A syntactic filter on second-position clitics in Tagalog
Daniel Kaufman
Cornell University & The CUNY Graduate Center - dak37@cornell.edu

0.1 Intro

2P clitics have played a considerable role in shaping our understanding of the phonology-syntax interface. Non-syntactic placement of 2P clitics requires reference to phonology in some capacity but accounts differ widely on how this reference is made.

A typology of approaches (adapted from Bošković 2001):

Strong syntax: Phonology is negligible in the placement of 2P clitics. It all takes place in regular syntax before PF

Strong Phonology: Placement is entirely a PF matter and has little or nothing to do with regular syntax

Weak Syntax: Syntax offers several options for clitic placement and PF adjudicates between these by means of a phonological filter

Weak Phonology: A limited number of phonological repair mechanisms are able to move clitics from their base position to 2P to satisfy a clitics phonological needs

Focusing here on the phenomenon of “obligatory non-pre-enclitics” (Schachter & Otanes 1972:187) AKA “fortresses” (Halpern 1995), I will argue for the existence of a syntax-phonology interaction which has not yet been considered and which is not easily subsumed under any of the above approaches. Specifically, the Tagalog facts argue for a counter-derivational interaction: phonological placement with a syntactic filter.

1.1 Tagalog 2P clitics


Tagalog clitics consist of two broad types - pronominal (T1) and adverbial (T2):

<p>| Table 1. Tagalog pronominals |
|-----------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th>Trad. labels</th>
<th>Gloss</th>
<th>Features</th>
<th>NOM</th>
<th>GEN</th>
<th>NOM</th>
<th>GEN</th>
<th>OBL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st sing.</td>
<td>1s</td>
<td>[1]</td>
<td>=ako</td>
<td>=ko</td>
<td>ako</td>
<td>ákin</td>
<td>sa ákin</td>
</tr>
<tr>
<td>2nd sing.</td>
<td>2s</td>
<td>[2]</td>
<td>=ka</td>
<td>=mo</td>
<td>ikaw</td>
<td>iyo</td>
<td>sa iyo</td>
</tr>
<tr>
<td>3rd sing.</td>
<td>3s</td>
<td>[∅]</td>
<td>=siya</td>
<td>=niya</td>
<td>siya</td>
<td>kaniya</td>
<td>sa kaniya</td>
</tr>
<tr>
<td>1st excl. pl.</td>
<td>1+3</td>
<td>[1, p]</td>
<td>=kami</td>
<td>=námin</td>
<td>kami</td>
<td>ámin</td>
<td>sa ámin</td>
</tr>
<tr>
<td>(1st dual)</td>
<td>1+2</td>
<td>[1,2]</td>
<td>=kata/kita</td>
<td>=ta</td>
<td>kata/kita</td>
<td>kanita</td>
<td>sa kanita</td>
</tr>
<tr>
<td>1st incl. pl.</td>
<td>1+2P</td>
<td>[1,2, p]</td>
<td>=táyo</td>
<td>=nátin</td>
<td>táyo</td>
<td>átin</td>
<td>sa átin</td>
</tr>
<tr>
<td>2nd pl.</td>
<td>2p</td>
<td>[2, p]</td>
<td>=kayo</td>
<td>=nín</td>
<td>kayo</td>
<td>inyo</td>
<td>sa inyo</td>
</tr>
<tr>
<td>3rd pl.</td>
<td>3p</td>
<td>[∅, p]</td>
<td>=síla</td>
<td>=nila</td>
<td>síla</td>
<td>kaniya</td>
<td>sa kaniya</td>
</tr>
</tbody>
</table>

