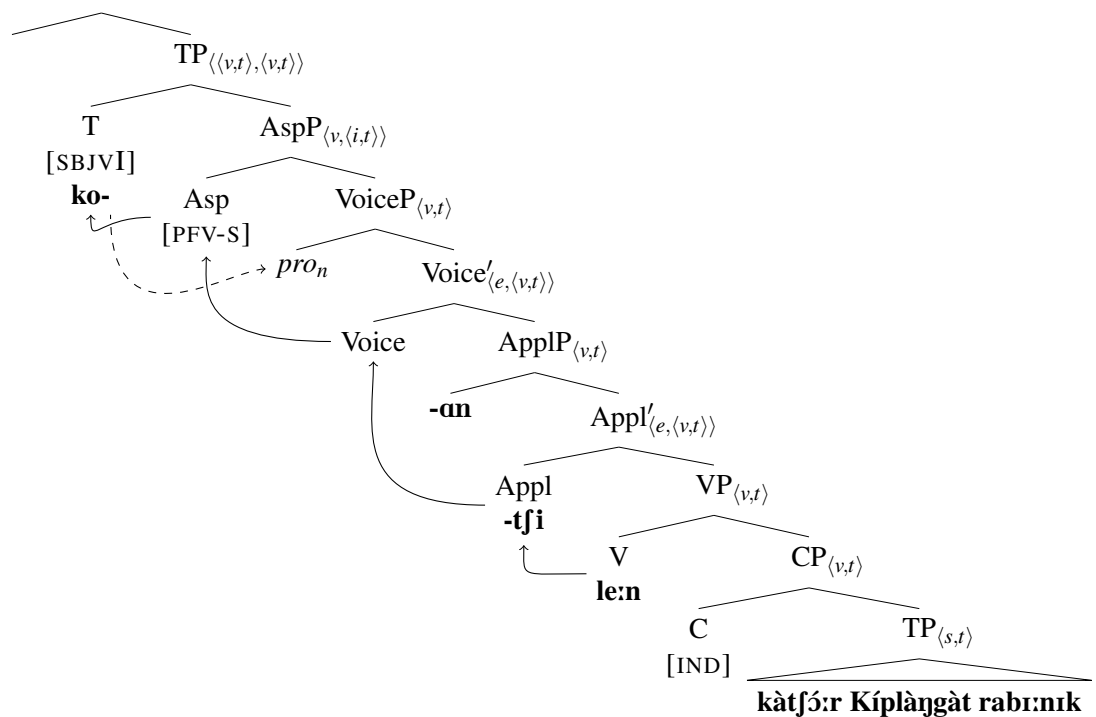
We provide the denotation of matrix VoiceP in (46). Suffixal agreement, decomposed into APPL and a 1SG object clitic, introduces a goal argument for the embedded saying

---

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.

event, matching the goal argument of the matrix whispering event. Since the subject of *le* co-refers with the matrix subject, the sentence receives an attitude report reading.

(46) $[\![$ *matrix VoiceP in (45)* $]\!]^{w,g}$
$= \lambda e''.\exists e'\exists t[\text{CAUSE}(e',e'') \wedge \tau(e') \subseteq t \wedge say(e') \wedge \text{AG}(e') = g(n) \wedge \text{GOAL}(e') = g(i) \wedge \forall w' \in \text{CONT}(e') : \text{Kiplangat steals the money in } w' \wedge whisper(e'') \wedge \text{GOAL}(e'') = g(i) \wedge \text{AG}(e'') = Kibeet_n]$, 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 *ɲoːɲ* 'complain' and *siːr* 'write', shown in (47) and (48).

(47) Koː-a-ɲoːɲ   α:-leːn-**tʃi**   Kibeːt koː-jαːtʃ-eːn àmìtwáːgík.
      PST-1SG-complain 1SG-LE-APPL.SBJVI Kibeet PST-bad-PL food.NOM
      'I complained to Kibeet that the food was bad.'

(48) Koː-a-siːr   α:-leːn-**tʃi**   Tʃèːbêt a-tʃεlεwani.
      PST-1SG-write 1SG-LE-APPL.SBJVI Chebet 1SG-be.late
      'I wrote to Chebet that I will be late.'

This section provided our main account of *le* as a clausal embedder under attitude predicates. In the next section, we will address the syntactic claims about the clause size we have made so far.

