

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

Imke Driemel & Maria Kouneli

January 3, 2021

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 logophoric 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, Koopman & Sportiche 1989, Major & Torrence 2020), which means that some instances of what has been described as C-Agree may instantiate standard verbal agreement.

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. Chomsky 2000, 2001), upwards-oriented complementizer agreement raises a number of questions about the directionality and locality of Agree, with some studies arguing

*We are grateful to Enock Kirui, Wesley Kirui, Hillary Mosonik, Philemon Ronoh, and Nathan Rotich for their valuable work as linguistic consultants. We’d also like to thank Mike Diercks, Doreen Georgi, Travis Major, Deniz Özyıldız, Malte Zimmermann, and the audiences at the University of Potsdam, NYU, Universität Leipzig, and the *Brussels Conference on Generative Linguistics 13* for useful feedback.

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

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 2020).

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. \hat{u} :-ŋgén **â**:-lé \emptyset -rú-è Kíbê:t.
 1SG-know 1SG-C 3-sleep-IPFV Kibeet.NOM
 ‘I know that Kibeet is sleeping.’
- b. \hat{i} :-ŋgén **î**:-lé \emptyset -rú-è Kíbê:t.
 2SG-know 2SG-C 3-sleep-IPFV Kibeet.NOM
 ‘You know that Kibeet is sleeping.’
- c. í-ŋgèn Kíplàngàt **kò**-lé \emptyset -rú-è Kíbê:t.
 3-know Kiplangat.NOM 3-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 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 logophoric center. We therefore present an analysis according to which what looks like C-Agree in (1) is an instance of logophoric agreement between the lexical verb ‘say’ and its locally introduced (often covert) subject. Downward Agree can straightforwardly account for instances of subject-verb agreement, and our analysis thus solves the locality and directionality problems posed by the (apparent) upwards-oriented nature of C-Agree. We also provide a semantic analysis in which the verbal category of the “complementizer” is reflected in its semantics, building on recent eventuality-based models of attitude and speech reports (e.g. Kratzer 2013, Elliott 2016, 2017, Moulton 2019).

The remainder of the paper is structured as follows: in Section 2, we provide a brief

²Kipsigis is the major variety of Kalenjin, a cluster of dialects of the Southern Nilotic branch of Nilo-Saharan, and it is spoken by about 2 million people in Western Kenya (Eberhard et al. 2020). Unless indicated otherwise, data in this paper come from the authors’ fieldwork. The authors had a series of Skype elicitations in the last year 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.

³Glossing 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.


overview of previous theories of upwards-oriented complementizer agreement, before presenting the Kipsigis pattern in Section 3. We then develop our analysis in three steps in Section 4: we argue that the Kipsigis “complementizer” is the lexical verb ‘say’ in 4.1, we provide arguments in favor of logophoricity as the relevant factor driving agreement in 4.2, and we provide a complete analysis in 4.3. In Section 5, we conclude and discuss avenues for further research.

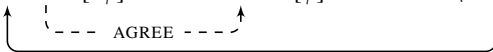
2 Previous theories of upwards-oriented C agreement

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 2020). 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 2020), 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*

$$[{}_{\text{VP}} \text{SUBJECT}_{[\phi]} \dots [{}_{\text{ForceP}} \text{FORCE}_{[u\phi]} \dots [{}_{\text{FinP}} \dots [{}_{\text{TP}} \text{SUBJECT} \dots]]]]$$

- (3) *Downward Agree account*

$$[{}_{\text{VP}} \text{FORCE}_{[u\phi]} [{}_{\text{VP}} \text{SUBJECT}_{[\phi]} \dots [{}_{\text{ForceP}} \langle \text{Force} \rangle \dots [{}_{\text{FinP}} \dots [{}_{\text{TP}} \text{SUBJECT} \dots]]]]]]$$


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 (2020) 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. \hat{a} :-ŋgén à:-lé Ø-rú-è Kíbê:t.
 1SG-know 1SG-LE 3-sleep-IPFV Kibeet.NOM
 ‘I know that Kibeet is sleeping.’
 b. Kà-ŋmwá ò:-lé Ø-rú-è Kíbê:t.
 PST-2PL-say 2PL-LE 3-sleep-IPFV Kibeet.NOM
 ‘You(pl) said that Kibeet is sleeping.’
 c. Kí:-ŋgèn kè:-lé Ø-rú-è Kíbê:t.
 IMP-know IMP-LE 3-sleep-IPFV Kibeet.NOM
 ‘It is known that Kibeet is sleeping.’ (impersonal)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) α -ŋgɛn *(α -le/kɔlɛ) ko-Ø-ruuja tuya amut
 1SG-know 1SG-C/that PST-3-sleep cows yesterday
 ‘I know (that) the cows slept yesterday.’ ([Diercks & Rao 2019: 372](#))

The 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 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 with verbs from a variety of lexical classes (e.g. *ja:n* ‘to believe’, *mwa* ‘to say’, *ruatit* ‘to dream’, *ta:m* ‘to falsely accuse’, *neretf* ‘to be angry (about)’).

It is clear from our data, however, that the complementizer may agree with non-

⁴The 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. In the subjunctive of CV verbs (such as *le*) there is no tonal difference between 1/2nd and 3rd person forms.

⁵Our [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 \(2019\)](#).

