Variable nominative anaphora resolution in Cabo-Verdean Creole

Adrián Rodríguez Ricccelli
University at Buffalo (SUNY)
Cabo-Verdean Creole (CVC) – Kriolu – Língua berdiana

• Portuguese-based Creole language spoken since mid-XV century on Cabo Verde Archipelago

• Primary source languages:
  • Late Medieval/Early Classical Portuguese
  • Wolof (Atlantic)
  • Mandinka (Mande)

• Diglossia with European Portuguese since the late colonial era
  • Independence in 1975
    • EP remains the H variety

• Large diaspora (almost double home population) mostly in USA & Europe
Three nominative referential expressions in CVC
Subject clitic = SC

(1) \( \textbf{N} = \text{txoma} \quad \text{Jas} \)
   1sg.sbj=call J.
   ‘I am called Jas.’
Double-subject pronoun construction (tonic + clitic)

(2) \textbf{Anos Nu}=ta \quad \textbf{fla “ag”}, \textbf{Aes} \quad \textbf{Es}=ta

\begin{align*}
1\text{pl.} & \quad 1\text{pl.} & \quad \text{say water} & \quad 3\text{pl.} \\
\text{SBJ} & \quad \text{=TMA} & & \text{SBJ} \\
& & & \text{=TMA}
\end{align*}

\textit{fla “agu”}

say water

“We say “ag”, they say “agu”.”
Null/zero subject (contested)

(3)  

a. *Nho Lion_i ka fazi almusu, dja Ø_i ba fazi pastoria.*

mister lion  NEG make lunch,    TMA ø.3sg go  make field

‘Mr. Lion didn’t make lunch, he had gone to the fields.’ (Baptista 2002:268)

b. *(N/Bu/E)  

papia txeu

1.sg.nom.cl/2.sg.nom.cl/3.sg.nom.cl speak  a lot

‘I/you/he/she speak a lot.’  

(Costa & Pratas 2013:39)
No person-number ‘agreement/inflection’ suffixes

N ta kanta  I sing
Bu/nhu/nha ta kanta  You sing
E ta kanta  He sings
Nu ta kanta  We sing
Nhos ta kanta  You sing
Es ta kanta  They sing

Eu canto
Tu cantas
Ele/ela/você/o senhor/a senhora canta
Nós cantamos
Vós cantais
Eles/elas/vocês/os senhores/as senhoras cantam
Objectives

• Examine 5 of the most important/predictive variable constraints (observed across three parallel quantitative analyses) that conditioned the expression of (1) – (3a)
  • What environments promote zero subject expression (3a)?
  • What is the discursive function of double subject pronoun constructions (2)?
  • Can the results for the above two referential expression tell us anything about formal-functional status of subject clitics?
Overt subject pronouns in CVC
### The tripartite analysis

<table>
<thead>
<tr>
<th>Person/#</th>
<th>Strong</th>
<th>Weak</th>
<th>Clitic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>Ami</td>
<td>Mi</td>
<td>=N= / =M=</td>
</tr>
<tr>
<td>2SG</td>
<td>Abo</td>
<td>Bo</td>
<td>Bu= / =u</td>
</tr>
<tr>
<td>2SG (POLITE)</td>
<td>Anho (M) / Anha (F)</td>
<td>Nho (M) / Nha (F)</td>
<td>=Nhu= [M] =Nha= [F]</td>
</tr>
<tr>
<td>3SG</td>
<td>Ael</td>
<td>El</td>
<td>=E(l)=</td>
</tr>
<tr>
<td>1PL</td>
<td>Anos</td>
<td>Nos</td>
<td>=Nu=</td>
</tr>
<tr>
<td>2PL</td>
<td>Anhos</td>
<td>Nhos</td>
<td>=Nhos=</td>
</tr>
<tr>
<td>3PL</td>
<td>Aes</td>
<td>Es</td>
<td>=Es=</td>
</tr>
</tbody>
</table>

- Pratas (2004) (see also Baptista 2002) based on the “the Typology of Structural Deficiency” (Cardinaletti & Starke 1999)
The bipartite analysis

- Subject clitic attachment to verb root, TMA particle, negation, preceding complementizer
- In double subject pronoun constructions, the tonic pronoun can be adjacent or separated from the subject clitic by a range of intervening material

