Abstract:
This paper examines the likely diachronic sources from which the nominative case markers tóg and ra developed in the Kaingáng languages, and Kísêdjê (Suyá) and Tapayuna respectively, in which they have been referred to generically as ‘mark of subject’. The starting point of the study is the pioneering work of Professor Ludoviko Carnasciali dos Santos and, at the same time, explores this property in the first two languages, to which he dedicated a large part of his academic production. Data analysis is based on grammaticalization theory, according to which it was allowed to establish distinct diachronic sources for nominative case markers, although they are related to the conceptual domain of spatial deixis. Given the different grammatical functions presented by nominative case markers, it is argued that this multiplicity of uses constitutes an exemplary case of polygrammaticalization. Although these case markers have emerged from different lexical sources, the results obtained connect the Kaingáng, Kísêdjê (Suyá) and Tapayuna languages in terms of this grammaticalization process and shed light on grammatical characteristics still little explored in synchronic and diachronic studies of Jê languages (Macro-Jê).

Keywords:
‘Subject Marker’ in Some Jê Languages (Macro-Jê):
Diachronic Sources and Grammaticalization

Maxwell Miranda

INTRODUCTION

Grammatical relations and their formal expression in different case marking patterns have been a reasonably well-explored topic from languages of the Jê family (Macro-Jê). Although these languages are close from a lexical point of view, there is a notable variation between them regarding the grammatical mechanisms in which syntactic functions are coded and marked in each clausal pattern. Among these, the subject function has received the most attention, due to the fact that in certain types of predicates it is the marked constituent, whose differential treatment is generally based on the transitivity and finiteness of the sentence. Alongside these more typical cases, some languages, such as Laklãnõ (Xokléng) (GAKRAN, 2005, 2015), Kaingáng (WIESEMANN, 1986, 2002), Kisêdjê (Suyá) (SANTOS, 1997, 1999a, 1999b) and Tapayuna (CAMARGO, 2015), exhibit a pattern in which nominal arguments in the role of subject receive a specific grammatical marker, regardless of the transitivity of the sentence and the type of predicate, as well as the grammatical properties that regulate it can vary from one language to another. Furthermore, these languages differ in terms of form and number of grammatical markers, which have been conventionally referred to in the literature on Jê languages by the terms ‘subject indicator’ (WIESEMANN, 2002, p. 159), ‘suffixal particle’ (GUEDES, 1993, p. 102), ‘subject mark’ (CAMARGO, 2015, p. 152; GAKRAN, 2015, p. 163; GONÇALVES, 2011, p. 13-14; NASCIMENTO, 2013, p. 38; SANTOS, 1997, p. 100) or ‘nominative case’ (D’ANGELIS, 2004, p. 73; NONATO 2014, p. 13).

The starting point (not to mention the restlessness) of this study was the description of the Kisêdjê (Suyá) language by Professor Ludoviko Carnasciali dos Santos1, as I have sought to relate certain grammatical properties to as many languages as possible in recent works (CABRAL; MIRANDA; GAKRAN, 2018; MIRANDA, 2019, 2020, 2021). The work of Santos (1997) is characterized by its pioneering spirit, by exploring and addressing some of the more typical grammatical characteristics of these languages and, at the same time, incorporating other topics that have become essential in Jê Linguistics, for example, the nominal nature of verbal forms distinguishable on the basis of their morphological and syntactic behavior in the clause2 or the case marking system and their splits that manifest themselves in different grammatical configurations3. Subsequent linguistic descriptions directly or indirectly began to include such topics within the specificities of each language.

The purpose of this article is to examine the diachronic origin and, consequently, the path of grammaticalization of the morphemes tóg and ra (~ ta) in the Kaingáng, Kisêdjê (Suyá) and Tapayuna languages, respectively, known as ‘subject mark’. Interestingly, the two first languages were those to which

1 The linguistic description was presented as a doctoral thesis, in 1997, at the Universidade Federal de Santa Catarina, under the guidance of Prof. Dr. Lucy Seki. Before Santos (1997), the Suyá language had only the description of Guedes (1993), which deals with phonology and grammatical aspects.

2 Shell and Quain (1952, p. 119, 121) were the first to mention the existence of a suffixation process in verbal stems from the Krahô language. Popjes and Popjes (1986), in their analysis of the varieties Canela (Ramkôkamêria) and Krahô, introduced the terminology long form and short form of the verb, without attributing a morphological status to such contrast in the latter case. In both analyses, the long form is commonly associated with the past tense (POPJES; POPJES, 1986, p. 192) and with the non-final position of the clause, when the verb is not followed by other elements.