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

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

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

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 Kíplàngàt kobun **ìpê:** kò-lé/ì-lé kà-∅-tjór Kíbê:t
 PST-3-hear Kiplangat.NOM from 2SG 3-LE/ 2SG-LE PST-3-steal Kibeet.NOM
 rabɪ:nɪk.
 money
 ‘Kiplangat heard from you that Kibeet stole the money.’
- (7) Ko:-ɑ-mwai-te:-tʃi **Tfèbê:t** ɛ:n tɔ:jɛ:t à:-lé/kò-lé kò:-∅-tjór
 PST-1SG-say-IT-APPL Cheebeet at meeting 1SG-LE/3-LE PST-3-steal
 Kíbê:t rabɪ:nɪk.
 Kibeet.NOM money
 ‘At the meeting, I said on Cheebeet’s behalf that Kibeet stole the money.’

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 (8) below.

- (8) Kà-∅-kás Kíplàngàt kè:-lé kà-∅-tjór Kíbê:t rabɪ:nɪk.
 PST-3-hear Kiplangat.NOM IMP-LE PST-3-steal Kibeet.NOM money
 ‘Kiplangat heard (a rumour) that Kibeet stole the money.’

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

- (9) ko-ɑ-mwaa-un ɑ-lɛ-ndʒin ko-∅-it tuya amut
 PST-1SG-tell-2SG.OBJ 1SG-C-2SG.OBJ PST-3-arrive cows yesterday
 ‘I DID tell you (sg) that the cows arrived yesterday.’ ([Diercks & Rao 2019](#): 371)

We henceforth term this pattern *suffixal agreement* since our data reveal two types of object agreement: prefixal object agreement for objects that act as the logophoric center (as in (6) and (7)) and suffixal object agreement for indirect objects of speech verbs.

4 Analysis: 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 logophoric source, and not necessarily the matrix subject (*contra* Diercks & Rao 2019). Finally, in 4.3, we put these pieces together, and analyze the phenomenon as agreement between the verb *le* ‘say’ and a locally merged subject, which in most cases is *pro*.

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). We provide here three 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, 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 (10). Crucially, the “complementizer” is ungrammatical in this case.

- (10) **kà-∅-lé** Kíbê:t (***kò-lé**) ∅-rú-è là:kwè:t.
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 (10). Matrix uses of *le* are also reported in Diercks & Rao (2019), but Diercks et al. (2020) take this as evidence in favor of an analysis in which the C head (*le*) overtly raises to the matrix clause. More specifically, they argue that a silent speech verb occupies the matrix verb position, and *le* (which is base-generated in C) moves to this position (see (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.⁷ Morphologically, the subjunctive differs from the indicative in the

⁷While various tense and aspect distinctions are made in the indicative, only two forms are distinguished in the subjunctive: the perfective and imperfective. So far we have not identified matrix uses

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 (11) below, the verb *ru* ‘sleep’ has a short-voweled subject agreement prefix in its indicative (matrix) form in (11-a), but a long-voweled prefix in its subjunctive (embedded) form in (11-b). For 3rd person subjects, the prefix is \emptyset in most cells of the paradigm, while it is always *ko-* in the subjunctive.⁸

- (11) a. Kì:-[á]-rú.
PST-1SG-sleep(IND)
‘I slept.’
b. á-mátǫ-é [à:]-rú.
1SG-want-IPFV 1SG-sleep(SBJV)
‘I want to sleep.’

We observe in (12) that the inflection of *le* ‘say’ in matrix vs. complementation contexts shows the same contrast between indicative and subjunctive that we see in lexical verbs like *ru* ‘sleep’ in (11).

- (12) a. Kì:-[á]-lé kì:- \emptyset -tǫ́:r Kíbê:t rab:rnik.
PST-1SG-LE(IND) PST-3-steal Kibeet.NOM money
‘I said that Kibeet stole the money.’
b. Kì:-á-mwá [à:]-lé kì:- \emptyset -tǫ́:r Kíbê:t rab:rnik.
PST-1SG-say 1SG-LE(SBJV) PST-3-steal Kibeet.NOM money
‘I said that Kibeet stole the money.’

In a C-raising account ([Diercks & Rao 2019](#), [Diercks et al. 2020](#)), it is an accident that the complementizer has the subjunctive form. 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 (13), the imperfective form of *le* ‘say’ has the form *le:len*, which exhibits irregular stem allomorphy.⁹ In the verbal analysis pursued here, *le* is a lexical verb and is thus

of the subjunctive in Kipsigis, nor have we observed modality effects often present with subjunctives in other languages (see [Quer 2017](#) for an overview). This indicates that the term ‘subjunctive’ probably describes different categories in different languages ([Matthewson 2010](#)). 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 \(1989\)](#), but we leave a detailed investigation of this mood in Kipsigis as a topic for further research.

⁸The 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 \(1982\)](#), and [Creider & Creider \(1989\)](#) for a complete description and sample conjugation paradigms.

⁹The 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.

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.