<table>
<thead>
<tr>
<th>Person/#</th>
<th>Tonic</th>
<th>Atonic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>(A)mi</td>
<td>=N= / =M=</td>
</tr>
<tr>
<td>2SG</td>
<td>(A)bo</td>
<td>Bu= / =u</td>
</tr>
<tr>
<td>2SG (POLITE)</td>
<td>(A)nho (M) / (A)nha (F)</td>
<td>=Nhu= [M]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>=Nha= [F]</td>
</tr>
<tr>
<td>3SG</td>
<td>(A)el</td>
<td>=E(I)=</td>
</tr>
<tr>
<td>1PL</td>
<td>(A)nos</td>
<td>=Nu=</td>
</tr>
<tr>
<td>2PL</td>
<td>(A)nhos</td>
<td>=Nhos=</td>
</tr>
<tr>
<td>3PL</td>
<td>(A)es</td>
<td>=Es=</td>
</tr>
</tbody>
</table>

The disputed formal/functional status of subject clitics

In CVC and cross-linguistically
The disputed status of subject clitics cross-linguistically

a. Inflectional affix hypothesis

b. Phonological clitic hypothesis

→ CVC (Baptista 1995, 2002)

Colloquial French (Auger 1994; Culbertson 2010; Palais 2013; *inter alia*)
Krèyol Ayisyen (DéGraff 1993)
Northern Italian Dialects, Rhaeto-Romance, Southern Gallo-Romance (Rizzi 1986a; Vanelli 1987; Brandi & Cordin 1989; Renzi 1992; Zribi-Hertz 1994; [for some of these varieties] Poletto 2000; Goria 2004; Poletto & Totora 2016; *inter alia*)
Several European languages (Cardinaletti & Starke 1999)
Wolof (Zribi-Hertz & Diagne 2002; Torrence 2013; Dione 2013)

→ CVC (Pratas 2004; Costa & Pratas 2008, 2013)

French (Kayne 1975; Rizzi 1986a; Brandi & Cordin 1989; Roberge 1990 De Cat 2005; *inter alia*)
Krèyol Ayisyen (Déprez 1994)
Degema, Igbo, Naro, Swahili, and Kalibari (Kari 2017)
Wolof (Dunigan 1994)
Walpiri (Jelinek 1984)
Mohawk (Baker 1996)
Subject-doubling structures (French; Culbertson 2010)

**Inflectional affix hypothesis**

```
       IP
       /\  
      /   
   DP  I'  
   /    
Jean I   VP
    /  
   il V
    
parle
```

→ [IP [DP \textit{El}ₐ] I’ \textit{e}ₐ [VP \textit{papia}]]

‘He/She said’

**Phonological clitic hypothesis**

```
       Top
       /
    DP  
   Jean
    /
   I    
    /
   il  I'  
    /
   V    
    
parle
```

→ [CP [DP \textit{El}ₐ][IP [DP \textit{e}ₐ][VP \textit{papia}]]]

‘He/She said’
Cardinaletti & Starke (1999) – “the Typology of Structural Deficiency”

<table>
<thead>
<tr>
<th>Strong</th>
<th>Weak</th>
<th>Clitic</th>
</tr>
</thead>
<tbody>
<tr>
<td>• +animate</td>
<td>• +/-animate</td>
<td>• +/-animate</td>
</tr>
<tr>
<td>• +specific</td>
<td>• +/- specific</td>
<td>• +/- specific</td>
</tr>
<tr>
<td>• +definite</td>
<td>• +/-definite</td>
<td>• +/-definite</td>
</tr>
<tr>
<td>• +tonic</td>
<td>• +tonic</td>
<td>• -tonic</td>
</tr>
<tr>
<td>• XP</td>
<td>• XP (but ‘deficient’, i.e. lacking a functional projection =CP)</td>
<td>• X°</td>
</tr>
<tr>
<td>• does not receive specification from antecedent (has “functional case features”)</td>
<td>• receive specification from antecedent</td>
<td>• receive specification from antecedent</td>
</tr>
</tbody>
</table>

→ pro is a deficient pronoun (Chomsky 1981; Rizzi 1986a)

I assume (at least in discourse) double subject pronoun constructions are equivalent to strong forms, I don’t distinguish when the involve a di- or monosyllabic tonic pronoun
Breaking the impasse

• Classic **diagnostic tests** or **taxonomies** (such as Kayne 1975; Zwicky & Pullum 1983; Cardinaletti & Starke 1999; among many others) have not provided a resolution to the protracted debate over the formal status and clausal position of subject clitics cross-linguistically (Haspelmath 2013, 2015).

