ANOTHER LOOK AT THE PROBLEM OF THE COPULA IN THE SANTIAGO VARIETY OF CAPE VERDEAN CREOLE

Nicté-Ha Itzel Salas Barrena

0. Introduction

The distinctive characteristics of the copula in the Santiago variety (Sg.V) of Cape Verdean Creole (CVC) have been discussed since the first studies of Sg.V in the nineteenth century (cf. Brito 1887, Costa & Duarte 1886, Coelho 1886). Contemporary creolists (eg. Quint 2000, Baptista 2002, Lang 2002 and Pratas 2004) propose rather divergent explanations of this copula behavior. However, most of this recent work ignored the data attested in the earlier studies when it did not conform to current theory.

The aim of the present paper is to critically evaluate the existing literature on the copula in Sg.V from the nineteenth century works up to the most recent ones in order to present a more complete picture of this issue. Therefore, this paper examines in detail a number of recent grammars and texts, paying special attention the puzzling nature of the e morpheme, given its different analyses, i.e as a copula, a 3s. pronoun or an amalgam of both, as in:

E ten paxenxa ku mi pamodi e nh’amigu (Pratas 2004:106)
he have patience with me because ?? my+friend
‘He is patient with me because he is my friend.’

This aims to cast light on the origin and the properties of both the copula and the clitic third person subject pronoun. For this purpose, Baptista’s (1999) proposal of a continuum through which the e morpheme would have acquired nominal and verbal properties will be re-analysed in light of the copula genesis hypothesis of CVC’s counterpart in Guinea-Bissau Creole (Ichinose 1993), and contrasted with Pratas’s (2004) arguments about copula deletion and her analysis of CVC as a semi-pro-drop language. The specific properties of the morpheme e are then examined by testing current theoretical views on copulas genesis (Michaelis 2000), which make a case for third person singular pronouns as the origin for copulas, in light of the nineteenth-cen-
tury grammars mentioned above, which present data that contemporary researchers need to take into account.

Finally, this paper provides a renewed perspective on both the substrate and the superstrate influences that may have acted in shaping the synchronic features of the Sg.V copula in an effort to improve our understanding of its complex behavior.

1. Copula use in CVC

As stated in (Holm1999: 98):

“Copula has come to be used in creole studies for a number of words that either link subjects and predicates or serve to emphasize the following word. This extension of the Latin term (used only for the copula before NPs) began with Labov (1969), who used the term to cover any word (or zero) corresponding to a form of English *be*.”

From this standpoint, we will consider the instances of copula use in CVC given by Baptista (1999) before we turn to specific questions arising from this use.

1.1 Sta and e

As pointed out by Baptista (2002:102), in CVC both nominal and adjectival predicates may be introduced either by the copula *sta* (1, 3), when referring to temporary states, or by the morpheme *e* (2,4), when expressing a temporary state.

(1) **Gossi, Vieira sta director di skola, kada anu, e ta troka trabadju** (Baptista 2002:102)

'Now Vieira is director of school each year CL ASP change job'

(2) **Vieira e director di skola** (idem)

'Vieira is a school director'

(3) **Joao sta duenti** (Baptista 1999:27)

'Joao is sick’
(4) \textit{Joao e mao} (idem.)

\textit{Joao cop evil}

‘Joao is evil’

Other instances of \textit{be} in CVC are found in locative constructions, which are exclusively introduced by the copula \textit{sta} (5), and in focalized constructions, where the morpheme \textit{e} emphasizes the word it occurs next to (6).

(5) \textit{Valerie sta na skola} (idem.)

\textit{Valerie is at school}

‘Valerie is at school’

(6) \textit{E mi ki ta fika ku kes minizu} (Baptista 2002:103)

\textit{FOC me comp asp stay with the children}

‘It’s I who stay with the children’

1.2 Semi-copulas in CVC

Within the scope of copula studies, it seemed pertinent to note that unlike the reported analysis of copular forms in CVC, no attention has been given to the analysis of verbs such as \textit{kontinua} (7), \textit{bira} (8), \textit{fika} (9,10) and \textit{parse} (11), which seem to share some of the characteristics of copulas in terms of their compatibility with nouns, adjectives and verbs.