Portmanteau forms: [1.GEN+2.NOM] =kita, kita
Table 2. Tagalog adverbial clitics

<table>
<thead>
<tr>
<th>CLITIC</th>
<th>FREE</th>
</tr>
</thead>
<tbody>
<tr>
<td>aspect</td>
<td></td>
</tr>
<tr>
<td>=na ‘already’</td>
<td>ø</td>
</tr>
<tr>
<td>=pa ‘still’</td>
<td>ø</td>
</tr>
<tr>
<td>focus</td>
<td></td>
</tr>
<tr>
<td>=din ‘also’</td>
<td>ø</td>
</tr>
<tr>
<td>=man ‘even’</td>
<td>ø</td>
</tr>
<tr>
<td>=naman ‘switch topic’</td>
<td>(naman)</td>
</tr>
<tr>
<td>=nág ‘emphasis’</td>
<td>ø</td>
</tr>
<tr>
<td>=lang ‘only’</td>
<td>lámang</td>
</tr>
<tr>
<td>=talaga ‘emphasis’</td>
<td>talaga</td>
</tr>
<tr>
<td>politeness</td>
<td></td>
</tr>
<tr>
<td>=pò, =hó ‘politeness’</td>
<td>ø</td>
</tr>
<tr>
<td>mood</td>
<td></td>
</tr>
<tr>
<td>=pala ‘surprise’</td>
<td>ø</td>
</tr>
<tr>
<td>=yátà ‘perhaps’</td>
<td>ø</td>
</tr>
<tr>
<td>=sána ‘hopefully’</td>
<td>sána</td>
</tr>
<tr>
<td>=náwa ‘hopefully’</td>
<td>náwa</td>
</tr>
<tr>
<td>=ba ‘question marker’</td>
<td>(baga)</td>
</tr>
</tbody>
</table>

In simple cases, both types of clitics cluster together in 2P:

(1)  Gúrè=ŋá=siya  ni=Mao
teacher=EMP=3S.NOM P.GEN=Mao
‘She’s Mao’s teacher.’ (N host)

(2)  Na-túto=ŋá=siya  naŋ=wíka=ŋ  Instsik
AV.BEG-learn=EMP=3S.NOM GEN=language=LNK Chinese
‘She really learned Chinese’ (V host)

(3)  Hindì=ŋá=siya  na-túto  naŋ=wíka=ŋ  Instsik
NEG=EMP=3S.NOM AV.BEG-learn GEN=language=LNK Chinese
‘She really didn’t learn Chinese’ (Neg host)

(4)  Saan=ŋá=siya  na-túto  naŋ=wíka=ŋ  Instsik?
NEG=EMP=3S.NOM AV.BEG-learn GEN=language=LNK Chinese
‘Where did she really learn Chinese?’ (Wh- host)

(5)  Kay=Yao=ŋá=siya  na-túto  naŋ=wíka=ŋ  Instsik
P.OBL=Yao=EMP=3S.NOM AV.BEG-learn GEN=language=LNK Chinese
‘She really learned Chinese from Yao.’ (Focus host)

(6)  Dápat=ŋá=siya  na-túto  naŋ=wíka=ŋ  Instsik
NEG=EMP=3S.NOM AV.BEG-learn GEN=language=LNK Chinese
‘She really should have learned Chinese.’ (Modal host)

(7)  Ma-bilis=ŋá=siya=ŋ  na-túto  naŋ=wíka=ŋ  Instsik
NEG=EMP=3S.NOM=LNK AV.BEG-learn GEN=language=LNK Chinese
‘She really learned Chinese quickly.’ (Adverbial host)
1.2 The untenability of the strong and weak syntax approaches

Unlike in languages like Serbo-Croatian, non-clitic arguments are not licensed in any of the clitic positions in (1)-(7). Likewise, clitics are not licensed in argument positions:

(8)  Gûrô [=ako] ni=Mao[*=ako] teacher=15.NOM P.GEN=Mao
     ‘I’m Mao’s teacher.’

(9)  Gûrô [*aŋ=babae] ni=Mao [aŋ=babae] teacher P.GEN=Mao NOM=woman
     ‘The woman is Mao’s teacher.’

As Chung (2003:558) shows, a syntactic account for Chamorro 2P clitics, requires illicit extraction in examples like (8) (among many others). Tagalog (and Austronesian languages more generally) disallow the extraction from NP which would be required for a syntactic account of (8):

(10)  *Nino/kaniño, aŋ=babae  nP[gûrô  t1]?
     whose NOM=woman teacher
     (For, ‘Whose teacher is the woman?’)

