Different Kinds of Second-position Clitics in Caquinte

Nicholas Rolle & Zachary O’Hagan
University of California, Berkeley

∼

WSCLA 23∗
University of Ottawa
April 14, 2018

1 Introduction

• In this presentation we explore the distribution of different sorts of second-position clitics at the phrase and clausal level in Caquinte (Arawak, Peru)

• Second-position clitics (2P) are also known as Wackernagel clitics

[...]. In current language descriptions, a Wackernagel clitic is a clitic that can be positioned either after the first word or after the first phrase [...]. the property of promiscuous attachment or low selectivity (Spencer and Luís 2012:41)

\[\left\| \begin{array}{c} \text{[X]} = \text{CL} \\ 1 \end{array} \right\| \text{[Y]} \right\|_\text{DOMAIN} 2 3 \]

• Questions – division of labor between syntax, morphology, and phonology

  – Constituency

    * What kind of constituent is [X] to which the clitic attaches?

      . phonological: foot \sim prosodic word \sim phonological phrase

      . syntactic: grammatical word \sim \textit{X}\circ \sim phrase XP

    * What kind of constituent is the resulting \([X]=\text{CL}\)?

    * What kind of constituency holds between [X] and [Y]?

1 See also: Tobler (1875); Mussafia (1898); Wackernagel (1892); Schachter and Otanes (1972); Schachter (1973); Kaisse (1982); Klavans (1985); Black (1992); Anderson (1993, 2005, 2008, 2011); Halpern (1995); Halpern and Zwicky (1996); Bošković (2001); Noyer (2001); Chung (2003); Fischer (2003); Pancheva (2005); Legate (2008).
– Derivation
  * Where are [X], [Y], and =CL merged in the syntax?
  * By which principles are they positioned resulting in surface 2P?
– Domain
  * What determines the domain of second-positionhood?
  * For example, Anderson (2008) describes different sorts of clitics, ones that must be non-initial in CP versus in IP (see Appendix)

• Analysis
  – Claim 1 – Caquinte clitics are prosodically deficient
  – Claim 2 – [X]=CL structure is formed post-syntactically
  – Claim 3 – clausal clitics are syntactically merged at different points in the clause

2 Caquinte Language
• Caquinte (ISO 639-3:cot) is a Kampa Arawak language spoken in southeastern Peru by fewer than 500 people (Castillo Ramírez 2017; Swift 1988)
  – polysynthetic: large number of morphemes per word
  – strongly headmarking, frequent pro-drop
  – VSO with subject agreement, reality status, and voice as obligatory categories
• The verbal template includes three prefixal and upwards of twenty suffixal slots, as well as at least seven clausal second-position clitic slots

(1) a. [Irira iranianishite] ikenkejanake, ikanti: “Imaikampani nojokeneri kachojari kameetsanjite namenabakeri impeanakempageti.”
  iri- ra iri- anianishi -te i- kenkej -an -k -i i- 3M- D:MED 3M.P- brother-in-law -POSS 3M.S- think -ABL -PFV -R:ACT 3M.S-
  kan -i imaika =Npani no- ojok -e -nV -ri kachojari kameetsa
  say -R:ACT now -CT 1S- give -IRR -TH -3M.O manioc.bear PURP
  =niji -te no- amen -ab -k -e -ri i- N- peg -an -k
  =PURP =CE 1S- see -DIR -PFV -IRR -3M.O 3M.S- IRR- transform -ABL -PFV
  -e -Npa =geti
  -IRR -MID =when

  [His brother-in-law] thought, and said: “Now I’m going to give him manioc beer so I can watch him when he transforms.”

