

Finiteness across languages: a case study of the Jê family*

Bernat Bardagil
University of Groningen

Kim A. Groothuis
Ghent University

Abstract: In this paper, we take Jê languages as a case study to investigate the problematic notion of finiteness, a frequently used but still poorly understood notion within linguistics. It has been proposed that finiteness relates to the presence or absence of anchoring (Roussou 2001; Bianchi 2003; Ritter & Wiltschko 2014; Wiltschko 2014; Groothuis 2020). Following Ritter and Wiltschko (2014), we take INFL to be the anchoring category which can have different substantive content cross-linguistically. This paper makes two main claims: first, contrary to previous analyses of Jê, which consider the so-called long form as non-finite, we argue that the form of the verb (i.e. short or long) is not indicative of the level of finiteness (intended as presence or absence of direct anchoring) of the clause. Instead, we extend Nonato’s (2014) analysis of Kisêdjê modal particles as the realization of infl to all Jê languages. Second, we argue that there is more than one type of non-finite (i.e. indirectly anchored) clause in Jê, which differ in terms of subject licensing. We analyze one type as an instance of subject raising.

Keywords: Jê languages, finiteness, complementation, anchoring

1 Introduction

Finiteness is a long-honoured notion in linguistics, but it still remains poorly understood (Joseph 1983; Ledgeway 2007; Nikolaeva 2007). In this paper, we want to take a close look at finiteness as it relates to Jê languages by adopting a non-binary approach. The indigenous South American languages in the Jê family all present a complex interaction between finiteness, on the one hand, and other syntactic categories such as nominalization, clause type, and case marking, on the other. When added to a traditional conflation of the notions of nominal and non-finite in their grammar, the Jê languages are an ideal choice for testing theoretical approaches to the notion of finiteness.

Section 2 briefly introduces the languages of the Jê family (Section 2.1), with a special focus on the syntactic patterns of the short verbal and the long nominal forms of the verb (Section 2.2). Section 3 dives into the discussion of what finiteness is, which in this paper we relate to deictic anchoring (Section 3.1). In Section 3.2, we discuss the presence of finiteness as represented by INFL in Jê languages, looking both at its substantive content and its morphosyntactic reflexes. In Section 4, with the novel observation that Jê languages present a class of raising constructions (Section 4.1), we propose a unified analysis of the degrees of finiteness displayed by Jê clauses (Section 4.2). Section 5 concludes the paper.

* We would like to thank the audience of WSCLA 25, as well as the members of the Ghent Generative Grammar Group, where previous versions of this paper were discussed. We also thank the Endangered Languages Documentation Programme, the Jacobs Fund, and the Endangered Languages Fund for providing financial support to various stages of Bardagil’s fieldwork in Brazil. We would also like to thank the Panará people of Nãnsêpotiti and in particular Perankô Panará, Kypakjã Panará, and Tepakriti Panará, as well as Bepêre Kayapó, Kunity Metyktire Panará, and Takakpe Tapayuna Metyktire for their information on Mëbêngôkre. Finally, we are also thankful to the organizers of WSCLA 25 and the editors of this volume.

□ Contact: bbardagil@gmail.com kim.groothuis@ugent.be.

2 Jê verb forms

The Jê languages form a modest-sized family spoken in eastern Brazil. Generally considered split accusative-ergative languages, in Jê languages, ergative case marking is restricted to a “long form” of the verb, associated with dependent clauses and analyzed as being nominal or nonfinite (Urban 1985; Salanova 2011; Nonato 2014; Bardagil 2018). A different, finite verbal “short form” of the verb that is associated to main clauses correlates with a nominative-accusative case marking with a morphologically marked nominative — in all Jê languages but one, Panará.

2.1 Jê languages

The languages of the Jê family are spoken in eastern Amazonia and the *cerrado* savanna, in what is today Brazil. Their geographic distribution spans from south-eastern Amazonia to the southernmost states of Brazil, on the basin of the Paraná, itself a tributary of the Río de la Plata. The family includes some of the indigenous languages with the most speakers in Brazil, such as Kaingang (over 20,000), Xavante, and Mëbêngôkre (over 10,000 each). Jê languages have also been related to other languages, in what is called the larger Macro-Jê family (Rodrigues 1999; Ribeiro & Voort 2010; Nikulin 2020), but as of today, the extant core Jê languages are ten, shown in Figure 1 below.

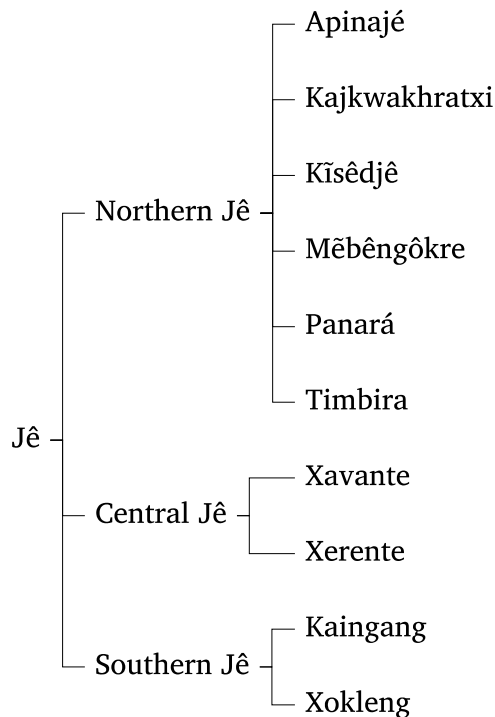


Figure 1: The Jê languages, adapted from Davis (1966) and Rodrigues (1999).