  • As an alternative, I propose a quantitative analysis of the distribution of the referential expressions that ‘compete’ with subject clitics for the ‘subject space’ in CVC

    • This will ultimately lead me to propose that CVC subject clitics are “ambiguous person agreement markers” as in Siewierska (2004)
Zero subjects

With no local person marking, under referential accessibility accounts, in variationist sociolinguistics, ‘exceptions’ to referential accessibility
‘True’ zero subjects

(4) Si  M=bem  teni  fidju gósi  ∅  ta  fika mutu kansadu
COMP  1sg.SBJ=come  have kid  now  1sg.SBJ  TMA  COP very  tired
‘If I were to go have kids now I would become very tired.’  (P17)

(5) Ta  atxa [munti alunus]i  pa  li,  li  na  béra mar li
TMA  find  many students  PREP  here  here  PREP  seaside  here
[...]  ∅  ta  fuma  ganza
3pl.SBJ  TMA  smoke weed
‘One finds a lot of students around here, here by the seashore here [...] [they] smoke weed.’  (P3)

(6) [Kel sapu]i  ta  fudji  d-el,  ∅  ta  bai fika riba d-un tronku
DEM  toad  TMA  escape  PREP-3sg.NS  3sg.SBJ  TMA  go  stay  PREP-DET  log
‘That toad escapes from him, [he] goes on top of a log.’  (P6)
Other languages with attested ‘true’ zero subjects

Cognitively oriented discourse analysis

• Reference as ‘mental processing instructions’
  • Referential accessibility → mental representation of anaphoric or deictic referent in a ‘mental map’ of the discourse
    • Enhanced by greater referential coherence / salience / continuity / activation → when antecedent or last mention was:
      • subject of prior clause
      • anaphorically proximate
      • 1st & 2nd > 3rd person
      • +animate +definite +specific
      • a high continuity device
      • Topical, focused
      • etc...

Contributions from quantitative methods and variationist sociolinguistics

• Variable constraints related to **persistence (priming)** of morphological form across congruent syntactic contexts and **referential accessibility** may be active cross-linguistically, regardless of overall rates of zero/overt subject pronoun expression (across Romance: Duarte & Soares da Silva 2016; **Santomean Portuguese**: Bouchard 2018; **across varieties of Spanish**: Otheguy & Zentella 2012; Carvalho, Orozco, & Lapidus Shin eds. 2015, *inter alia*; **across varieties of English**: Wagner 2016; Torres Cacoullos & Travis 2014, 2019; Schröter 2019; **Mandarin**: Chen 1986, Pu 1997, Huang 2000; *inter alia*)
  - Languages differ in the envelope of variation, magnitude of effect & hierarchical ordering of constraints.
  - **Prosodic & Syntactic Linking** between Intonational Units (proxy for referential accessibility) promotes zero subjects in English (Torres Cacoullos & Travis 2019)
Possible (language-specific) exceptions to referential accessibility

• Mandarin zero anaphora often resume inanimate and indefinite antecedents (Chen 1986; Pu 1997)

• Semantic referential deficiency promote zero subjects in Brazilian Portuguese → “avoid referentially deficient pronoun constraint” (Kato 1999; Kato & Duarte 2003, 2005; Duarte & Soares da Silva 2016)
Methodology

Data collection, transcription, the envelope of variation, coding predictors, statistical analysis
Data collection

- Sociolinguistic interviews & picture description narrative (‘The Frog Story’) collected over 3 summers
- Word of mouth snowball sampling lead by local informants/collaborators
- 33 native speakers of CVC born and raised in CVC from Santiago (3 dialect zones) & Maio (1 dialect zone) spanning a range of socioeconomic backgrounds, urban v. rural residents, male v. female.
- Summers 2 & 3 – local community member informants aided in conducting interviews
- Zoom H2n Handy Recorder
Transcription

• 3 transcriptions assistants
  • 2 native speakers from São Salvador do Mundo
  • 1 near-native L2 learner of CVC, 15+ years resident in CV

• Prosodic Transcription (Du Bois et al. 1993; Chafe 1993, 1994; Torres Cacoullos & Travis 2019) → Intonational Unit (IUs; more on this when discussing coding for LINKING)
The envelope of variation

**Finiteness not a discrete category in CVC**, not marked morphologically → occurrence of all 3 referential expression must be observable in corpus for admission into the envelope

- Excluded contexts:
  - Copula *e* ‘to be’ only admits singleton tonic pronoun or zero, never subject clitics
  - Singleton tonic subject pronouns only 7 out of ~8,500 isolated IUs with all other verb types
  - The second ‘subject slot’ of serializing constructions
    
    \[ E = \text{fika} \quad \text{ta} \quad \text{djobi sapu} \]
    