3 Urban (1985, p. 186-187) had reached similar conclusions when calling attention to the particularities of the Xokléng language with regard to split ergativity in it and in other Jê languages.
Professor Ludoviko Carnasciali dos Santos devoted more time in his studies and academic production (CABRAL, in this volume, brings together Santos’ main works). I will argue that these marks come from distinct spatial deixic demonstratives, from which several grammatical functions developed and still coexist in most of these languages. Although the multiple attested functions cannot be intertwined in a chain of grammaticalization (HEINE, 1992), the linguistic data examined point to a scenario of polygrammaticalization (CRAIG, 1991).

The paper is organized as follows: in section 2, I examine the clause constructions in which nominal arguments receive the ‘subject mark’ from the Kĩsêdjê (Suyá) and Tapayuna (2.1) and Kaingáng (2.2) languages. Section 3, in turn, is intended to identify the likely diachronic sources and the grammatical functions they play in the analyzed languages. In section 4, I argue that the marked nominative constitutes an example of polygrammaticalization (CRAIG, 1991), in which the different attested functions cannot be arranged in a linear order, despite all being related to the domain of spatial deixis.

1. **Subject-marked Constructions: Morphological and Syntactic Properties**

In the description of some languages of the Jê family (Macro-Jê), ‘subject mark’ (and its variations) is a term adopted referring to grammatical morphemes that accompany nominal arguments and, to a lesser extent, also pronominal ones. These morphemes manifest themselves in different ways in a reduced number of languages of this family, as is the case of Lakkânoto (Xoklêng), Kaingáng, Kĩsêdjê (Suyá) and Tapayuna. For the purposes of this article, the clause constructions and syntactic restrictions that these languages, except Lakkânoto (Xoklêng)⁴, have regarding the occurrence of such grammatical morphemes are examined. First, these constructions are analyzed in Kĩsêdjê (Suyá) and Tapayuna (2.1), and, subsequently, in Kaingáng (2.2).

1.1. ‘Subject mark’ in Kĩsêdjê (Suyá) and Tapayuna

Kĩsêdjê (Suyá) and Tapayuna are the only languages of the Northern branch of the Jê family in which the nominal subject of verbal and non-verbal predicates is marked differently from that of constructions in which it is expressed by a personal pronoun. In them, the subject noun phrase is followed by the marker *ra* (CAMARGO, 2015, p. 152; SANTOS, 1997, p. 100), which presents the allomorph *ta* only in Kĩsêdjê (Suyá), after nouns ending in a voiceless consonant. In this case, the marking of the nominal subject is in complementary distribution with that of the pronominal one, in which different patterns of case marking are triggered for the coding of core arguments, depending or not on the nominalization of the clause. In (1)⁵, the subject noun phrase of intransitive and transitive sentences is marked by *ra*, in contrast to those in (2) and (3), where the subject is a personal pronoun and does not occur with it. For the sake of clarity, data from Santos (1997) have been adapted and reinterpreted by me.

(1) a. karupi ra ɲ̃ɡʌ mã tẽ
   N.PESS SM men’s house POSP go
   ‘Karupi went to the men’s house.’ (SANTOS, 1997, p. 27)

⁴ Although the Lakkânoto (Xoklêng) and Kaingáng languages have the morpheme *ву/vỹ* respectively, which has also been referred to as ‘subject mark’, it is not discussed in this article because its diachronic origin is less evident than the one formulated for the morpheme *togh*.

⁵ Abbreviations: 1 = First person; 2 = Second person; 3 = Third person; *ADV* = Adverb; *ANAF* = Anaphoric (Pronoun); *ASP* = Aspect; *AUX* = Auxiliary; *DAT* = Dative; *DEIT* = Deictic; *DEM* = Demonstrative; *DM* = Discursive marker; *ENF* = Emphatic; *ERG* = Ergative; *FEM* = Female; *PURP* = Purpose; *FUT* = Future; *IMPERF* = Imperfective (Aspect); *INCL* = Inclusive; *INDEF* = Undefined; *INST* = Instrumental; *LOC* = Locative; *MASC* = Male; *SM* = Subject mark; *NEG* = Negation; *N.PESS* = Personal name; *NMLZ* = Nominalizer; *NOM* = Nominative; *OBJ* = Oblique; *PERF* = Perfective (Aspect); *PL* = Plural; *POSS* = Possessive; *POSP* = Postposition; *R*¹ = Contiguity relational prefix; *R*² = Non-contiguity relational prefix; *REFL* = Reflexive; *SG* = Singular; *SUB* = Subordinator; *TOP* = Topic.
In addition to subject marking, other formal contrasts are notable, especially in relation to verb form, when the core of main clauses is modified or structurally dependent on post-verbal elements. These elements include certain types of adverbs (mood, intensity and negation) (3b) or the morpheme mą to express future tense (2b). The dependency relation with such elements triggers the nominalization of the clause, whose syntactic functions expressed by independent personal pronouns (nominatives) are readjusted to those typical of possessive phrases, in which the S/O arguments are directly indexed in the verbal head as the possessor of inalienable nouns, while the argument A is marked by the morpheme ɛr in an oblique phrase. In other languages, such as Krahô (Timbira), Xavante and Xerente, this morpheme corresponds to the genitive postposition tɛ/te with an analogous grammatical function in genitive phrases, possessive predicates and nominalized sentences (MIRANDA; COSTA, 2019).