Jê languages present relatively small consonantal inventories but larger vocalic distinctions, with a phonologically contrastive nasal dimension. They typically display an analytic morphological profile, with rigid clausal positions. See Salanova (to appear) for more information.

2.2 Long and short verbal forms

A characteristic trait of Jê languages is the interplay between verbal forms, clause type, and case marking alignment. Jê languages, such as Mëbêngôkre (1), exhibit an opposition between a short form (1a) and a long form (1b) of the verb, traditionally analyzed as finite and non-finite, respectively (Alves 2004; de Oliveira 2005; Salanova 2007; Estevam 2011; Nonato 2014; Bardagil 2018).

- (1) a. Ga nê ga ku= **bĩ**.^{1,2}
 2NOM NFUT 2NOM 3ACC kill.V
 ‘You killed it.’
- b. Ba nê ba [aje Ø= **bĩn**] pumũ.
 1NOM NFUT 1NOM 2ERG 3ABS kill.N see.V
 ‘I saw you kill it.’ (Mëbêngôkre)

As can be seen in (1), the long form triggers a case marking alignment shift from nominative-accusative to ergative-absolutive. Embedded clauses obligatorily require a long form verb, with the corresponding ergative-absolutive case marking, as in the sentences in (2).

- (2) a. *[Ba tep **krẽ**] kêt.
 1SG.NOM fish eat.V NEG
 Intended: ‘I didn’t eat fish.’
- b. *[Ije tep **krẽ**] kêt.
 1SG.ERG fish eat.V NEG
 Intended: ‘I didn’t eat fish.’
- c. [Ije tep **krẽn**] kêt.
 1SG.ERG fish eat.N NEG
 ‘I didn’t eat fish.’ (Mëbêngôkre)

In most Jê languages, the long form can also be used in main clauses, where it expresses a specific Tense/Aspect/Modality/Evidentiality-related value, which varies from language to

¹ Abbreviations used in this article: ABS = absolutive, ACC = accusative, ASP = aspect, CP = complementizer phrase, ERG = ergative, FACT = factive, FUT = future, H = honorific, INDF = indefinite, INES = inessive, INFL = inflection, INTR = intransitive, IP = inflectional phrase, IRR = irrealis, N = nominalization, NEG = negation, NFUT = non-future, NOM = nominative, PL = plural, Q = question, RLS = realis, SG = singular, TAME = tense/aspect/mood/evidentiality, TR = transitivizer, V = verb, VOC = vocative, VP = verb phrase, vP = little verb phrase.

² If no source is cited, the examples are from Bernat Bardagil’s fieldwork. Annotations have been adapted to fit the analysis presented in the paper and, where relevant, orthographic representations have been altered.

language. For instance, in Měbêngôkre (3b), the long form expresses perfective aspect (Salanova 2007:17).

(3) a. \emptyset ba ku= **krě**.
 NFUT 1SG.NOM 3SG.ACC eat.V
 ‘I eat it / I ate it.’

b. \emptyset ba \emptyset = **krěn**.
 NFUT 1SG.ERG 3SG.ABS eat.N
 ‘I have eaten it.’

(Měbêngôkre; Salanova 2007:17)

The so-called non-finite long form can be used in both finite (3b) and non-finite (2c) contexts. In contrast, the short form can never appear in embedded contexts. There is thus no one-to-one correspondence between morphological form and syntactic context in which the verb can be used. A purely morphological view which considers the long form as non-finite and the short form as finite is therefore untenable. We pick up the issue of what constitutes finiteness in Jê languages in Section 3.

3 INFL and anchoring in Jê

3.1 Finiteness as anchoring

Since the distinction between long and short verb forms cannot be equated with a difference between finite and non-finite, or even embedded vs. non-embedded clauses, we need to adopt a different approach to finiteness and embeddedness. In the last few decades, syntacticians started considering oppositions of finiteness as differences in the logophoric anchoring of the clause (Holmberg & Platzack 1995; Roussou 2001; Bianchi 2003; Ritter & Wiltschko 2014; Wiltschko 2014; Groothuis 2020).

On this view, the syntactic property of finiteness encodes the logophoric anchoring of the clause (Bianchi 2003). The logophoric centre of a clause is defined as “a speech or mental event, with its own participants and temporal coordinates, which constitutes the centre of deixis” (Bianchi 2003:15). The event reported in the clause is anchored to this logophoric centre by finiteness. Non-finite forms, on the other hand, are anchored indirectly through a higher clause.

In this article, we will follow Ritter and Wiltschko (2014) who adopt a similar notion of anchoring when analyzing the properties of INFL. They put forward the Parametric Substantiation Hypothesis, which makes a distinction between the universal spine, a set of hierarchically organized functional categories made available by UG, and the language-specific substantive content that is associated with these functional categories. Within the universal spine, INFL is the functional category which links the event with the utterance. Its substantive content can differ crosslinguistically: INFL may be associated with temporal, personal, and spatial marking, which anchors the event situation to the utterance situation. The anchoring mechanism is therefore universal, but it need not be mediated by Tense (Ritter & Wiltschko 2014:1339). Mood has also been proposed as a fourth anchoring category in Dravidian (Amritavalli 2014) and Upper Austrian German (Wiltschko 2014: Chapter 4). The substantive content of INFL is normally deictic, since the interpretation of this content is dependent on the context.