    3sg.SBJ=stay TMA stare toad
    
    ‘He stays starring at the toad ~ He stays [and] stares at the toad.’
  - Presentatives/expletives with *ten* ‘have’, weather predicates, *parsi* ‘seem’ and other seem- and experiencer-type predicates
  - Impersonal ‘*ta*’ construction *ta kumi txeu pexi na CV* ‘one eats a lot of fish in CV/a lot of fish is eaten in CV’
  - Otherwise nonreferential, impersonal, generic referents
  - Modal/auxiliary verb + bare verb clause followed by ‘list’ of bare verbs
  - Bare non-stative (dynamic/actives verbs) that don’t get a default past reading when bare
  - Certain collocations and cognition verbs with superstrate inflection *axu* ‘I think’ and *sei* ‘I know’

→ Double subject pronoun construction never resumed an **inanimate** or **indefinite** antecedent ←
The envelope of variation

- 3rd person and 1st person plural → identifiable discourse antecedent introduced by lexical DP/NP (could be nonspecific):

  \[\text{Kes gafanhoto} \quad es=sta \quad mas \é \quad na \quad tenpu \quad di \quad azagua\]
  \[\text{DEM cricket} \quad 3\text{pl.SBJ}=\text{COP} \quad \text{more COP PREP} \quad \text{time} \quad \text{PREP} \quad \text{rainy season}\]

  
  [...] 

  \[\text{Kriansa} \_ i \quad \text{ta} \quad \text{panha} \quad \emptyset,\]
  \[\text{kid} \quad \text{TMA} \quad \text{catch} \quad 3\text{pl.OBJ}\]

  \[e_i=ta \quad \text{brinka} \quad \text{ku-el},\]
  \[3\text{sg.sbj}=\text{tma} \quad \text{play} \quad \text{prep-3sg.obj}\]

  \[\emptyset \_ i \quad \text{ta} \quad \text{sta} \quad \text{na} \quad \text{rua} \quad \text{ta} \quad \text{korri} \quad \text{di} \quad \text{sédu} \quad \text{a} \quad \text{tardi}!\]
  \[3\text{sg.SBJ} \quad \text{TMA COP PREP road} \quad \text{TMA run} \quad \text{PREP} \quad \text{early PREP late}\]

  ‘Those crickets are there more in the rainy season. [...] Kids catch [them], he plays with it, [he] is in the street running from early until late!’

- First-person singular → unambiguously referential/argumental contexts
Five predictors

• Some predictors that emerged as significant/predictive in at least 2 of 3 analyses (except SPECIFICITY)
  • SURFACE FORM OF THE ANTECEDENT
  • LINKING (between Intonational Units)
  • PERSON/NUMBER
  • ANTECEDENT ACCESSIBILITY PATTERN
  • SPECIFICITY (of the antecedent, part of semantic referential deficiency)
<table>
<thead>
<tr>
<th>Predictor variables</th>
<th>levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANTECEDENT ACCESSIBILITY</td>
<td>PATTERN A, PATTERN B, PATTERN C, PATTERN D, PATTERN E, PATTERN F, PATTERN G (modification of Duarte and Soares da Silva 2016; see next slide)</td>
</tr>
<tr>
<td>PATTERN</td>
<td></td>
</tr>
<tr>
<td>SURFACE FORM OF THE ANTECEDENT</td>
<td>LEXICAL DP, DP + INTERVENING MATERIAL + Ø, TONIC PRONOUN, CLITIC, Ø, DOUBLE SUBJECT CONSTRUCTION, DI OBLIQUE, POSSESSIVE PRONOUN, INFLECTION, INCORPORATED TA MEDIAL SUBJECT</td>
</tr>
<tr>
<td>LINKING</td>
<td>PROSODIC, SYNTACTIC, BOTH, NO LINK</td>
</tr>
<tr>
<td>PERSON/NUMBER</td>
<td>1SG, 1PL, 3SG, 3PL</td>
</tr>
<tr>
<td>SPECIFICITY*</td>
<td>SPECIFIC, NONSPECIFIC</td>
</tr>
</tbody>
</table>
LINKING between Intonational Units

• Prosodic linking
  
  * E ta pila kana,
  * Ø ta fazi si groringhu
  * ‘He presses the sugar cane, makes his moonshine’

• Syntactic linking
  
  * Rapzinhu buâ n-el.
  * i li e=sa ba ku redi p=e panha-l
  * ‘The boy jumped on him. And here he is going to with his net to catch him.’

• Both
  
  * Minis di Praia é so si,
  * ma=es ta arma bon karaka pâ
  * ‘Guys from Praia are all like that, but they put on a good party bro’

• No linking
  
  * N=sta pa pega autukarru númuru 3.
  * 40 minutu N=ta txiga na bo
  * ‘I’m catching bus number 3. 40 minutes I’ll be there.’