2Epenthetic segments /t/ and /a/, which repair vowel and consonant hiatus, respectively, are not represented in the segmentation. Abbreviations: A = applicative; ABL = ablative; ACT = active; ALL = allative; ALT = alternative; AM = associated motion; AUG = augmentative; CE = counter-expectational; CF = counterfactual; CL = classifier; CNGR = congruent; COP = copula; D = demonstrative; DIR = directional; DISTR = distributive; DUR = durative; EPST = epistemic; F = feminine; FOC = focus; FRUST = frustrative; IDEO = ideophone; INCL = inclusive; INCNGR = incongruent; INSTR = instrumental; IRR = irrealis; LOC = locative; M = masculine; MAL = malefactive; ME = male ego; MED = medial; MID = middle; NEG = negation; O = object; P = possessor; PFV = perfective; PERSP = perspectival; PL = plural; POSS = possessive; PRO = pro-form; R = realis; REG = regressive; REL = relativizer; S = subject; TOP = topic.
b. Yoanaji itsobironakiteki.

\[
\begin{array}{ll}
  i- & \text{go} \\
  -\text{an} & \text{ABL} \\
  -\text{aj} & \text{-R:ACT} \\
  -i & \text{3M.P-} \\
  \text{tsobironaki} & \text{house} \\
  \text{-te} & \text{-POSS} \\
  =\text{ki} & \text{LOC}
\end{array}
\]

He went back to his house.

c. \text{Iri iri\text{o\text{ro}ki} tee irogeji, arigenti ichokotaji amperitaki amashaitake.}

\[
\begin{array}{ll}
  \text{iri-} & \text{ra} \\
  \text{irioiroki} & \text{collared.peccary} \\
  \text{tee} & \text{neg:r} \\
  \text{iri-} & \text{3M.s.irr} \\
  \text{og-} & \text{neg:r} \\
  \text{e-} & \text{irr} \\
  \text{ji} & \text{neg:r} \\
  \text{arige} & \text{foc:pred} \\
  \text{ani} & \text{3M.s} \\
  \text{perita} & \text{river.rocks} \\
  =\text{ki} & \text{LOC} \\
  \text{amashai} & \text{sing} \\
  -\text{i} & -\text{Ø} \\
  \text{- REG -R:ACT} \\
  \text{river.rocks} & \text{=LOC} \\
  \text{sing} & -\text{PFV -R:ACT} \\
  -3S
\end{array}
\]

\text{The collared peccary} didn’t go, he sat on the river rocks singing.

- The preverbal field includes positions for topics and foci

<table>
<thead>
<tr>
<th>TOPIC FOCUS</th>
<th>NEGATION</th>
<th>FOCUS</th>
<th>VERB</th>
<th>SUBJECT</th>
<th>OBJECT</th>
</tr>
</thead>
</table>

Figure 1: Schematic Syntactic Structure

3 Second-position Clitics

3.1 Clausal Second-position Clitics

- Clausal second-position clitics express meanings often associated with CP\(^3\)
  - information structure, clause linking, modality, evidentiality, adverbials, emotion

<table>
<thead>
<tr>
<th>Form</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>(=ga)</td>
<td>contrastive topic</td>
</tr>
<tr>
<td>(=N\text{pani})</td>
<td>contrastive topic</td>
</tr>
<tr>
<td>(=N\text{pa})</td>
<td>incongruence</td>
</tr>
<tr>
<td>(=\text{geti})</td>
<td>‘when’, ‘if’, ‘where’, ‘until’</td>
</tr>
<tr>
<td>(=\text{tari})</td>
<td>congruence (“because”)</td>
</tr>
<tr>
<td>(=\text{nika})</td>
<td>congruence under negation</td>
</tr>
<tr>
<td>(=\text{ka})</td>
<td>epistemic</td>
</tr>
<tr>
<td>(=\text{kea})</td>
<td>nonwitnessed evidential</td>
</tr>
<tr>
<td>(=\text{sa})</td>
<td>inferential</td>
</tr>
<tr>
<td>(=\text{me})</td>
<td>counterfactual, frustrative</td>
</tr>
<tr>
<td>(=\text{te})</td>
<td>counterexpectational</td>
</tr>
<tr>
<td>(=\text{sakanika})</td>
<td>mirative</td>
</tr>
<tr>
<td>(=\text{ja})</td>
<td>modal strengthener</td>
</tr>
<tr>
<td>(=\text{riji})</td>
<td>deontic</td>
</tr>
<tr>
<td>(=\text{ta})</td>
<td>‘soon’</td>
</tr>
<tr>
<td>(=\text{sano})</td>
<td>‘truly’</td>
</tr>
<tr>
<td>(=\text{keti})</td>
<td>‘ahead’</td>
</tr>
<tr>
<td>(=\text{ketiro})</td>
<td>‘ahead’</td>
</tr>
<tr>
<td>(=\text{shia})</td>
<td>anticipated apprehensiveness</td>
</tr>
<tr>
<td>(=\text{shiatsu})</td>
<td>current apprehensiveness</td>
</tr>
<tr>
<td>(=\text{ka})</td>
<td>relativizer</td>
</tr>
<tr>
<td>(=\text{shine})</td>
<td>irritation</td>
</tr>
<tr>
<td>(=\text{gitatsi})</td>
<td>confirmation</td>
</tr>
<tr>
<td>(=\text{satine})</td>
<td>request for confirmation</td>
</tr>
</tbody>
</table>

