Multi-CAST

Arta
annotation notes

Yukinori Kimoto

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

Kimoto, Yukinori. 2021. Multi-CAST Arta annotation notes. In Haig, Geoffrey & Schnell,
Stefan (eds.), *Multi-CAST: Multilingual corpus of annotated spoken texts.*
(`multicast.aspra.uni-bamberg.de/#arta`) (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)

Contents

1   Notes on the GRAID annotations

This document is a revised reproduction of Kimoto (2018).[1] It collects selected notes on the GRAID (Haig & Schnell 2014) annotation conventions in the Multi-CAST Arta corpus, corresponding to version `2101` of the annotations, published in January 2021. Unless a more recent version of this document exists, it also applies to any later versions of the annotations.

Arta is a Philippine language spoken by the Philippine Negrito people inhabiting the Quirino and Aurora provinces in Northeastern Luzon, the Philippines. The language was first reported on by Reid (1989) and is currently being documented by the author (see Kimoto 2014; 2017b; a).

1.1   Clause structure and grammatical relations

In GRAID, annotators employ symbols such as ⟨:s⟩, ⟨:a⟩, ⟨:p⟩, and ⟨:obl⟩ to identify arguments. 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.

In Arta, verbs may show voice alternation to confer absolutive/nominative on different arguments. The availability of voice alternation differs with respect to the valency that the given verbs have. First, monovalent verb stems such as *taddyor* 'to stand up', *pati* 'to die', and *tubu* 'to mature (of plant)' take so-called actor voice by default; therefore, the single argument is given absolutive case. The verb *t<in><um>addyor* in (1), for example, the infix *<um>* (the actor voice affix) is responsible for the actor voice construction. Actor voice may be constructed by different affixes such as *maN-* (*N* indicates nasal assimilation), *maC-* (*C* indicates consonant gemination), *ma-*, and ⟨<um>⟩.

(1)   *Tinumaddyor ti Sanuwa:tèng.*

T<in><um>addyor	*ti*	*Sanuwa:tèng*
<PST><AV>stand	PSN	Sanuwateng
v:pred	ln	np.h:s_a

'Sanuwateng stood up.'

In the Arta GRAID annotations, the absolutive NP in this case is glossed ⟨:s_a⟩, indicating an 'S argument in Actor voice clause'.

Bivalent verb stems such as *papati* 'killing' in (2) and (3) show voice alternation between actor voice (AV) and undergoer voice (UV). The actor voice marks the actor role NP with the absolutive, whereas the undergoer voice marks the undergoer role NP in the clause as absolutive. Note that "undergoer voice" is an umbrella term covering patient voice (PV, *-èn*), location voice (LV, *-an*) and conveyance voice (CV, *i-*), and the specific choice of affix in the undergoer voice is not fully predictable.

(2)   *Mampapati ti Sanuwatèng ta arta.*

mam-papati	*ti*	*Sanuwatèng*	*ta*	*arta*
AV-killing	PSN	Sanuwateng	OBL	person
v:pred	ln	np.h:a_a	ln	np.h:p_a

'Sanuwateng will kill a person/people.'

---

1 Kimoto, Yukinori.   2018.   *Operationalizing Philippine-type syntax for GRAID system: Clause structure, case marking, and verb class in Arta.*   Asian and African Languages and Linguistics *12.   17–35.* (http://hdl.handle.net/10108/91147).

verb	participant
AV (*maN-*, *maC-*, *ma-*, *<um>*	⟨:s_a⟩ (ABS)

Table 1   Voice and case marking in monovalent clauses

verb	actor	undergoer
AV (*um*, *maC-*, *maN-*	⟨:a_a⟩ (ABS)	⟨:p_a⟩ (OBL)
UV (PV *-en*, LV *-an*, CV *i-*)	⟨:a_u⟩ (GEN)	⟨:p_u⟩ (ABS)

Table 2   Voice and case marking in bivalent clauses

(3)   *Papatièn ni Sanuwatèng i artay.*

papati-èn	*ni*	*Sanuwatèng*	*i*	*arta*	*=y*
killingPV	GEN.DEF	Sanuwateng	DEF	person	=SPC
v:pred	ln	np.h:a_u	ln	np.h:p_u	=rn

'Sanuwateng will kill the person/people.'

In the actor voice, the undergoer role is marked as oblique (OBL), and in the undergoer voice, the actor role is marked as genitive (GEN), as shown in Table 2. The two arguments in the actor voice in (2) are tagged as ⟨:a_a⟩ and ⟨:p_a⟩, indicating "A argument in Actor voice clause" and "P argument in Actor voice clause". The two arguments in undergoer voice in (3) are tagged as ⟨:a_u⟩ and ⟨:p_u⟩, indicating "A argument in Undergoer voice clause" and "P argument in Undergoer voice clause".