- (13) **Le:le:n** ləyɔ́jwè:k kà:-Ø-tfó:r Kíbê:t rabı:nı:k.
 LE.IPFV news.NOM PST-3-steal Kibeet.NOM money
 ‘The news say that Kibeet stole the money’

In (13), *le* is in the matrix verb position. What is more striking, however, is that *le* can inflect for aspect even when used in complementation contexts (as a reminder, verbs in the subjunctive only make a perfective vs. imperfective distinction). We see in (14) that when the matrix verb is inflected in the past imperfective, *le* can be appear in either its perfective or imperfective form.¹⁰

- (14) Kα-α-mwa-e à:-lé/α:-le:le:n kà:-Ø-tfó:r Kíbê:t rabı:nı:k.
 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 (9), which are repeated below as (15). Diercks & Rao (2019) give a list of *le* forms with object agreement, shown in Table 2.

- (15) ko-α-mwaa-un α-lɛ-ndʒin ko-Ø-ıt tuya amut
 PST-1SG-tell-2SG.OBJ 1SG-C-2SG.OBJ PST-3-arrive cows yesterday
 ‘I DID tell you (sg) that the cows arrived yesterday.’ (Diercks & Rao 2019: 371)

	SG	PL
1	-lɛ-ndʒ-an	-lɛ-ndʒ-ɛtʃ
2	-lɛ-ndʒ-in	-lɛ-ndʒ-ɔ:ɣ
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* – *le.n* –, 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. 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

¹⁰It is not clear what the interpretational difference between the two is, which is a question for further research. A possibly relevant observation is that imperfective forms of *le* (when used in complementation) seem to be impossible with non-past matrix verbs.

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

Table 3: Suffixal agreement decomposed into APPL and object clitics

thematic roles (e.g. recipient, beneficiary).^{11,12} An example is given in (16).

- (16) a. Kà-∅-tʃáp Kíbê:t kímpé:t.
PST-3-make Kibeet.NOM ugali
‘Kibeet made ugali (type of food).’
b. Kà-∅-tʃap-tʃi Kíbê:t Tʃè:bê:t kímpé:t.
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 Towett (1979: p.209).¹³

	SG	PL
1	-an	-ε:tʃ
2	-in	-a:k
3	∅	

Table 4: Object clitics

The last piece of the reanalysis is the claim that the verb *le* has an allomorph *le:n*. We have already seen in (13) 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 *-kε:* used to form reflexives and reciprocals, illustrated in (17) below.¹⁴

- (17) Ki-ke:r-e-kε:
1PL-look-IPFV-REFL
‘We are looking at ourselves/at each other.’

¹¹There is another applicative suffix *-em*, which is mostly used for instruments and sources (Towett 1979, Rottland 1982).

¹²The applicative *-tʃi* has an allomorph *-ʃi* 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).

¹³The 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) (Towett 1979, Creider & Creider 1989).

¹⁴This suffix is unique in being outside of the [ATR] harmony domain of the verb.

The suffix *-kɛ:* 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 *-kɛ:*, suffixal agreement on *le* can include both the applicative and the reflexive/reciprocal suffix, as shown in (18).

- (18) Ko:-Ø-tʃa:m-tʃi-kɛ: Kíbê:t ko-le:n-tʃi-kɛ: ŋɑ:m.
 PST-3-whisper-APPL-REFL Kibeet.NOM 3SG-say-APPL-REFL clever
 ‘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 (19-a), but not when *le* is used in complementation, irrespective of whether there is matrix negation present, as shown in (19-b) and (19-c). [Diercks et al. \(2020\)](#) argue that the ungrammaticality of negation in complementation uses indicates that *le* is a complementizer, and not a verb.

- (19) a. Mɑ-a-le Ø-rú-è là:kwè:t.
 NEG-1SG-LE 3-sleep-IPFV child.NOM
 ‘I didn’t say that the child is sleeping.’
 b. Ma-a-mwa (*mɑ-)ɑ:-le Ø-rú-è là:kwè:t.
 NEG-1SG-say NEG-1SG-LE 3-sleep-IPFV child.NOM
 ‘I didn’t say that the child is sleeping.’
 c. Ka-a-mwa (*mɑ-)ɑ:-le Ø-rú-è là:kwè:t.
 PST-1SG-say NEG-1SG-LE 3-sleep-IPFV child.NOM
 ‘I didn’t say that the child is sleeping.’

However, all that (19) 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, there is nothing in the data suggesting that the explanation lies in the verb 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 (20) below. What we see in (20) 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 (19) with respect to negation: the negative prefix *ma-* is ungrammatical when attached to the embedded (subjunctive) verb, as shown in (20-b)-(20-c).¹⁵ Thus, we see that there is a class of subjunctives in the language that

¹⁵The examples in (20) 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 (20) 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

does not tolerate negation. Whatever the reason for this might be, what data like (20) show is that unavailability of negation in complementation uses of *le* does not constitute an argument against its analysis as a verb.

- (20) a. â:-ŋgén à:-pír pè:k.
 1 SG-know 1 SG-hit water
 ‘I know how to swim (*lit*: to hit water).’
- b. Mǎ-a(:)-ŋgen (*mǎ)-à:-pír pè:k.
 1 SG-know NEG-1 SG-hit water
 ‘I don’t know how (not) to swim.’
- c. *â:-ŋgén mǎ-à:-pír pè:k.
 1 SG-know NEG-1 SG-hit water
 Intended: ‘I don’t know how to swim.’ OR ‘I know how not to swim.’

Finally, if the explanation for the ungrammaticality of negation in (19-b)-(19-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 (20).