- It is ungrammatical for second-position clitics to appear in initial position

\(^3\)They resemble non-pronominal clitics in Tagalog (Schachter and Otanes 1972; Schachter 1973; Anderson 2011).
• When the host is a verb or negator, clausal clitics behave similarly, as in (2) & (3)

(2)  a. Yaanakenatari...
    \[i\quad ag\quad -an\quad -k\quad -i\quad =tari\]
    3M.S- take -ABL -PFV -R:ACT -1O =CNGR
   It’s that he took me...

   b. “...teetari nonkejetempipi.”
    \[tee\quad =tari\quad no\quad N\quad keje\quad -e\quad -Npi\quad -ji\]
    NEG:R =CNGR 1S- IRR- be.like -IRR -2O -NEG:R
   “...because we’re not like you.”

(3)  a. Oposatanakegeti nokemi bero bero.
    \[o\quad posa\quad -an\quad -k\quad -i\quad =geti\]
    3F.S- be.cooked -ABL -PFV -R:ACT =when 1S- hear -R:ACT IDEO:voices
   When it cooked, I heard voices.

   b. Napatsaabantake teegeti nontsateroji.
    \[no\quad -apatsaaban\quad -ak\quad -i\quad =geti\]
    1S- read -PFV -R:ACT NEG:R =when 1S- IRR- know -IRR -3F.O -NEG:R
   I would read when I didn’t know it.

• However, when the possible host is a topic expression, two sorts of clausal clitics emerge

• Those like \(=tari\) (“T”-type) may attach either to the demonstrative (4a) or to the constituent following the noun (4b)

(4)  a. Iriratari Taatakini tee inkajemeji...
    \[iri\quad ra\quad =tari\quad Taatakini =ni\quad tee\quad i\quad N\quad kajem\quad -e\quad -ji\]
    3M.S.IRR- D:MED =CNGR Taatakini =DEC NEG:R 3M.S- IRR- shout -IRR -NEG:R
   Taatakini didn’t shout...

   b. Ora noninanite teetari ontsatakotempaji...
    \[o\quad ra\quad non\quad iiinani\quad -te\quad =tari\quad o\quad N\quad tsa\quad -ako\quad -e\]
    3F- D:MED 1P- mother -POSS NEG:R =CNGR 3F.S- IRR- know -A:INDR -IRR
    -Npa\quad -ji
    -MID -NEG:R
   My mother didn’t know...

• Those like \(=geti\) (“G”-type) may attach only to the constituent following the entire topic expression (5)

(5)  Irira pinchinchi itsatabakerogeti...
    \[iri\quad ra\quad p\text{iNchiNchi}\quad i\quad tsa\quad -ab\quad -k\quad -i\quad -ro\quad =geti\]
    3M- D:MED vampire.bat 3M.S- kow -DIR -PFV -R:ACT -3F.O =when
   When the vampire bat heard her...
• In neither case may the clitic attach to the noun, as summarized in Table 4

| Table 2: Distribution of Clausal Second-position Clitics ("T" & "G" Types) |
|-----------------|---|---|---|---|
|                 | V | Neg | FocP | TopP-D N |
| =tari "because"| V=CL | Neg=CL V | FocP=CL V | D(=CL) N V(=CL) |
| =geti 'when, if, where' | V=CL | Neg=CL V | FocP=CL V | D N V=CL |
| GENERALIZATIONS | 2P | 2P | 2P | clitic dependent |

