Multi-CAST

Tulil

annotation notes

Chenxi Meng

July 2019
v1.0

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

Meng, Chenxi. 2019. Multi-CAST Tulil 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The Multi-CAST collection has been archived at the *University of Bamberg*, Germany, and is freely accessible online at `multicast.aspra.uni-bamberg.de/`.

Multi-CAST Tulil annotation notes v1.0 last updated 5 July 2019
This document was typeset by NNS with XeLaTeX and the *multicast3* class (v3.1.9001).

Contents

1 Notes on the GRAID annotations

The following comprises selected notes on the GRAID (Haig & Schnell 2014) and RefIND (Schiborr et al. 2018) annotations of Tulil. It corresponds to version 1907 of the annotations, published in July 2019. Unless a more recent version of this document exists, it also applies to any later versions of the annotations.

1.1 Verbal clause types and marking of arguments

Verbal clauses can be classified in terms of the transitivity of the predicate, as either transitive, intransitive or ditransitive. Intransitive clause shows split-s marking, differentiating between active intransitive and stative intransitive predicates. Transitive impersonal clauses can be identified in the language, as a special case of the transitive clauses.

Structurally, verbal clauses normally either take A/S_A argument person markings on verbs and/or show patientive case on P/S_P arguments, depending on the subcategories of the verbs. The person indexes are obligatory on verbs, specify the person, number, gender of the argument, as well as tense. They are not specified in the GRAID annotation, but taken as part of the verbs, such as *və-* SG.M.PST- in (1).

(1) *vənua bəvətton vareot.*

	vənu	*=a*	*bə=*	*və-t~ton*	*vareot*
	sun	=SG.CL:MASC	ASP=	3SG.M.PST-RED~look	long_time
##	np:s	=rn	=other	v:pred	rv

'The sun was shining for a long time.' [mc_tulil_alrm_0010]

On the other hand, NPs and free pronouns as the indexed arguments are optional if they are recoverable from the context. In the case when there is no overt arguments in the clause, a zero-argument marking is assumed and placed at the beginning of the clause, as shown in (2):

(2) *təmət idə.*

		tə-mət	*idə*
	0	3PL.PST-eat	3N
##	0.h:a	v:pred	pro:p

'They ate them [the pigs].' [mc_tulil_sves_0006]

Usually a topic specifies the identity of the arguments in a clause, and is annotated as ⟨:dt_a⟩ or ⟨:dt_s⟩, with zero-marking of the subject (3). Topics are recognized by their overt marking with *o*.

(3) *tudəka o vəre təvai.*

		tudək	*=a*	*o*	*və-re*	*təvai*
	0	wallaby	=SG.CL:MASC	TOP	3SG.M.PST-carry	fern.sp.
##	0.h:a	np.d:dt_a	=rn	other	v:pred	pro:p

'The wallaby carried ferns.' [mc_tulil_lrdw_0010]

1.1.1 Transitive impersonal clauses

Transitive impersonal clauses also take person markings, but do not involve a substantial A argument. In a transitive impersonal construction, only the third person neuter prefixes *idə-* PST and *it-* NPST are used, and they do not refer to any arguments (dummy arguments). Thus the person indexes in transitive impersonal clauses only represent the tense of the predicate, but not person or gender. In example (4), the argument *kəguing* following the transitive impersonal verb *(idə)mang* 'burnt' is taken as the subject, and is annotated with ⟨np:s_iv⟩, where the specifier ⟨_iv⟩ stands for 'impersonal verb'.

(4) *bidəmmanga kəguing.*

	b=	*idə-m~mang*	*=a*	*kəguing*
	ASP=	3N.PST-RED~burn	=PAT	grass.sp.
##	other=	v:pred	=rv	np:s_iv

'Grass was burnt / dried out.' [mc_tulil_alrm_0022]

In the case of (5), there is no overt marking of a topic, so the argument before the verb complex *idil=a* is annotated as having the same function as the actual argument *=a* inside the verb complex structure. The same structure is observed in stative intransitive clauses, and is treated the same there.