There are still cases that clearly deviate from what would be expected, albeit less frequently, such as, for instance, the possibility of the direct object also receiving such a ‘subject mark’, as in (4).

The Tapayuna language behaves similarly to Kísêdjê (Suyá) from the point of view of the distribution of the ‘subject mark’ in verbal predicates, such as (5a) and (5b), and non-verbal (5c), but not in relation to the marking of the direct object. In addition, there are situations in which the ‘subject mark’ also does not occur, as in (6).
Because they are very close languages and, in fact, the only ones in the northern branch to exhibit this case-marking pattern, Kĩsêdjê (Suyá) and Tapayuna have been characterized as nominative case-marked languages (CAMARGO, 2015, p. 193; SANTOS, 1997, p. 169). In typological terms, it is not uncommon for languages to have formal means of marking S/A arguments as opposed to O (see, for example, König’s (2006) discussion of marked nominatives in African languages). The issue to be considered here is which internal and/or external mechanisms would have favored the emergence and development of this pattern in these languages, thus distinguishing them from the other members of the Northern branch.

1.2. Kaingáng

The Kaingáng language has several grammatical elements that can mark the subject of the clause. In the descriptive tradition of this language, such elements have been generically called ‘subject indicators’ (WIESEMANN, 1986, p. 362; 2002, p. 159), ‘subject marker’ (ALMEIDA; SANTOS, 2008, p. 234; GONÇALVES, 2011, p. 13-14; SANTOS, 2007, p. 147) sometimes as ‘nominative case markers’ (D’ANGELIS, 2004, p. 73). From the lexical point of view, these markers (henceforth SM) would correspond to ‘nominative postpositions’ (WIESEMANN, 1986, p. 362), which, for the purposes of this analysis, is of particular interest to us in the form tóg.

Unlike other subject marks⁶, the tóg form exhibits greater syntactic flexibility and much broader uses, for example, the constituent marked by it follows the verb and, in this case, does not emphasize the subject (7a), as it is the only one element with which third-person pronouns can occur (7b) or even combine with other markers like vỹ, indicating that the subject is the topic of the sentence (7c). Kaingáng data come from the variety spoken in Paraná (WIESEMANN, 2002) and were reinterpreted by me.

(7a) wǐtfi ra thi
    alligator SM die
    ‘The alligator died.’ (CAMARGO, 2015, p. 85)

(7b) kè ra wèthō ra a-kura
    stone INST someone SM 2sg-hit
    ‘Someone hit you with a stone.’ (CAMARGO, 2015, p. 87)

(7c) ŋgojra ra kahri
    cup SM full
    ‘The cup is full.’ (CAMARGO, 2015, p. 167)

Because they are very close languages and, in fact, the only ones in the northern branch to exhibit this case-marking pattern, Kĩsêdjê (Suyá) and Tapayuna have been characterized as nominative case-marked languages (CAMARGO, 2015, p. 193; SANTOS, 1997, p. 169). In typological terms, it is not uncommon for languages to have formal means of marking S/A arguments as opposed to O (see, for example, König’s (2006) discussion of marked nominatives in African languages). The issue to be considered here is which internal and/or external mechanisms would have favored the emergence and development of this pattern in these languages, thus distinguishing them from the other members of the Northern branch.

1.2. Kaingáng

The Kaingáng language has several grammatical elements that can mark the subject of the clause. In the descriptive tradition of this language, such elements have been generically called ‘subject indicators’ (WIESEMANN, 1986, p. 362; 2002, p. 159), ‘subject marker’ (ALMEIDA; SANTOS, 2008, p. 234; GONÇALVES, 2011, p. 13-14; SANTOS, 2007, p. 147) sometimes as ‘nominative case markers’ (D’ANGELIS, 2004, p. 73). From the lexical point of view, these markers (henceforth SM) would correspond to ‘nominative postpositions’ (WIESEMANN, 1986, p. 362), which, for the purposes of this analysis, is of particular interest to us in the form tóg.

Unlike other subject marks⁶, the tóg form exhibits greater syntactic flexibility and much broader uses, for example, the constituent marked by it follows the verb and, in this case, does not emphasize the subject (7a), as it is the only one element with which third-person pronouns can occur (7b) or even combine with other markers like vỹ, indicating that the subject is the topic of the sentence (7c). Kaingáng data come from the variety spoken in Paraná (WIESEMANN, 2002) and were reinterpreted by me.