4.2 Logophoric and not subject agreement

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 a logophoric center (recall (6) and (7)). The sensitivity to logophoricity for contexts of ‘say’-based complementation should not come as a surprise. Sells (1987: 456-457, 475) identifies the subject of *say* as “the fully logophoric case” – a SOURCE in Sells’ terms – thus providing the criterial licensing condition for true logophoric pronoun systems. This line of thought is actively pursued in the accounts for logophoric pronouns and ‘say’-based complementation in the Kwa languages *Abe* (Koopman & Sportiche 1989) and *Ewe* (Clements 1987, Pearson 2015). In this section we provide two further arguments in favor of logophoric agreement and four arguments in favor of treating the local subject as a pronoun that establishes co-reference with a matrix or discourse antecedent.

First, a property of logophoric elements that is often discussed in the literature is their sensitivity to animacy, and it has been argued that logophoric anaphors or pronouns can only refer to animate antecedents (e.g. Charnavel & Sportiche 2016, Charnavel 2020a,b). Agreement on *le* is also subject to an animacy restriction, as shown by the contrast in (21). In both (21-a) and (21-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 *-en*. The source of information is animate in (21-a), but inanimate in (21-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 (21-b)] because the door cannot talk and *kole* is for

alike with respect to negation.

living things”. This is in line with our arguments in favor of logophoricity.¹⁶

- (21) a. **Ka-a-kas-ɛ:n** **Alice** à:l-**lé/kò-lé** ka-kɔ-Ø-it là:gô:k.
 PST-1SG-hear-APPL Alice 1SG-LE/3-LE PST-PRF-3-arrive children.NOM
 ‘I heard from Alice that the children have arrived.’
- b. **Ka-a-kas-ɛ:n** **kurget** à:l-**lé/*kò-lé** ka-kɔ-Ø-it
 PST-1SG-hear-APPL door 1SG-LE/3-LE PST-PRF-3-arrive
 là:gô:k.
 children.NOM
 ‘I heard from the door that the children have arrived.’

Finally, *le* can agree with benefactive arguments introduced by the applicative *-tʃi*, but only if they can act as a logophoric center. Thus, we see that agreement is possible in (7), repeated here as (22), but not in (23), where the benefactive argument of the predicate *kas* ‘hear’ cannot be construed as a logophoric center.

- (22) **Ko:-a-mwai-te:-tʃi** **Tʃèbê:t** ɛ:n tɔ:ʒɛ:t à:l-**lé/kò-lé** kà:-Ø-tʃó:r
 PST-1SG-say-IT-APPL Cheebet at meeting 1SG-LE/3-LE PST-3-steal
Kíbê:t rabɪ:nɪk.
 Kibeet.NOM money
 ‘At the meeting, I said on Cheebet’s behalf that Kibeet stole the money.’
- (23) ***Ka-a-kas-ɛ̀i** **Kíbê:t kò-lé** Ø-jà:tʃ-é kò-wá Nairobi.
 PST-1SG-hear-APPL Kibeet 3-LE 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 a logophoric source, 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 logophoric 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 logophoric relation is established via co-reference or binding. A binding relation is questioned by the fact that c-command is not necessary for logophoric agreement. The verb *le* can agree with the source even if the source is embedded in a PP, see (24).¹⁷

- (24) **Ka-I-kas** [PP kobun **Kiplàngàt**] **kò-lé/** ì:l-**lé** kà-Ø-tʃó:r **Kíbê:t**
 PST-2SG-hear from Kiplangát 3-LE/ 2SG-LE PST-3-steal Kibeet.NOM

¹⁶For some speakers, agreement with DPs denoting the source of information is not only sensitive to animacy, but also to how reliable the source is judged to be by the speaker (Culy 1994, Speas 2004). For example, in a context where Alice in (21-a) is known to be an unreliable person (e.g. someone who lies often), one consultant reports that 3rd person agreement on *le* is no longer possible. Thanks to Deniz Özyıldız for creating the ‘unreliable Alice’ context.

¹⁷In (24), as well as (29) later in the paper, the transcription for 2SG past forms shows the underlying representation of the morphemes, but a regular phonological process of vowel coalescence (Kouneli 2019: Chapter 2) applies to *ka-* and *ɪ-*, resulting in the form *kɛ:-* on the surface. Similarly, we give underlying forms in (35), where vowel coalescence applies between the ventive and 2SG object clitic.

rabr:nik.
 money
 ‘You heard from Kiplangat that Kibeet stole the money.’

Further support for this claim comes from *le*-clauses which take the impersonal form *kè:l-lé* indicating a rumour interpretation, as we saw in (8), 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 (25) for 3rd person and (26) for 2nd person.¹⁸

(25) 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-Ø-tfó:r Kíbê:t rabr:nik.
 PST-1SG-hear 3-LE PST-3-steal Kibeet.NOM money
 ‘I heard that Kibeet stole the money.’

(26) 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:*

α-ja:n-i ì:l-é kà-Ø-tfó:r Kíbê:t rabr:nik.
 1SG-believe-IPFV 2SG-LE PST-3-steal Kibeet.NOM money
 ‘I believe you that Kibeet stole the money.’

Since the contexts given in (24)-(25) 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 2020b) 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, Sundaesan 2016, Murugesan 2020). As was shown in (17), reflexivization in Kipsigis takes place through the ϕ -invariant verbal suffix *-kɛ:*, a strategy which is in complementary distribution with cliticization in non-anaphoric contexts, shown in (27).