This deictic interpretation of INFL is formalized through a feature [\pm coin(cedence)], that in indicative main clauses is valued based on the m(orphological)-marking. This feature encodes whether the event situation coincides (or not) with the utterance situation. For example, if anchoring

is mediated through Tense, this means that [+coin] indicates that the event takes place at the same time as the utterance (i.e. present tense); [-coin], on the other hand, indicates past tense.

In certain (non-finite) contexts, the [\pm coin] feature is valued in different ways. This can happen through so-called *predicate valuation*, when a higher (finite) predicate values the embedded INFL, e.g. the INFL of infinitival complements to aspectual verbs in English. Another option is valuation by Comp, as happens in the case of English imperatives (obligatorily present tense [+coin]) or in counterfactuals ([-coin]). With these types of valuation, INFL is not deictic anymore, as [\pm coin] does not mark the relation with the utterance situation but with a higher predicate or the C-head. In fact, in these cases, we can see the absence of m-marking (for instance with English infinitives) or so-called ‘fake’ m-marking, e.g. past tense marking in English counterfactuals:

(4) If I **were** rich...

The past tense morphology in (4) does not express past tense but rather counterfactuality.

In sum, we adopt the idea that the function of finiteness is to anchor the event to the moment of speech via the logophoric centre, which contains also the speech participants and is therefore crucial for subject/person interpretation (Bianchi 2003). We follow Ritter and Wiltschko (2014) in assuming that INFL is the anchoring head, which can be substantiated by different features cross-linguistically.³ It is thus necessary to investigate INFL in Jê more closely in order to understand how finiteness oppositions are realized.

3.2 INFL in Jê

Jê languages display the following basic clause structure, as shown in (5), which can be translated in generative terms as in (6):

(5) emphatic | TAME | NOM/ERG | ABS/ACC | [ACC/ABS=verb] (Bardagil 2018:129)

(6) [_{CP} left-peripheral items [_{IP} TAME NOM/ERG [_{VP} ABS/ACC V]]

In main clauses, there is a particle that marks a TAME-related category. This particle is located in the left field of the clause, preceding the subject position but following the position for left-dislocated elements. As seen in (7) and (8) for Kĩsêdjê, this particle is obligatory in main clauses, both with long and short verb forms (Nonato 2014:16):

(7) *(kôt) ka thãmã.
 INF.FUT you.SG.NOM fall.V
 Intended: ‘You may fall.’ (Kĩsêdjê; Nonato 2014:20)

(8) Nhũm kôt ngô thyk nhihwêrê?
 who INF.FUT water black make.N
 ‘Who would make the coffee?’ (Kĩsêdjê; Nonato 2014:5)

In all languages, there can only be one inflectional particle per clause, exactly like how there can only be one instance of tense inflection in English or other well-known Indo-European languages

³ For Romance, Groothuis (2020) argues that Fin is the anchoring head. At this stage, we do not have conclusive diagnostics to distinguish between Fin and higher IP heads in Jê, so further research is needed to confirm the exact anchoring head.

(Nonato 2014:18). This is an argument for interpreting them as lexicalizations of functional heads. We will consider these as INFL in all Jê languages, as already suggested by Nonato (2014) for Kĩsêdjê and by Bardagil (2018) for Panará.

The question arises what substantive content (Tense, Person, Location, or Mood) is associated with the INFL head in Jê. Contrastive m(orphological)-marking serves as the initial diagnosis for identifying the language-specific feature associated with INFL (Ritter & Wiltschko 2014:1341). When tense marking is not contrastive in a language, i.e. the absence of tense marking does not imply present tense, this means that Tense is not the relevant feature. If we apply this diagnostic to Jê languages, we see that in fact Tense marking is not obligatory:

(9) Ba ku= bĩ.
1SG.NOM 3SG.ACC kill.V
'I killed it.'
(Mêbêngôkre; Salanova 2007:32)

(10) Ba ku= by.
1SG.NOM 3SG.ACC grab.V
'I will grab it.'
(Mêbêngôkre; Salanova 2017:23)

(11) Ø Pasi=ra thẽ.
FACT Pasi=NOM go.V
'P. is gone/going.'
(Kĩsêdjê; Nonato 2014:16)

(12) Jy= ra= jõtĩ.
RLS.INTR 1PL.ABS sleep
'We slept / are sleeping.'
(Panará)

The absence of tense marking does not imply a non-past interpretation, as seen in (9) to (12), where the short verb form in Mêbêngôkre and Kĩsêdjê main clauses is compatible with both past and future interpretations. Tense is thus not the substantiating feature of INFL.

Instead, the function of the particles that lexicalize INFL varies in the descriptions of different Jê languages. In many of them, the substantive content of the anchoring head could correspond to Mood and/or Evidentiality, as the particles express meanings such as 'irrealis', 'factual future', or 'counterfactual' (Dourado 2001; Oliveira 2005:170; Nonato 2014:5). For instance, Panará clauses marked for irrealis mood can encode not only future (13), but also hypotheticals and counterfactuals (14).