(7) \textit{Pedru kontinua obi si bós klaru} (Lang 2000:35)

‘Pedru still heard his voice clearly’

(8) \textit{Nhu Lobo bira más feiu kusa di mundu} (Lang 2000:162)

‘Sir Wolf became the ugliest thing in the world’

(9) \textit{E fika kunpridu dimás} (Lang 2000:43)

‘It was too loose’
Given their particular status, such verbs have sometimes been called semi-copulas or quasi-copulas (Pustet 2003)\textsuperscript{19}. However, it seems difficult to choose with a specific terminology at this stage since only an in-depth study of such forms could enlighten us on their nature, which falls out of the scope of the present work.

2. On the nature of the morpheme $e$

Unlike the unproblematic nature of the copula $sta$, the CVC morpheme $e$ has repeatedly been treated as a puzzling particle, as has its Guiné-Bissau Creole counterpart $i$. The conflicting interpretations arise from the homophony of the copular construction and the third person singular clitic pronominal in both languages (Kihm 2006). In order to achieve a better understanding of the problem, we will briefly consider the main factors that lead to confusion, namely the behavior of this morpheme regarding negation, pronominal selection and tense.

2.1 CVC $e$ and negation

Contrary to the pre-Neg position of verbs in CVC (12), $e$ obligatorily precedes the negator (13).

\textsuperscript{19} Pustet (2003) elaborates on a cross-linguistic study of copulas, which aims at elucidating universals in the relationship of copularization, i.e. use of copula and the categorization of lexicon from a semantic standpoint. This study also aims at solving cross-linguistic variation with respect to the compatibility of copulas with lexical items, as well as the intimate connection between patterns of copula usage and the division of the lexicon into the major parts of speech: nouns, verbs and adjectives. In this work, she proposes a generic distinction of a particular category of verbs which she defines as semi-copulas: “In this context, two additional types of elements which participate in the formation of predicates, and are therefore related to copulas to some extent, must be addressed: semi-copulas — or quasi-copulas, as they are sometimes called — and auxiliaries. However, unlike copulas, both semi-copulas and auxiliaries add meaning to the predicate phrases in which they are contained.” (Pustet 2003:5)
(12) *Joao* ka *bai* skola (Baptista 1999:35)
  Joao Neg went school
  ‘Joao did not go to school’

(13) *Joao* e ka *inteligenti* (idib.)
  Joao e Neg intelligent
  ‘Joao is not intelligent’

### 2.2 CVC e and pronominal selection

To bring to the fore the syntactic properties that these seemingly two different morphemes share, we now turn to the pronominal system of CVC, especially the description of the 3sg.subject pronoun.

#### 2.2.1 CVC pronominals for subject predicates

<table>
<thead>
<tr>
<th>Subject pronouns</th>
<th>Non-clitic strong pronouns (XP)</th>
<th>Weak pronouns (XP)/Tonic Clitics (Xº)</th>
<th>Clitics (atonic) (Xº)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.sg</td>
<td>ami</td>
<td>mi</td>
<td>N</td>
</tr>
<tr>
<td>2.sg (informal)</td>
<td>abo</td>
<td>bo</td>
<td>bu</td>
</tr>
<tr>
<td>2.sg (formal, masc)</td>
<td>anhu</td>
<td>nhu</td>
<td>nhu</td>
</tr>
<tr>
<td>2.sg (formal, fem)</td>
<td>anha</td>
<td>nha</td>
<td>nha</td>
</tr>
<tr>
<td>3.sg (fem and masc)</td>
<td><em>ael</em></td>
<td>el</td>
<td>e</td>
</tr>
<tr>
<td>1.pl</td>
<td>anos</td>
<td>nos</td>
<td>nu</td>
</tr>
<tr>
<td>2.pl</td>
<td>anhos</td>
<td>anhos</td>
<td>nhos</td>
</tr>
<tr>
<td>3.pl</td>
<td>aes</td>
<td>aes</td>
<td>es</td>
</tr>
</tbody>
</table>

