The Multi-CAST collection has been archived at the University of Bamberg, Germany, and is freely accessible online at multicast.aspra.uni-bamberg.de/.

The entirety of Multi-CAST, including this document, is published under the Creative Commons Attribution 4.0 International Licence (CC BY 4.0), unless noted otherwise. The licence can be reviewed online at creativecommons.org/licenses/by/4.0/.

Multi-CAST Tondano annotation notes v1.1 last updated 8 May 2019
This document was typeset by NNS with Xpackages and the multicast3 class (v3.2.1).
# Contents

1. **Notes on the GRAID annotations** .................................................................................. 1
   1.1 Symmetrical voice system .................................................................................. 1
   1.2 Clitic pronouns ................................................................................................. 2
   1.3 Demonstratives .................................................................................................. 2
   1.4 Numerals ............................................................................................................ 3

**References** .................................................................................................................. 3

**Appendices** ................................................................................................................ 4
   A List of corpus-specific GRAID symbols ............................................................... 4
   B List of abbreviated morphological glosses ......................................................... 5
Notes on the GRAID annotations

This document contains notes on the implementation of the GRAID annotation conventions (Haig & Schnell 2014) in the Multi-CAST Tondano corpus. It corresponds to version 1905 of the annotations, published in May 2019. Unless a more recent version of this document exists, it also applies to any later versions of the annotations.

1.1 Symmetrical voice system

Tondano is a so-called Philippine-type language (Ross 2002; Himmelmann 2005; Blust 2013) that has a symmetrical voice system (Himmelmann 2005; Foley 2007; Riesberg 2014; Riesberg & Primus 2015). This system is realised morphologically via a paradigm of obligatory verbal affixes which mark the verb for either actor voice (AV) or one of three undergoer voices (UV), namely patient voice (PV), location voice (LV), and conveyance voice (CV). The voice marking indicates the semantic macro-role of the so-called “nominative” argument, which is the privileged pivotal syntactic argument. Crucially, both AV and UV constructions are transitive, which precludes application of standard GRAID conventions to the identification of A and P (see Haig & Schnell 2014: 13). We therefore applied a two-step approach to the glossing of S, A, and P in Tondano, as follows.

We first distinguish between intransitive and transitive clauses on constructional grounds, glossing a the single nominative argument of an intransitive clause as S. In transitive clauses, we identify either of the two core arguments as A or P according to their semantic macro-role, indicated also by the voice marking. Thus, we identify the nominative argument in an actor voice construction as A and the so-called “accusative” argument (a post-verbal NP without case marking) as P. We then combine the glossing of A and P with a gloss for the respective voice construction these occur in, thus yielding ⟨:a_a⟩ for an A argument in an actor voice construction, and ⟨:p_a⟩ for a P argument in an actor voice construction, as in (1):

(1) komèdomoulaniaku
   ko= <um>èdo =mow =la niaku
2SG.NOM= <AV>take =COMP =DIR.PROX 1SG
# =pro.2: a_a v:pred =rv =rv pro.1: p_a
   ‘You would take me away.’

In undergoer voice constructions, we identify the nominative as the P argument and the so-called genitive argument (a post-verbal clitic or NP marked for genitive case) as A. Thus, the function gloss for a nominative P argument in an undergoer-voice construction is ⟨:p_u⟩, and that of a genitive A argument in an undergoer-voice construction is ⟨:a_u⟩, as shown in examples (2) and (4):

(2) èpaketorenamou
   sè= pa-ketor-en =na =mow
3PL.NOM= DYN-slice-PV =3SG.GEN =COMP
# =pro:p_u v:pred =pro.h:a_u =rv
   ‘He slices them up.’

[mc_tondano_kiniar03_0038]
S, A, and P are clearly distinct from oblique arguments, which receive flagging by prepositions. All oblique arguments are glossed according to the standard GRAID conventions.

1.2 Clitic pronouns

In Tondano arguments may be realised as NPs, proforms, or zero anaphora. In transitive clauses which contain NPs or independent personal pronouns, the identification of an argument’s syntactic function is primarily achieved through word order. In transitive clauses which contain bound pronominals, such as (1) and (4) above, the type of clitic pronominal which is used indicates the syntactic function. In both AV and UV marked clauses the pre-verbal proclitic expresses the pivot, while in UV marked clauses enclitics always express the non-pivot actor argument. The Tondano GRAID annotation for pronominal forms does not differentiate between those with and without specific syntactic function. However, the notation used to express clitics, ⟨⟩, when used with pronominal forms, denotes that this form has a specific syntactic function.

1.3 Demonstratives

Demonstratives in Tondano may function as modifiers within NPs or as proforms which have full argument status. In cases where demonstratives have the former function the standard GRAID notation of ⟨rn⟩ is used. When demonstratives functions as arguments they are annotated in the same way as other independent pronouns, that is as ⟨pro⟩. The following two examples show demonstratives as modifiers (5) and demonstratives as pronominal arguments (6):

(5)  kina’atoan nêtu’a rior entimpa’ ye’i
    k<in>a-ato-an nê= tu’a rior N= timpa’ ye’i
    POT.PST-look-LV AN.PL.GEN= fast INAN= sap DEM1
    # v:pred =ln np.h:a_u rn =ln sap rn
    ‘The elders from before saw that palm sugar sap.’ [mc_tondano_kiniar02_0096]

(6)  linungala ni’tu
    l<in>unga-en =la ni’tu
    <PST>firewood-PV 0 =DIR.PROX DEM2
    # v:pred =DIR.PROX h:a_u =rv pro:p_u
    ‘They have lit that [the fire].’ [mc_tondano_kiniar01_0096]
1.4 Numerals

The characteristics of Tondano numerals and their corresponding GRAID notation mirror that of the demonstratives. In their more common function as modifiers within an NP, numerals are annotated as either \( \langle \text{ln} \rangle \) or \( \langle \text{rn} \rangle \) depending on their position in the phrase. In some cases numerals occur without the presence of a head noun and express a previously identified participant. In this situation they are considered proforms and are labelled as \( \langle \text{pro} \rangle \).