The total lack of a correlation between extractable constituents and clitic hosts makes both the weak and strong syntax approaches untenable.

1.3 The untenability of the weak phonology approach

Halpern (1995) proposes that 2P may be achieved by two separate mechanisms. Clitics are adjoined to a position in the left periphery and syntactic movement can place material to their left, thereby satisfying their prosodic dependency. Or, if the clitics reach PF without a clitic host, an operation dubbed Prosodic Inversion inverts clitics with the prosodic word to their right.

If clitics are uniformly adjoined to the left edge of their positioning domain, then a weak phonology approach could account for their placement via Prosodic Inversion:

(11)  a.  Fp[=ná=sïya  np[na-túto  naŋ=wíka=ŋ  Instsik]]
       =EMP=3S.NOM AV.BEG-learn GEN=language=LNK Chinese

       b.  ______ na-túto=ŋá=sïya  naŋ=wíka=ŋ  Instsik
            AV.BEG-learn=EMP=3S.NOM GEN=language=LNK Chinese
            ‘She really learned Chinese’

This could also explain the impenetrability of oblique focus phrases to clitics as focus phrases can felicitously be analyzed as moving above the base position of the 2P clitics:

(12)  Focp[[Sa=dalawa=ŋ  malaki=ŋ  palabas], Fp[=sila  np[li~litaw  t1]]]
     OBL=two=LNK big=LNK show =3P.NOM AV.PROSA~appear
     ‘They will appear in two big shows’
But, two serious problems arise:

- Pronominal and adverbial clitics do not follow precisely the same patterns. Focus phrases are impenetrable to pronominal clitics but not to adverbial clitics:

  \[(13) \quad \text{[Sa=dalawa=ba=η malaki=η palabas]=silā lī˘litaw?} \quad \text{OBL=two=QM=LNK \big=LNK \ show=3P.NOM \ AV.PROS\^{appear}} \]

  ‘Will they appear in two big shows?’

  This could be handled by adjoining/generating adverbial clitics above the focus phrase and pronominal clitics below the focus phrase:

  \[(14) \quad \text{QM} \quad \Downarrow \quad \text{PRON} \]

  \[\text{[Sa=dalawa=ba=η malaki=η palabas]=silā lī˘litaw t} \]

- However, when a host appears even higher than the focus phrase, it still attracts clitics. If the base/adjunction position is below the focus position, there is no motivation for the pronominal clitic to keep moving.

  \[(15) \quad \text{Hindi=ba=silā [sa=dalawa=η malaki=η palabas] lī˘litaw?} \quad \text{NEG=QM=3P.NOM OBL=two=LNK \ big=LNK \ show \ AV.PROS\^{appear}} \]

  ‘Won’t they appear in two big shows?’

1.4 Problems for a strong phonology approach

The above facts are also problematic for a strong phonology approach:

- It is not clear how to characterize impenetrable constituents in prosodic terms.

- It is not possible to characterize pronouns as prosodic phrase clitics and adverbials as prosodic word clitics, as in Chamorro (Chung 2003):

  (i) 1σ pron. clitics occur internally to 1σ adverbial clitics within the clitic cluster

  (ii) Pronominal clitics cannot be separated from the left edge of their domain in simple constructions by pre-modifiers, as is the case in Chamorro.

  \[(16) \quad \text{ma-bilis\[^{=ka}\]=ba=η} \quad \text{t<um>ákbo\[^{=ka}\]?} \quad \text{ADJ-quick=2S.NOM=QM=LNK <AV>run} \]

  ‘Did you run quickly?’

2.1 The solution: phonological placement + syntactic filter

As shown in (1)-(7), 2P Clitics are positioned after the first available prosodic word within their domain. Following Anderson (1996, 2005) Legendre (, this can be handled in OT as the conflict of a constraint which requires clitics to appear leftmost within their domain and a constraint which prohibits them from appearing initially:
ALIGN-L (F_{person}; CP)
The spell-out of a person feature is aligned to the left edge of CP

ALIGN-L (F_{adverbial}; ForceP)
The spell-out of an adverbial feature is aligned to the left edge of ForceP