5.3 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 (36). In this section, we provide empirical arguments in favor of these choices.

    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 over topic marker *ko* (Driemel & Kouneli 2021), illustrated in (49). Following previous work on Nilotic (van Urk 2015), we assume that the pre-verbal topic position in Kipsigis is SpecCP.

(49) Kìbêːt **kó**   kà-∅-ám kímɲéːt.
      Kibeet TOP PST-3-eat ugali
      'Kibeet ate ugali.'

As shown in (50), *le* can introduce clauses with an overt topic marker, indicating that the embedded clause is a CP.

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

A similar argument can be made on the basis of embedded questions. Kipsigis is gen-

erally wh-in-situ, as shown in (51), which displays the standard VSO order.[27]

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

We see in (52) 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.

(52) Mɑ-ɑ(ː)-ŋgen àː-lé [kà-∅-tʃɔ́ːr **ŋàː** rabɪːnik].
      NEG-1SG-know 1SG-LE.SBJVI PST-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 (36) 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 (31)) 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 (53) that the subject of *le* cannot be topicalized (in contrast to the subject of the embedded proposition, shown in (49) above).

(53) *Ka-a-mwa [anɛː ko àː-lé [kà-∅-tʃɔ́ːr Kíbêːt rabiːnik]].
      PST-1SG-say 1SG TOP 1SG-LE.SBJVI PST-3-steal 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).[28, 29] According to the same work, however, the languages under investigation (Greek, Romanian, and Ndebele) distinguish between CP and TP subjunctives. As was already hinted at in in section 4.1,

---

[27]However, 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.

[28]*le* does not show tense distinctions and it is incompatible with negation, as discussed in 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.

[29]In their grammar of the related dialect Nandi, Creider & Creider (1989) claim that full clauses cannot be coordinated in the language, and we have already seen in (13-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 the Kipsigis, examples such as (13-a) could be interpreted as evidence for the lack of a C layer in subjunctive Type I.

Kipsigis does distinguish between two types of subjunctive, which are morphologically different in the first person only. The discussion so far has focused on what we called Type I subjunctive, which is the inflectional form of *le* when used in complementation. In the next section, we provide a brief discussion of the two types of subjunctive.

5.4 More on the subjunctive

As has already been mentioned, Kipsigis lacks infinitives, which means that the subjunctive is widely used in complementation contexts. 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 (indirectly) select for indicative complements, and those that can select either. Representative examples are provided in (54)-(56) below.[30] The crucial difference between Kipsigis and European languages lies in the "indirect" nature of indicative selection: under our analysis, verbs select for a *le*-clause (which is of Type I subjunctive) which then introduces the indicative complement. 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).[31] For those verbs that can select for a subjunctive complement, see (54) and (56), subjunctive is always of Type II. This is the most common use of subjunctive Type II. The difference in mood choice in (56) is reflected in the interpretation of the matrix predicate, where *le* + indicative triggers an assertive/reportative reading of the complement clause and subjunctive II conveys a directive meaning.

(54)   Subjunctive only: *matʃ* 'want'
   á-mátʃ-é   à-rú.
   1SG-want-IPFV 1SG-sleep.SBJVII
   'I want to sleep.'

(55)   *Le* + indicative only: *jaːn* 'believe'
   aː-jaːn-i   àː-lé   ∅-rú-è   Kíbêːt.
   1SG-believe-IPFV 1SG-LE.SBJVI 3-sleep-IPFV Kibeet.NOM
   'I believe that Kibeet is sleeping.'

(56)   Either subjunctive or *le* + indicative: *tʃaːm* 'whisper'
   a.   Ka-∅-tʃaːm   Kíbêːt   ko-le   ka-∅-tʃap kímɲéːt.
        PST-3-whisper Kibeet.NOM 3-LE.SBJVI PST-3-make ugali
        'Kibeet whispered that he made ugali.'