(5) *idila bidəmat na da ...*

	idil	*=a*	*b=*	*idə-mat*	*n*	*=a*	*da*
	small	=SG.CL:MASC	ASP=	3N-get	APPL	=3SG.M	PURP
##	np.h:dt_s	=rn	=other	v:pred	rv	=pro.h:s_iv	# other

'The son wanted to …' [lit. 'it gets to him that...'] [mc_tulil_sves_0008]

1.1.2 Stative-intransitive clauses

Stative intransitive clauses take stative verbs as the heads of their predicates, which however are not marked by person indexes. The subject follows the verb, and is annotated as ⟨:s_sv⟩, where ⟨_sv⟩ stands for 'stative verb'.

(6) *toktoka kilap do.*

	toktok	*=a*	*kilap*	*do*
	end	=PAT	mud	here
##	v:pred	=ln	np:s_sv	other:l

'The mud is about this height.' [lit. 'the mud stops here.'] [mc_tulil_jkpp_0066]

As with (5) above, in stative-intransitive clauses such as (7) that lack overt topic marking, the argument before the verb complex, here *ianem*, is annotated as having the same function as the argument inside the verb complex structure, here *=idə*:

(7) *ianem bələlət nidə.*

	ianem	*bə=*	*lə~lət*	*n*	*=idə*
	lightning	ASP=	IPFV~arrive	APPL	=3N
##	np:s	=other	v:pred	rv	=pro:s_sv

'The lightning started [lit. arrived].' [mc_tulil_alrm_0083]

1.1.3 Patientive case

Patientive case shows on NPs as a preposed case marker *=a*, and causes personal pronouns to take their patientive forms (only for 3rd person singular). Patientive case extends its use to mark all non-A/S_A arguments, including: (i) objects (P) and obliques in transitive clauses and (ii) subjects of stative-intransitive clauses (S_P). The patientive case marker *=a* is phonologically detached from the marked argument and encliticized to the preceding element. It is annotated as ⟨=ln⟩. See (6) for an example.

1.1.4 Omission of P/S_P arguments

P/S_P arguments can only be omitted when they are 3rd person neuter pronouns (i.e. *idə*). The second clause in (8) takes a zero patient that could be expressed overtly as *idə*:

(8) *vətak be nangə idə itən, me kori ngatak bem.*

	və-tak	*be*	*nangə*	*idə*	*itən,*
0	3SG.M.PST-call	at	tree	3N	name
## 0.h:a	v:pred	rv	ln_np:poss	ln_pro:poss	np:p

		me	*kori*	*nga-tak*	*bem*
0	0	and	NEG	1SG.NPST-call	at
##neg 0.1:a	0:p	other	other	v:pred	rv

'He called the names of the herbs, and I'm not going to call (these names).'

[mc_tulil_alrm_0035]

1.1.5 Oblique arguments

Oblique arguments are primarily marked by prepositions. Like P arguments, oblique arguments also take patientive case, either marked by a patientive marker, or by the patientive form of a personal pronoun.

(9) *vəlata o vətuka kəvop na.*

	və=	*lat*	*=a*	*o*	*və-tuk*	*=a*
0	3SG.M.POSS.INAL=	head	=SG.CL:MASC	TOP	3SG.M.PST-give	=PAT
## 0.h:a	=ln_pro:poss	np:dt_obl	=rn	other	v:pred	=ln

kəvop	*n*	*=a*
dog	APPL	=3SG.M.PAT
np:p	adp	=pro:obl

'He gave his [the pig's] head to the dogs.' [mc_tulil_sves_0015]

Locative PPs usually receive the ⟨:l⟩ function gloss, as in (10). Arguments expressing goals are glossed ⟨:g⟩.

(10) *da nguteltel ba viuve.*

	da	*ngu-teltel*	*ba*	*viuv*	*=e*
0	PURP	1DL.NPST-wander	in	sea	=SG.CL:FEM
## 0.1:s	other	v:pred	adp	np:l	=rn

'[Dog said to wallaby,] "Let's wander in the sea."'