Unlike previous works which identified only two types of pronominals in CVC (non-clitics and clitics), Pratas (2004) proposes the division of pronouns into three categories given the double nature of weak pronouns, which can also behave like clitics, namely in the context of a copular construction. Thus, Pratas emphasises that only this category of pronouns is compatible with the copula. That is why we consider pertinent her differentiation of non-clitics and weak pronominal

---

20 Adapted from Pratas (2004).
2.2.2 CVC non-clitic pronouns, tonic and atonic ciltics

Since verbs in CVC may either select non-clitics (15a), weak pronouns (15b) or atonic clitics (14), we can identify the different status of the morpheme e from its incompatibility with the clitic pronouns (15a, b)

(14) N bai (Baptista 1999:36)
    Cl. left
    'I left'

(15) a. Amí bai (idem.)
    Non-Cl. left
b. Mí bai
    W.pron left
    'I left'

(16) a. Mí e inteligenti21 (idib.)
    W.pron e intelligent
b. *N e inteligenti
    Cl. e intelligent
    'I am intelligent'

2.2.3 CVC e and tense: e different in nature from past era

As to its behavior regarding tense, we can see that e differs from its past counterpart era from its position vis-à-vis negation (17a, b). In this era behaves like a real verb, as well as in its compatibility with clitic pronouns (18a, b).

(17) a. Joao ka era inteligenti (idib.)
    Joao Neg was intelligent
b. *Joao era ka inteligenti
    Joao was Neg intelligent
    'Joao was not intelligent'

21 Pratas’s framework is used to re-analyse the data whenever it seemed pertinent to do so.
These characteristics portray an apparent lack of verbal properties of the morpheme e, which could lead us to consider it a pronoun. However, Baptista (1999) refers to theoretical reasons that challenge such assumptions, namely the problematic status of topicalized constructions (19, 20).

(19) El, el e timozu (Baptista 1999:41)
Non-Cl. Non-Cl. Cl. stubborn
’S/he is (really) stubborn’

(20) E el (Baptista 1999:42)
E Him/her
‘It’s him/her’

She asserts that in such cases, analysing e as a pronoun would cause major problems for the case theory and theta-criterion, since these constructions would then be characterized by three third person singular pronouns.

Because of this, the conflicting nature of e is still unresolved. Therefore, we will now consider another parameter which may elucidate the real nature of this morpheme.

3. The pro-drop parameter

Having introduced the problem of analysing the behavior of this particular morpheme, we will now turn to an analysis of its nature from the standpoint of the pro-drop parameter discussed in Baptista (1999, 2002) and Pratas (2004), which will provide a better understanding of the Sg.V subject pronominal paradigm.

22 Pratas considers this form impossible for Sg.V: ‘N e/era spertu (Pratas 2004:72), she estimates that both e and era (copula) are incompatible with atonous clitics, which is used to build on the previously described categorization of pronouns in Sg.V.
3.1 Arguments for the analysis of CVC as a pro-drop language

3.1.1 Null non-argumental pronouns

Baptista (1999, 2002) states that expletive constructions as they appear with weather (21a,b) predicates, existential predicates (22a, b) and raising predicates (23a, b) are obligatorily null in CVC.

(21) a. Sta fazi friu (Baptista 1999:28)
   is making cold

   b. *El/e sta fazi friu
   it is making cold
   ‘It’s cold’

(22) a. Fika dois dia antes di nos viaji (idem.)
   remain two days before of our trip
   ‘There remain two days before our trip’

   b. *El/e fika dos dia antis di nos viaji
   It remains two days before of our trip
   ‘Two days remain before our trip’

(23) a. Parse Ki Joao sta na ospital (Baptista 1999:29)
   seems that John is in hospital
   ‘It seems that John is in hospital’

   b. *El/e parse ki joao sta na ospital
   It seems that John is in hospital
   ‘It seems that John is in the hospital’

3.1.2 Null argumental pronouns

As to null argumental pronouns Baptista (1999) affirms that the 3sg. argumental pronoun el is possibly overt but preferably null with individual-level (permanent state) predicates (24a) and obligatorily overt with stage-level (temporary state) predicates (24b, c).
(24) a. (El) E azul  (Baptista 1999:30)
    (it) Is blue
b. El sta azul
    it Is blue
c. *Sta azul
    is/blue

‘It is blue’

However, in Baptista (2002), she develops argumentation for the pro-drop nature of CVC providing the following examples, which seem to contradict the ungrammaticality of the previous parallel construction (24c).

