This paper concerns the grammatical and annotation notes on the GRAID annotation system, which was extended for the discourse data of Arta, a language spoken in Northern part of Luzon, the Philippines. GRAID (Grammatical Relations and Animacy in Discourse) is an annotation system developed by Geoffrey Haig and Stefan Schnell to explore the relation between argument realization patterns and possible discourse-functional motivations across languages through a qualitative approach. When applying it to Arta and other Philippine languages, it is necessary to calibrate the cross-linguistic concepts by providing some empirical data and evidence. This article first provides evidence for determining the grammatical functions of each argument (S, A, P, and other roles) within each clause. Possible realization patterns of referential expressions are discussed, and some additional information is tagged for capturing these patterns in Arta. The structures of complex sentences, especially those involving a relative clause, are also noted.

Keywords: Arta, GRAID annotation, argument realization patterns, ergativity, gap construction

1. Introduction
2. Clause structure
3. Referential expressions
4. Complex sentence: the treatment of gap constructions
5. Concluding remarks

1. Introduction

This article aims to document the implementation of the Grammatical Relations and Animacy in Discourse annotation system (GRAID) to Arta discourse data. The GRAID glossing conventions are “a system of symbols and conventions for glossing..."
the grammatical relations and overt forms (noun phrases, pronouns etc.) of major clause constituents in texts” developed by Geoffrey Haig and Stefan Schnell (Haig and Schnell 2014).² Arta is a Philippine language spoken by the Philippine Negrito people who inhabit Quirino and Aurora provinces in Northeastern Luzon in the Philippines. The language was first reported by Reid (1989) and is being documented by the present author (Kimoto 2014, 2017a, b). This paper begins with an overview of GRAID annotations as applied to Arta, and annotations on clause structure (Section 2), followed by its implementation to nominal structure (Section 3), and sentence (Section 4).

2. Clause structure

2.1. Grammatical relations of core arguments

In GRAID, each annotator is expected to annotate such tags as ⟨:s⟩, ⟨:a⟩, ⟨:p⟩, and ⟨:obl⟩ onto each argument. However, as is widely recognized, there has been a long controversy over the nature of case marking system in Philippine languages, or more broadly, Philippine-type languages, including most Formosan languages in Taiwan and some languages in Sulawesi and Borneo. Although there are admittedly many Philippine-type languages that do not show a clear ergative-absolutive alignment, it is relatively clear that Arta does have the ergative-absolutive alignment in the case marking of full nominals, personal forms, and demonstratives.³ The following elicited sentences exemplify the typical encoding patterns of single-participant and two-participant events (note that the label genitive gen is used, and has been used in Austronesian linguistics in general, both for possessor role and transitive subject because of their homonymy):

(1) \( T<in><um>addyor=di \ [i \ babakat=i]. \)  
\(<\text{pST}><\text{INTR}>\text{stand}=\text{COMP} \ \text{DEF} \ \text{old.woman}=\text{SPC}\)  
‘The old woman stood up.’ (elicited)

(2) \( B<in>isag=di \ [ni \ babakat=i] \ [i \ bo:te]. \)  
\(<\text{pST}>\text{break}=\text{COMP} \ \text{GEN.DEF} \ \text{old.woman}=\text{SPC} \ \text{DEF} \ \text{bottle}\)  
‘The old woman/women broke the bottle(s).’ (elicited)

Regardless of whether the single core argument of the intransitive clause designates an agent or undergoer (cf. split-Š system), the nominal is introduced by the determiner \( i \) in (1), which is identical to the determiner of the undergoer role in the two-participant event in (2), whereas the agent-like argument of (2) receives a distinct marking \( ni \).

Note, however, that there is another encoding strategy in which the same prototypical transitive situation is expressed. This kind of construction, which is called “actor-focus

---

² The list of symbols employed in the GRAID convention is given in the appendix of Schnell and Schiborr (2018).
construction,” has been problematic in Philippine languages in that the actor argument receives the marking identical to the single core argument of the intransitive clause. In Arta, however, this construction is evidently not considered to be a transitive clause on the grounds that the undergoer role is marked by the oblique determiner \textit{ti}. Compare (3) with other examples (4)–(5) shown below, in which the determiner \textit{ti} is used to introduce the temporal phrase ‘in March’ and the locative phrase ‘in the mountains’. Thus, the apparent transitive clause with so-called actor-focus construction is not a true transitive clause in Arta, just like the “conative” construction in English I kicked at him where the patient role is encoded by the prepositional phrase.\footnote{This encoding pattern is not preferred in the language in such typical transitive situations such as that of the change-of-state event towards a single undergoer individual, with punctual and completive aspect. This actor-voice construction with a patientive argument appears in such cases as ‘to eat food’ and ‘to hunt for wild pigs’, with the object as a mass entity or non-specific/indefinite entity.}

(3) \textit{Nam-bisag=di} \ [i \ babakat=i] \ [ti \ bo:te].
\textsc{pst.intr-break} = \textsc{comp} \ \textsc{def} \ \textsc{old.woman}=\textsc{spc} \ \textsc{obl.def} \ \textsc{bottle}
‘The old woman/women broke the bottle(s).’ (elicited)

(4) \textit{Pam-mula} \ =ami \ ta \ pagay \ [ti \ Marso].
\textsc{prg.intr-plant} \ =1\textsc{pl} \ \textsc{obl} \ \textsc{rice} \ \textsc{obl.def} \ \textsc{March}
‘We will be planting rice in March.’ (arta0505)