[mc_tulil_lrdw_0006]

Adverbial demonstratives are likewise glossed either as ⟨:l⟩ or ⟨:g⟩, as in (11). When a PP expressing location or goal is present after the demonstrative, it is the PP that receives the ⟨:l⟩ or ⟨:g⟩ gloss while the demonstrative does not receive a function gloss, as shown in (12).

(11) *duppi nəmumə, …*

	du-p~pi	*nə-mumə*
0	1PL.PST-RED~go	LOC-DOWN.FAR
## 0.1:s	v:pred	dem_other:g

'We went down there, …' [mc_tulil_jkpp_0066]

(12) *kavar vəkərtang o vəppi nəmumə təba məngəda.*

		kavar	*və-kərtang*		*o*		*və-p~pi*
	0	just_then	3SG.M.PST-arise		TOP	0	3SG.M.PST-IPFV~go
##	#ac 0.h:s	other	v:pred	%	other	0.h:s	v:pred

nə-mumə	*tə=*	*ba*	*məngəd*	*=a*
LOC-DOWN.FAR	to=	in	home	=SG.CL:MASC
dem_other	adp=	adp	np:g	=rn

'Just then he got up and he went down to the village.' [mc_tulil_alrm_0070]

1.2 Demonstratives

There are four types of demonstratives that are relevant to GRAID: adnominal demonstratives, pronominal demonstratives, predicative demonstratives, and locative adverbial demonstratives. The first three types typically have the same form (but different syntactic functions), and are glossed respectively as ⟨rn_dem⟩ (13), ⟨dem_pro⟩ (14), and ⟨dem_other(:pred)⟩ (15). Adverbial demonstratives are glossed as ⟨dem_other⟩ (16) and function variously as predicates or adjuncts.

(13) *atiriva abo o məreka.*

a-tiriv	*=a*	*a-bo*	*o*	*mərek=a*
ART-kulau	=SG.CL:MASC	3SG.M-UP	TOP	good=SG.CL.PAT
## np:s	=rn	rn_dem	other	np:pred

'This kulau I mentioned is good.' [mc_tulil_alrm_0171]

(14) *ava itira evi.*

	ava	*i-tir*	*=a*	*e-vi*
0	also	3SG.F.PST-ask	=PAT	3SG.F-PROX1
## 0.d:a	other	v:pred	=ln	dem_pro.d:p

'She [the turtle] also asked her [the shell].' [mc_tulil_lnsl_0014]

(15) *ləmat mukəm ipbo.*

ləmat	*mukəm*	*ip-bo*
coconut	two.CL:MASC	3DL.M-UP
## np:s	rn	dem_other:pred_l

'Two coconut trees are up there.' [mc_tulil_jkpp_0046]

(16) *ava ngang o nəbət konəng davə.*

	ava	*ngang*	*o*	*nə-bət*	*konəng*	*davə*
	also	0	1SG	TOP	LOC-PROX2	only	up
##	other	0.1:s	pro.1:dt	other	dem_other:pred	other	other

"'I just live up there."' [mc_tulil_lnsl_0009]

1.3 PPs as predicates

PPs can function as either non-verbal predicates (17) or verb-like predicates (18). In both cases, the complement NP of the preposition receives one of the function glosses ⟨:pred⟩, ⟨:pred_l⟩, or ⟨:pred_g⟩.

(17) *ngang o təpma iaor, ti voin.*

	ngang	*o*	*təpm*	*=a*	*iaor*	*ti*	*voin*
	1SG	TOP	with	=PAT	fear	lest	ghost.PL
##	pro.1:s	other	adp	=ln	np:pred	adp	np:obl

'I was with fear of ghosts.' [mc_tulil_jkpp_0021]

(18) *... dava ip katum itəba tipur.*

	da-va	*ip*	*katum*	*i-tə=*	*ba*	*tipur*
	PURP-again	3DL.M	all	3DL.M.NPST-GOAL=	in	bush
##	other	pro.h:s	rn	=rn_pro.h	adp	np:pred_g

'[The two wanted] to go to the bush again.' [mc_tulil_sves_0008]

1.4 Possessive NPs

In a possessive construction, the possessed noun, as the head of the possessive NP, is preceded by possessive indexes. Possessive indexes specify the number, person, and gender of the possessor. The possessor as a NP can either be present (19) or omitted (20). In the first case, both the possessor NP and the possessive pronoun receive the ⟨:poss⟩ function, while the the second case, only the possessive pronoun receives the glossing.