(25) a. Bu sta livri  (Baptista 2002:255)
    you are free
    ‘you are free’
b. Sta livri
    is/are free

Later, the author states that the 3sg.pronoun is preferably dropped with individual-level predicates and that the first and second person arguments with stage-level predicates may be dropped with the condition that their antecedent may be recoverable in the discourse (Baptista 2002: 263).

The potentiality of both non-argumental and argumental pronouns to be dropped constitutes the main argument for the author to affirm the pro-drop nature of CVC. This argument is further supported by “assuming that subject clitics occur in AGR and their non-clitic counterpart in SpecAgrP or higher up in the tree. In the event that AGR is morphologically empty, abstract features in AGR are recoverable from antecedent context” (idem.).

3.2 Arguments against the pro-drop nature of CVC

Pratas (2004) refutes the arguments for the pro-drop nature of CVC by pointing to the specificity of the co-occurrence of the 3sg. pronoun and what she considers to be the copula e. Thus, she undertakes a detailed analysis of this particular context by proposing hypotheses for an anaphoric process, copula deletion and finally a phonological resolution. As we will see, these arguments offer a new perspective on the treatment of the morpheme e.
3.2.1 Anaphoric hypothesis

The first hypothesis raised by Pratas is that in embedded clauses an anaphoric process could allow a null pronoun, since its referential value may be easily recovered, even without a morphological agreement operating in the verb. However, as we can see in sentences (26a, b, c), we can confirm that only the 3sg. subject allows a null subject.

(26) a. E tem paxenxa ku mi pamodi e nh’amigu (Pratas 2004:106)
   S/he have patience with me because ?? my+friend
   ‘S/he is patient with me because s/he is my friend’

   b. Bu ten paxenxa ku mi pamodi bu E nh’amigu
      you have patience with me because you ?? my+friend
      ‘You are patient with me because you are my friend’

   c. *Bu ten paxenxa ku mi pamodi bu nh’amigu

3.2.2 Negation contexts

The second hypothesis given by Pratas is that in negation contexts the null category might be the copula and not the pronoun: given the “specially functional nature of the copula”, it would be more likely to be dropped. This could explain a sentence like (27a).

(27) a. E ka gosta di mi pamodi e ka nh’amigu (idem.)
   S/he Neg like of me because ?? Neg my+friend
   ‘S/he does not like me because s/he is not my friend’

   b. Es ka gosta di mi pamodi es ka nh’amigu
      They Neg like of mi because they Neg my+friends
      ‘They do not like me because they are not my friends’

   c. *Es ka gosta di mi pamodi pro ka nh’amigu
      They Neg like of me because Neg my+friends

However, this explanation seems implausible since this phenomenon is only apparent in negative sentences. Indeed, the author refers to copula deletion as being the preferable option in Sg.V negation constructions as we can see in (28a, b, c, d).
Thus, copula deletion cannot be the explanation for sentence (26a), which can be confirmed by the ungrammaticality of (26c)

5.2.3 Phonological resolution

Given the failure of sustaining the pro-drop argument in contexts other than 3sg. pronouns in a copular constructions, Pratas maintains that CVC is not a pro-drop language and proposes that this specific context could be interpreted as a phonological coalescence of the 3sg. clitic subject pronoun and the copula.

4. Copula genesis

Up to this point we have analysed the synchronic features of this morpheme in an effort to elucidate the two possible analyses of it, either as a copula or as a pronoun, yet its real nature remains unclear so far. Therefore, we will now turn to a diachronic analysis of this problem. Here we will not treat it as an isolated phenomenon; instead we will look for parallels in other creole languages and in language universals.