(5) \textit{Man-di:madima} \ =te: \ =tid \ [ti \ talatalun \ =i] \ i \ be:kut \ =\textit{na}.
\textsc{intr-walk} \ =only \ =3\textsc{pl} \ \textsc{obl.def} \ \textsc{mountain}=\textsc{spc} \ \textsc{def} \ \textsc{ghost}=3\textsc{sg.gen}
‘Their ghosts will be walking in the mountains.’ (arta0111-06)

Pronominal indexes on predicates show the same pattern. The pronominal counterparts of the examples (1–3) are shown in (6–8). Note that prototypical transitive events encoded by actor-voice constructions, as in (8), are not permitted when the arguments are realized pronominally:

(6) \textit{T<in><um>addyor} \ =\textit{de}: \ =\textit{tid}.
\textsc{<pst><intr>stand} \ =\textsc{comp} \ =3\textsc{pl}
‘They stood up.’ (elicited)

(7) \textit{B<in>isag} \ =\textit{na} \ =d \ =\textit{tid}.
\textsc{<pst>break} \ =3\textsc{sg.gen} \ =\textsc{comp} \ =3\textsc{pl}
‘He/she broke them.’ (elicited)

(8) ??\textit{Nam-bisag} \ =\textit{de}: \ =\textit{tid} \ \textit{did}.
\textsc{pst.intr-break} \ =\textsc{comp} \ =3\textsc{pl} \ 3\textsc{pl.obl}
‘The old woman/women broke the bottle(s).’ (elicited)

These data provide a basis for determining S, A, P, and oblique in this language (see Kimoto 2017b for a more extensive discussion). Construction patterns (2, 7) are
regarded as transitive; thus, the actor and undergoer arguments are tagged as A and P, respectively. Construction patterns (1, 6) are regarded as intransitive; thus, the single core arguments are tagged as S. If the clause has two arguments, as in (3), the agentive argument is identified as S, and the undergoer-like argument is identified as oblique. The GRAID annotations are thus:

(9)  T<in><um>addyor=di [i babakat =i].
     <pst><intr>stand=comp def old.woman =spc
     # v:pred ln np.h:s rn
     ‘The old woman stood up.’ (elicited)

(10) B<in>isag=di [ni babakat =i] [i bo:te].
     <pst>break=comp gen.def old.woman =spc def bottle
     # v:pred ln np.h:a rn ln np.h:p
     ‘The old woman broke the bottle.’ (elicited)

(11) ?Nam-bisag =di [i babakat =i] [ti bo:te].
     pst.intr-break =comp def old.woman =spc obl.def bottle
     # v:pred other ln np.h:s rn ln np.h:obl
     ‘The old woman broke the bottle.’ (elicited)

Many Philippine-type languages, most notably those spoken in the Southern Philippine and in Sulawesi, appear to lack the ergative alignment (see, for example, Shibatani 1988 for Cebuano). In fact, Brickell (2016), on annotating the Tondano discourse in the GRAID system, rejects the ergative analysis, analyzing both actor-voice constructions and undergoer-voice constructions as transitive with A and P arguments within them. To differentiate both constructions, the tag of each argument includes the information of the predicate as in ⟨A_a⟩ (A argument of actor voice), ⟨:a_u⟩ (A argument of undergoer voice), ⟨:p_a⟩ (P argument of actor voice), ⟨p_u⟩ (P argument of undergoer voice). The annotation system of Tondano is incorporated in the present annotation on Arta to enable the comparison of Philippine-type languages that may or may not have clear ergative characteristics. Thus, the modified version of the annotation is shown below:

(12) T<in><um>addyor =di [i babakat =i].
     <pst><intr>stand =comp def old.woman =spc
     # v:pred rv ln np.h:s_a rn
     ‘The old woman stood up.’ (elicited)

(13) B<in>isag =di [ni babakat =i] [i bo:te].
     <pst>break =comp gen.def old.woman =spc def bottle
     # v:pred rv ln np.h:a_u rn ln np.h:p_u
     ‘The old woman broke the bottle.’ (elicited)
(14) Nam-bisag =di [i babakat =i] [ti bo:te].
PST.INTR-break =COMP DEF old.woman =SPC OBL.DEF bottle
## v:pred rv ln np.h:s.a rn ln np.h:obl.a
‘The old woman broke the bottle.’ (elicited)

We have so far used elicited examples for the sake of clarity of presentation. Below are actual discourse data annotated in which S, A, P, and oblique, location, and goal phrases are found:

(15) Pang-u:ser-ën didi ama =mi =ti ta dutul.
PRG-USE-TR PL.GEN.DEF father =1PL.GEN =SPC OBL first
## v:pred ln np.h:a.u =pro.1:poss rn ln other
‘Our fathers were using them in those days.’ (arta0002-43)

(16) Saya n-inta =ku ti bebbe: =m.
DEM.DIST <PST.TR>-see =1SG.GEN PSN aunt =2SG.GEN
## other v:pred =pro.1:a.u ln np.h:p.u =pro.2:poss
‘Then I met your aunt.’ (arta0601-45)

(17) Amma atti: um-angay =ti,
if exist INTR-GO =DEM.PROX.OBL
## other other:predex #rc gap.h:s.a v:pred =dem:g.a
d<um>aretyo =tid dën
<INTR>straight =3PL 1SG.OBL
% v:pred =pro.h:s.a pro.1:g.a
‘If there are people who are coming here, (that is, if) they come straightly to me,’ (arta0601-57)

(18) Atti: =ami ti Aglipay, wa,
exist =1PL OBL.DEF Aglipay PLH
## other:predex =pro.1:s.predex ln np:l.predex other
Disubu.
Disubu other
‘We were in Aglipay, whatchamacallit, Disubu.’ (arta0601-06)

The list and exposition of predicate tags are provided in Section 2.4.