3.2 Nominal Second-position Clitics

• Clausal clitics can be contrasted with nominal clitics, which sometimes attach to a noun
  =pae PLURAL
  =ki LOCATIVE

• Both attach to an unmodified noun (6a,b) and stack in the order in (6) when they co-occur

(6) a. otsempipae
    otseNpi =pae
    mountain =PL
    mountains

    ibakopaeki

b. otsempiki
    otseNpi =ki
    mountain =LOC
    in the mountain

c. i- bako =pae =ki
    3M.P- hand =PL =LOC
    in his hands

• However, when a prenominal adjective is present, the adjective hosts a single clitic (7a,b)

(7) a. otsipaki otsempi
    o- tsipa =ki otseNpi
    3F- other =LOC mountain
    in another mountain

b. iririjegipae pamakabiri
    iririjegi =pae pamakabiri
    large =PL fish
    large fish

• When both clitics are present in the context of a prenominal adjective, they separate (8)

(8) onirojegipae chomoki
    onirojegi =pae chomo =ki
    large =PL clay.pot =LOC
    in large clay pots
• With any non-adjectival modifier, the noun serves as the host of all clitics

(9) **Demonstrative**

a. irira iraajanirepa[^1]  
   *iri- ra iri- aajanire*[^2]  
   3M- D:MED 3M.P- employee  
   his employees

b. onta otsep[^3]  
   *o- Nta otseNpi*[^4]  
   3F- D:DIST mountain  
   in that mountain

(10) **Quantifier**

a. maasano nojaat[^5]  
   *maasano no- jaai*[^6]  
   all 1P- brother  
   all my brothers

b. osheki kotsiron[^7]  
   *osheki kotsironaki*[^8]  
   many pot  
   in many pots

c. maasano tsobiron[^9]  
   *maasano tsobironaki*[^10]  
   all house  
   in all the houses

• The distribution of nominal clitics is summarized in Table 3

<table>
<thead>
<tr>
<th>Form</th>
<th>[N]</th>
<th>[Adj] [N]</th>
<th>Dem/Quant/Num [N]</th>
</tr>
</thead>
<tbody>
<tr>
<td>=pae</td>
<td>N=CL</td>
<td>Adj=CL N</td>
<td>Dem/Quant/Num N=CL</td>
</tr>
<tr>
<td>=ki</td>
<td>N=CL</td>
<td>Adj=CL N</td>
<td>Dem/Quant/Num N=CL</td>
</tr>
<tr>
<td>GENERALIZATIONS</td>
<td>2P</td>
<td>2P</td>
<td>never 2nd position</td>
</tr>
</tbody>
</table>

• The sequence of a demonstrative, quantifier, or numeral and a noun is impenetrable with respect to the nominal clitics, unlike what is observed for clausal clitics (11)

(11) [Orakea mankiparentsip[^11] yaajiakero...  

   *o- ra mankiparentsi*[^12]  
   3F- D:MED  
   woman  
   The women they took [as wives]...

• **A conspiracy?** Nouns are attested in naturalistic data only with a single modifier

  – Thus adjectives are not attested with demonstratives, quantifiers, or numerals

• The ban on demonstratives hosting nominal clitics leaves that position open for clausal clitics
4 Analysis

4.1 Claim 1 – Caquinte Clitics are Prosodically Deficient

- Like affixes, clitics are phonologically bound and require a host

  (12) a. [ibétsata] ‘he speaks’
      b. [ibetsáta=geti] ‘when he speaks’

  (13) a. [imetojake] ‘he died’
      b. [imetojáke=tari] ‘because he died’

  (14) a. [ábíro] ‘(it’s) you’
      b. [ábíro=sa] ‘(it’s) you’ (inferential)

- **Subcategorization** A morpheme can be prespecified for prosodic requirements (Inkelas 1990; Raffelsiefen 1999; Itô and Mester 2013; Bennett et al. to appear), requiring a particular host