(13) Pykkôdmã mãra hẽ ka= ti= a= sisryi ka.
morning 3SG ERG IRR 3SG.IRR 2ABS hit 2SG
'Tomorrow he will hit you.'
(Panará; Bardagil 2015:6)

(14) Ka= ti= rō kypa amã.
IRR 3SG.IRR fall ground INES
'It has sunk to the ground [*pretending to drop a stone*].'
(Panará; Bardagil 2015:6)

Just like in other Jê languages, the Panará modal proclitics (*jy* for realis intransitive and *ka* for irrealis) can only occur one time in any given clause. An overview of the INFL particles in the Northern Jê languages and their meaning as described in the literature can be found in Table 1:

Table 1: Northern Jê INFL particles

| Language | Particle | Meaning |
|---|----------------|------------------------|
| Apinajé (Oliveira 2005:170–171) | <i>na</i> | realis |
| | <i>kot</i> | irrealis |
| | <i>pre</i> | past |
| | <i>ra</i> | perfective |
| | <i>te</i> | habitual |
| Kajkwakhratxi (Camargo 2015:131–133) | <i>hên</i> | past |
| | <i>wã</i> | future |
| | <i>kwã~kaw</i> | habitual asp. |
| Kĩsêdjê (Nonato 2014:16) | <i>arân</i> | counterfactual |
| | <i>hên / Ø</i> | factual non-future |
| | <i>kê / Ø</i> | factual future |
| | <i>kôt</i> | inferential future |
| | <i>man</i> | witnessed |
| | <i>waj</i> | inferential non-future |
| Mêbêngôkre (Salanova 2007:131) | <i>nê / Ø</i> | non-future |
| | <i>dja</i> | future/irrealis |
| | <i>rãnh</i> | counterfactual |
| | <i>we</i> | evidential |
| Panará (Bardagil 2018:33) | <i>jy</i> | realis (intransitive) |
| | <i>ka</i> | irrealis |
| Timbira (Alves 2004:89–91) | <i>Ø</i> | realis |
| | <i>ha</i> | irrealis |
| | <i>pe</i> | past |
| | <i>jamã</i> | evidential |

As can be seen from the table, many of these particles have cognates in the other languages, such as Apinajé *na* and Mêbêngôkre *nê*; Apinajé *kot* and Kĩsêdjê *kôt*; Kajkwakhratxi *wã*, Kĩsêdjê *waj*, and Mêbêngôkre *we*, or Mêbêngôkre *dja*, Timbira *ha*, and Panará *ka*. We take this as an argument in favour of analyzing all these as modality/evidentiality-expressing realizations of INFL. These particles relate the reported event to the utterance in terms of *worlds* (cf. Nonato 2014:20–21).

Another characteristic of functional heads, pointed out also by Nonato (2014:19) is that these heads can be null, typically when they express an ‘unmarked’ feature. We expect thus that the realis particles can be optional or null. This is confirmed in Kĩsêdjê, where the particle *kê* is obligatorily null with 1st and 2nd person subjects and optional with 3rd person subjects (15), and in Mêbêngôkre, where the less marked non-future particle is only realized if the argument position to its left is occupied (16a). Otherwise, it is morphologically unrealized (16b).

(15) a. (Kê) wa khu=ku.
 FACT.FUT 1SG.NOM it.ACC=eat.V
 ‘I will eat it’.

b. (Kê) Ø khu=ku.
 FACT.FUT 3SG.NOM it.ACC=eat.V
 ‘He/she will eat it.’

(Kĩsêdjê; Nonato 2014:19)

(16) a. Kukryt nê ba ku= bĩ.
 tapir NFUT 1SG.NOM 3SG.ACC kill
 ‘A tapir, I killed.’

b. Ba kukryt bĩ.
 1SG.NOM tapir kill
 ‘I killed a tapir.’

(Mêbêngôkre)

In Timbira, the particle for the unmarked mood of realis is null (see Table 1). The (optionally) null realization of realis mood confirms our hypothesis that modality is the feature substantiating INFL in Jê languages.

Contrary to main clauses, the Jê INFL particles are banned from embedded non-finite (nominalized) clauses, as shown in (17) and (18):

(17) (*kukryt) (*nê) (*ije) arým ije Ø= bĩn.
 tapir NFUT 1SG.ERG already 1SG.ERG 3SG.ABS kill.N
 ‘... that I already killed tapir.’ (Mêbêngôkre; Salanova 2011:52)

(18) Hên wa [(*kôt) a= thêm] mba.
 FACT 1NOM (*INF.FUT) 2ABS fall.N know.V
 ‘I know you (*may) fall.’ (Kîsêdjê; Nonato 2014:5)

Thus, main clauses are solid candidates for finiteness: they are anchored to the speech act via the mood particles (lexicalizing INFL), both when they contain short and long verb forms. Clauses which are not anchored via these particles in INFL can hence be considered non-finite.

The absence of m-marking (in Jê, the absence of INFL particles) implies a non-deictic valuation of the [coin] feature. This raises the question of how these features are valued in that case. As discussed above, there are two alternative valuation mechanisms: valuation by Comp and predicate valuation. In Ritter and Wiltschko’s (2014) analysis, imperatives are one of the clause types where the [\pm coin] feature is not valued regularly but through the Complementizer head. This shows up as absence of the regular m-marking, or so-called ‘fake’ marking. The particles are indeed absent in Kîsêdjê imperatives, as is expected (19).