---

[30]While a complete investigation of the lexical semantics of the verbs that select for subjunctive vs. *le* + indicative is beyond the scope of the paper, our data so far point towards significant similarities between Kipsigis and European languages, as can already be seen in (54)-(56).
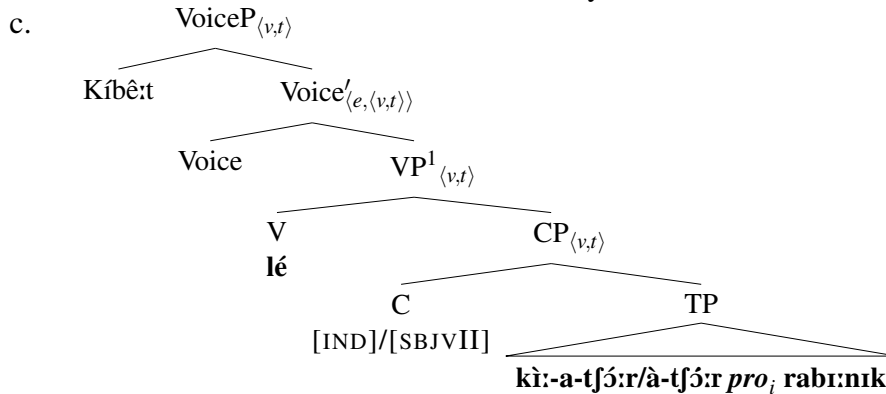
[31]Creider & Creider (1989: 129-131) only mention two strategies of clausal complementation in the related dialect Nandi, which coincide with the picture presented in this paper: subjunctive or *le* + indicative (Diercks & Rao 2019 make a similar claim, arguing that *le* cannot be dropped in Kipsigis). Bossi (2022), however, reports that the lexical predicate *par* 'think (with negative bias)' can directly select for indicative complements. Our analysis does not rule out lexical predicates other than *le* selecting for indicative, and it is left as a question for further research how widespread direct indicative selection is in the language.

b. Ka-∅-tʃaːm-u-an   Kíbêːt   à-tʃáp   kímɲéːt.
PST-3-whisper-VENT-1SG Kibeet.NOM 1SG-make.SBJVII ugali
'Kibeet whispered to me to make ugali.'

Additional data are needed to determine whether the difference between Type I and Type II subjunctives is one of size (TP vs. CP, respectively), as has been argued for different types of subjunctives in other languages (e.g. Alexiadou et al. 2012). For presentation purposes we will be assuming that subjunctive Type I is a TP and subjunctive Type II is a CP.[32] It should be stressed, however, that nothing hinges on this choice: the semantic analyses provided for *le*-clauses and the two types of subjunctives do not depend on the T/C distinction.

We will now extend our account of the Type I subjunctive to matrix predicates that are able to select for subjunctive II clauses and *le* + indicative clauses, as shown in (56). Since the focus of this paper is on the verb *le*, our discussion will be based on *le* in matrix uses, see (57). Our analysis of Type I subjunctive, which occurs with *le* in subordinate position throughout this paper, includes event abstraction and a causal component, recall section 5.2. The Type II subjunctive is morphologically distinct from the Type I subjunctive only in the first person singular in terms of vowel length, i.e. a short vowel for Type II and a long vowel for Type I. The semantic contrast in (56) can also be shown for *le* in (57) where *le* acts as a matrix predicate. In (57-a), *le* provides a reportative flavour, whereas (57-b) conveys intention-based/directive semantics. Both are subsumed structurally with the tree in (57-c).

(57)   a.   Kìː-∅-lé   Kíbêːt   kìː-a-tʃɔ́ːr   rabıːnɪk.
            PST-3-LE Kibeet.NOM PST-1SG-steal(IND) money
            'Kibeet said that I stole the money.'
   b.   Kìː-∅-lé   Kíbêːt   à-tʃɔ́ːr   rabıːnɪk.
            PST-3-LE Kibeet.NOM 1SG-steal(SBJVII) money
            'Kibeet said that I should steal the money.'
   c.

VoiceP$_{\langle v,t \rangle}$

Kíbêːt   Voice'$_{\langle e,\langle v,t \rangle \rangle}$

     Voice   VP^1$_{\langle v,t \rangle}$

         V   CP$_{\langle v,t \rangle}$
         lé

           C   TP
           [IND]/[SBJVII]

           kìː-a-tʃɔ́ːr/à-tʃɔ́ːr *pro$_i$* **rabıːnɪk**

The data presented in the previous sections has shown that *le* selects for indicative CPs, independent of whether *le* is in matrix or subordinate position. Following the account

---

[32]A question that might arise given these categories is how subjunctive morphology is calculated. 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: it arises in configurations with deficient inflectional features (e.g. a deficient T).

presented for (35), the analysis for (57-a) is straightforward. We provide the denotation for the VoiceP projected by *le* in (58). Since *le* selects for an indicative CP like in (35), the VoiceP denotations in (58) and (37-d) are essentially identical except for the different agents. In prose, (58) 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.