(Du Bois et al. 1993; Chafe 1993, 1994; Torres Cacoulllos & Travis 2019)
<table>
<thead>
<tr>
<th>Pattern</th>
<th>Description</th>
<th>Generalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>PATTERN A</td>
<td>The anaphor is in an embedded clause and is c-commanded by its antecedent</td>
<td>C-commanding antecedent, adjacent clauses, embedded target subject</td>
</tr>
<tr>
<td>PATTERN B</td>
<td>The antecedent is the subject of the immediately prior clause</td>
<td>Same subject, adjacent independent clauses/IUs</td>
</tr>
<tr>
<td>PATTERN C</td>
<td>The antecedent is in the prior clause, but has some role other than the subject</td>
<td>Switch-reference, adjacent clauses</td>
</tr>
<tr>
<td>PATTERN D</td>
<td>The antecedent is the subject of a clause that is separated from the clause containing the anaphor by at least one clause (like PATTERN B but in a nonadjacent clause)</td>
<td>Switch reference, antecedent is subject in nonadjacent clause</td>
</tr>
<tr>
<td>PATTERN E</td>
<td>The antecedent plays a role in the clause other than the subject and is separated from the clause containing the target anaphor by at least one intervening clause (like PATTERN C but in a nonadjacent clause)</td>
<td>Switch-reference, non-subject antecedent in nonadjacent clause</td>
</tr>
<tr>
<td>PATTERN F</td>
<td>The anaphor is in a main clause following a fronted adverbial subordinate, <em>si</em>-clause, or <em>pa</em>-clause, containing the antecedent</td>
<td>Fronted embedded clause</td>
</tr>
<tr>
<td>PATTERN G</td>
<td>The antecedent is in a prior, separate, and distinct clause-chain or ‘discourse chunk’ (the referent is an old discourse topic)</td>
<td>Distinct discourse ‘chunk’</td>
</tr>
</tbody>
</table>
Statistical procedure

- Of ~8,500 Intonational Units isolated, 3,651 were retained for analysis
- 3 parallel analyses
  - Analysis 1. fixed effects multinomial logistic regression → Ø / SUBJECT CLITIC / DOUBLE SUBJECT PRONOUN CONSTRUCTION
  - Analysis 2. mixed effects binomial logistic regression → Ø / OVERT
  - Analysis 3. mixed effects binomial logistic regression → DOUBLE SUBJECT PRONOUN CONSTRUCTION / ALL ELSE
- Descriptive visualizations with ggplot
- Stepwise comparison of nested regression models for AIC score
- Conditional inference trees and random forests
- Inferential visualization with sjplot (Lüdecke 2019)
- Tests for multicollinearity and singularity
Results
Distribution of nominative referential expressions

- Clitic: 81.9%
- DBL: 7.3%
- Null: 10.8%
The effects SURFACE FORM OF THE ANTECEDENT on SUBJECT EXPRESSION – Analysis 1
Response variable application value = SUBJECT CLITIC; Predictor application value = CLITIC
Distribution SUBJECT EXPRESSION by PERSON/NUMBER
The effects LINKING and PERSON/NUMBER on Ø SUBJECTS – Analysis 2
Response variable application value = OVERT; Predictor application values = NO LINK; 1SG

**Log-Odds**

- **Linking: both**
  - 0.77 ***

- **Linking: prosodic**
  - 1.14 ***

- **Linking: syntax**
  - 0.13

- **Pers.numb: 1p**
  - 0.28

- **Pers.numb: 3p**
  - 1.11 ***

- **Pers.numb: 3s**
  - 0.67 **
Summary of results for ZERO SUBJECTS

• Inter-clausal LINKING and short ANAPHORIC DISTANCES ‘unlock’ the:

  • Ø-to-Ø persistence (priming) effect

• The favoring effect from NONSPECIFIC (referentially deficient) LEXICAL DP antecedents on Ø SUBJECTS
  • Presumably from inanimate and indefinite antecedents too (needs further testing!)
  → The “avoid referentially deficient pronoun” constraint, like in Brazilian Portuguese (Kato 1999, Kato & Duarte 2003, 2005; Duarte & Soares da Silva 2016)
    • This may be a consequence of having person marking forms that can resume semantically referentially deficient antecedents (and the extent to which they have grammaticalized into subject clitics)
The effects of ANTECEDENT ACCESSIBILITY PATTERN on X2SBJ – Analysis 3
Response variable application value = ALL ELSE; Predictor application value = PATTERN B (same subject prior clause)
Summary of results for DOUBLE SUBJECT PRONOUN CONSTRUCTIONS

• **DOUBLE SUBJECT PRONOUN CONSTRUCTIONS** show a persistence (priming effect) from other DOUBLE SUBJECTS and singleton TONIC PRONOUNS
  • But only across non-adjacent clauses

• **DOUBLE SUBJECT PRONOUN CONSTRUCTIONS** are switch-reference devices and used to establish contrastive focus
What can this tell us about CVC subject clitics?