4.1 In Creole and Non-Creole languages: evolution of subject pronouns

Pratas's resolution of a phonological amalgam is the first hint to look for a conjunct study of the copula and the pronoun. Following this line of thought, Michaelis (2000) makes explicit universal diachronic processes by which 3sg. subject pronouns can develop into a category of copulas, called pronominal copulas. She elaborates on previous cross-linguistic studies (Stassen 1997) and based on evidence from three creole languages, namely CVC, Guiné-Bissau Creole and Tok Pisin, as well as non-creole languages, to show that such a restructuring process is by no means creole-specific. Such an approach clarifies why pronominal copulas still show some of the morphosyntactic characteristics of their original category, i.e. they occupy the syntactic positions typical of pronouns, such as the pre-Neg position, and they select a specific kind of

Pratas uses an interrogation mark to attest to the fact that this sentence is not the most common form in negative sentences and might be considered nearly ungrammatical by some native speakers.
Another look at the problem of the copula in the Santiago variety of Cape Verdean pronominal. Later, with progressive grammaticalization pronominal copulas may lose their pronominal features altogether and acquire verbal ones, eventually reaching morphological and/or syntactic alignment with auxiliaries, verbal particles and verbs.

The main motivation for this phenomenon resides in the higher token frequency\(^{24}\) of the 3sg. contexts, which increasingly promotes the use of this pronoun as a linking element between two entities in equative constructions.

Finally, the previously discussed restructuring process presenting homonymies between copulas and pronouns, has been widely studied in non-creole languages such as Maltese (Borg and Azzopardi-Alexander 1997), Kenya Luo (Tucker 1993), Lango (Noonan 1992), Logbara (Crazzolara 1960).\(^{25}\) This aims to show that even though this process in by no means a universal fate, when it does occur, such a phenomenon cannot be seen as creole-specific. To look for those specificities one needs to include also non-creole languages, especially those that are neither lexifier nor substratate languages (Michaelis 2000: 181).

To further this argument, we could say that the speed of the grammaticalization process in creole languages is increased by their intrinsic characteristics, namely the intensive language contact during their genesis and development. In other words, if category conversion takes several centuries in (old) non-creole languages, it is likely that in intensive contact situations such as those that lead to the genesis and development of creoles, there is greater likelihood of category change (cf. Thomason & Kaufman 1988).

4.2 GBC copula genesis: Ichinose (1993), (Kihm 2006)

Guiné-Bissau Creole (GBC) has been widely reported in the literature as historically linked to CVC. It seems to present the same problem in analysing the morpheme \(i\), which seems to be both a 3sg. pronoun and a copula. As we will see in this section, the treatment that this problem has received provides different perspectives in approaching this question. Therefore, we will evaluate each contribution for its adequacy and applicability to our problem in Sg.V.

4.2.1 Ichinose (1993)

Ichinose claims that GBC \(i\), the counterpart of CVC \(e\), evolved from the 3sg. pronoun. This would explain the lack of verbal characteristics associated with \(i\), namely the fact that this particle cannot follow an aspect marker (29a, b) or precede the tense marker \(ba\) (30a, b).

\(^{24}\) Michaelis states that this is not so much caused by the unmarkedness of the third person but rather by its high token frequency that speakers prefer 3sg. markers to be zero-marked and that they overgeneralize this to the rest of the verbal paradigm.

\(^{25}\) cited in (Pustet 2003)
(29) a. *Ami \text{N} na i pursor \quad (\text{Ichinose 1993:24})
   \begin{tabular}{lll}
   1SG/\text{DISJ} & 1SG/\text{CLI} & TMA \text{teacher} \\
   
   \end{tabular}

b. Ami \text{N} na sedu pursor
   \begin{tabular}{lll}
   1SG/\text{DISJ} & 1SG/\text{CLI} & TMA \text{teacher} \\
   \end{tabular}

   ‘I will be a teacher’

(30)a. El i pursor \text{ba} \quad (\text{Ichinose 1993:25})
   \begin{tabular}{l}
   TMA \\
   \end{tabular}

   ‘S/he was a teacher’

b. *El i \text{ba} pursor

The morpheme $i$ cannot appear sentence-initially in imperative constructions, since (31) cannot be considered an imperative sentence.