(58) $⟦$ *matrix VoiceP in (57-a)* $⟧^{w,g}$
$= \lambda e_v.say(e) \wedge \text{AG}(e) = kibeet \wedge \forall w' \in \text{CONT}(e) : g(i)$ steals the money
in w', defined iff $g(i)$ is speaker

We will now turn to the analysis of Type II subjunctive. As shown in (57), the choice of mood in the embedded clause affects the interpretation of *le*. The indicative C head in (57-a) triggers the semantics in (58), which leads to a reportative reading. For the analysis of (57-b) where the C head comes with Type II subjunctive mood, we follow Grano (2020, 2021) in arguing that the mood type introduces causally self-referential content, which results in the directive speech act flavour when combined with 'say'. This split in interpretation is reminiscent of Romance languages where indicative morphology in the clause embedded under 'say' triggers a reportative reading and subjunctive morphology a directive reading, see e.g. discussions in Ahern & Leonetti (2004) and Farkas (1992) for Spanish and Romanian. Key to our analysis is the content function encoded by the C head that comes with Type II subjunctive. The denotation is given in (59-a). As with Type I subjunctive, the analysis contains a causal component which in turn requires abstraction over eventualities (and times), indicating that the aspect denotations in (39) are at play in the embedded clause. In (59-b), we provide the denotation of the embedded TP. Combined with (59-a), we arrive at the meaning in (59-c).

(59) a. $⟦$ SBJVII $⟧^{w,g} = \lambda P_{⟨v,⟨i,⟨s,t⟩⟩⟩} \lambda e'.\forall⟨w',t'⟩ \in \text{CONT}(e') : \exists e'' \text{CAUSE}(e',e'') :$
$\exists t > t' \wedge P(e'')(t)(w')$

b. $⟦$ *TP in (57-b)* $⟧^{g}$
$= \lambda e \lambda t \lambda w [\tau(e) \subseteq t \wedge steal(e,w) \wedge \text{AG}(e,w) = g(i) \wedge \text{TH}(e,w) =$
$\iota x.money(x)]$, defined iff $g(i)$ is speaker

c. $⟦$ *CP in (57-b)* $⟧^{w,g}$
$= \lambda e'.\forall⟨w',t'⟩ \in \text{CONT}(e') : \exists e'' \text{CAUSE}(e',e'') : \exists t > t' \wedge \tau(e'') \subseteq t$
$\wedge steal(e'',w') \wedge \text{AG}(e'',w') = g(i) \wedge \text{TH}(e'',w') = \iota x.money(x)]$,
defined iff $g(i)$ is speaker

As always, *le* denotes a set of saying events, which combines with the CP via Predicate Modification. The addition of the agent via a Voice layer results in the denotation given in (60).

(60) $⟦$ *matrix VoiceP in (57-b)* $⟧^{w,g}$
$= \lambda e.say(e) \wedge \text{AG}(e) = kibeet \wedge \forall⟨w',t'⟩ \in \text{CONT}(e) : \exists e'' \text{CAUSE}(e,e'') :$
$\exists t > t' \wedge \tau(e'') \subseteq t \wedge steal(e'',w') \wedge \text{AG}(e'',w') = g(i) \wedge \text{TH}(e'',w') =$
$\iota x.money(x)]$, defined iff $g(i)$ is speaker

We can paraphrase (60) as a set of saying events whose agent is Kibeet and for all world-time pairs $⟨w',t'⟩$ compatible with the content of such saying events there is a money-stealing event in the future of t' whose agent is the speaker and which is caused

by the saying event.[33] Note that the Neo-Davidsonian approach allows us to have only one lexical entry for *le*, while the different semantic flavours come about via the choice of mood, i.e. the selection of indicative vs. Type II subjunctive on the one hand and the possibility of Type I subjunctive inflection for *le* on the other hand.

We have now provided an analysis for different types of subjunctive as those relate to the verb *le* 'say', the focus of our paper. The subjunctive appears in more environments in Kipsigis (e.g. after bouletic predicates, see (54)), which are summarized in the Appendix. How our theory extends to the other uses of the subjunctive is left as a topic for further research.