Referential expressions may be located in the sentence-initial position, serving as a topic NP. Following the annotation rule, these NPs are glossed as ⟨dt⟩ ‘dislocated topic’. The nominals may be co-referential with a clause-internal argument, which is specified, for example, as ⟨dt.s⟩ (dislocated topic corresponding to the S function clause-internally), as shown in:

5 The location phrase found in (18) constitutes part of an existential construction, whose predicate atti: (‘exist’ other:predex) is the morphologically irregular verb in that it does not inflect for tense or aspect. The predicate is indexed as ⟨l.predex⟩ in this case. See Section 2.3 for further exposition.
2.2. Argument realization patterns

In Arta, an argument is normally realized by either a person index or a full NP; that is, the occurrences of person indexes and full NPs are distributed almost complimentarily. For example, in (21), plural third-person referents are encoded by the person index on the predicate without an independent nominal phrase, and, in (22), by an independent nominal without an index on the predicate:

(21) \[\text{Pab-bi:r`e-n} \quad =di \quad =d \quad tidi \quad a:h:a: \quad =di,\]
prg-search-tr \quad =3pl.gen \quad =comp \quad pl.def \quad children \quad =3pl.gen
# v:pred \quad =pro.h:a.u \quad rv \quad ln \quad np.h:p.u \quad =pro.h:poss
amma \quad nap-pati \quad =d \quad =tid.
if \quad pst.intr-die \quad =comp \quad =3pl
#ac \quad other \quad v:pred \quad rv \quad =pro.h:s:a

‘They are looking for their children if they died.’ (arta0111-25)

(22) \[\text{Nappati} \quad =d \quad [tidi \quad ama \quad =na \quad =y \quad aydi:}\]
pst.intr-die \quad =comp \quad pl.def \quad father \quad =3pl.gen \quad =spc \quad and
# v:pred \quad rv \quad ln \quad np.h:s:a \quad =pro.h:poss \quad rn \quad rn
ina \quad =na].
mother \quad =3sg.gen
rn \quad rn

‘His father and mother died.’ (arta0502-22)

These realizations are, according to Haig and Schnell (2014), tagged as \(<=\text{pro}>\) and \(<\text{np}>\) respectively. The equal sign \(<=\) shown in the first case indicates that it is an enclitic attaching to the predicate.

A careful examination reveals that another pattern is occasionally observed; this is the case in which the same role is encoded both by a person index and a full NP. In the example below, the actor role is doubly instantiated by the person index and the independent nominal phrase:
(23)  Saya  iggam-an  =na  a:yi:  [ni  kanakannak  =i].
        DEM.DIST  hold-TR  =3SG.GEN  DEM.PROX  GEN.DEF  child  =SPC
   ‘Then the child held this.’ (arta0110-114)

This type of argument encoding appears at a relatively low frequency. If this pattern is observed, it is annotated as ⟨bpi⟩ (bound person index) on the gloss of the NP, with the pronominal index left unglossed, as shown in (23).

2.3. Predicate types

This subsection examines the syntactic categories that may function as predicates. First, nouns, adjectives, and verbs can occupy the predicate slot in the language without any formal device such as a copulative formative. In this annotation, these predicates are tagged as ⟨np:pred⟩, ⟨adj:pred⟩, ⟨v:pred⟩ respectively. See the following examples:

(24)  ⟨np:pred⟩ (nominal predicates)
   a.  Saya  ina  =de:  Mulo,  ay  ti
       DEM.DIST  mother  =3PL.GEN  Mulo  PLH  PSN
       ##  other  np.h:dt_s  other  np.h:poss  0.h:s_np  other  ln
   Brida.
   Brida
   np:pred
   ‘So, as for Mulo’s mother, (she is) Brida.’ (arta0601-41)
   b.  siye:,  wa  =m,  kwartó  =m.
       DEM.PROX  PLH  =2SG.GEN  money  =2SG.GEN
       ##  dem:s_np  other  other  np:pred  =pro.2:poss
   ‘This is your what-cha-ma-call-it (placeholder), your money.’ (arta0601-94)

(25)  ⟨adj:pred⟩ (adjective predicates)
   a.  Asawa  =ku  =ti,  (...)  apitti
       spouse  =1SG.GEN  =SPC  short
       ##  np.h:dt_s  =pro.1:poss  rn  0.h:s_adj  adj:pred
       =te.
       =only
       other
   ‘As for my husband, (he) was short.’ (arta0502-02)
       ADJ-bad  DEF  habit  =1SG.GEN  =SPC
       ##  adj:pred  ln  np:s_adj  =pro.1:poss  rn
   ‘My habit was bad.’ (arta0601-74)