(7a) inh=fa vỹ nũ-v ky tóg kānhpar mü
    1sg=leg top twist-NMLZ DM SM swollen AUX.PERF
    ‘I twisted my leg, that’s why it’s swollen.’ (Lit. My leg (is) twisted, so it swelled)
    (WIESEMANN, 2002, p. 74)

---

⁶ The number of elements that mark the subject may vary according to the author. In addition to tóg, Wiesemann (2002, p. 160), for example, registers nine more forms (jé, mỹ, ne, né, nỹ, pijé, tỹ, vé and vỹ), while D’Angelis (2004, p. 73) presents a much smaller set that includes the forms tag, tỹ, vỹ and ne.
Another syntactic property of tóg not mentioned by Wiesemann (2002) is related to the possibility of marking circumstantial adjuncts, as observed in the examples in (10). However, the form tóg does not occur with direct and indirect objects.

(8) a. fênhêng tóg gâr kagra=ja nì
tântó gûnì
armadillo SM corn eat=PERF be.sitting.AUX
‘The armadillo ate all the corn.’ (WIESEMANN, 2002, p. 18)

b. mru tóg sî pê nì
woodpecker SM small ADV be.sitting.AUX
‘The woodpecker is very small.’ (WIESEMANN, 2002, p. 62)

(9) a. ūn ter kîy ẽg tóg fy tî
INDEF die.NMLZ SUB 1PL SM cry AUX.IMPERF
‘When someone dies, we cry.’ (WIESEMANN, 2002, p. 19)

b. vâju kamê ā=tóg nì
smoke ADV 2SG=SM be.sitting.AUX
‘You smoke a lot.’ (WIESEMANN, 2002, p. 93)

The form tóg appears in both verbal and non-verbal predicates in Kaingâng. Besides the nominal subject, as in (8), a distinctive feature of this language in contrast to Kîsêdjê (Suyá) and Tapayuna lies in its occurrence with personal pronouns, according to example (9).

(10) a. ūri tóg, inh=prûfêr tî
today SM 1SG=hunger.for.meat AUX.IMPERF
‘Today I’m hungry for meat.’ (WIESEMANN, 2002, p. 75)

b. goj nig ki tóg, mro hâ nì
water well LOC SM bathe ADV be.sitting.AUX
‘It’s nice to swim in the lagoon.’ (WIESEMANN, 2002, p. 65)

The clauses provided in (10) have in common the characteristic of presenting the syntactic constituent marked by tóg in a sentence topic position. By observing the distribution of tóg in different syntactic environments, it is expected that the function of marking the subject, specifying a syntactic role in the sentence, being more grammatical, still coexists with those less grammatical functions as an element marking topicalized constituents, such as in (10). The next section examines the probable diachronic sources that resulted in the grammaticalization of specific morphemes to mark the S/A subject in the Kaingâng, Kîsêdjê (Suyá) and Tapayuna languages.
In Kaingang, Kisédjê (Suyá) and Tapayuna, the forms tòg and ra (~ ta) described as ‘subject mark’ correspond to nominative case markers in their most grammatical uses. From a diachronic point of view, case markers develop more frequently from adpositions, adverbial words with semantic properties similar to adpositions, demonstratives or articles, the latter being related to pronouns (HARRIS; CAMPBELL, 1995, p. 341; KULIKOV, 2009, p. 440-447). In most of the Jê languages, several argument markers (S/A) of verbal and non-verbal predicates originated from postpositions that originally express genitive, dative, instrumental and locative case (MIRANDA; COSTA, 2019; MIRANDA, 2021).

From a comparative point of view, several Northern Jê languages have the form ta, whose lexical status and grammatical uses vary widely. This is the case of the languages spoken by the Apinajé, Canela (Ramkôkamekra) and Krahô peoples, for which Oliveira (2005, p. 410) and Popjes and Popjes (1986, p. 175) described the forms tam and ta as ‘third person emphatic’, respectively, while Miranda (2014, p. 107) interpreted it for the latter language as an ‘emphatic’ form, as in (11).