6 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. The analysis pursued in this paper also highlights that 'say'-based complementizers might be of category V, and not C, in more languages than previously thought (e.g. Koopman 1984, Koopman & Sportiche 1989, Major 2021), which means 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 2020a,b) instead.

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. In other words, *le* denotes a set of eventualities like any other verb.

References

Ahern, A. & M. Leonetti. 2004. The Spanish subjunctive: Procedural semantics and pragmatic inference. In R. Márquez-Reiter & M. E. Placencia (eds.), *Current Trends in the Pragmatics of Spanish*, 35–56. Amsterdam: John Benjamins.

van Alem, Astrid. 2020a. Complementizer agreement in West-Germanic is clitic doubling. Ms., Leiden University.

van Alem, Astrid. 2020b. Complementizer agreement is not allomorphy: A reply to Weisser (2019). *Glossa* 5. 44.

---

[33]We also adopt Grano's (2020, 2021) notation of $>$ as a relation between two times t and t' iff t is later than t'.

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 & 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.

Baker, Mark. 2008. *The Syntax of Agreement and Concord*. Cambridge: Cambridge University Press.

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. 2022. Negative bias and pragmatic reasoning in Kipsigis belief reports. Handout from talk given at MIT LingLunch.

Bossi, Madeline & Michael Diercks. 2019. V1 in Kipsigis: Head movement and discourse-based scrambling. *Glossa: a journal of general linguistics* 4. 65.

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 Resher (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. A dedicated topic position in Kipsigis. Handout from talk given at TripleA8 at the University of Singapore.

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.

Farkas, Donka. 1992. On the semantics of subjunctive complements. In P. Hirschbuhler & K. Koerner (eds.), *Romance Languages and Modern Linguistic Theory: Papers from the 20th Linguistic Symposium on Romance Languages*, 69–104. Amsterdam/Philadelphia: John Benjamins.

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.

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. Cherono. 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.

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.

Kinyalolo, Kasangati Kikuni W. 1993. On some syntactic properties of ɖɔ̀. *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. 2021. Low (in)transitivity: Evidence from Kipsigis. In Alessa Farinella & Angelica Hill (eds.), *Proceedings of the Fifty-First Annual Meeting of the North East Linguistic Society*, 53–66. Amherst, MA: GLSA.

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 Jacques Moeschler Stephen R. Anderson & Fabienne Reboul (eds.), *L'interface Langage-Cognition/The Language-Cognition Interface: Actes du 19e Congrès International des Linguistes Genève*, 179–199. Librarie Droz.

Kratzer, Angelika. 2016. Evidential moods in attitude and speech reports. Slides, available at https://works.bepress.com/angelika$_k$ratzer/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.

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, Illinoise: 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. Handout from talk given at DGfS 2022 at Universität Tübingen.

Major, Travis & Harold Torrence. 2020. "Say"-chains, not "say"-complementation. To appear in Proceedings of WCCFL 38.

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.

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. 2020. Deriving the Anaphor Agreement Effect and its Violations. To appear in *Syntax*, lingbuzz/005135.

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.

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 inmood 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 structuresy*, 58–75. Oxford: Oxford University Press.

Toweett, Taaitta. 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.

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 5.4), Type II subjunctive is also used for purpose clauses and after modals.

(61)   Purpose clauses (Toweett 1979: 199)

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

    b.   (ɑ)si àː-kát
         so   1SG-greet.SBJVII
         'so that I greet'

(62)   Modals (Toweett 1979: 225)

    a.   ∅-ɲɑːl-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 (63), 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 (64).

(63)   Iŋgot àː-rú...
       if   1SG-sleep(SBJVI)
       'if I sleep...'

(64)   Koː-∅-tʃap-e   kímɲéːt Tʃeːbeːt   koːn ɑ-ɲoː   kɑːt.
       PST-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 hinted at in 4.1). As shown in (65), 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.

(65)   Kɑ-∅-puːtʃ Tʃéːbêːt kɑːt   (ak) ɑ-tʃap   tʃaːɪːk.
       PST-3-sweep Cheebeet house   and 1SG-make.SBJVI tea
       'Cheebeet cleaned 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. We leave a complete investigation of verbal mood

in the language as a topic for further research.