(31) aluno (idem.)
   \begin{tabular}{l}
   student \\
   \end{tabular}

   ‘S/he is a student’

(32) Sedu garandi (idem.)
   \begin{tabular}{ll}
   \text{COP} & \text{big} \\
   \end{tabular}

   ‘Be an adult!’

This led Ichinose to conclude that $i$ is not a verb unlike sedu, which behaves like a real verb.

Following Ferguson (1971), Ichinose defends the idea that languages without a copula in equative constructions may use 3sg. pronouns, which serves to topicalize the subject of the clause. He elaborates on this to argue for the transformation of the 3sg. pronoun into a copula. Indeed, the author asserts that the synchronic copular characteristics of $i$ coexist with its topicalization value, recognizable by the small pause that speakers produce between the subject and the copula.

According to Ichinose the equative construction has evolved from:

\begin{align*}
Y \rightarrow XY & \rightarrow i \quad Y \rightarrow X \ Y
\end{align*}

26 Ferguson attests that in pidgins, equative expression are most frequently $XY$. Ichinose elaborates on this to postulate $XY$ as an earlier stage on GBC equative constructions.

27 Finally, the introduction of the previously referred copula sedu into equative constructions, seems to reinforce the predicate and not the subject.
The construction «i Y» for non explicit subject (31) in GBC is also found in West African languages like Hausa:

(33) Shi yaro ne (Schachter1986:55)
    3SG boy COP
    ‘He is a boy’

(34) Audu ne (idem)
    Audu COP
    ‘He is Audu’

Ichinose also refers to the use of a *predicator* to indicate a nominal predicate with non-explicit subjects in languages such as Bambara28.

Having advocated substrate influence in the process of grammaticalization does not prevent the author from arguing in favour of universals in the emergence of copular elements in languages originally lacking copular constructions, especially after analyzing such processes within non-creole languages such as Mandarin and Russian.

Again, his approach to the genesis and development of copular constructions rejoins the line of argumentation of universal processes in grammaticalization.

Before evaluating the adequacy of this approach for Sg.V, we will turn to another analysis of GBC copular expressions for a broader understanding of this question.

4.2.2 Kihm (2006)

In the most recent study of this phenomenon, Kihm (2006) presents a different proposal to attest for the evolution of the properties of *i* from its pidgin status to its synchronic state. To do this, Kihm identifies two stages of GBC grammar, an Ancient (A) and a Modern (M), and attests that the particle *i* is present in both the A (35) and M grammars (36). The difference between them is that in A, *i* is not a copula since it does not attract the enclitic past particle *ba*. Whereas in M, it behaves like a copula given its paradigmatic past form *yera*. Finally he infers that this morpheme also functioned as a 3sg. subject pronoun in A and M grammars.

28 Corne (1977) attests the same phenomenon in Seychelles Creole.
From this standpoint, the author elaborates on the passage from A to M grammars by refraining the pronominal characteristics associated with i in its earlier stage as an explanation for the alleged pronominal characteristics that this morpheme would have preserved in its modern stage. Kihm argues that i in copular sentences was a pronoun only during GBC pidgin stage and never after it creolized. In this view, i in the A grammar was (and still is to a certain point) a predicate marker, which he defines as a functional morpheme whose only function is to linearly separate the subject domain from the predicate domain.

5. Copula genesis in CVC

5.1 A diachronic perspective on the development of the copula in CVC

The diachronic analysis of the particle i in GBC suggests a point of departure in retracing the genesis of the CVC copula. Baptista (1999) contends that the CVC copula is derived from the 3sg. pronoun, which went through some type of category conversion following the pattern:

$\emptyset > e \text{ (PRON)} > e \text{ (COP)} > e \text{ (FOC)}$

This means that at an early stage CVC (like GBC) had copulaless constructions such as XY, then a pronoun was introduced in equative constructions before it assumed the role of the copula and finally that of a focalizer.

This evolution would account for the pronominal properties reminiscent of a copula with regard to negation, tense and pronominal selection discussed above.