(27) Kα-α-ke:r(*-an)-kɛ: / Kα-α-ke:r-kɛ:(*-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 *le*, we conclude that the subject does not instantiate an anaphor. Instead we propose that logophoric agreement takes place between *le*

¹⁸There is variation in our consultants’ judgments regarding these examples. Three speakers find (26), but not (25), acceptable, while one speaker shows the opposite pattern accepting (25), but not (26).

and a covert pronoun introduced by *le*. This *pro* is coindexed with the matrix/discourse antecedent via the assignment function, thereby avoiding the need for c-command by the antecedent. We refrain from assuming a logophoric operator, as we encode logophoricity as a presupposition on *le* itself.

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 (28).

(28) 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 ì:-lé **anɛ:** kà-∅-tʃó:r Kíbê:t rabi:nik.
 PST-1SG-say 1SG-LE 1SG.NOM PST-3-steal Kibeet.NOM 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.

4.3 *le*-clauses as sets of contentful eventualities

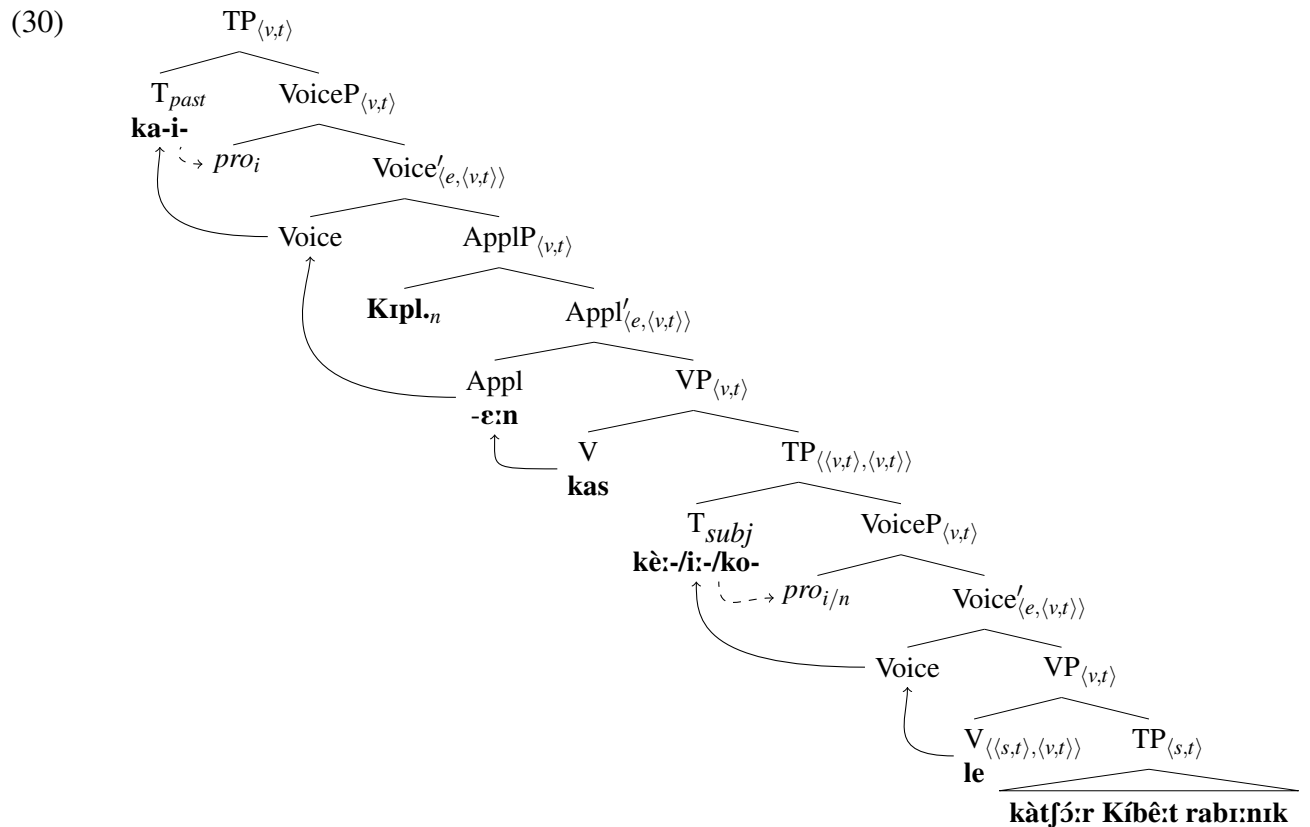
We propose that embedded clauses headed by agreeing forms of *le* constitute sets of contentful saying events, where the verbal nature of *le* 'say' is reflected in its semantics. In order to do so, we adopt an eventuality-based framework where the relation between the attitude holder and the proposition is mediated by *contentful eventualities* (Kratzer 2013, Elliott 2016, 2017, Moulton 2019, Özyıldız et al. 2018, Demirok et al. 2020, Bondarenko 2020). This idea has its roots in the *CP-property hypothesis* under which *that*-clauses combine with content nouns like *theory*, *belief*, and *rumour* via adjunction, rather than as clausal arguments (Stowell 1981, Grimshaw 1990, Kayne 2010). The complementizer *that* expresses that the content of the proposition identifies the content of individuals such as rumours/beliefs/theories, that is complementizers are functions from propositions to sets of *contentful individuals* (Kratzer 2006, 2013, Moulton 2009, 2015). In order to get from individuals to their content, Kratzer introduces a content function CONT which if applied to an individual returns a set of possible worlds. We will adopt CONT but claim that *le* introduces events instead of individuals, leading to *le*-clauses denoting sets of contentful saying events.