(11) ta, wa nê a=mã hô-r nare
    ENF 1SG NEG 2SG=DAT give-NMLZ NEG
    ‘This one, I don’t give it to you.’ (MIRANDA, 2014, p. 107)

In the Xikrin language of Catetê, Costa (2015, p. 78) includes the form ta in the set of personal pronouns as a third-person expression, and can exercise both a nominative and absolutive function. A similar situation is the Laklânõ language (Xokléng), whose third-person masculine singular pronoun (nominative series) corresponds to the form ta (GAKRAN, 2015, p. 176). In Xavante, Lachnitt (2003, p. 68) registers ta as a “demonstrative-prefix”, meaning ‘this one’. Other languages and, in some cases, the same language seem to combine the form ta with emphatic, deictic (locative) or indefinite elements, resulting in the formation of demonstratives, such as ta ne ‘that’, ‘he’, that Jefferson (2013, p. 248) describes for the Mêbêngôkre language spoken by the Mêtykire people (Txucarramãe). In Xerente, ta bô expresses ‘that’, ‘there’ (SOUZA FILHO, 2007, p. 134), while in Laklânô (Xokléng) the combinations hâ ta (-close to the speaker, +visible) and ô ta (- close, +/- visible) are more grammaticalized and integrate the demonstrative system of the language (GAKRAN, 2015, p. 107). For the Xikrin do Catetê language, Costa (2015, p. 85) mentions the possibility of ta combining with the demonstratives ja ‘this(a)’ and ôa ‘that’, producing the compound forms ta ja ‘he, who is close to the speaker’ and ta ôa ‘he, who is far from the speaker’. Kaingang, in turn, presents the demonstrative tô that Wiesemann (2002, p. 84) registers as “over there.” In this set of languages, Tapayuna is the only member in which the (indefinite) morpheme tôô joins the form ra, obtaining the combination tôôa ‘far from the listener (far from both)’ (CAMARGO, 2015, p. 109). Thus, it is conceivable that the form ra was originally a deictic-spatial demonstrative, which corresponds to the proto-form *ta, *tam of Proto-Jê (DAVIS, 1966, p. 23) and to the form *tû (kô) reconstructed by Jolkesky (2010, p. 233) for Southern Proto-Jê, meaning ‘localized invisible distal deictic’.

In Kisédjê (Suyá), the forms ta/yta and the demonstrative pronouns ila (close to the speaker), ata (close to the listener) and nira (far from the speaker and listener) are used in reference to the third person (singular/plural), which constitute a specific pronominal series (Series IV), as shown in Table 1. The pronouns

From the form yta one can deduce the number morpheme ay, with which the form ta combines. The morpheme ay expresses paucal/plural number and is the result of the grammaticalization of the collective word *ari ‘limited group’ common to other Jê languages, such as Mêbêngôkre and Panará (MIRANDA, 2020, p. 259-263). The current form is therefore the product of eliminating of the consonant /r/, producing the monosyllabic form ai, but registered in the orthographic writing of the language as <ay>.

The suffix -ye is added to these demonstratives and thus the plural forms itaye, ataye and niraye are obtained (SANTOS, 1997, p. 61).
of this series differ from the others “because they participate in emphatic constructions as a topological element” (my emphasis) (SANTOS, 1997, p. 58). These are, in fact, constructions in which the pronoun is topologicalized by the particle *n(a) ‘topic’.

Table 1 - Series IV personal pronouns in Kĩsêdjê (Suyá).

<table>
<thead>
<tr>
<th></th>
<th><strong>Singular</strong></th>
<th><strong>Plural</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><em>pa</em></td>
<td><em>aypa</em></td>
</tr>
<tr>
<td>2</td>
<td><em>ka</em></td>
<td><em>ayka</em></td>
</tr>
<tr>
<td>1+2</td>
<td><em>kupa</em></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td><em>ta (ita, ata, nirá)</em></td>
<td><em>ayta (itaye, niraye)</em></td>
</tr>
</tbody>
</table>

**Source**: Santos (1997, p. 46).

Based on these facts, it is plausible that the subject mark *ra (~ *ta*) in Kĩsêdjê (Suyá) and Tapayuna originated from a spatial deictic demonstrative, which corresponds to the forms and functions attested in languages such as Kaingáng and Xavante. Demonstratives are one of the main sources for the grammaticalization of third-person pronouns (DIESSEL, 1999, p. 119; HEINE; SONG, 2011, p. 595; KUTEVA et al., 2019, p. 142). As a personal pronoun, the form *ta* is found only in Kĩsêdjê (Suyá), where it can replace a nominal referent and be followed by the topic marker *n(a)*, as in (12). However, when the reference to the third person is made by demonstratives, they receive the marker *ra* both in that language and in Tapayuna, as indicated by the examples in (13) and (14) respectively.

(12) **ta=n asi kake**  
3SG=TOP REFL scratch  
‘He scratched himself.’ (SANTOS, 1997, p. 49)

(13) **ita ra wa-mũ**  
3SG SM 1PL.INCL-see  
‘He saw us.’ (SANTOS, 1997, p. 53)

(14) **kere, atha ra ku-re wîtî wîrî kere**  
NEG DEM SM 3SG-ERG alligator kill NEG  
‘No, he didn’t kill the alligator.’ (CAMARGO, 2015, p. 104)

The form *tóg* in the Kaingáng language presents a more complex situation in relation to the cases examined so far, due to the different syntactic contexts in which it can appear and the grammatical functions it can perform in each of them (Section 2). It is likely that *tóg* is also related historically to an ancient deictic-spatial demonstrative similar to that proposed for the form *ra (~ *ta*)* in Kĩsêdjê (Suyá) and Tapayuna. In the variety of Kaingáng spoken in Paraná, there is evidence that such an element existed at some historical stage, as indicated in the examples in (15), alongside the most common demonstratives recorded by Wiesemann (2002, p. 161).