Existential predicates are one of the items that appear most frequently in the discourse. The positive existential is atti: ‘exist, there is’, and the negative counterpart
is awan ‘not exist, there is no’. They are tagged as ⟨other:predex⟩ in this annotation system, signaling a non-verbal class predicate functioning as existential (see Haig and Schnell 2014: 21). The negative existential awan should be followed by the obliquely-marked subject, annotated as ⟨:ncs_predex⟩:

(26) Positive existential predicate

a. Atti konta atti: ti ayti exist but exist obl.def dem.prox.obl
   # other:predex other other:predex ln ln
   Sinabagan.
   Sinabagan np:1_predex
   ‘They were (here), but they were here in Sinabagan.’ (arta0002-32)

b. Atti: =tep i gilangan =i ta exist =still def man =spc obl
   # other:predex other ln np:h:s_predex rn ln
   ayta Danak.
   dem.dist.obl Danak
   ln np:1_predex
   ‘The man was still there in Danak.’ (arta0106-15)

(27) Negative existential predicate

a. Awan ta Ilongot, awan ta not.exist obl Ilongot not.exist obl
   #neg other:predex ln np:ncs_predex other:predex ln
   agani:.
   non-Negrito np:ncs_predex
   ‘There were no Ilongot people (in that place); there were no non-Negrito people (there).’ (arta0601-24)

b. Ayta dutul, awan ta a:nus.
   obl first not.exist obl kindness
   #neg ln np:other other:predex ln np:ncs_predex
   ‘Before, there was no kindness (i.e., no one was kind to others).’ (arta0601-53)

Finally, nonfinite verb forms in Arta, defined as non-inflected verb forms for tense or aspect, may appear in some of the subordinate clauses. These are annotated as ⟨vother:pred⟩:

(28) ⟨vother:pred⟩ (nonfinite verb as a predicate)

a. Saya n-inta =ku a
   dem.dist <pst.tr>-see =1sg.gen lig
   # dem:p_u v:pred =pro.1:a_u rc gap:p_u other
K.imoto, Yukinori: Operationalizing Philippine-type syntax for GRAID system

\[ \text{pang-a:n-ën} = \text{mi.} \]
\[ \text{NMZ-eat-tr} = \text{1PL Gen} \]
\[ \text{vother:pred} = \text{pro.1:a u} \]

‘That is what I saw that we ate,’ (arta0002-01)

\[ \text{Sa:bit-én} = \text{di} = \text{d, \# ngay ti} \]
\[ \text{carry-tr} = \text{3PL Gen} = \text{comp \# go obl.def} \]
\[ \text{# 0:p u v:pred} = \text{pro.h:a u} \text{ other \# vother:pred ln} \]

b. \[ \text{bunbun} = \text{mi.} \]
\[ \text{house} = \text{1PL Gen} \]
\[ \text{np: other} = \text{pro.1:poss} \]

‘They carry it, going (back) to our house.’ (arta0002-62, 63)

2.4. Predicate tag on nominals

<table>
<thead>
<tr>
<th>tag</th>
<th>predicate class</th>
<th>voice</th>
<th>example</th>
<th>translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>⟨a⟩</td>
<td>dynamic verb</td>
<td>actor voice</td>
<td>man-lutu</td>
<td>‘cook’</td>
</tr>
<tr>
<td>⟨u⟩</td>
<td>dynamic verb</td>
<td>undergoer voice</td>
<td>i-lutu</td>
<td>‘be cooked’</td>
</tr>
<tr>
<td>⟨ap⟩</td>
<td>potentive verb</td>
<td>actor voice</td>
<td>maka-tim</td>
<td>‘can drink’</td>
</tr>
<tr>
<td>⟨up⟩</td>
<td>potentive verb</td>
<td>undergoer voice</td>
<td>ma-tim</td>
<td>‘can be drunk’</td>
</tr>
<tr>
<td>⟨stv⟩</td>
<td>stative verb</td>
<td>n.a.</td>
<td>tit-tuttud</td>
<td>‘be sitting’</td>
</tr>
<tr>
<td>⟨adj⟩</td>
<td>adjective</td>
<td>n.a.</td>
<td>mep-pullaw</td>
<td>‘be white’</td>
</tr>
<tr>
<td>⟨predex⟩</td>
<td>existential</td>
<td>n.a.</td>
<td>atti:</td>
<td>‘there is, exist’</td>
</tr>
<tr>
<td>⟨np⟩</td>
<td>nominal predicate</td>
<td>n.a.</td>
<td>buka:gan</td>
<td>‘be a woman’</td>
</tr>
</tbody>
</table>

Table 2 Dynamic, potentive, stative verbs, and adjectives

<table>
<thead>
<tr>
<th></th>
<th>dynamic verb class</th>
<th>potenti(ve) verb class</th>
<th>static(e) verb class</th>
<th>adjective class</th>
</tr>
</thead>
<tbody>
<tr>
<td>progressive</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>tense (non-past vs. past)</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>voice (actor vs. undergoer)</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>temporality implication</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>comparative construction</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>intensifying reduplication ‘very X’</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>

In Arta, almost every verb are marked in nature; i.e. morphological roots should take various kinds of verbal (and/or adjectival) affixes to formulate predicates (Table 1). Depending on the possible morphosyntactic behavior, these affixes fall into three verbal classes (dynamic verb, potentive verb, stative verb) and one adjective class. As
shown in Table 2, these predicate classes differ as to (i) whether the class may inflect for progressive, (ii) whether the class has tense distinction between non-past (or present) and past, (iii) whether the class has a productive voice alternation between actor and undergoer voices, (iv) whether the class has an implication of temporality, (v) whether the class can appear in the comparative construction, and (vi) whether the class has the reduplicative morphology signaling such intensification as ‘very, too X’, all of which are briefly summarized in Table 2 (see Kimoto 2017b for a further examination). Dynamic and potentive verb classes have voice distinction: actor voice and undergoer voice.