(19) *məlanga va kuvit idəkəreut do.*

	məlang	*=a*	*va*	*kuvit*	*idə-kəreut*	*do*
	lizard	=SG.CL:MASC	3SG.M.POSS.AL	claw.PL	3N.PST-scratch	here
##	ln_np.d:poss	=ln	ln_pro.d:poss	np:s	v:pred	other:l

'The lizard's claws scratched here.' [mc_tulil_lnsl_0039]

(20) *nga vove o bingəp.*

	nga	*vov*	*=e*	*o*	*b=*	*i-ngəp*
0	1SG.POSS	grandma	=SG.CL:FEM	TOP	ASP=	3SG.F.PST-die
##	0.h:s	ln_pro.1:poss	np.h:dt_s	=rn	other	=other	v:pred

'My aunt died.' [mc_tulil_jkpp_0030]

1.5 Comitatives

The comitative *n(ə)* "with" is usually prefixed by person markings, and is commonly used for coordinating two NPs. In this case, the person marking indexes the first conjunct. The comitative *n(ə)* is glossed as a part of the NP, ⟨rn⟩:

(21) *io vəre bəli tuka va dərnge ena ipvila.*

	io	*və-re*	*bəli*	*tuk*	*=a*	*va*
0	then	3SG.M.PST-carry	pig.PL	give/RECIP	=PAT	3SG.M.POSS
## 0.h:a	other	v:pred	np:p	adp	=ln	ln_pro.h:poss

dərng	*=e*	*e-n*	*=a*	*ip=*	*vil*	*=a*
old	=SG.CL:FEM	3SG.F-with	=PAT	3DL.M.POSS=	son	=SG.CL:MASC
np.h:g	=rn	rn	=rn	=rn_pro.h:poss	rn_np.h	=rn

'Then he carried pigs to his wife and their son.' [mc_tulil_sves_0005]

As with verbal and possessive constructions, the indexed conjunct can be omitted (22). In this case, the comitative receives the ⟨pro⟩ form gloss and argument function, and the complement is glossed as ⟨rn⟩:

(22) *ngana nga məlna avi o, ngunu kərtang məndo.*

	nga-n	*=a*	*nga*	*məln*	*=a*	*a-vi*
0	1SG-COM	=PAT	1SG.POSS	cousin	=SG.CL:MASC	3SG.M-PROX1
## 0.h:s	pro.h:dt_s	=rn	rn_pro.1:poss	rn_np.h	=rn	other:dt

o	*ngunu-kər*	*tang*	*mən=*	*do*
TOP	1DL.PST-rise	REFL	from=	here
other	v:pred	=pro.1:p	adp=	other:l

'Me with my cousin, we got up from here.' [mc_tulil_alrm_0005]

1.6 Quotatives

The quotative *-e* can be used either as a complementizer (23), or as a quotative predicator (24). In both cases it takes a person index. In the first case, the quotative *-e* is glossed as ⟨other⟩. When it is used as a predicate, it receives the glossing ⟨other:pred⟩, similar to a verb.