  - In this case a preceding phonological word, as evidenced by stress shifts

  (15) **Underlying representation for** \(=\text{geti}\)
  \[\Sigma: [+\text{CONGRUENT}]\]
  \[\Phi: [\omega[...]/\text{geti}/]\]
  where \(\omega[...]=\text{phonological word}\)

- The presence of stress shift suggests that the clitic and its host form a single phonological word to which stress is assigned

![Prosodic Structure Types](image)

- We refer to these as internal clitics, following Selkirk (1996); Peperkamp (1997); Vogel (2009); Cardinaletti and Repetti (2009)

4.2 Claim 2 – \([X]=\text{CL}\) Structure is Formed Post-syntactically

- There are two logical possibilities for deriving the \([X]=\text{CL}\) (host plus clitic) constituent
• When =tari appears with a topicalized phrase, it appears between the D element and the N
  
  – [D=CL N]_TopP [X]
  – [D N]_TopP [X]=CL

  cf. *[D N=CL]_TopP [X]

• **Argument 1** Syntax is phonology-blind
  
  – It is the clitics, not the host (e.g., noun, verb, etc.) that are prosodically deficient
  – We assume that syntactic movement is strictly sensitive to syntactic features, not phonological features (*contra* Richards (2010, 2016))
    
    * We rule out a syntactic EPP condition on CLP with a phonological motivation

• **Argument 2** Evidence from topicalized clauses
  
  – As shown above, topic expressions precede the verb and canonically consist of a demonstrative followed by a noun ([D N V] in the same field as the clausal clitics
  – When a clausal clitic co-occurs with a topic expression, the clitic can only attach to the demonstrative or some element lower than the noun (e.g., verb), but never the noun
  – This would be unexpected under the alternative syntactic movement upwards
    
    * Under the syntactic movement account, the order D=CL N would entail that D has moved around the clitic, forming a discontinuous constituent with the noun
    
    * And it would be unexpected that the entire DP could not move around the clitic (cf. *[D N=CL] V)

• **Schematic post-syntactic analysis**
  
  syntactic structure: [CL]_TopP [DP D N]_TopP Top [VP V] ...

  spell-out/vocabulary insertion: /CL/ /D/ /N/ /V/ ...

  post-syntactic clitic movement: /D/ /CL/ /N/ /V/ (cf. X/D/ /N/ /CL/ /V/)

Figure 3: Deriving [X]=CL
4.3 Claim 3 – Clausal Clitics are Syntactically Merged at Different Points in the Clausal Spine

- Recall the structure of the clausal spine: TopicP > {FocusP; NegP} > [...VP...]

<table>
<thead>
<tr>
<th></th>
<th>[V]</th>
<th>[Neg]</th>
<th>[V]</th>
<th>[FocP]</th>
<th>[V/X]</th>
<th>[TopP; D; N; V/X]</th>
</tr>
</thead>
<tbody>
<tr>
<td>=tari “because”</td>
<td>V=CL</td>
<td>Neg=CL V</td>
<td>FocP=CL V</td>
<td>D(=CL) N V(=CL)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>=geti ‘when, if, where’</td>
<td>V=CL</td>
<td>Neg=CL V</td>
<td>FocP=CL V</td>
<td>D N V=CL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GENERALIZATIONS</td>
<td>2P</td>
<td>2P</td>
<td>2P</td>
<td>clitic dependent</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Because both =tari and =geti always appear in second position with [V], [Neg] [V], and [FocP] [V], we capture this by always merging the clitics above these positions

(16) **Verb, Negation, Focus**

<table>
<thead>
<tr>
<th>Syntax</th>
<th>Surface</th>
<th>Syntax</th>
<th>Surface</th>
</tr>
</thead>
<tbody>
<tr>
<td>tari &gt; V</td>
<td>V=tari</td>
<td>geti &gt; V</td>
<td>V=geti</td>
</tr>
<tr>
<td>tari &gt; Neg &gt; V</td>
<td>Neg=tari V</td>
<td>geti &gt; Neg &gt; V</td>
<td>Neg=geti V</td>
</tr>
<tr>
<td>tari &gt; Foc &gt; V</td>
<td>Foc=tari V</td>
<td>geti &gt; Foc &gt; V</td>
<td>Foc=geti V</td>
</tr>
</tbody>
</table>