As mentioned in Section 2.1, GRAID annotations on Arta discourses include a cross-reference tag on nominals about the relevant predicate type. Dynamic verb class, a morphologically unmarked category, is tagged simply by (\texttt{\langle a\rangle}) (actor voice (of dynamic verb)) or (\texttt{\langle u\rangle}) (undergoer voice (of dynamic verb)), and potentive verb class, a morphologically marked category, is tagged either by (\texttt{\langle ap\rangle}) (actor voice of potentive verb) or (\texttt{\langle up\rangle}) (undergoer voice of potentive verb). Other predicate categories, which lack voice distinction, are specified simply as abbreviations as shown in Table 1.

(29) \texttt{\langle :s\_a\rangle} (S argument of a dynamic-verb predicate of actor voice)

\begin{verbatim}
Amma mam-purab tidi ama =mi =ti, if intr-hunt pl.def father =1pl.gen =spc
\end{verbatim}

`If our fathers go hunting’ (arta0002-06)

(30) \texttt{\langle :a\_u\rangle}/\texttt{\langle :p\_u\rangle} (A/P argument of a dynamic-verb predicate of undergoer voice)

\begin{verbatim}
Sa:bit-\`en =di =d, # ngay ti
carry-tr =3pl.gen =comp go obl.def
\end{verbatim}

`They carry it, going (back) to our house.’ (arta0002-62, 63)

(31) \texttt{\langle :s\_ap\rangle} (S argument of a potentive-verb predicate of actor voice)

\begin{verbatim}
Awan =tep maka-angay ta ayta lugar.
\end{verbatim}

`They could not come there to the place yet.’ (arta0002-33, 34)

(32) \texttt{\langle :s\_up\rangle} (S argument of a potentive-verb predicate of undergoer voice)

\begin{verbatim}
Saya na-pi:piya =d i pamilia =mi.
\end{verbatim}

`They'
'Then our family became better.' (arta0601-46)

(33) \(\langle:s_{\text{stv}}\rangle\) (S argument of a stative-verb predicate)

\[
\begin{align*}
\text{Tit-tékêk} & =a & ta & Dios. \\
\text{stv-wish} & =2sg & \text{obl} & \text{God} \\
\# & \text{v:pred} & =\text{pro:s_{stv}} & \ln & \text{np:obl_{stv}}
\end{align*}
\]

‘You should be praying to God.’ (arta0601-89)

(34) \(\langle:s_{\text{adj}}\rangle\) (S argument of an adjective predicate)

\[
\begin{align*}
\text{Med-dès} & i & \text{uga:li} & =k\text{u} & =y. \\
\text{adj-bad} & \text{def} & \text{habit} & =1sg,\text{gen} & =\text{spc} \\
\# & \text{adj:pred} & \ln & \text{np:s_{adj}} & =\text{pro.1:poss} & \text{rn}
\end{align*}
\]

‘My habit was bad.’ (arta0601-74)

(35) \(\langle:s_{\text{predex}}\rangle\) (S argument of an existential predicate)

\[
\begin{align*}
\text{Att}: & =\text{tep} & i & \text{gilangan} & =i & ta & ayta \\
\text{exist} & =\text{still} & \text{def} & \text{man} & =\text{spc} & \text{obl} & \text{dem.dist.obl} \\
\# & \text{other:predex} & \text{other} & \ln & \text{np.h:s_{predex}} & \text{rn} & \ln & \ln
\end{align*}
\]

Danak.

Danak

np:1_{\text{predex}}

‘The man was still there in Danak.’ (arta0106-15)

(36) \(\langle:s_{\text{np}}\rangle\) (S argument of a nominal predicate)

\[
\begin{align*}
\text{siye}: & \text{wa} =m & \text{kuvarto} & =m. \\
\text{dem.prox} & \text{plh} =2sg,\text{gen} & \text{money} & =2sg,\text{gen} \\
\# & \text{dem:s_{np}} & \text{other} & \text{other} & \text{np:pred} & =\text{pro.2:poss}
\end{align*}
\]

‘This is yours, your money.’ (arta0601-94)

3. Referential expressions

3.1. Nominal structure

When an independent referential expression is headed by a lexical noun, the noun should be preceded by a determiner that inflects for the number, case, and definiteness. The noun may be followed by a specificity marker, which signals that the referent is a specific object known to the speaker. A determiner is tagged as \(\langle 1n \rangle\), and a specificity marker is, when it appears after a noun, tagged as \(\langle \text{rn} \rangle\):\(^6\)

---

\(^6\) In GRAID annotations, \(\langle 1n \rangle\) is defined as “NP-internal subconstituent occurring to the left of NP head”, and \(\langle \text{rn} \rangle\) as “NP-internal subconstituent occurring to the right of NP head” (Haig and Schnell 2014: 9).
When a specificity marker appears within a noun phrase, it should occupy the slot immediately after the first lexical element. For example, when a modifier appears before a head noun, the specificity marker should no longer follow the noun but should be encliticized to the modifier:

(38) Tidi tallip =i a buka:gan, awan =tid
    pl.def two =spc lig woman neg =3pl
    ln ln ln np.h:dt.s other =pro.h:s.ap
    naka-panga:dal
    pst.pot-learn
    v:pred

    ‘As for the two women, they were not able to go to school.’ (arta0110-046)

Some optional elements may modify a head noun with the intervening connective a (ligature), as shown by the numeral quantifier in (38). Such elements within the nominal are also annotated as ⟨ln⟩ or ⟨rn⟩ based on the relative position to the head noun. The examples below illustrate the cases in which the adjectives, quantifiers, and/or demonstratives modify the head nouns.7

(39) (adjective+noun)

    Basta in-an’anu:s-an =mi =tèddi ay [ka:man =i
    just pst-tolerate-tr =1pl.gen =only big =spc
    ln ln ln np.p:u

    ‘We just tolerated a big tent (instead of their own houses).’ (arta0007-08)

(40) (quantifier+adjective+noun)

    Um-angay =de: =tèn =ti, man
    intr-go comp =1sg dem.prox.oobl as.if
    v:pred other =pro.1:s.a =dem:g:a #ac 0.1:a_up other

7 It is sometimes difficult to determine which element is the head of the nominal among several words because alternative ordering of elements is possible in Arta, and, in fact, in Philippine languages in general.
na ne:but [attanan a med-dès a uga:li].
gen pst.pot.lose all lig adj-bad lig custom
other v:pred ln ln ln ln np:p_up

‘After I came here, it seems that (I) have lost every bad custom.’ (arta0601-55)

(41) (quantifier+noun), (demonstrative+adjective+noun)

ènsi:na di:sa:n =na [i gissa =y a lingo] [aynina a
so.that reach =3sg.gen def one =spc lig week dem.med lig
## other v:pred ln ln ln ln np:p_u ln ln me”a:du a bagbat].
adi-plenty lig rice
ln ln dpi_np:a_u

‘so that this plenty of rice will last for one week.’ (arta0515-107)

Note that the same constructional template [modifier a head] or [head a modifier]
is employed both for adjective modifications and relative clauses; i.e., both
of them could be described as instances of the single constructional template
“adnominal modification.” For the purpose of cross-linguistic comparisons with
non-Philippine-type languages, adnominal modifications exclusively by means of
adjectives, quantifiers, and demonstratives are treated as ⟨⟩ or ⟨⟩, while adnominal
modifications by means of verbs (or more precisely verb-headed clauses) are treated as
relative clauses and annotated separately as ⟨#rc⟩, which will be noted in Section 4.

3.2. Possessive construction

Possessive constructions, composed of the possessed (or possessum) and the
possessor, are structurally parallel to argument realizations composed of a transitive
verb and an A argument in that both the possessor and the A argument receive a genitive
marking. The possessor may be encoded by (i) a bound person form on the possessed
noun, (ii) a (bound) demonstrative form, (iii) a full NP, (iv) a demonstrative form with a
bound person index on the possessed noun, or (v) a full NP with a bound person index
on the possessed noun. In this annotation, the possessive constructions are tagged as ⟨:poss⟩, and those different formal strategies are glossed as follows:

<table>
<thead>
<tr>
<th>tag</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(=pro:poss)</td>
<td>bound person form</td>
</tr>
<tr>
<td>(=(=)dem:poss)</td>
<td>(bound) demonstrative form</td>
</tr>
<tr>
<td>(np:poss)</td>
<td>full NP</td>
</tr>
<tr>
<td>(bpi_dem:poss)</td>
<td>demonstrative form with bound person index</td>
</tr>
<tr>
<td>(bpi_np:poss)</td>
<td>full NP with bound person index</td>
</tr>
</tbody>
</table>
Following the glossing rules of argument realization patterns, if the possessor role is doubly marked by a bound person index as well as a demonstrative or a full NP, the bound person index is glossed as \(<\text{bpi.np:poss}\>). Some of the actual annotations of possessive constructions are shown below.

(42) \(<\text{pro:poss}\>) (the person index only)

\[
\text{Amma mam-purab tidi ama } =\text{mi } =\text{ti},
\]

if \text{INTR-hunt} PL.\text{DEF} father \text{=} 1\text{PL.\text{GEN}} =\text{SPC}  

\(\text{# # ac other v:pred ln np.h:s.a } =\text{pro.1:poss} \text{ rn}\)

‘If our fathers go hunting’ (arta0002-06)

(43) \(<\text{np:poss}\>) (the genitive NP only)

\[
\text{Punan } =\text{na a} \text{yni babakat } =\text{i, a}
\]

say \text{=} 3\text{SG.\text{GEN}} GEN.\text{DEF} old.\text{woman} \text{=} \text{SPC \text{LIG}}  

\(\text{# # 0:p.u v:pred ln } =\text{pro.np.h:a.u} \text{ rn} \text{ rn}\)

\text{ina ni buka:gan } =\text{i.}

\text{mother GEN.\text{DEF} woman } =\text{SPC}

\text{np.h:appos ln np.h:poss} \text{ rn}

‘this old woman, (who is) the mother of the woman, said it.’ (arta0106-25)

