

Multi-CAST

Tondano *annotation notes*

Timothy Brickell

May 2019
v1.1



ARC CENTRE OF EXCELLENCE FOR
THE DYNAMICS OF LANGUAGE



Australian Government
Australian Research Council



University of Bamberg

DFG

Multi-CAST

*Multilingual Corpus of
Annotated Spoken Texts*

Citation for this document

Brickell, Timothy. 2019. Multi-CAST Tondano annotation notes. In Haig, Geoffrey & Schnell, Stefan (eds.), *Multi-CAST: Multilingual corpus of annotated spoken texts*. (multicast.aspra.uni-bamberg.de/) (date accessed)

Citation for the Multi-CAST collection

Haig, Geoffrey & Schnell, Stefan (eds.). 2015. *Multi-CAST: Multilingual corpus of annotated spoken texts*. (multicast.aspra.uni-bamberg.de/) (date accessed)

The Multi-CAST collection has been archived at the *University of Bamberg*, Germany, and is freely accessible online at multicast.aspra.uni-bamberg.de/.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 $\text{X}_{\text{L}}^{\text{A}}\text{T}_{\text{E}}\text{X}$ 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

1 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 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) *komèdomoula niaku*
- | | | | | |
|--------------|----------------------|-------------|------------|--------------|
| <i>ko=</i> | <i><um>èdo</i> | <i>=mow</i> | <i>=la</i> | <i>niaku</i> |
| 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.’ (observed/elicited)

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) *èpaketorenāmou*
- | | | | |
|------------|--------------------|------------|-------------|
| <i>sè=</i> | <i>pa-ketor-en</i> | <i>=na</i> | <i>=mow</i> |
| 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]

(3) *kotaanganèamou*

ko= taang-an =nèa =mow
 2SG.NOM= intercept-LV =3PL.GEN COMP
 # =pro.2:p_u v:pred pro.h:a_u =rv

‘They would stop you (at that particular place).’

[mc_tondano_mapalus_0019]

(4) *lia’ empamèanatè*

lia’ N= i-pa-mèan =na =itè
 ginger 3SG.NOM= CV-DYN-hit =3SG.GEN =LIM
 # np:dt_p_u =pro:p_u v:pred =pro.h:a_u =rv

‘The ginger, he just strikes it.’

[mc_tondano_kiniar03_0062]

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 function 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 nètu’a rior entimpa’ ye’i*

k<in>a-ato-an nè= tu’a rior N= timpa’ ye’i
 POT.PST-look-LV AN.PL.GEN= old fast INAN= sap DEM1
 # v:pred =1n np.h:a_u rn =1n 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 0.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 ⟨1n⟩ or ⟨rn⟩ 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 ⟨pro⟩. Examples of the GRAID annotation of numerals in different functions are as follows:

(7) *paèdonou nituama esa*

	<i>pa-èdo-en</i>	=moew	<i>ni=</i>	<i>tuama</i>	<i>esa</i>	
0	DYN-take-PV	=COMP	AN.SG.GEN=	man	one	
#	0:p_u	v:pred	=rv	=1n	np.h:a_u	rn

'The first man takes (the bats).'

[mc_tondano_kiniar01_0015]

(8) *rua nèi èdo*

	<i>rua</i>	<i>nèy</i>	<i>èdo</i>	
	two	CV.PST	take	0
#	np:p_u	lv	v:pred	0.h:a_u

'(He) took two [of them].'

[mc_tondano_water_0025]

References

- Blust, Robert. 2013. *The Austronesian languages*. Revised edition. Canberra: Australian National University. (<http://hdl.handle.net/1885/10191>) (Accessed 2016-05-19).
- Brickell, Timothy C. 2016. Multi-CAST Tondano. In Haig, Geoffrey & Schnell, Stefan (eds.), *Multi-CAST: Multilingual corpus of annotated spoken texts*. (<http://multicast.aspra.uni-bamberg.de/#tondano>) (Accessed 2019-03-08).
- Foley, William A. 2007. A typology of information packaging in the clause. In Shopen, Timothy (ed.), *Language typology and syntactic description*. Volume 1: Clause structure, 362–446. Cambridge: Cambridge University Press.
- Haig, Geoffrey & Schnell, Stefan. 2014. *Annotations using GRAID (Grammatical Relations and Animacy in Discourse): Introduction and guidelines for annotators (version 7.0)*. (<http://multicast.aspra.uni-bamberg.de/#annotations>) (Accessed 2019-03-08).
- Haig, Geoffrey & Schnell, Stefan (eds.). 2016. *Multi-CAST: Multilingual Corpus of Annotated Spoken Texts*. (<http://multicast.aspra.uni-bamberg.de/>) (Accessed 2019-03-08).
- Himmelman, Nikolaus P. 2005. The Austronesian languages of Asia and Madagascar: Typological characteristics. In Adelaar, Alexander K. & Himmelman, Nikolaus P. (eds.), *The Austronesian languages of Asia and Madagascar*, 110–181. London: Psychology Press.
- Riesberg, Sonja. 2014. *Symmetrical voice and linking in Western Austronesian languages*. Berlin: Mouton de Gruyter.
- Riesberg, Sonja & Primus, Beatrice. 2015. Agent prominence in symmetrical voice languages. *Language Typology and Universals (STUF)* 68(4). 551–564.
- Ross, Malcolm D. 2002. The history and transitivity of Western Austronesian voice and voice-marking. In Wouk, Fay & Ross, Malcolm D. (eds.), *The history and typology of Western Austronesian voice systems*, 17–62. Canberra: Australian National University.

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

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

Multi-CAST

Multilingual Corpus of Annotated Spoken Texts



multicast.aspra.uni-bamberg.de/