- With [TopP] [V], however, =tari and =geti behave differently, which we capture by imposing different restrictions on where each of these two clitic types can merge

(17) **Topic**

<table>
<thead>
<tr>
<th>Syntax</th>
<th>Surface</th>
<th>Syntax</th>
<th>Surface</th>
</tr>
</thead>
<tbody>
<tr>
<td>tari &gt; Top &gt; V</td>
<td>Top=tari V</td>
<td>*geti &gt; Top &gt; V</td>
<td>*Top=geti V</td>
</tr>
<tr>
<td>Top &gt; tari &gt; V</td>
<td>Top V=tari V</td>
<td>Top &gt; geti &gt; V</td>
<td>Top V=geti V</td>
</tr>
</tbody>
</table>

- We end by showing that merge order is sensitive only to the immediately lower node

- First, adverbs may precede the verb and serve as a host to =geti

(18) **geti > Adv > V**

...iroakergeti ikorakeke nogoestate.

*iroakera  =geti i- korake -k -i no- goesta -te
first.time =when 3m.s- come -PFV -R:ACT 1P- linguist -POSS

...when our linguist came for the first time.

- Specific evidence for merge order comes from the distribution of clitics with topic expressions occurring with adverbs

(19) **Ordering possibilities of topics and adverbs**

Top > Adv > V
Adv > Top > V
Recall that *geti > Top > V is unattested

- However, =geti can be merged above Top if there is an intervening Adv (20)

\[(20) \quad \text{geti} > \text{Adv} > \text{Top} > \text{V} \quad \text{Adv=cl \ Top \ V} \]

Arigeti naatimpa irimetojiten\(a\)...

\[\begin{array}{lllll}
\text{ari} & \quad \text{=geti} & \text{naat} & \text{Npa} & \text{iri-} \\
\text{thus} & \quad \text{=if} & \text{1:TOP} & \text{3M.S.IRR-} & \text{kill} \\
\end{array}\]

If they kill me...

An explanation of this pattern might require enriching the selectional properties of the =geti-class of clausal clitics to only combining with certain sorts of hosts, independent of syntax

5 Final Observations

- A clitic and its host form a single word phonologically, formed post-syntactically
- Clausal clitics fall into two classes
- The sequence [D N] cannot be broken up by nominal clitics or G-type clausal clitics
- Restrictions on G-type clitics and topic expressions disappear when an adverb intervenes

References


BENNETT, RYAN; BORIS HARIZANOV; and ROBERT HENDERSON. to appear. Prosodic Smothering in Macedonian and Kaqchikel. *Linguistic Inquiry* 49(2).


A Is the Domain of Second Positionhood Self-imposed?

- || [X]=cl [Y] ||Domain

- What determines the domain?

- Logical possibilities:
  - external imposition – element outside of domain: head directly above =cl

- Domain self-imposition – at the point of linearization, a morpheme X which is realized with a vocabulary item /]/X/ imposes a lefthand domain boundary

- D > N > Top > cl > V
  /D/ /N/ /]/=cl/ /V/ \rightarrow (D) (N) || (V=cl)

- cl > D > N > Top > V
  /]/=cl/ /D/ /N/ /V/ \rightarrow || (D=cl) (N) (V)

- Alternative with external imposition: Top is a phase head that spells out its sister

(21) Aisa ikantajititari itsigajatakogetajiri...  EXAMPLE WITH ADVERB
    aisa i- kan -ji -i [=tari] i- tsigaja -ako -ge -aj
    also 3M.S- say -NREF -R:ACT =CNGR 3M.S- place.in.fish.trap -A:INDR -DISTR -REG
    -i -ri
    -R:ACT -3M.O

They also used to say that he placed them [the dead] in a fish trap...

- Under external imposition, we would be forced to say in these examples that the adverb (or its head) spells out its sister