(23) *vəngar ate, ngikutəng mata go rovənəm.*

	v-əngar	*at-e,*		*ngi-kutəng*	*mat*	*=a*	*go*
0	3SG.M.PST-say	3SG.M-QUOT	0	2SG.NPST-cut	get	=PAT	NSPEC
## 0.h:s_ds	v:pred	other	##ds 0.2:a	v:pred	rv	=ln	ln

ro	*=vənəm.*
bamboo	=SG.CL:SHORT
np:p	=rn

'He said, "You cut a bamboo cutting".' [mc_tulil_alrm_0047]

(24) *ngane, o! laikta o təpma kəmak.*

	ngan-e,	*’o*	*laik*	*=ta*	*o*	*təpm*	*=a*	*kəmak*
0	1SG-QUOT	INTERJ	big	=PL:HUM	TOP	with	=PAT	lie
## 0.1:s	other:pred	##ds	other	np.h:s	=rn	other	adp	=ln	np:pred

‘I thought, “Oh! The ancestors were full of lies”.’ [mc_tulil_jkpp_0053]

References

Haig, Geoffrey & Schnell, Stefan. 2014. *Annotations using GRAID (Grammatical Relations and Animacy in Discourse): Introduction and guidelines for annotators (version 7.0)*. (https://multicast.aspra.uni-bamberg.de/#annotations) (Accessed 2019-03-08).

Haig, Geoffrey & Schnell, Stefan (eds.). 2016. *Multi-CAST: Multilingual Corpus of Annotated Spoken Texts*. (https://multicast.aspra.uni-bamberg.de/) (Accessed 2019-03-08).

Meng, Chenxi. 2018. *A grammar of Tulil*. Melbourne: La Trobe University Ph.D. dissertation.

Meng, Chenxi. 2019. Multi-CAST Tulil. In Haig, Geoffrey & Schnell, Stefan (eds.), *Multi-CAST: Multilingual corpus of annotated spoken texts*.

Schiborr, Nils N. & Schnell, Stefan & Thiele, Hanna. 2018. *RefIND — Referent Indexing in Natural-language Discourse: Annotation guidelines (v1.1)*. Bamberg / Melbourne: University of Bamberg / University of Melbourne. (https://multicast.aspra.uni-bamberg.de/#annotations) (Accessed 2019-03-08).

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 Tulil corpus. Please refer to the *GRAID manual* (Haig & Schnell 2014: 54–55) for an inventory of basic GRAID symbols.

Form symbols and specifiers

⟨cl_pro⟩	classifier used pronominally
⟨dem_pro⟩	demonstrative used pronominally
⟨pn_np⟩	proper name
⟨dem_other⟩	demonstrative used adverbially
⟨intrg_other⟩	interrogative form

Function symbols and specifiers

⟨:s_iv⟩	subject of a transitive impersonal clause
⟨:s_sv⟩	subject of a stative-intransitive clause

Clause boundary symbols

⟨purp⟩, ⟨_purp⟩ *tag:* purposive clause

Subconstituent symbols

⟨_dem⟩	demonstrative used adnominally; attaches to ⟨ln⟩ and ⟨rn⟩

Other symbols

⟨nc_⟩	*specifier:* marks form glosses with RefIND indices in segments otherwise not considered (i.e. those marked ⟨#nc⟩)

B List of abbreviated morphological glosses

1	first person
2	second person
3	third person
AL	alienable possession
APPL	applicative
ART	article (borrowed from Tolai)
ASP	aspect
AUG	augmentative (class marking)
CL	class marking
CLU	cluster (class marking)
COM	comitative
DIM	diminutive
DL	dual
F	feminine
FEM	feminine (class marking)
FLAT	flat (class marking)
HUM	human (class marking)
IPFV	imperfective
INAL	inalienable possession
INSTR	instrumental
INTERJ	interjection
LOC	locative
LON	long (class marking)
M	masculine
MASC	masculine (class marking)
N	neuter
NEG	negation
NHUM	non-human
NPST	non-past
NSPEC	non-specific
PAT	patientive
PIECE	piece (class marking)
PL	plural
PN	proper noun
POSS	possessive pronoun
PROX1	speaker proximal
PROX2	hearer proximal
PST	past
PURP	purpose
QUOT	quotative
REFL	reflexive
SEG	segment (class marking)
SG	singular
SHORT	short (class marking)
TOP	topic marking
NC	not classified

Multi-CAST

Multilingual Corpus of Annotated Spoken Texts

multicast.aspra.uni-bamberg.de/