(44) \(<\text{dem:poss}\>) (bound demonstrative form) and \(<\text{dem:poss}\>) (demonstrative NP)

\[
\text{Satidi:na ne:nan } =\text{mi } =\text{ta a pare:ho}
\]

\text{DEM.\text{MED} PST-go-TR } =\text{1PL.\text{GEN} =DEM.\text{DIST.OBL LIG same}}  

\(\text{# other v:pred } =\text{pro.1:a.u} =\text{dem.g.u other other}\)

\(=\text{mi tidi:na arta aydi ina } =\text{ni, ama}\)

\(=\text{1PL.\text{GEN} DEM.\text{MED} person and mother } =\text{DEM.\text{PROX.GEN} father}\)

\(=\text{pro.1:poss ln np.h:p.u} \text{ rn} \text{ rn } =\text{dem.h:poss} \text{ rn}\)

\text{ni ayni. GEN.\text{DEF DEM.\text{PROX.GEN}}}

\text{ln dem.h:poss}

‘We went to those people and to the mother of this (guy) and the father of this (guy).’ (arta0114-052)

(45) \(<\text{bpi.np:poss}\>) (a demonstrative form with a bound person index)

\[
\text{Ti Lenin ama } =\text{na ni ayni.}
\]

\text{PSN (personal.name) father } =\text{3SG.\text{GEN GEN.\text{DEF DEM.\text{PROX.GEN}}}  

\(\text{# ln np:pred np.h:s.np} \text{ ln } \text{bpi.np.h:poss}\)

‘The father of this (guy) is Lenin.’ (arta0114-053)

(46) \(<\text{bpi.np:poss}\>) (full NP with a bound person index)

\[
\text{konta ay ngadin } =\text{na } =\text{te } /\text{ni} \text{ wanga:r}
\]

\text{but name } =\text{3SG.\text{GEN =only GEN.\text{DEF stream}}  

\(\text{# other other np:s.np} \text{ rn ln } =\text{bpi.np:poss}\)
But as for the name of the stream, it is Dikerawyan.’ (arta0114-049)

4. Complex sentence: the treatment of gap constructions

The structure of complex sentences is relatively simple, so it is unproblematic to apply the GRAID annotation rules to our data. However, the treatment of gap constructions employed for relative clauses is worth noting. Consider the following excerpt from a discourse, in which the complex nominal phrase is headed by the head noun \textit{ka:huy} ‘sweet potato’ and immediately followed by a relative clause \textit{a nimulamula=mi} ‘that we planted’:

\begin{verbatim}
(...)
  (47) /* #ac
   Maski
   adin
   [S ARGUMENT: i e:n-an =mu
   even where
   def go-tr =2sg.gen
   ## #ac other
   np:pred #rc:s np gap:p_u ln v:pred =pro.2:a_u
   =y],
   =spc
   rn
   'Wherever you go’ (lit. ‘Even where the place you go to is’) (arta0601-90)
\end{verbatim}

Although the transitive verb \textit{nimulamula} ‘planted’ within the relative clause creates the expectation for two arguments to occur, the undergoer argument cannot appear within the embedded clause. This is not because the argument is pragmatically inferable but because the construction does not allow overt appearance of the argument. In GRAID annotations, the zero realization caused by grammatical suppressions should be specified differently from the one caused by pragmatic conditions. It is thus specified as \textlangle\textit{gap:p_u}\textrangle, implying a gap argument with \textbf{no pronominal index} within a relative clause, in which the gap functions as the \textbf{P argument of the undergoer voice} of a dynamic verb.

One of the prominent features of Philippine languages is that a relative clause itself may function as an argument of another clause (i.e. the headless relative clause). As illustrated in the examples below, the relative clauses serve as the \textit{S argument} (48), the \textit{A argument} (49), the \textit{P argument} (50), the oblique argument (51), and the predicate (52) of the higher-order clauses respectively:

\begin{verbatim}
(48) /* #ac
   Maski
   adin
   [S ARGUMENT: i e:n-an =mu
   even where
   def go-tr =2sg.gen
   ## #ac other
   np:pred #rc:s np gap:p_u ln v:pred =pro.2:a_u
   =y],
   =spc
   rn
   'Wherever you go’ (lit. ‘Even where the place you go to is’) (arta0601-90)
\end{verbatim}
In this GRAID annotation, this type of headless relative clause in Arta is differentiated by specifying the grammatical relation with the predicate type after ⟨#rc⟩, as in ⟨#rc:a_u⟩ (the relative clause functions as the A argument of the undergoer voice of a dynamic verb). This annotation is exactly the same as that of other simpler referential expressions, which makes it possible to compare the distributions of virtually all kinds of referential strategies in the language.

5. Concluding remarks

In this short paper, the annotation notes were provided for implementing the general GRAID glossing system (Haig and Schnell 2014) for the particular purpose of annotating Arta discourse. First, some exposition was made concerning the case marking system and argument structure, which have been problematic in Philippine languages and necessary for identifying S, A, and P roles in this glossing system. Different realization patterns of referential expressions in this language were also
documented, and a particular way to annotate them was proposed. A predicate tag on nominals is added to this annotation for comparing it with other Philippine-type languages with particular reference to the glossing convention in Tondano. Some exposition was also given of complex sentences, especially of gap constructions, which is crucial for understanding relative clauses and complex nominal structure.

