

‘Say’-based complementation: Insights from Kipsigis

Imke Driemel, Maria Kouneli ❖ Leipzig University

1. Introduction. One of the central questions in research on clausal complementation is how clausal complements are formed and combine with the verb. For European languages, a range of solutions have been proposed in the literature (Stowell 1981, Kayne 2005, Kratzer 2006, Moulton 2009, Roussou 2010, Angelopoulos 2019 a.o.), but less attention has been given to languages that employ ‘say’-based complementizers (Koopman 1984, Koopman & Sportiche 1989, Özyıldız et al. 2018, Major & Torrence 2020 a.o.). In this talk, we analyze two types of ‘say’-based complementation in Kipsigis (Nilotic, Kenya; VSO). The starting point of our investigation is what has been described as upwards-oriented complementizer agreement with a matrix subject (Diercks & Rao 2019, Diercks et al. 2020): the complementizer consists of the root of the verb *le* ‘say’ and an agreement prefix, as in (1)-(2).

- (1) Kà-á-tfá:m á:-lé rú-è Kíbê:t. (2) Kà-í-tfá:m ì:-lé rú-è Kíbê:t.
 PST-1SG-whisp. 1SG-C sleep-IPFV K.NOM PST-2SG-whisp. 2SG-C sleep-IPFV K.NOM
 ‘I whispered that Kibet is sleeping.’ ‘You whispered that Kibet is sleeping.’

We argue that what has been described as an agreeing ‘say’-based complementizer in Kipsigis is in fact the lexical verb ‘say’, which is embedded under the matrix verb and is the element introducing a TP complement, showcasing another possible strategy to introduce clausal complements cross-linguistically. However, we also find a non-agreeing form that contrasts with the agreeing forms in creating clausal complements with nominal properties. This indicates that there is significant variation – even within the same language – in the properties of ‘say’-based complementizers. Data come from original fieldwork with 8 native speakers. **2. Agreeing *le* is a verb.** We present two pieces of evidence: A) *le* can be a matrix verb. The language’s VSO order makes it clear that *le* occupies the verb position in (3). Crucially, the “complementizer” is ungrammatical in this case. Diercks & Rao (2019) (henceforth D&R) report examples like (3), but Diercks et al. (2020) take them as evidence that *le* raises to the matrix clause. However, such an analysis cannot account for the difference in mood inflection that we observe: *le* is in the indicative in (3), but in the subjunctive in (1) and (2); mood is reflected in the form of the subject agreement prefix, and Toweett (1979) and Creider & Creider (1989) claim that subjunctive is used in Kipsigis when a verb is embedded under another verb (the language lacks infinitives). If *le* is a verb, it follows that subjunctive will be used in complementation, but indicative in matrix clauses.

- (3) kâ-∅-lé Kíbê:t (*kò-lé) ∅-rú-è là:kwè:t.
 PST-3.IND-LE Kibet.NOM (*3.SUBJ-LE) 3.IND-sleep-IPFV child.NOM
 ‘Kibet said that the child is sleeping.’

B) applicative and reflexive morphology on *le*. With speech verbs, *le* optionally displays suffixal agreement with the indirect object of the matrix verb (in addition to subject prefixal agreement). While D&R call this agreement, our data show that suffixal agreement consists of the applicative *-tʃi* (Toweett 1979, Rottland 1982) followed by an object clitic. Further support comes from (4), where *-tʃi* and the reflexive *-kɛ:* are present (*le:m* is the allomorph of *le* before *-tʃi*; this is an allomorphy rule targeting a class of CV verbs in Kalenjin dialects (Zwarts 2004)).

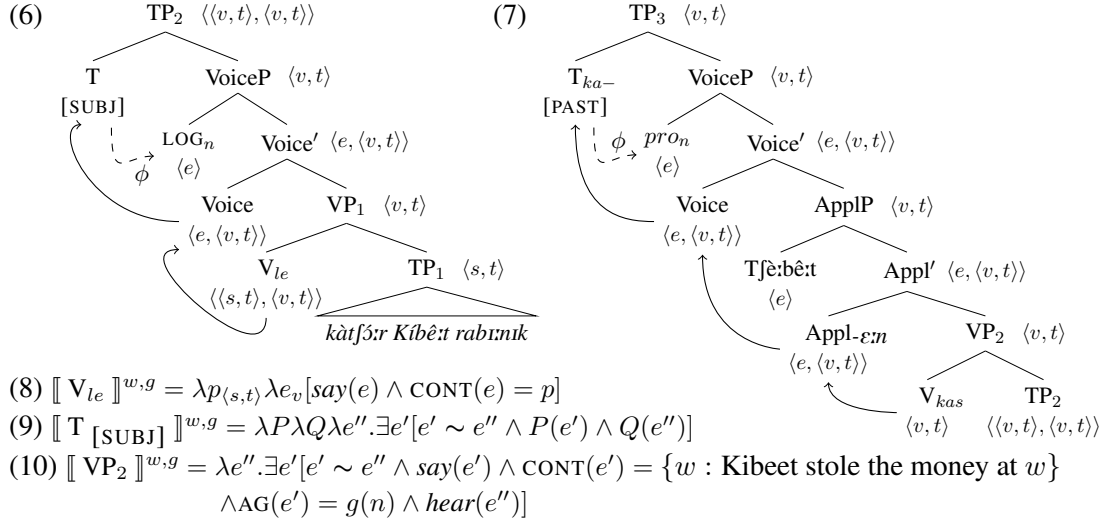
- (4) Kà-∅-tʃa:m-tʃi-kɛ: Kíbê:t kò-le:n-tʃi-kɛ: ŋâ:m.
 PST-3-whisper-APPL-REFL Kibet.NOM 3-LE-APPL-REFL clever
 ‘Kibet whispered to himself that he’s intelligent.’