(19) (*kê) rik thẽ!
 FACT.FUT quick go.V
 ‘Go (quickly)!’ (Kîsêdjê; Nonato 2014:20)

As we will show in the next section, breaking down finiteness as a by-product of anchoring of the clause not only accounts for Jê verb forms, but it also sheds light in other areas of Jê morphosyntax, such as case marking. In fact, embedded clauses present a more complex challenge since they do not form a homogeneous category, as will be shown in the next section.

Before delving into the analysis section, we want to point out the presence of an additional functional element in the Jê clause. In (20) we can see the familiar INFL realis particle *na* that appears to the left of the lowest subject position, but there is also an Aspect particle *ra* to the right of the lower subject position.

(20) Na pa ra mō.
 RLS 1SG.NOM ASP go.V
 ‘I’ve come.’ (Apinajé; Oliveira 2005:172)

With different morphemes, cognate or not, this lower Aspect position is also consistent across Jê languages, as we see with perfective aspect for Mëbêngôkre in (21) and Xavante in (22).

(21) Ba nê ba arým ku= ma.
 1SG.NOM NFUT 1SG.NOM ASP 3SG.ACC hear.V
 ‘I have already heard it.’ (Mëbêngôkre)

(22) Bödi, e ma aa= wisi ni?
 grandchild.VOC Q ASP 2H.ABS arrive.V INDF
 ‘Grandchild, have you arrived?’ (Xavante; Estevam 2011:283)

4 Finiteness in Jê

So far, we have looked at the nature of INFL anchoring in Jê languages. We have seen that its substantive content can be other than Tense, such as Mood or Evidentiality, and that it is morphologically represented by a series of particles. In this section, we put forward our syntactic analysis of Jê clauses in relation with the anchoring category INFL, and present evidence to support the notion that, rather than a clear-cut distinction between finite and non-finite clauses, what we observe is a finiteness gradience.

4.1 Raising verbs in Jê

As seen in Section 2, the long forms used in Jê embedded clauses are non-finite, as diagnosed by the absence of a left-peripheral layer of functional projections that correspond to finiteness as clausal anchoring. However, not all embedded non-finite clauses behave in the same way in terms of their subject licensing. As we are going to see, Jê languages present two classes of embedded clauses which differ with respect to the case assigned to their subject.

As mentioned in Section 2, the use of a long form causes an alignment shift, whereby the long form assigns absolutive and ergative case. In most embedded clauses, the subject is therefore marked with ergative case, as can be seen in (23):

(23) [_{IP} Ø_{INFL} Ka [ire Ø= khuru] mũ].
 FUT 2SG.NOM 1SG.ERG 3SG.ABS eat.N see.V
 ‘You are going to see me eat.’ (Kĩsêdjê; Nonato 2014:7)

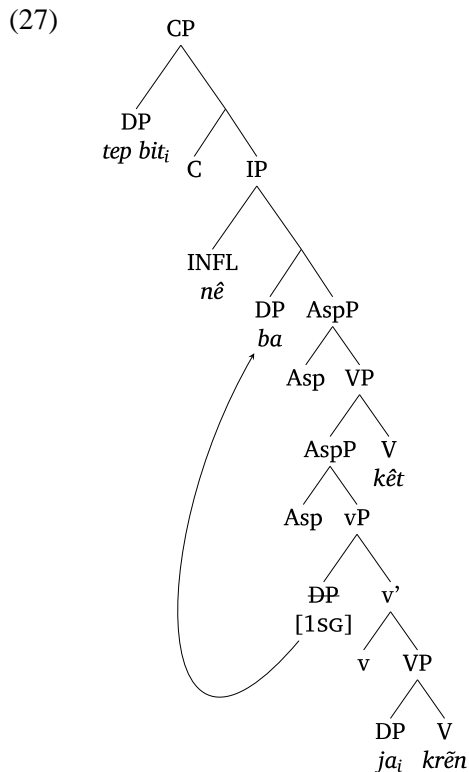
As opposed to (23), in (24), a long form is embedded under a negative verb (cf. Salanova 2007:58). The subject receives its thematic role from the lexical verb (viz. the embedded long verb) but raises to the higher clause. This can be seen from the nominative case marking on the subject argument of the lower clause, which can only be assigned by short verb forms (i.e. in a matrix clause). The same structure is also attested in Mëbêngôkre (25) and Apinajé (26).