• A cross-linguistically common grammaticalization cline:

  independent person marker > weak form > clitic > agglutinative affix > fusional form > Ø

  (Siewierska 2004)

• The functional properties of CVC double subject pronoun constructions reflect those of tonic pronoun + bound inflectional suffixes in Canonical Null Subject languages (switch reference, contrastive focus)

  **Ami N=gosta di fazi kaldu** = **Eu gosto de fazer caldo** ‘I like to make soup.’
  **Ma nos nu=ka ta studa** = **Mas nós não estudamos** ‘But we don’t study’
  **El e=sta riba di pó** = **Ele está acima da rama** ‘He is on top of the stick.’
  **Aes es=ta vivi sabi** = **Eles vivem bem** ‘They live well.’

• Zero subjects are promoted in the 3rd person
  • Respond to properties of the antecedent
CVC subject clitics are ‘ambiguous person agreement markers’

• The grammaticalization cline also implies changes in functional properties
  pronominal agreement marker > ambiguous agreement marker > syntactic agreement marker

• ‘Grammatical agreement’ (with a local conominal)
  \[Nha\] pai \(e\) bem dipôs \(k=N\)
  1sg.POSS father 3sg.SBJ=come after COMP=1sg.SBJ
  tinha onzi anu
  have.IMPERF eleven year
  ‘My dad [he] came after I was eleven years old.’ (P34)

• ‘Anaphoric agreement’ (with a nonlocal disocurse antecedent)
  \(É\) [Nha irmôan] \(ki\) \(é\) político
  COP 1sg.POSS brother COMP COP politician
  [\ldots] \(E\) sai di político
  \ldots 1sg.SBJ=left PREP politics
  ‘[It] is my brother who is a politician [...] he left politics.’ (P38)

For Haspelmath (2013, 2015), these are ~ cross-indexes
A ‘split-paradigm’ system

**Person/#** | **Tonic** | **Atonic**
---|---|---
1sg | (A)mi | =N= / =M= / Ø (low rates)
2sg | (A)bo | Bu= / =u / ?
2sg (polite) | (A)nho [M] | =Nhu= [M] / ?
3sg | (A)el | =E(l)= / Ø
1pl | (A)nos | =Nu= / Ø (low rates)
2pl | (A)nhos | =Nhos= / ?
3pl | (A)es | =Es= / Ø

• The results here provide quantitative empirical support to Wratil’s (2009) proposal that CVC **subject clitics are ‘in-between’ forms** developing **along a grammaticalization cline** from independent pronoun to bound inflectional affix
  - This is the reason that the phonological clitic vs. inflectional affix debate has stalled at an impasse for so long → it’s not an either/or binary, **clitics are by their nature diachronically ‘transitory’ forms**

• **Grammaticalization** can obtain **to different degrees across the pronominal paradigm**, resulting in **different functional roles for clitics** (e.g. they may be ‘more pronominal’ or ‘more affixal’ in different person/number instantiations) (see Vainikka & Levy 1999; Meyerhoff 2000; Wratil 2009, 2011; inter alia)
  - **CVC 1sg is the most affixal**
    - Largest portion of double subject pronoun constructions
    - Low rate of zero subjects → becoming invariable/obligatory
    - Most phonologically reduced
  - **3sg retains more pronominal properties**
    - Responds to properties of discourse antecedent
    - Only partial phonological reduction (loss of lateral coda)
  - **3pl is the most pronominal**
    - Largest portion zero-subjects
    - Homonymy-but-for-stress between monosyllabic tonic and atonic forms
null subject languages

minimization of the morph. and phon. representation

optimization of the perception

partial null subject languages

(Wratil 2011)
Dja kaba!

A special thank you to:

Sandro Sessarego, Jacqueline Toribio, Chiyo Nishida, Nicholas Quint, Lars Hinrichs, Edmir Semedo, Fáron Peckham, Clever Cunha, Zita Vieira, Jacob Cullen, António ‘Jupa’ Varela, all the participants, and tudo kauberdianu ku kauberdiana, pa kel morabeza ki nhos mostram
Brigadu nhos tudu pa sisti nha palestra!
References


• Bresnan, Joan & Mchombo, Sam A. 1987. Topic, Pronoun, and Agreement in Chichewa. Language 63(4): 741


• Corbett, Greville. 2006. Agreement. Cambridge University Press.