Let us introduce the main analysis by illustrating our proposal with the structure in (30), based on the example in (29). 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.

(29) Ka-i-kas-ɛ:n Kìplàngàt kè:-lé/ ì:-lé/ kò:-lé kà-∅-tʃó:r
 PST-2SG-hear-APPL Kiplangat IMP-LE/ 2SG-LE/ 3-LE PST-3-steal
 Kíbê:t rabi:nik.
 Kibeet.NOM money
 'You heard from Kiplangat that Kibeet stole the money.'

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 (30), where the embedded proposition (a TP) is a sister to the verb *le* ‘say’, which itself is part of a subjunctive TP embedded under the matrix predicate (see Alexiadou et al. 2012, Pietraszko 2017, 2020 for arguments against a C layer for at least some subjunctives).¹⁹ 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).²⁰ To account for the verb initiality of Kipsigis, we assume that V moves via Voice to T (or a higher projection, see Bossi & Diercks (2019)), indicated by the arrows in (30). The dashed arrows show that the subjunctive T head probes for the ϕ -features of the logophoric subject – 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 *ile* is chosen if *pro* points to the addressee of the utterance, whereas *kole* appears if *pro* is co-indexed with *Kiplàngàt*, that is the source argument from the matrix clause. Another option is the impersonal form *kè:lé* which leads to a rumour interpretation, recall (8). In this case, *pro* co-refers with an impersonal antecedent in the discourse.



¹⁹One might find the absence of a C layer suspicious. Since there is no evidence for the presence of such a layer, however, we choose to not postulate an unnecessary projection. If our analysis is correct, this means that in Kipsigis *le* is the only verb that can directly select for indicative TPs, while other verbs can only select for subjunctive TPs. See also fn. 7 for discussion on the subjunctive.

²⁰In 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 *-e:n*, 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.

We will now present the key points of the semantic analysis. Since *le* is not a complementizer but a verbal category, it introduces an eventuality and a content argument, whilst also presupposing that the agent of the eventuality is the logophoric SOURCE (Sells 1987) of the proposition *le* embeds. The lexical entry is given in (31).

$$(31) \quad \llbracket le \rrbracket^{w,g} = \lambda p_{\langle s,t \rangle} \lambda e_v [\text{say}(e) \wedge \text{CONT}(e) = p],$$

defined iff AG(*e*) qualifies as the logophoric SOURCE of *p*

The result of *le* applied to TP is a set of saying events the content of which is that Kibeet stole the money. Both Voice and Appl combine with their complements via *Event Identification* (Kratzer 1996). Thus, the embedded VoiceP outputs the denotation in (32), where the agent is now added to the saying event. We choose to illustrate the *kole* derivation, where *pro* is co-indexed with the applied argument *Kiplàŋgàt*.

$$(32) \quad \llbracket \text{VoiceP} \rrbracket^{w,g} = \lambda e_v [\text{say}(e) \wedge \text{CONT}(e) = \{w : \text{Kibeet stole the money at } w\} \\ \wedge \text{AG}(e) = g(n)]$$

Subjunctive is the mood used in subordinate clauses, and we take this to mean that the subjunctive T head serves as a causal linker between the event introduced by the embedded predicate and the event introduced by the matrix predicate. We give the denotation in (33-a) for subjunctive, a function from a set of events to another set of events where \sim indicates a bidirectional causal relation. This denotation is adopted from Özyıldız et al. (2018), who provide this entry for a gerundive affix serving a similar linking function in Turkish complementation. Since we introduce the subjunctive on the embedded T head, the first argument it takes are the saying events in (32), while the second argument are the hearing events in (33-b), the result of which is the denotation of matrix VP, shown in (33-c).

$$(33) \quad \begin{array}{l} \text{a. } \llbracket \text{SUBJ} \rrbracket^{w,g} = \lambda P \lambda Q \lambda e'' . \exists e' [e' \sim e'' \wedge P(e') \wedge Q(e'')] \\ \text{b. } \llbracket kas \rrbracket^{w,g} = \lambda e_v [\text{hear}(e)] \\ \text{c. } \llbracket kole \ kàt\acute{s}òr \ Kíbê:t \ rabɪ:nɪk \rrbracket^{w,g} (\llbracket kas \rrbracket^{w,g}) \\ = \lambda e'' . \exists e' [e' \sim e'' \wedge \text{say}(e') \wedge \text{CONT}(e') = \{w : \text{Kibeet stole the money at } w\} \\ \wedge \text{AG}(e') = g(n) \wedge \text{hear}(e'')] \end{array}$$

Finally, both the experiencer and the source of the hearing event are added via Event Identification in the matrix clause, resulting in the denotation in (34), where presuppositions are repeated for completion and tense semantics is ignored.