**Abbreviations**

<table>
<thead>
<tr>
<th>1</th>
<th>first person</th>
<th>NMZ</th>
<th>nominalizer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1+2</td>
<td>first-second person</td>
<td>O</td>
<td>transitive object</td>
</tr>
<tr>
<td>2</td>
<td>second person</td>
<td>OBL</td>
<td>oblique case</td>
</tr>
<tr>
<td>3</td>
<td>third person</td>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>A</td>
<td>transitive subject</td>
<td>PLH</td>
<td>placeholder (filler-like element)</td>
</tr>
<tr>
<td>ABS</td>
<td>absolutive</td>
<td>POT</td>
<td>potentive verb</td>
</tr>
<tr>
<td>ADJ</td>
<td>adjective</td>
<td>PRG</td>
<td>progressive aspect</td>
</tr>
<tr>
<td>COMP</td>
<td>posterior phase ‘already’</td>
<td>PROX</td>
<td>proximal (demonstrative)</td>
</tr>
<tr>
<td>DEF</td>
<td>definite</td>
<td>PSN</td>
<td>personal nominal</td>
</tr>
<tr>
<td>DIST</td>
<td>distal (demonstrative)</td>
<td>PST</td>
<td>past tense</td>
</tr>
<tr>
<td>GEN</td>
<td>genitive case</td>
<td>S</td>
<td>intransitive subject</td>
</tr>
<tr>
<td>INTR</td>
<td>intransitive verb</td>
<td>SG</td>
<td>singular</td>
</tr>
<tr>
<td>LIG</td>
<td>ligature (linker)</td>
<td>SPC</td>
<td>nominal specificity</td>
</tr>
<tr>
<td>MED</td>
<td>medial (demonstrative)</td>
<td>STV</td>
<td>stative verb</td>
</tr>
<tr>
<td>NEG</td>
<td>negation</td>
<td>TR</td>
<td>transitive verb</td>
</tr>
</tbody>
</table>

**Appendix: Person forms, determiners, and demonstratives**

The following tables show the paradigm of person forms (including enclitic forms and independent pronouns), determiners, and demonstratives (including enclitic and independent forms). Note that in the grammatical glosses in the second line of each example, “absolutive” and “singular” are omitted for the sake of simplicity.
Table A  Person forms

<table>
<thead>
<tr>
<th>PERSON</th>
<th>TOPICAL</th>
<th>ABSOLUTIVE</th>
<th>GENITIVE</th>
<th>OBLIQUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>tən</td>
<td>=tən</td>
<td>=ku</td>
<td>dən</td>
</tr>
<tr>
<td>1PL</td>
<td>tami</td>
<td>=ami</td>
<td>=mi</td>
<td>dami</td>
</tr>
<tr>
<td>2SG</td>
<td>taw</td>
<td>=a, =taw</td>
<td>=mu</td>
<td>daw</td>
</tr>
<tr>
<td>2PL</td>
<td>tam</td>
<td>=am</td>
<td>=muyu</td>
<td>dam</td>
</tr>
<tr>
<td>1+2SG</td>
<td>tita</td>
<td>=ita</td>
<td>=ta</td>
<td>dita</td>
</tr>
<tr>
<td>1+2PL</td>
<td>titam</td>
<td>=itam</td>
<td>=tam</td>
<td>ditam</td>
</tr>
<tr>
<td>3SG</td>
<td>siya</td>
<td>Ø</td>
<td>=na</td>
<td>dya</td>
</tr>
<tr>
<td>3PL</td>
<td>tidi</td>
<td>=tid</td>
<td>=di</td>
<td>did</td>
</tr>
</tbody>
</table>

Note: “1+2SG/PL” is the person category “we” which includes “both speaker and hearer (and other(s) in plural).” It differs from “1PL” which includes “speaker and other(s), not hearer,” or “2PL” which includes “hearer and others (not speaker).”

Table B  Determiners

<table>
<thead>
<tr>
<th>ABSOLUTIVE</th>
<th>GENITIVE</th>
<th>OBLIQUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>INDEFINITE</td>
<td>Ø</td>
<td>na</td>
</tr>
<tr>
<td>DEFINITE</td>
<td>SINGULAR</td>
<td>(COMMON) i</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(PERSONAL) ti</td>
</tr>
<tr>
<td>PLURAL</td>
<td>tidi</td>
<td>didi</td>
</tr>
</tbody>
</table>

Table C  Demonstratives

<table>
<thead>
<tr>
<th>TOPIAL</th>
<th>ABS</th>
<th>GEN/ERG</th>
<th>OBL</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROXIMAL</td>
<td>SG</td>
<td>si:yəy</td>
<td>a:yi:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>=i</td>
<td>=ni</td>
</tr>
<tr>
<td></td>
<td>PL</td>
<td>satidi:</td>
<td>(ay)tididi a:yi:</td>
</tr>
<tr>
<td>MEDIAL</td>
<td>SG</td>
<td>sayna</td>
<td>a:yina,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>=ina</td>
<td>=nina</td>
</tr>
<tr>
<td></td>
<td>PL</td>
<td>satidi:na</td>
<td>(ay)tididi:na</td>
</tr>
<tr>
<td>DISTAL</td>
<td>SG</td>
<td>saya</td>
<td>a:ya:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>=ya:</td>
<td>=nə</td>
</tr>
<tr>
<td></td>
<td>PL</td>
<td>satiddya:</td>
<td>(ay)tiddiya</td>
</tr>
</tbody>
</table>

References