Finally Baptista hypothesizes that the use of e as a pronoun in a copula position may have emerged under the influence of substrate languages, and that the outcome of the evolution of this morpheme in its double function both as a copula and a focalizer is due to Portuguese influence, which may have created a shift from a pronoun to a copula and a focalizer (Baptista 1999:44).
5.2 A reconstruction of copula use in the 19th century CVC (Brito 1887)

In this section, all the previous proposals for the genesis of copulas will be re-examined in light of data from 19th century grammars of CVC, especially of Sg.V. As stated in the introduction, such data has been ignored in the literature, possibly for its not fitting within the parameters of current theory on copula research. However, we find that these documents offer a valuable testimony of earlier forms of copula use, which is highly revealing in our search for a scenario of Sg.V copula genesis and development.

5.2.1 Phonetic differentiation of copula and pronoun

In the 19th century, Sg.V did differentiate the copula and pronoun as revealed by Brito’s use of two different graphemes: ê for the clitic subject pronoun and é for the copula.

(37) Ess é sînku (Brito 1887 [1967]: 355)
   they COP five
   ‘they are five’

(38) ê Tâ- sirbi (idem.)
   Cl. TMA serve
   She/he/it serves

The diacritics represent the phonetic quality of these vowels: the clitic pronoun ê is the closed vowel /e/ and the copula é is an open vowel /ɛ/. Thus, we can clearly observe in (37) and (38) the difference in their distribution.

This phonetic distinction reveals the 19th century CVC speaker’s consciousness of the two being distinct morphemes, which coexisted during this period. This finding poses serious problems for the hypothesis of a pronoun evolving into a copula in CVC. Indeed, if we accept that two distinct morphemes occurred in 19th century speech, this does not fit the previous theory of the pronominal origin for the copula in CVC, and thus leaves the problem of its genesis unresolved. As to its development, we could postulate that those two forms might have been competing, so that their phonetic proximity may have led to their fusion into what is today the 3sg. clitic pronoun with a closed vowel. However this explanation fails to provide an insightful account of why
the pronoun prevailed over the copula. Therefore, we will now turn to other areas of morpho-syntax which might cast light on this question.

5.2.2 é and pronominal selection

Brito’s (1887) account of the pronominal paradigm in Sg.V suggests that verbs may have only selected atonic clitic pronouns in subject position at this stage of the language, namely: í, bu, ê, nu, ess. As to the weak pronouns (mi, bó, el, nu, es) they seem to be restricted to object position.29

From this standpoint, we can assert that unlike its modern stage, CVC did not admit variation of the weak and the clitic pronouns in subject position, which might have influenced the need to differentiate the ê from é, since there was no other way to avoid ambiguity in sentences with a 3sg. subject.

As for the rest of Brito’s paradigm, strong pronouns such as ael are analysed as dative markers not as a subject pronouns. This means that ael was still analysed as a (preposition) and el (object pronoun), as we can see in (39a). These might have later been reanalysed by the speakers as (39b).

(39) a. A el ê ka acha nada (Brito 1887 [1967]: 356)
   To her/him s/he Neg Find anything
   b. Non-Cl Cl. Neg Find anything
   ‘She/he (really)did not find anything’

However, we note that with the copula é, pronominal selection is altered since the copula selected weak pronouns exclusively, which were not even presented as possible subject pronouns by the author.

(40) Mi é riku (Brito 1887 [1967]: 372)
   W.pron cop rich
   ‘I am rich’

(41) Bó ê kônfiadu (idem.)
   W.pron cop shameless
   ‘you are shameless’

---

29 As Dominika Swolkien pointed out, the object form which reached São Vicente was the clitic bu and not the weak pronoun bo which led to the closing of the vowel u and opening of the vowel preceding the pronoun as in aoió-b. It seems likely that Brito referred to cases preceded by a preposition such as pa, in an effort to differentiate the clitic from the weak pronoun as so to show the predominance of the clitic over the weak pronoun in subject position.
Even if the author does not give an example for the 3sg. pronoun we can postulate that the construction in (42a) was ungrammatical and was rendered by the weak pronoun *el as in (42b).