$$(34) \quad \llbracket kaikasɛ:n \ Kìplàŋgàt \ kole \ kàt\acute{s}òr \ Kíbê:t \ rabɪ:nɪk \rrbracket^{w,g}$$

$$= \lambda e'' . \exists e' [e' \sim e'' \wedge \text{say}(e') \wedge \text{CONT}(e') = \{w : \text{Kibeet stole the money at } w\} \\ \wedge \text{AG}(e') = g(n) \wedge \text{hear}(e'') \wedge \text{SOURCE}(e'') = \text{kiplangat}_n \wedge \text{EXP}(e'') = g(i)],$$

defined iff $g(i)$ is addressee²¹ and AG(*e'*) qualifies as the logophoric SOURCE of $\{w : \text{Kibeet stole the money at } w\}$

Since the agent of the saying event co-refers with the source of the hearing event in

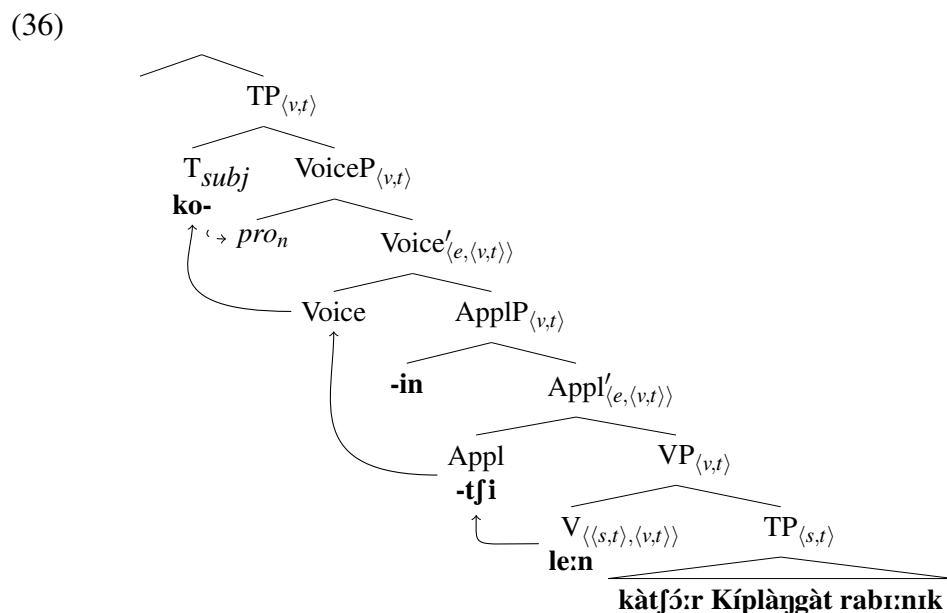
²¹ ϕ -features on pronouns denote partial identity functions of type $\langle e, e \rangle$ (Sauerland 2003, 2008, Heim 2008); for free pronouns the relevant assignment is given by the utterance context.

(34), we will interpret \sim to indicate 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 holds in case the agent co-refers with the subject of the matrix predicate, that is if *le* inflects for 2SG in (29). In this case, the hearing event causes the addressee to become the agent of the saying event, where the latter 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 (29) 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.

Let us now turn to suffixal agreement, which was discussed in section 4.1. We provide another example in (35) where *le* not only shows prefixal agreement with the matrix subject but also an object clitic introduced by APPL.

- (35) Ka-mwa-u-in Kíbê:t ko-le:n-tji-in ka-tʃɔ:r Kíplàngàt
 PST-say-VENT-2SG Kibeet.NOM 3-LE-APPL-2SG PST-steal Kiplangat.NOM
 rabɪ:nɪk.
 money
 ‘Kibeet told you that Kiplangat stole the money.’

The occurrence of suffixal agreement is predicted under an account that treats *le* as a verb. In such cases, *le* introduces an applied argument in addition to a subject, shown for the partial derivation in (36).



We provide the denotation of matrix VoiceP in (37). Suffixal agreement, decomposed into APPL and a 2SG object clitic, introduces a goal argument for the embedded saying event, matching the goal argument of the matrix telling event. Since the subject of *le* co-refers with the matrix subject, the sentence receives an attitude report reading.

- (37) $\llbracket \textit{kamwauin Kíbê:t kole:ntfin katfɔ:r Kíplàngàt rabı:nık} \rrbracket^{w,g}$
 $= \lambda e''. \exists e' [e' \sim e'' \wedge \textit{say}(e') \wedge \text{CONT}(e') = \{w : \text{Kiplangat stole the money at } w\}$
 $\wedge \text{GOAL}(e') = g(i) \wedge \text{AG}(e') = g(n) \wedge \textit{tell}(e'') \wedge \text{GOAL}(e'') = g(i) \wedge \text{AG}(e'') =$
 $\textit{Kibet}_n]$, defined iff $g(i)$ is addressee and $\text{AG}(e')$ qualifies as the logophoric
SOURCE of $\{w : \text{Kiplangat stole the money at } w\}$

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. Such a verb is *ɲo:ɲ* ‘complain’, see (38), but also *sir* ‘write’.

- (38) Ko:-a-ɲo:ɲ a:-le:n-tfi Kibet ko:-ja:tʃ-e:n àmítwá:gìk.
PST-1SG-complain 1SG-LE-APPL Kibeet PST-bad-PL food.NOM
‘I complained to Kibeet that the food was bad.’