---

9 This spatial deictic demonstrative should not be confused with the current homonymous form *tóg* in Laklânò (Xokléng), since this is the synchronic reflection of the protoform *tóg* ‘visible proximal deictic’ that Jolkesky (2010, p. 233) reconstructs to the Southern Proto-Jê.
Another grammatical function of tóg is that of a localized/definite third-person pronoun, whose use is dependent on the discursive-pragmatic context, as is typical of cases involving deictics in a pronominal function. The clauses in (16) show the use of tóg with this grammatical function, in which the third-person subject is not expressed in the subordinate clause, but its reference is expressed in the main clause by the form tóg.

(16) a. ti=sâpe kugrŷ-r vâm jé tóg výr
   3SG.MASC=hat sew,PL-NMLZ sell PURP 3SG go,PERF
   ‘He went to sell his sewn hats.’ (Lit. To sell his sewn hats, he went) (WIESEMANN, 2002, p. 54)

b. kŷ tóg inh=nïjë ki tân th mû
   DM 3SG=1SG=nose LOC hit AUX,PERF
   ‘Then he hit me hard on the nose.’ (WIESEMANN, 2002, p. 84)

As an extension of this function, the form tóg can be used to anaphorically resume topicalized syntactic constituents in the subject function of main clauses (17a) and S/A/O arguments of subordinate clauses (17b). As an anaphoric pronoun, sentence (17a) can be rephrased as His nose, he’s bleeding.

(17) a. ti=nïje tóg, kyvénh mû
   3SG.MASC=nose 3.ANAF bleed AUX,PERF
   ‘His nose, it’s bleeding.’ (WIESEMANN, 2002, p. 66)

b. ëg=tŷ ku-r, kânhkrûnh kŷ tóg, nûgnûj nî
   1PL=OBL wear-NMLZ loosen SUB 3.ANAF loose be.sitting, AUX
   ‘When we loosen clothes, they become loose.’ (WIESEMANN, 2002, p. 43)

Additional evidence of an anaphoric function performed by tóg can be obtained by observing the ŝ marker, whose occurrence is restricted to nominal arguments, indicating that the “subject is topic” of the clause (WIESEMANN, 2002, p. 100, 160). Even though both markers can co-occur side by side, in a single intonation contour, as in (18b), from the syntactic point of view, they would belong to different constituents.

(18) a. inh=manûnŷ ŝ tóg kógnâ-j nî
   1SG=banana TOP 3.ANAF mash,NMLZ be.sitting, AUX
   ‘My banana mashed.’ (Lit. My banana, it’s mashed) (WIESEMANN, 2002, p. 48)

b. ūn kaga=ag ŝ tóg, krônhrôj nû ŭ tî
   INDEF sick=PL TOP 3.ANAF weak,PL be.lying AUX,IMPERF
   ‘The sick are weak.’ (Lit. Some sick, they’re weak) (WIESEMANN, 2002, p. 52)

For Wiesemann (1986, p. 378), it is the demonstrative ën ‘that there’ that corresponds to the anaphoric form.

In this example, I keep the original translation as it appears in the work of Wiesemann (2002, p. 66), from which the data was extracted, but which could also be translated as His nose is bleeding.
Another even more grammatical function of *tóg* is characterized by its use as a nominative case marker. Here, it is important to highlight the action of semantic-pragmatic processes (extension and desemantization) and morphosyntactic processes (decategorization), which were fundamental for the development of this new function and for the grammaticalization process as a whole. From the point of view of extension, the case marker *tóg* now covers both nominal (19) and pronominal (20) arguments. On the side of desemantization, there is the loss of semantic content, such as its deictic-spatial properties, generalizing and fulfilling a strictly grammatical function in marking arguments of verbal and non-verbal predicates. In terms of decategorization, one observes the loss of its syntactic freedom regarding its occurrence in different positions, as in its deictic uses in (15), becoming restricted to the subject of the sentence.