(24) [_{CP} Thep_i wit [_{IP} nã_{INFL} wa [_{VP} Ø_i= khuru] khêrê]].
 fish only FACT 1SG.NOM 3SG.ABS eat.N NEG.V
 ‘Only fish didn’t I eat.’ (Kĩsêdjê; Nonato 2014:7)

(25) [_{CP} Tep_i bit [_{IP} nê_{INFL} ba [_{VP} ja_i krên] kêt]].
 fish only NFUT 1SG.NOM DET eat.N NEG.V
 ‘Only fish didn’t I eat.’ (Mëbêngôkre)

- (26) [_{IP} Kət paj [_{VP} a prɛ] ketnẽ].
 IRLS 1.NOM.IRLS 2ABS= tie.up.N NEG
 ‘I won’t tie you up.’ (Apinajé; Oliveira 2005:405 *apud* Gildea & Spike 2010:181)

We propose that these are raising verbs, which cause A-movement of the subject out of an embedded clause with a nominal verb to the main clause, where it is assigned nominative case. We can exclude an analysis of the Jê raising construction as \bar{A} -movement of the subject on the basis of its clause-internal position, following the INFL particle (21–22). See the following tree diagram for the proposed derivation:



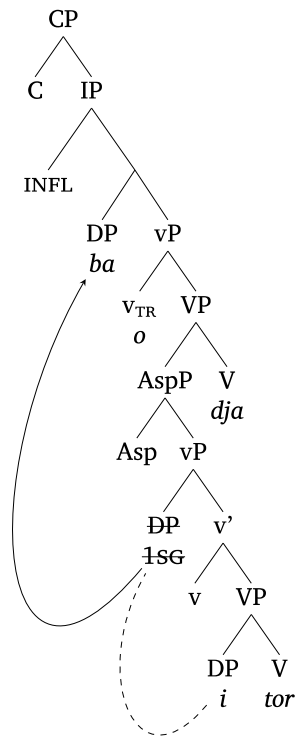
This type of construction in Jê is sometimes called nominative-absolutive (Gildea & Alves 2010; 2020), since it combines nominative marking of the subject of the embedded verb with absolutive marking of the embedded object. It is found with negation and focus structures, as in examples (23) to (26) above, but also with some posture-indicating verbs that are used, together with the transitivizing particle⁴ *o* to express progressive aspect (Salanova 2007:59). As we can see in (28), this Apinajé sentence also presents the subject of the verb in the embedded clause as surfacing with nominative case, just as it would present in the simple main clause version of the same sentence (29). In the case of (28), we additionally see that the raised-to-nominative subject is couched between INFL (realis *na*) and Aspect (progressive *ra*), the latter most likely marking the edge of the embedded clause (see discussion below).

⁴ This particle has been variously analyzed as a light verb (Oliveira 2005:295), a transitivizer (Bardagil 2018:152), and an applicative head (Salanova 2014; Bardagil 2018:152).

(28) Na **pa** ra [ASPP ik= tyk] o mō.
 RLS 1SG.NOM ASP 1.ABS= die.N do go.V
 ‘I’m dying.’ (Apinajé; Oliveira 2005:293)

(29) Na pa ty.
 RLS 1SG die.V
 ‘I died.’ (Apinajé; Oliveira 2005:293)

(30) [IP Ba [ASPP i= tor] o dja.
 1SG.NOM 1SG.ABS dance.N TR stand.V
 ‘I’m dancing.’ (Měbêngôkre; Salanova 2007:61)



In sum, we have proposed a raising analysis for a type of constructions in Jê languages that involve the subject of an embedded clause with a long form verb appearing on the matrix clause. The raising analysis is supported by the otherwise unexpected nominative case marking on the subject, and the fact that it appears to the left of an Aspect particle. Besides the novelty of this analysis, this is relevant for the present discussion in that it instantiates a type of Jê clause that cannot license its subject. As we discuss in the next section, this relates directly to the notion of finiteness in this language family.

4.2 Different types of non-finite complementation

The goal of this paper is to explore the notion of finiteness in the Jê family. Due to their morphosyntactic profile, finiteness in Jê languages is directly intertwined with nominality. The two notions have often been bundled together (Koptjevskaja-Tamm 1993). As we pointed out in Section

2.1, the form of the Jê verb is not indicative of the level of finiteness of the clause. In contrast, we would like to propose that Jê languages do not display a binary distinction between finite and non-finite clauses, but instead, Jê finiteness should be approached as a gradual notion with clauses that are more or less finite than others.

Finiteness is the presence of deictic anchoring in an utterance (Roussou 2001; Bianchi 2003; Ritter & Wiltschko 2014; Wiltschko 2014; Groothuis 2020). Jê languages present two types of deictic anchoring. On the one hand, we have TAME anchoring, as mediated by INFL (cf. Section 3.2). On the other hand, person anchoring also relates the clause to the context of the utterance and, as such, it too participates in the notion of finiteness. In fact, for Bianchi (2003), participants are part of the deictic centre to which clauses are anchored (cf. Section 3.1).

Based on the type of evidence presented in this paper, we propose the existence of three types of clauses in Jê languages based on their finiteness. **Clause type I** are main clauses, which are maximally finite in Jê languages. Both INFL and person anchoring are present, as seen by the fact that these clauses license a TAME particle realizing the INFL head. Main clauses can have either a verbal or a nominal verb as their main lexical verb (31). In this clause type subjects are licensed, even when raised from an embedded clause.

(31) [_{CP} left-peripheral items [_{IP} TAME subject.NOM/ERG [_{VP} ABS/ACC verb.N/V]

Clause type II are regular embedded clauses (32). They lack INFL anchoring, diagnosed by their inability to license the TAME particle corresponding to the INFL head. Jê embedded clauses present a reduced structure vis à vis main clauses, lacking a left periphery. Despite not having an INFL anchoring projection, they are capable of licensing their subject, with ergative-absolutive case marking.

(32) [[subject.ERG/ABS verb.N]]

These embedded clauses can be both complement and relative clauses. There is structurally no difference between these two clause types (cf. Salanova 2011:46).