5 Outlook

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 & Torrence 2020](#)), 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 supports the recent claim that the semantic type of embedded clauses varies cross-linguistically: $\langle e, t \rangle$ vs. $\langle v, t \rangle$ (e.g. [Moulton 2019](#), [Demirok et al. 2020](#)). Our account also suggests an interesting avenue for further research, where the semantic type of the embedded proposition is reflected in the syntactic category of the embedder ([Moulton 2019](#), [Bondarenko 2020](#)).

References

- 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.

- 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 & 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.
- Bondarenko, Tatiana. 2020. Hyperraising and Logical Form: evidence from Buryat. GLOW 43 Poster, tbond.scripts.mit.edu.
- 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. 2020a. *Locality and logophoricity: A theory of exempt anaphora*. Oxford: Oxford University Press.
- Charnavel, Isabelle. 2020b. Logophoricity and Locality: A View from French Anaphors. *Linguistic Inquiry* 51. 671–723.
- Charnavel, Isabelle & Dominique Sportiche. 2016. Anaphor Binding: What French Inanimate Anaphors Show. *Linguistic Inquiry* 47. 35–87.
- 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.
- Clements, George. 1987. The logophoric pronoun in ewe: Its role in discourse. *Journal of West African Languages* 10. 141–177.
- Creider, Chet & Jane Tapsuei Creider. 1989. *A grammar of Nandi*. Hamburg: Helmut Buske Verlag Hamburg.
- Culy, Christopher. 1994. Aspects of logophoric marking. *Linguistics* 32. 1055–1094.
- 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.
- Duncan, Philip T. & Harold Torrence. 2017. Verbal complementizers and the Indirect Agree relation in Ibibio. Handout from talk given at the 91st Meeting of the LSA.
- Eberhard, David M., Gary F. Simons & Charles D. Fennig (eds.). 2020. *Ethnologue: Languages of the world*. Dallas, Texas: SIL International 23rd edn.
- Elliott, Patrick D. 2016. Explaining DPs vs. CPs without syntax. In *Proceedings of the 52nd annual meeting of the Chicago Linguistics Society*, 171–185. Chicago: Chicago Linguistic Society.
- Elliott, Patrick D. 2017. *Elements of clausal embedding*: University College London dissertation. Revised version, January 2020.
- Fuß, Eric. 2008. Multiple agreement and the representation of inflection in the C-domain. *Linguistische Berichte* 213. 78–108.
- Fuß, Eric. 2014. Complementizer agreement (in Bavarian): Feature inheritance or feature insertion? In Günther Grewendorf & Helmut Weiß (eds.), *Bavarian Syntax: Contributions to the theory of syntax*, vol. 220, 51–82. Amsterdam/Philadelphia: John Benjamins.
- Grimshaw, Jane. 1990. *Argument structure*, vol. 18 Linguistic Inquiry Monographs. Cambridge, MA: The MIT Press.
- 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.
- Haegeman, Liliane & Marjo van Koppen. 2012. Complementizer agreement and the relation between C^o and T^o. *Linguistic Inquiry* 43. 441–454.
- Hall, Beatrice, R.M.R. Hall, Martin D. Pam, Amy Myers, Stephen A. Antell & Godfrey K. Cheron. 1974. African vowel harmony systems from the vantage point of kalenjin. *Afrika und Übersee* 241–267. 241–267.
- Halle, Morris & Jean-Roger Vergnaud. 1981. Harmony processes. In *Crossing the boundaries in linguistics: Studies presented to Manfred Bierwisch*, 1–22. Dordrecht: Springer.
- 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.
- 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.
- Kayne, Richard. 2010. *Comparisons and contrasts*. Oxford: Oxford University Press.
- Kinyalolo, Kasangati Kikuni W. 1993. On some syntactic properties of \varnothing . *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* 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.
- 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. 2006. Decomposing attitude verbs. Handout from a talk in honor of Anita Mittwoch on her 80th birthday.
- Kratzer, Angelika. 2013. 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. Librairie Droz.
- 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, Illinois: University of Chicago.
- 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.
- Matthewson, Lisa. 2010. Cross-linguistic variation in modality systems: The role of mood. *Semantics and Pragmatics* 3. 1–74.
- McFadden, Thomas & Sandhya Sundaresan. 2020. Unifying species of C-agreement. Flash Talk, NELS 51.

- 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. 2015. CPs: copies and compositionality. *Linguistic Inquiry* 46. 305–342.
- 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.
- Pearson, Hazel. 2015. The interpretation of the logophoric pronoun in Ewe. *Natural Language & Linguistic Theory* 23. 77–118.
- 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*.
- Pylkkänen, Liina. 2008. *Introducing arguments*. Cambridge, MA: MIT Press.
- Quer, Josep. 2017. Subjunctives. In Martin Everaert & Henk C. van Riemsdijk (eds.), *The Wiley Blackwell Companion to Syntax, Second Edition*, 658–681. Oxford: Blackwell.
- 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.
- 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.
- Sells, Peter. 1987. Aspects of logophoricity. *Linguistic Inquiry* 18. 445–479.
- 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.
- Stowell, Timothy. 1981. *Origins of phrase structure*: MIT dissertation.
- 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.
- Toweett, Taaita. 1979. *A study of Kalenjin linguistics*. Nairobi: Kenya Literature Bureau.
- 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.
- 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.