Examples of the GRAID annotation of numerals in different functions are as follows:

(7) \[ \text{paèdo} \text{nituama} \text{esa} \]
\[
\begin{array}{llllll}
\text{pa-ðo-en} & \Rightarrow & \text{moew} & \text{ni=} & \text{tuama} & \text{esa} \\
\text{0} & \text{DYN-take-PV} & \Rightarrow & \text{COMP} & \text{AN.SG.GEN=} & \text{man one} \\
\# & \text{0:p_u v:pred} & \Rightarrow & \text{rv} & \text{ln} & \text{np:h:a_u} \\
\end{array}
\]

‘The first man takes (the bats).’

(8) \[ \text{rua nèy èdo} \]
\[
\begin{array}{llllll}
\text{rua} & \text{nèy} & \text{èdo} \\
\text{two} & \text{cv.pst} & \text{take 0} \\
\# & \text{np:p_u} & \text{lv} & \text{v:pred} & \text{0.h:a_u} \\
\end{array}
\]

‘(He) took two [of them].’

References


Appendices

A List of corpus-specific GRAID symbols

The following is a list of the non-standard GRAID symbols used in the annotation of the Multi-CAST Tondano corpus. Please refer to the GRAID manual (Haig & Schnell 2014: 54–55) for an inventory of basic GRAID symbols.

Form symbols and specifiers

⟨pro⟩ independent pronominal; also includes demonstratives and numerals in addition to its basic definition

Function symbols and specifiers

⟨: a_a⟩ A argument in a clause marked with actor voice
⟨: a_u⟩ A argument in a clause marked with a type of undergoer voice
⟨: p_a⟩ P argument in a clause marked with actor voice
⟨: p_u⟩ P argument in a clause marked with a type of undergoer voice
⟨_ds⟩ specifier: subject of a verb of speech; attaches to ⟨: s⟩ and ⟨: a⟩
## B List of abbreviated morphological glosses

<table>
<thead>
<tr>
<th>Code</th>
<th>Meaning</th>
<th>Code</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>first person</td>
<td>LV</td>
<td>location voice</td>
</tr>
<tr>
<td>2</td>
<td>second person</td>
<td>MANN</td>
<td>manner</td>
</tr>
<tr>
<td>3</td>
<td>third person</td>
<td>MED</td>
<td>medial</td>
</tr>
<tr>
<td>AFF</td>
<td>affirmative</td>
<td>MULT</td>
<td>multiple</td>
</tr>
<tr>
<td>AUX</td>
<td>auxiliary</td>
<td>NEG</td>
<td>negation</td>
</tr>
<tr>
<td>AV</td>
<td>actor voice</td>
<td>NOM</td>
<td>nominative</td>
</tr>
<tr>
<td>CAUS</td>
<td>causative</td>
<td>ORD</td>
<td>ordinal</td>
</tr>
<tr>
<td>CMP</td>
<td>completive</td>
<td>PART</td>
<td>partitive</td>
</tr>
<tr>
<td>COMP</td>
<td>completive aspect</td>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>CV</td>
<td>conveyance voice</td>
<td>PN</td>
<td>proper name</td>
</tr>
<tr>
<td>DEM</td>
<td>demonstrative</td>
<td>POT</td>
<td>potentiive</td>
</tr>
<tr>
<td>DIR</td>
<td>directional</td>
<td>PREP</td>
<td>preposition</td>
</tr>
<tr>
<td>DIST</td>
<td>distal</td>
<td>PROX</td>
<td>proximal</td>
</tr>
<tr>
<td>DYN</td>
<td>dynamic aktionsart</td>
<td>PST</td>
<td>past</td>
</tr>
<tr>
<td>EX</td>
<td>exclusive</td>
<td>PV</td>
<td>patient voice</td>
</tr>
<tr>
<td>EXIST</td>
<td>existential</td>
<td>RDP</td>
<td>reduplication</td>
</tr>
<tr>
<td>GEN</td>
<td>genitive</td>
<td>REL</td>
<td>relativizer</td>
</tr>
<tr>
<td>HES</td>
<td>hesitation</td>
<td>REQ</td>
<td>requesitive</td>
</tr>
<tr>
<td>IN</td>
<td>inclusive</td>
<td>SG</td>
<td>singular</td>
</tr>
<tr>
<td>INAN</td>
<td>inanimate</td>
<td>SPEC</td>
<td>specific</td>
</tr>
<tr>
<td>INCMP</td>
<td>incompletive (aspect)</td>
<td>STAT</td>
<td>stative</td>
</tr>
<tr>
<td>IRR</td>
<td>irrealis</td>
<td>NC</td>
<td>not classified</td>
</tr>
<tr>
<td>LIM</td>
<td>limitative?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Multilingual Corpus of Annotated Spoken Texts

multicast.aspra.uni-bamberg.de/