(42)a. *ê é riku
   Cl. cop rich
   ‘S/he is rich’

b. el é riku
   S/he cop rich
   ‘S/he is rich

5.2.3 *é and negation

Brito states that *é only occurs in affirmative and interrogative sentences, which seems to indicate that in negative constructions the copula was deleted. This point confirms the arguments of Pratas (2004:106) and Quint (2000:254) who assert that in negative sentences the copula *e is preferably omitted.

(43) A-mi M ka buru (Quint 2000: 254)
   Non-Cl Cl Neg fool
   ‘I am not a fool’

(44) A-mi ka mutu frientu (idem.)
   Non-Cl Neg very chilly
   ‘I am not very chilly person’

Following the reported pre-Neg behavior of pronominal copulas, the absence of a copular element in both 19th century and modern Sg.V grammars seems to disagree with the pronominal origin of the copula. Thus, a sentence like (45) would only convey the pronoun and not the copula.

(45) ê ka buru
    Cl. Neg fool
    ‘S/he is not a fool’
5.2.4 Pronominals and the pro-drop hypothesis

Both Brito (1889) and Costa & Duarte’s (1886) analyses demonstrate that the only instance of a subject pronoun being dropped occurs in the context of a 3sg. subject preceding a copula. In such cases, the pronoun may be dropped because it is subsumed in the copula itself. Such a phenomenon, which is exclusive to this context, may be explained by the phonetic proximity of the ê and ê, which in rapid speech might have blurred the distinction in height between these two vowels.

The analysis of copula use in the 19th century grammar, following the parameters which allowed contemporary researchers to advocate a pronominal origin of the CVC copula (i.e. the behavior of the copula in negation and pronominal selection) permits us to observe the inadequacies of such a hypothesis for Sg.V. However, we can also take into account the fact that this study was realized by a scholar who was bilingual in both CVC and Portuguese, and whose theoretical frame for describing Sg.V were 19th century Portuguese grammars, which certainly oriented him towards a particular type of analysis and description30.

6. Conclusions

The synchronic properties of the CVC morpheme e used in copular constructions were discussed in the first part of this paper, raising a number of difficulties in distinguishing this particle from the homophonous 3sg. clitic pronoun. We went on to study two different analyses of this problem. On the one hand, a disjunctive perspective was concerned with the synchronic state of the copula and the 3sg. pronoun. On the other hand, a conjunctive perspective was concerned with a diachronic study of the copula by providing patterns of category conversion motivated by universals.

As we saw in the last section, these proposals seem to conflict with data regarding copula use in an earlier stage of CVC grammar, which shows that unlike the universal perspective, a distinction between the copula and the pronoun was present at an earlier stage. This implies that the present state of both the copula and the pronoun, which merged into the same form /e/, is a recent phenomenon in CVC and not the original state of affairs, i.e. a pronominal origin of the CVC copula not longer seems a tenable analysis.

Both substrate and superstrate languages apparently had an influence in shaping the copula at some point. The difficulty arises when we try to situate these influences chronologically. From Brito’s data, the 19th-century copula’s form provides evidence

30 Brito’s work was supervised by the 19th-century Portuguese philologist Gonçalves Viana.
for superstrate influence since the use and form of the copula é was parallel to that of Portuguese. However, the synchronic form of the copula is likely to reflect the influence of substrate languages, which might have led to the merger of the pronoun and the copula in the pronoun’s form, given that many Niger-Congo languages lack equative copulas.

Ultimately, what this paper attempted to show is that postulating a linear evolution of the copula (as the universal hypothesis proposed by Michaelis does) is unlikely to elucidate the CVC copula’s synchronic form sufficiently. Although the process of grammaticalization is not completely invalid for the CVC case, there seems to have been several turns in the shaping of the copula where substrate, superstrate and universal influences inter-twine.

Finally, despite the complex origin of the copula, it is clear that in its synchronic state, the CVC copula cannot be examined by exactly the same scenario proposed for the evolution of GBC i, namely the one proposed by Kihm (2006) since the CVC morpheme e cannot be considered a predicate marker, since its copular properties seem expressed in its earlier stage.

References