The different alignment of adverbial and pronominal clitics is only apparent with certain PWh complementizers and with topics. Adverbials may attach to these but pronominals may not:

(17) a. upanė=diŋ hindi=táyö ma-sisi naŋ=maŋa=há~halili sa=átin so.that=also NEG=1p.NOM PV.ABL-blame GEN=PL=AV. INCM~vote OBL=1p ‘...also so that we won’t be blamed by those who will vote for us.’ (Malvar 1991 via ileto 1997:164)

b. Bukas=ba naŋ=gabi ay ð-sá~sayaw=síla naŋ=pandango? tomorrow=QM GEN=night TOP AV- INCM~dance=3p.NOM GEN=fandango ‘Tomorrow night, will they dance a fandango?’ (S&O)

*WEAKSTART (PPh)
Violated by a prosodic phrase whose left edge is aligned to a non-PWh head (head status distinguishes between PWh clitics and PWh lexical items, see Kaufman 2008)

The result of the ranking *WEAKSTART (PPh) >> ALIGN-L (F; XP) results in the spell out of features avoiding PPh initial position, as clitics cannot constitute PWh heads.

To handle impenetrability Anderson (1993:16, 2005) proposes: “INTEGRITY constraints, which characterize uninterruptible sub-parts of a domain and thus define what counts as occupying ‘first’ position.”

But listing types of impenetrable phrases as INTEGRITY constraints lacks explanatory force:
• Misses the unifying syntactic generalization over the relevant environments (below)
• Has to be parochialized for pronominal clitics and adverbial clitics

Impenetrability stems follows from the following condition:

(18) Clitic Visibility Condition (CVC)
For an argument clitic \( \alpha \) assigned a thematic role by a predicate head \( \beta \), the minimal maximal projection linearly containing \( \alpha \) must dominate \( \beta \).

(19) Linear Containment
\( \alpha \) is linearly contained in \( \beta \) if \( \alpha \) is both preceded and followed by overt terminal nodes dominated by \( \beta \), i.e., in the configuration \( [X \alpha Y] \).

2.2 The CVC in action I: focus fronted & interrogative oblique phrases

(20) [F]-CP

(21) [F]-CP
(22) a. kaníno=ka=ŋ estudyánnte?
   OBL:who=2S.NOM=LNK student
   ‘Whose student are you?’

b. kaníno[*=ka]=ŋ estudyánnte[=ka] nag-bigay naŋ=pérâ?
   OBL:who=2S.NOM=LNK student=2S.NOM AV:BEG-give GEN=money
   ‘To which student did you give money?’

(23) [F]-CP
    \[ \text{ObiP} \]
    \[ \text{TP} \]
    \[ \text{[Wh...cl...Pred}^0_i \] ... t \} \]

NB: Attestations of CVC violations can be found but these are disfavored by most speakers:

(24) ...dáhiñ kuniŋ=kaníŋ=kaníno=ka=ŋ laláke l<um>á~lápít
    ...because COMP=PL~OBL:who=3S.NOM=LNK man <AV:BEG>INCM~approach
    ‘...because you approach just any man.’ <www.spunby.com/search/ch_porka.html>

2.3 The CVC in action II: modification structures

(25) a. masyádo=ŋ ma-bilis na mag-pa-loko
    overly=LNK ADJ-quick LNK AV.INF-CAU-crazy
    ‘too quick to allow oneself to be tricked.’ (non-finite)

b. masyádo=ŋ ma-bilis na nag-pa-loko
    overly=LNK ADJ-quick LNK AV.BEG-CAU-crazy
    ‘was too quick to allow himself to be tricked.’ (finite)

(26) b. (*mag-pa-loko naŋ=masyádo=ŋ ma-bilis
    AV.BEG-CAU-crazy GEN=overly=LNK ADJ-quick
    (can only be interpreted as imperative/hortative) (non-finite)

a. nag-pa-loko naŋ=masyádo=ŋ ma-bilis
    AV.BEG-CAU-crazy GEN=overly=LNK ADJ-quick
    ‘was too quick to allow himself to be tricked.’ (finite)

(27) A’
    \[ \text{A} \]
    \[ \text{VP} \]
    \[ \text{[Mod A] (Pred}^0_i) \]

(28) VP
    \[ \text{AP} \]
    \[ \text{VP} \]
    \[ \text{[Mod A] (Pred}^0_i) \]
Prediction: Pronominal clitics should be able to interrupt the adverb + adjective constituent in structures such as (27) but not in (28) (because clitics between Mod and A in (28) would be linearly contained by AP and invisible to Pred$^0$).