Finally, in **clause type III**, we have raising embedded clauses, the least finite type (33). Raising embedded clauses lack the capability of licensing either a TAME particle or a subject. As a consequence of the latter, the subject raises to the higher clause, where it is able to be licensed and receive case marking.

(33) [subject.NOM_i [Ø_i verb.N] verb.V]

It is clear that Jê clause types II and III are structurally smaller than IPs, since they do not present INFL anchoring. We propose that type II clauses are AspPs. This is backed up by the fact that both are compatible with the presence of an Aspect particle, as in (34), which would correspond to the highest projection in their structure:

(34) (*kukryt) (*nẽ) (*ije) [_{ASPP} arým ije Ø= bĩn].
 tapir NFUT 1SG.ERG already 1SG.ERG 3SG.ABS kill.N
 ‘... that I already killed tapir.’ (Mêbêngôkre; Salanova 2011:52)

This is also consistent with existing proposals in the Jê literature that connect ergative case with a low licenser (Coon & Salanova 2009; Nonato 2014; Bardagil 2018). Thus, nominative case would be licensed by a higher head, which corresponds to the INFL head in the present proposal and whose

absence in clause types II and III is consistent with the unavailability of nominative case in these clauses.

The question arises how much structure is present in the complements to raising verbs. If we look at the Apinajé example in (35), we note the presence of an aspectual particle *ra*:

- (35) [_{CP} Ka [_{IP} na ka [*ra* ik pe a= pikudor] o mō]].
2SG.NOM RLS 2SG.NOM ASP 1SG MAL 2SG.ABS disappear.N TR go.V
‘You’re already disappearing from me.’ (Apinajé; Oliveira 2005:294)

Given the currently available evidence, there is no way of determining whether *ra* ‘already’ is located in the matrix or in the embedded clause. Under both analyses, the conclusion is that the nominative subject (*ka* ‘you.NOM’) has raised to the matrix clause: if *ra* is in the matrix clause, *ka* has to be too, since it precedes *ra*; if *ra* is in the embedded clause, it determines the edge of this embedded clause (cf. (34)). Further research will have to show whether clause types II and III are both AspPs or whether clause type III is even more reduced in structure than type II, perhaps as small as a vP.

Since person anchoring is part of our notion of finiteness, we can say that the clause types II and III differ in finiteness. Whereas type II does not allow deictic modal anchoring, it still licenses its own referential subject, which is interpreted directly with respect to the speech act. In type III, the subject needs to raise to the matrix verb and is thus necessarily coreferential with the matrix clause subject (i.e. subject control). Type II occupies therefore an intermediate position between type I and type III in terms of the degree of finiteness.

Combining TAME anchoring via INFL with person anchoring via subject licensing, a second type of deictic anchoring, we can get a more fine-grained view of different degrees of finiteness: besides fully finite main clauses, Jê languages present different degrees of non-finiteness in embedded clauses. A similar interplay of two anchoring mechanisms has also been proposed for Romance languages (Groothuis 2020).

5 Conclusions and questions for further research

In this paper, we have investigated the notion of finiteness by looking at the Jê language family. We have extended Nonato’s (2014) analysis of Kĩsêdjê to the whole family, assuming that the preverbal particles are the lexicalization of the anchoring head INFL (cf. Ritter & Wiltschko 2014). The feature that substantiates this anchoring head seems to be Mood (or Evidentiality) in the languages of the Jê family.

Finite clauses, or type I clauses, are characterized by presenting both an INFL head and licensing their subject. In non-finite (embedded) contexts, the long verbal form is used. However, the long verbal form is not restricted to these contexts: it can also appear in main clauses with an INFL particle. Non-finiteness thus implies the use of the long verbal form, but the reverse does not hold: the long form can appear also in finite (i.e., main) clauses. We can conclude therefore that verbal morphology is not a reliable indication of the level of finiteness of a clause in Jê languages (cf. Groothuis 2020’s discussion of Romance).

There are two non-finite environments which differ in terms of subject licensing. In type II clauses, referential subjects are licensed and marked with ergative case by the nominal form of the verb. In contrast, type III clauses are better analyzed as complements to raising verbs, because they cannot case-license their own subject; it has to instead move to the matrix clause, where it is case-marked with nominative case in the context of a short verb form.

In sum, the data from Jê languages lend support to the idea that the substantive content of INFL can vary, and that finiteness oppositions are not to be understood as presence vs. absence of tense, but in terms of direct vs. indirect deictic anchoring, as instantiated by INFL and by person anchoring.