The presence of applicative and reflexive morphology on *le* strongly supports its analysis as a verb.

3. Analysis. While D&R only show that *le* agrees with the matrix subject, we observe in our data that *le* can additionally co-vary with an applied object if it qualifies as the logophoric source. Both options are shown in (5); indices indicate the respective co-reference relations, and *-ɛ:n* is the applicative introducing sources/instruments in the language (Toweett 1979). Thus, we conclude that *le* locally agrees with a logophoric pronoun. Further evidence for logophoricity comes from the fact that neither unreliable sources, e.g. politicians and children, nor inanimate sources (Charnavel 2019) can serve as antecedents for agreement (data not shown here).

- (5) Ka-i-kas-ɛ:n *pro*₁ Tʃè:bê:t₂ ì₁-lé/kò₂-lé LOG_{1/2} kà-∅-tʃó:r Kíbê:t rabɪ:nɪk.
 PST-2SG-hear-APPL Chebet 2SG-LE/3-LE PST-3-steal Kibet.NOM money
 ‘You heard from Chebet that Kibet stole the money.’

Since *le* is a verb and embeds a clause, it introduces an eventuality and a content argument (cf. Kratzer 2013), see (8). We provide a detailed derivation for (5), where (7) shows the derivation of the matrix clause, and (6) the internal structure of TP₂ in (7). Voice introduces the external argument (Kratzer 1996) and Appl the source; each combines with its complement via *Event Identification*. VSO order results from head movement (Bossi & Diercks 2019). Subjunctive is introduced in T and serves as a causal linker (\sim in 9) between the saying event and the event introduced by the matrix predicate (cf. Özyıldız et al. 2018), the result of which is shown in (10). As a free pronoun, the logophor comes with its own ϕ -features serving as a goal for agreement with T (via downward Agree). Co-reference via the assignment function with the antecedent of LOG determines the different forms of *le*: for (5), 2SG = co-reference with matrix *pro*, 3SG = co-reference with the source *Tfjê:bêt*.



4. Non-agreeing *le*. In addition to the agreeing forms discussed above, we also find a form of *le* (*keɪ-le*) with a default-like distribution. Three environments in which it appears indicate that complements with *keɪ-le* are of type $\langle e, t \rangle$, unlike complements with the agreeing form; *keɪ-le* has possibly grammaticalized into a C category introducing CPs with nominal properties. First, it is the only form of *le* that can directly combine with content nominals, such as *lɔgɔjwɛ:k* in (11), a diagnostic that singles out $\langle e, t \rangle$ CPs (Moulton 2019). Importantly, agreeing *le* forms are ungrammatical in this environment, which is predicted by their $\langle v, t \rangle$ type. Second, phrases introduced by *keɪ-le*, but not phrases introduced by agreeing forms, can appear in the pre-verbal position marked by the topic marker *ko* (12). This is a derived position that can only host nominals (data not shown here). The grammaticality of *keɪ-le* shows that it is part of a nominal constituent, while the ungrammaticality of the agreeing forms is consistent with their analysis as verbs. Further support comes from examples like (13), which show that only phrases introduced by *keɪ-le* can appear in subject position.

- (11) [lɔgɔjwɛ:k **keɪ-le**/***ko-le** ko:ki:tun Tfjê:bêt] ko ko:jaj Kíbê:t ko-ma-bajbaj.
 news KEE-LE/*3-LE PST-marry Chebet.NOM TOP PST-do Kibet ADV-NEG-happy
 ‘The news that Chebet got married made Kibet unhappy.’
- (12) [**Keɪ-le**/***ɑɪ-le** ko:si:r Tfjê:bêt] ko â:ɪgɛn
 KEE-LE/1SG-LE PST-pass Chebet.NOM TOP 1SG-know
 ‘That Chebeet passed the exams, I know.’
- (13) Já [**keɪ-le**/***ko-le**/***ɑɪ-le** kà-ø-tʃɔ:r Kíbê:t rabɪ:nɪk].
 bad KEE-LE/3-LE/1SG-LE PST-3-steal Kibet.NOM money
 ‘That Kibeet stole the money is bad.’

5. Conclusion. Our analysis supports the recent claim that the semantic type of CPs varies cross-linguistically ($\langle e, t \rangle$ vs. $\langle v, t \rangle$; Moulton 2019, Demirok et al. 2020 a.o.). Kipsigis is also added to a list of languages whose ‘say’-based complementizers are analyzed as verbs. However, we have shown that even in such a language, a ‘say’-based form introducing $\langle e, t \rangle$ CPs is also attested. Different ‘say’-based complementizers with verb-y and noun-y properties have also been described for Zulu (Halpert 2018); it is an open question whether both types are attested in all languages with ‘say’-complementation.