(19)  
\[ \text{a. } \text{fěněnh } tóg, \text{ gār kagra}=ja nǐ } \]  
\text{armadillo NOM corn eat=PERF be.sitting.AUX}  

‘The armadillo ate all the corn.’ (WISEMANN 2002, p. 18)

\[ \text{b. } ã=ku-r tóg, \text{ tor pē nǐ } \]  
\text{2SG=wear-NMLZ NOM dirty ADV be.sitting.AUX}  

‘Your clothes are very dirty.’ (WISEMANN, 2002, p. 87)

(20)  
\[ \text{a. } \text{kỹ sóg}^{12}, \text{ ti=ǐn tá kâkutē mũ } \]  
\text{DM 1SG,NOM 3SG,MASC=house LOC leave AUX.PRF}  

‘So I left his house.’ (WISEMANN, 2002, p. 42)

\[ \text{b. } ë=ku-r tóg, \text{ ë=nom } nǐ } \]  
\text{2SG=NOM be.sitting.AUX}  

‘You smoke a lot.’ (WISEMANN, 2002, p. 93)

Due to the different grammatical functions that ‘subject marks’ can play in Kaingáng, Kišêdjê (Suyá) and Tapayuna, it is perfectly possible and, to some extent expected, that they can coexist with those less grammatical functions in the languages in question. Examining the behavior of the morphemes *tóg* and *ra* (~ *tá*) in different types of constructions is illustrative of how they can be better understood by observing the diachronic sources from which they emerged and the different functions performed by them. In the following section, I argue that the emergence of new case markers, notably nominative case, constitutes a grammatical innovation in such languages and in the Jê family more broadly, characterizing such diachronic developments as a typical case of polygrammaticalization (CRAIG, 1991).

3. Discussion

Case marking patterns in languages of the Jê family are well known due to the multiple morphosyntactic configurations that clause constructions can present and the different factors that determine them in each language and even within the same language. Based on the data examined, at first glance, the sequence of grammatical developments could be conceived as a grammaticalization chain (HEINE, 1992), which we can schematize in (21).

---

12 The combination of the case marker *tóg* with the first person singular pronoun *inb* can result in the allomorphs *isóg* ~ *sóg* as a product of falling segments and phonetic changes (NASCIMENTO, 2013, p. 38), in which voiceless alveolar consonant /t/ is realized phonetically as voiceless alveolar fricative [ʃ] when preceded by nasal segments, as illustrated in the example below,

\[ \text{a. } kr=\text{věnhmŷ } isóg nĩ } \]  
\text{head=sad 1SG,NOM be.sitting.AUX}  

‘I’m worried.’ (WISEMANN, 2002, p. 52)
The proposition of such a grammaticalization chain raises the question of how to relate the case marker function to anaphoric pronouns, as in Kaingáng, since they share no common traits. Without denying the existence of different grammaticalization chains, the situation of the Jê languages examined here is more consistent with what Craig (1991, p. 486) calls polygrammaticalization, which is defined “as a multiplicity of grammaticalization chains that can originate in a particular lexical morpheme”. The justification for a polygrammaticalization scenario is based on the observation of the grammatical functions that can develop from demonstratives, such as personal and anaphoric pronouns (GIVÓN, 2001, p. 470), as shown in Figure 1, without them being necessarily intertwined in a linear order, as well as the performance of grammaticalization parameters in one domain, but absent in others.

**Figure 1** - Polygrammaticalization of demonstratives *ra* and *tóg* in Kaingáng, Kísêdjê (Suyá) and Tapayuna.

![Diagram](image)

Third person pronoun > Anaphoric pronoun

Deictic-spatial demonstratives *tóg* and *ra* (~ *ta*)

Case marker (nominative)

**Source**: The author.

Most Jê languages share a form *ta* that Davis (1966, p. 23) reconstructed into the protoform *ta, tam* as a third-person pronoun. In fact, some languages have grammaticalized a third-person pronoun (singular), such as Kísêdjê (Suyá) and Xikrin do Cateté, from the *ta* form, while in other languages, such as Krahô, this demonstrative is used for emphasis purposes, signaling that its use has been (and still is) dependent on the discursive-pragmatic context. Deictic properties, such as the location/spatial distance of the referent relative to the speaker and interlocutor and visibility, would have been conserved in its emphatic use.

The other way of grammaticalization would have been one in which the forms *tóg* and *ra* (~ *ta*) were used in apposition to a nominal referent, in a construction like *X, that (distant) makes/is Y*, which, in the sequence, is reinterpreted as *X NOM makes/is Y*, where the demonstrative becomes a nominative case marker. This scenario is similar to what McGregor (2008, p. 312) proposes for some Australian languages that developed case markers from indexical items such as pronominals and deictic-spatial demonstratives. In these languages, genitive and ergative markers derived from third-person pronouns or determiners (definite or indefinite/interrogative), used to highlight unexpected agents in the discursive context. In Kísêdjê (Suyá), some semantic-pragmatic properties of the diachronic source would have been maintained, such as spatial location and consequently extended to other more abstract conceptual domains, such as time, creating an effect of temporal distance, in contrast to constructions of topic marked by the morpheme *n(a)*, as pointed out by Santos (1997, p. 82) and shown in the following pair of examples in (22).