References

- Alves, Flávia de Castro. 2004. *O Timbira falado pelos Canela Apãniekrá: uma contribuição aos estudos da morfossintaxe de uma língua Jê*. Doctoral dissertation, Universidade Estadual de Campinas, Campinas, Brazil.
- Amritavalli, R. 2014. Separating tense and finiteness: anchoring in Dravidian. *Natural Language & Linguistic Theory* 32:283–306.
- Bardagil, Bernat. 2018. *Case and agreement in Panará*. Utrecht, Netherlands: LOT.
- Bianchi, Valentina. 2003. On finiteness as logophoric anchoring. In Jacqueline Guéron & Liliane Tasmowsky (eds.), *Temps et point de vue/Tense and point of view*. Paris, France: Nanterre, 213–246.
- Camargo, Nayara da Silva. 2015. *Tapayuna (Jê): aspectos morfossintáticos, históricos e sociolinguísticos*. Doctoral dissertation, Universidade Estadual de Campinas, Campinas, Brazil.
- Coon, Jessica & Andrés Pablo Salanova. 2009. Nominalization and predicate-fronting: Two sources of ergativity. *University of Pennsylvania Working Papers in Linguistics* 15(1):45–54.
- Davis, Irvine. 1966. Comparative Jê phonology. *Estudos Lingüísticos: Revista Brasileira de Lingüística Teórica e Aplicada* 1(2): 20–24.
- Dourado, Luciana Gonçalves. 2001. *Aspectos morfossintáticos da língua Panará (Jê)*. Doctoral dissertation, Universidade Estadual de Campinas, Campinas, Brazil.
- Estevam, Adriana Machado. 2011. *Morphosyntaxe du xavante. Langue jê du Mato Grosso (Brésil)*. Doctoral dissertation, Université Paris Diderot (Paris 7), Paris, France.
- Gildea, Spike & Flávia de Castro Alves. 2010. Nominative-Absolutive: Counter-Universal Split Ergativity in Jê and Cariban. In Spike Gildea & Francesc Queixalós (eds.), *Ergativity in Amazonia*. Amsterdam, Netherlands: John Benjamins, 159–199.
- Gildea, Spike & Flávia de Castro Alves. 2020. Reconstructing the source of nominative-absolutive alignment in two Amazonian language families. In Eugenio Luján, Jóhanna Barðdal, & Spike Gildea (eds.), *Reconstructing Syntax: Cognates and Directionality*. Leiden, Netherlands: Brill, 47–107.
- Groothuis, Kim A. 2020. *Reflexes of finiteness in Romance*. Doctoral dissertation, University of Cambridge, Cambridge, UK.
- Holmberg, Anders & Christer Platzack. 1995. *The role of inflection in Scandinavian syntax*. Oxford, UK: Oxford University Press.
- Joseph, Brian D. 1983. *The synchrony and diachrony of the Balkan infinitive: A study in areal, general, and historical linguistics*. Cambridge, UK: Cambridge University Press.
- Koptjevskaja-Tamm, Maria. 1993. *Nominalizations*. London, UK: Routledge.

- Ledgeway, Adam N. 2007. Diachrony and finiteness: subordination in the dialects of Southern Italy. In Irina A. Nikolaeva (ed.), *Finiteness: Theoretical and Empirical Foundations*. Oxford, UK: Oxford University Press, 335–365.
- Nikolaeva, Irina A. 2007. Introduction. In Irina A. Nikolaeva (ed.), *Finiteness. Theoretical and empirical foundations*. Oxford, UK: Oxford University Press, 1-19.
- Nikulin, Andrey. 2020. *Proto-Macro-Jê: um estudo reconstrutivo*. Doctoral dissertation, Universidade De Brasília, Brasília, Brazil.
- Nonato, Rafael. 2014. *Clause chaining, switch reference and coordination*. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.
- Oliveira, Christiane Cunha de. 2005. *The language of the Apinajé people of Central Brazil*. Doctoral dissertation, University of Oregon, Eugene, WA.
- Ribeiro, Eduardo Rivail & Hein van der Voort. 2010. Nimuendajú was right: The inclusion of the Jabutí language family in the Macro-Jê stock. *International Journal of American Linguistics* 76(4):517–570.
- Ritter, Elizabeth & Martina Wiltschko. 2014. The composition of INFL. An exploration of tense, tenseless languages and tenseless constructions. *Natural Language & Linguistic Theory* 32:1331–1386.
- Rodrigues, Aryon D. 1999. Macro-Jê. In Robert MW Dixon & Alexandra Aikhenvald (eds.), *The Amazonian languages*. Cambridge, UK: Cambridge University Press, 165–206.
- Roussou, Anna. 2001. Control and Raising in and out of Subjunctive Complements. In Maria-Luisa Rivera & Angela Ralli (eds.), *Comparative Syntax of Balkan Languages*. Oxford, UK: Oxford University Press.
- Salanova, Andrés Pablo. 2007. *Nominalizations and aspect*. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.
- Salanova, Andrés Pablo. 2011. Relative clauses in Mëbengokre. In Rik van Gijn, Katharina Haude, & Pieter Muysken (eds.), *Subordination in native South-American languages*. Bogotá, Colombia: John Benjamins, 45–78.
- Salanova, Andrés Pablo. 2014. Semántica causativa, sintaxis aplicativa. In Ana Carla Bruno, Francesc Queixalós, & Stella Telles (eds.), *Incremento de valencia en las lenguas amazónicas. Bogotá: Instituto Caro y Cuervo*. Bogotá, Colombia: Instituto Caro y Cuervo, 155–189.
- Salanova, Andrés Pablo. To appear. Macro-Jê. In Patience Epps & Lev Michael (eds.), *Amazonian Languages: a Handbook*, volume 3. Berlin, Germany: De Gruyter Mouton.
- Urban, Greg. 1985. Ergativity and accusativity in Shokleng (Gê). *International Journal of American Linguistics* 51(2):164–187.
- Wiltschko, Martina. 2014. *The Universal Structure of Categories. Towards a Formal Typology*. Cambridge, UK: Cambridge University Press.