This appears to be confirmed by (29) and (30):

(29) masyádo=$\text{ako}$=η ma-bilis mag-pa-lóko  
   overly=1S.NOM=LNK ADJ-quick AV-CAU-crazy  
   ‘I’m too quick to allow myself to be tricked’  
   <http://ayka08.multiply.com/journal/item/5/the_change_in_me>

(30) a. masyádo=η ma-bilis=$\text{ako}$=η nag-pa-lóko  
    overly=LNK ADJ-quick=1S.NOM=LNK AV.BEG-CAU-crazy  
    ‘I was too quick to allow myself to be tricked’

b. ?*masyádo=$\text{ako}$=η ma-bilis na nag-pa-lóko  
   overly=1S.NOM=LNK ADJ-quick LNK AV.BEG-CAU-crazy

(31) a. masyádo=η ma-dalas=$\text{ko}$=siya=η  
    overly=1S.NOM=LNK ADJ-frequent=1S.GEN=3S.NOM=LNK  
    <BEG>call-LV  
    ‘I called her too frequently.’

b. ?*masyádo=$\text{ko}$=siya=η ma-dalas na t<in>awág-an  
   overly=1S.GEN=3S.NOM=LNK ADJ-frequent LNK  
   <BEG>call-LV

(31) a. masyádo=ba=η ma-dalas=$\text{ko}$=siya=η  
    overly=QM=LNK ADJ-frequent=1S.GEN=3S.NOM=LNK  
    <BEG>call-LV  
    ‘Did I call her too frequently?’

3.1 Conclusion

- We’ve shown that placement of 2P clitics in Tagalog cannot rely on regular syntax in any capacity (i.e., neither via a strong syntax, weak syntax nor weak phonology account). Nonetheless, a filter referring to hierarchical structure is necessary to account for impenetrable constituents.

- This filter almost looks like the ECP at PF: clitics must be in a pseudo c-command relationship with their predicate but they are not sitting in a syntactic position.

- A surface syntactic filter suggests that syntax never becomes completely invisible, even at late stages of PF (i.e. after “prosodic positioning”). Note, however, that specific projections need not be referenced; only basic hierarchical structure.

- Although conceptually attractive, we should rethink the idea of trading syntactic relationships for linear relationships (Marantz 1988) and consider the possibility of prosodic and syntactic filters on an enriched output; not a radical idea but counter to the canonical interpretation of the T-model and its descendents.
3.2 Further issues

- Is this strictly a “morphological issue”, i.e., part of a syntactic filter operating at the point where morphosyntactic features are spelled-out?

  The “Everybody’s a clitic!” (EAC) problem
  Non-novel personal names can also appear in clitic position (Billings 2005) and clitics can coerce full NP conjuncts into clitic position. Lexical material, however, should be out of the purview of feature Spell-Out.

(32) Hindi = naman si= Ariel neg-ta~tagalog!
    NEG=SWITCH P.NOM=Ariel AV.BEG=INCM~Tagalog
    ‘Ariel doesn’t speak Tagalog!’  <http://pochanginamo.blogspot.com/2007_10_01_archive.html>

(33) Hindi = ko at na= ma= na= ka= samahan= ko= = n= Filipina kailanman
    NEG=INQ CONJ GEN=PL=colleague=1S.GEN=LNK Filipina ever

  s=in=Ir==a~ti=jin= na= i=ba sa=amin
  <BEG>destroy-PV NOM=view GEN=other OBL=1P.GEN
  ‘Neither I nor my Filipina colleagues ever destroyed the view of others towards us.’
  <www.nursesthoughts.com/2007_09_01_archive.html>

References


  Talk given at the Penn Linguistics Colloquium 32, U. Penn.