De Cat, Cécile. 2005. French subject clitics are not agreement markers. Lingua 115(9): 1195-1219.


• Kato, Mary Aizawa & Duarte, Maria Eugênia L. 2005. (Micro)parametric variation between European Portuguese (EP) and Brazilian Portuguese (BP): Similarities and Differences Related to Ongoing Changes in Latin American Spanish. In . Monterrey, MX.


50


Extra examples
‘Their kids that are born abroad, They don’t teach [them] to speak Kriolu. [They] don’t know how to speak Kriolu really.’
And the kids play with it, [they] make it so that [it] is like a cow, [they] play with it, [they] chase them, [they] catch a lot of them, [they] fill the bottle, [they] play with it again.

Now here he went away, [he] tired of catching frogs and [he] went on his way.

And the kids play with it, [they] make it so that [it] is like a cow, [they] play with it, [they] chase them, [they] catch a lot of them, [they] fill the bottle, [they] play with it again.

Now here he went away, [he] tired of catching frogs and [he] went on his way.
Most Cabo-Verdeans on the week they wash clothes because during the week they work.

People close their eyes to those thing there, they pretend that they aren’t seeing it.

In the oven now we put more cheese on it, it ends up tasting even better.
‘Once I used to play mancala. But now it’s like it’s a practice that is almost disappearing, like then there is also a lack of time, but I like mancala a lot.’

‘[It] could be very dangerous, a child alone, just him/her and his/her dog, [the] dog might not have the maturity, like, that intelligence, if [he/she] were to drown here, for [it] to help him/her.’

‘When I read it, I find that it is a typical story.’
‘Here we have an image, [it] could be a hunter, he has a dog, and here it seems that [it] is, that [it] is fish that they got (them) here we have a pond, that has a frog, it is in a forest. ... -TURNS PAGE- ... And here they run up behind the frog.’
‘First I would try to improve that issue of violence here, because it spoils the image of Cabo Verde.’

‘Kriolu does not really help them that much in speech, they/one have/has to, uhh... speak English but... [it] does garner help some, yeah.'
Then he fell ... He is in the sea.

I am twins with a female, yeah, [her] name is Nilda

He had his first child who died in the army, [he/she] died at 18 years old.

The boy is sitting, [he] went to take his bath.
'He remains starring at the toad ~ He remains [and] stares at the toad.'

'Then it becomes elongated [and] [it] dries'

'He remains [and] walks [and] stops.'

'They are always promising me.'

'We were good because we didn’t know what it was like to have light.' (P34)

'Orthography is when the professor speaks [and] dictates.'
Prosodic linking (only)
a.) \( E_i = ta \quad pila \quad kana, \)
3sg.SBJ=TMA press cane
\( \emptyset_i \quad ta \quad fazi \quad si \quad groginhu \)
3sg.SBJ TMA make 3sg.POSS sugarcane liquor
‘He presses the sugar cane, [he] makes his sugarcane liquor.’
b.) \( N = ta \quad ferbi \quad nha \quad kafé \quad pretu, \)
1sg.SBJ=TMA boil 1sg.POSS coffee black
\( N = ta \quad bebi \quad \emptyset \)
1sg.SBJ=TMA drink 3sg.OBJ
‘I boil my black coffee, I drink [it].’
c.) \( Es = es \quad stala, \)
3pl.SBJ=3pl.SBJ COP there
\( es = es \quad ten \quad ki \quad ntegra \quad na \quad sosiedadi! \)
3pl.SBJ=3pl.SBJ have.to integrate PREP society
‘They are there, they have to integrate into the society.’
Syntactic linking (only)

a.) \( iii\) \( dj=e_i\) \( ta\) sigi \( kel\) pasu.

CONJ TMA=3sg.SBJ TMA follow DEM footprint

\( i\) \( \emptyset\_i\) \( odja\) rä sta d-un ladu

CONJ 3sg.SBJ see frog COP on-DET side

‘And now he follows those footprints. And [he] sees that frog is to one side.’

b.) \( Rapzinha_i\) \( buâ\) n-el.

boy jump PREP-3sg.OBL

\( i\) \( li\) \( e_i=sa\) ba ku redi \( p=e_i\) panha=l

CONJ here 3sg.SBJ=TMA go PREP net COMP=3sg.SBJ catch=3sg.OBJ

‘The boy jumped on him. And here he is going with the net to catch it.’