(22) a. liana  **ra**  kafɛ  ɲ-ihwere
n.pess  sm  coffee  r₁-make
‘Liana made coffee.’ (just did a while ago) (SANTOS, 1997, p. 82)

b. liana=  **n**  kafɛ  ɲ-ihwere
n.pess=top  coffee  r₁-make
‘Liana made coffee.’ (just done now) (SANTOS, 1997, p. 82)

With respect to the nominative case marked in Kísêdjê (Suyá), it is also important to highlight what Santos (1997, p. 159) considered to be a split conditioned by the semantic nature of the nominal phrase, in which
“the clauses with nominal S or A have a pattern of case marking different from the one which occurs in clauses with pronominal S and A” (SANTOS, 1997, p. 159). Explanations for splits conditioned by the semantic nature of the noun phrase in the functions S, A and O generally resort to the referential/animacy hierarchy that would favor the explicit marking of certain types of noun phrases as opposed to others (see, for instance, Cristofaro (2013, 2019) for a broader discussion on this issue). Thus, Dixon (1994, p. 84) states that if pronouns and nouns exhibit different case marking, the pronominal system will be accusative and the nominal system ergative, and never the opposite. Based on this typological prediction, the facts reported by the Kĩsêdjê (Suyá) and Tapayuna languages would constitute an ‘exception’ to what would be usually expected, since nominal arguments are marked by the nominative case, while pronominal arguments are sometimes marked in a nominative-accusative case system, sometimes in an ergative-absolutive system.

In defense of a source-oriented approach to typological universals rather than a result-oriented one, Cristofaro (2019, p. 27) argues that the first is based on the idea that typological patterns, for example, alignment patterns, should be explained with regard to the actual diachronic processes that gave rise to them rather than the synchronic properties of the pattern itself. In other words, particularities and/or restrictions of a given grammatical pattern may be related to distributional restrictions of the diachronic source from which it was developed. In Kĩsêdjê (Suyá) and Tapayuna, the restriction of case marking on nominal arguments (S/A) is explained in reference to the deictic-spatial origin of the form ra (~ to), which does not occur with personal pronouns. Therefore, the split in case marking would not necessarily be linked to the semantic nature of the noun phrase, as justified by Santos (1997, p. 159), but it is closely related to semantic-pragmatic and morphosyntactic properties of the lexical item that served as an appropriate means for the development of different grammatical functions.

Alongside these cases, the Kaingáng language seems to present a more advanced stage in the grammaticalization process of case markers, not only due to the number of nominative case markers (D’ANGELIS, 2004, p. 73; WIESEMANN, 2002, p. 159-160), but also because of the scope of some of them, such as is the case of the tóg marker. The grammaticalization path, in turn, differs in some details from that reported for the Kĩsêdjê (Suyá) and Tapayuna languages. In addition to functioning as a third-person pronoun, the form tóg would have extended its use to anaphoric purposes. Despite examples (16) and (17) clearly showing a deictic and pronominal function of tóg, respectively, in the consulted sources, there is no mention under which situations they would be distinguishable. One hypothesis is that tóg has been (and probably still is) used in situations dependent on the discursive-pragmatic context in which objects and participants involved in the speech event are indicated and pointed out through certain referential parameters, such as spatial location, distance and visibility.

Alternatively, as a nominative case marker, one notes here the performance and interaction of semantic-pragmatic (extension and desemantization) and morphosyntactic (decatégorization) parameters responsible for the grammaticalization of tóg in Kaingáng, which made its use broader and even more grammatical, what does not occur in the Kĩsêdjê (Suyá) and Tapayuna languages. Although Wiesemann (2002, p. 87, 160) states that the use of tóg indicates that the subject is an agent, another effect of extension would have been the marking of this syntactic role in different types of non-verbal predicates.

A noteworthy issue concerns the existence of several nominative case markers in Kaingáng, making it a peculiar case in the Jê family. A viable explanation for such a situation may be directly or indirectly related to more general grammatical changes, in which one of them led to the reduction of its pronominal system and the maintenance of only the absolutive pronominal series (WIESEMANN, 1986). The consequence of this change would have been the restructuring of the case marking system through the development of new case markers to distinguish the syntactic roles in the S, A and O functions of different types of predicates, in contrast to case marking patterns restricted to certain configurations clauses or structural environments, such as nominalized clauses in the stative aspect, some types of subordinate clauses and verb moods. These changes would have promoted a typological reconfiguration of the Kaingáng language in the domain of case marking, moving it to a type of language with case increase, according to Kulikov’s (2009, p. 454-455) classification.
REFERENCES