\( Ali\) \( é\) abitasôn \( di\) sapu_i.

here COP inhabitance PREP toad_i

\( i\) \( el_i\) \( dj=e_i\) fika tristi, ne?

CONJ 3sg.SBJ_i TMA=3sg.SBJ_i become sad DM

‘Here is the frog’s place of inhabitance. And here he becomes sad, right?’
Both

a.) [**Kel sapu**]i

[that frog]**

TMA escape more DET time

\[i \quad \emptyset \quad ba fika \quad sukundidu \quad riba \quad d-un \quad pédra\]

CONJ 3sg.SBJ go stay hidden on top of-DET rock

‘That frog escapes one more time, and [he] goes to hide on top of a rock.’

b.) [**Minis di Praia**]i

[gguys from Praia]

COP only like that

\[m=es \quad ta \quad arma \quad bon \quad karaka \quad pa\]

COMP=3pl.SBJ TMA mount good beachside party DM

‘Guys from Praia are just like that, but they put on a good beachside party man.’

c.) [**Es**]

3pl.SBJ NEG TMA give a lot importance PREP violence COMP COP

na sosiedadi,

PREP society

\[pamodi \quad aes=es\]

3pl.SBJ=3pl.SBJ TMA go PREP 3pl.POSS security

\[na \quad karru\]

PREP car

‘They are not putting a lot of importance on the violence in society, because they go around safe in their cars.’
No linking

a.) *Kel* *otu* anu *M*=*ba* *trabadjia* MDR ...

DEM other year 1sg.SBJ=go work MDR

Ø *ba lóra* pédra, *koba* txón.

1sg.SBJ go mine rock dig earth COMP

‘That next year I went to work at MDR... [I] went to mine rocks, dig up earth’

b.) *Nì*=*sta* *pa* *pega* autukarru *númuru* tres.

1sg.SBJ=intend.to catch bus number three

*Kuaranta* minitu *Nì*=*ta* *txiga* na *bo*

forty minute 1sg.SBJ=TMA arrive PREP 2sg.OBL

‘I am planning to catch bus number three. Forty minutes, I’ll arrive where you are.’

c.) *Ulisis*, *ki* *elì* é *prizidentì* di kambia *li* di Praia.

U. COMP 3sg.SBJ COP president.of.chamber here PREP Praia

*Elì=* *eì* *kumesa* fazi un *bon* *trabadju*,

3sg.sbj=3sg.sbj begin do DET good job

kalseta rua, fazi prasas, fitines park, asi.

cobble street make plaza fitness park like that

‘Ulises, who is the municipal executive, here from Praia. He started to do a good job, cobbling the streets, making plazas, fitness parks, things like that.’
Si M=bem teni fidju gósí Ø ta fika mutu kansadu
COMP 1sg.SBJ=come have kid now 1sg.SBJ TMA COP very tired

‘If I were to go have kids now I would become very tired.’ (P17)

E po=m₁ la òrta, Ø₃ ta guberna animal
3sg.SBJ put=1sg.OBJ there garden 1sg.SBJ TMA govern animal

‘He put me there in the garden, [I] would tend to the animals.’ (P32, Mi=N)

1sg.SBJ=1sg.SBJ TMA think PREP language Kriola 1sg.SBJ TMA think
ma Ø₃ debi ser lingua ufisial
COMP 3sg.SBJ should COP language official

‘I think of the Creole language, [I] think that [it] should be an official language.’ (P8)

Ax-u k=N tinha pa ai’ unzzi anu,
think-1sg COMP=1sg.SBJ have.IMPF about DET eleven year
Ø ba-ba Fransa ku familia
1sg.SBJ go-TMA France PREP family

‘I think that I was about eleven years old, [I] went to France with family.’ (P1)

[Praia, Kauberdi], pudi ser sabi n-un pontu ma Ø₃ podi
Praia, Cabo Verde, can be nice PREP-DET point CNJ 3sg.SBJ can
ser ka sabi na kel otu pontu
be NEG nice PREP DEM other point

‘Praia, Cape Verde, can be nice on the one hand, but [it] can be not nice on the other hand.’ (p10, [Kel sapu], ta fudji d-el, Ø₃ to bai fika riba d-un tronku
DEM toad TMA escape PREP-3sg.NS 3sg.SBJ TMA go stay PREP-DET log

‘That toad escapes from him, [he] goes on top of a log.’ (P6, Funaná, é un ritimu mas faksi asin, Ø₃ to bai mas rápidu
Funaná COP DET rhythm more fast like that, 3sg.SBJ TMA go more fast

‘Funaná has a quicker rhythm, [it] goes faster.’ (P?, ??