Finally, trivalent verb stems such as *atèd* 'give' in (4–6) show a three-way voice alternation between actor voice and two undergoer voices (typically conveyance voice and location voice). Actor voice constructions, marked by *maN-* on the verb, confers the absolutive onto the agent role 'I' in (4), leaving the other two NPs as oblique. The theme role is given absolutive case via the conveyance voice prefix *i-* on the verb, as in (5), and the recipient role is given absolutive case via the location voice affix *-an* on the verb as in (6). This three-way alternation is schematically represented in Table 3.

(4)   *Mangatèdtèn ta arta ta agi.*

mang-atèd	*=tèn*	*ta*	*arta*	*ta*	*agi*
AV-give	=1SG	OBL	person	OBL	cloth
v:pred	=pro.1:a_a	ln	np.h:p_a	ln	np:g_a

'I will give cloth to a person/people.'

(5)   *Yatèdu i agi ta arta.*

y-atèd	*=u*	*i*	*agi*	*ta*	*arta*
CV-give	=1SG.GEN	DEF	cloth	OBL	person
v:pred	=pro.1:a_u	l	np:p_u	ln	np.h:g_u

'I will give the cloth to a person/people.'

verb	agent	theme	recipient
AV (*um, maC-, maN-*	⟨:a_a⟩ (ABS)	⟨:p_a⟩ (OBL)	⟨:g_a⟩ (OBL)
UV (CV *i-*)	⟨:a_u⟩ (GEN)	⟨:p_u⟩ (ABS)	⟨:g_u⟩ (OBL)
UV (LV *-an*)	⟨:a_u⟩ (GEN)	⟨:obl_u⟩ (OBL	⟨:p_u⟩ (ABS)

Table 3 Voice and case marking in trivalent clauses

(6) *Atdanu i arta ta agi.*

atd-an	*=u*	*i*	*arta*	*ta*	*agi*
give-LV	=1SG.GEN	DEF	person	OBL	cloth
v:pred	=pro.1:a_u	ln	np.h:p_u	ln	np.h:obl_u

'I will give cloth to the person/people.'

The GRAID annotation system for annotating trivalent clauses assumes that agent and theme roles are indicated as A and P, and the recipient as a goal. This principle is applied to the first two voice constructions. In actor voice constructions, the three arguments are tagged as ⟨:a_a⟩ for the agent, ⟨:p_a⟩ for the theme, and ⟨:g_a⟩ for the recipient. In the conveyance voice, The three NPs are annotated as ⟨:a_u⟩, ⟨:p_u⟩ and ⟨:g_u⟩, representing "A argument in Undergoer voice clause", "P argument in Undergoer voice clause", and "Goal in Undergoer voice clause", respectively. However, we do not apply the above rule to location voice whereby the recipient, not the theme, receives absolutive case. In this case, the P role is reserved for the promoted recipient, rather than the demoted theme. Following this rule, we annotate the obliquely marked theme as ⟨obl_u⟩ "Oblique argument in Undergoer voice clause", and the absolutively marked recipient as ⟨p_u⟩ "P argument in Undergoer voice clause".

1.2   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 (7), plural third-person referents are encoded by the person index on the predicate without an independent nominal phrase, and, in (8), by an independent nominal without an index on the predicate:

(7) *Pabbi:rèndid tidi a:na:di, amma nappatidtid.*

	pab-bi:rè-n	*=di*	*=d*	*tidi*	*a:na:*	*=di,*
	PRG-search-TR	=3PL.GEN	=COMP	PL.DEF	children	=3PL.GEN
##	v:pred	=pro.h:a_u	=rv	ln	np.h:p_u	=rn_pro.h:poss

	amma	*nap-pati*	*=d*	*=tid*
	if	PST.INTR-die	=COMP	=3PL
#ac	other	v:pred	=rv	=pro.h:s_a

'They are looking for their children, if they died.'   [mc_arta_t0111_0025]

GRAID specifier	predicate class	voice	example
⟨_a⟩	dynamic verb	actor voice	*man-lutu* 'cook'
⟨_u⟩	dynamic verb	undergoer voice	*i-lutu* 'be cooked'
⟨_ap⟩	potentive verb	actor voice	*maka-tim* 'can drink'
⟨_up⟩	potentive verb	undergoer voice	*ma-tim* 'can be drunk'
⟨_np⟩	nominal predicate	—	*buka:gan* 'be a woman'
⟨_other⟩	other	—	*atti:* 'there is, exist'

Table 4   Voice tags on nominals

(8)   *Nappatid tidi amanay aydi: inana.*

nappati	*=d*	*tidi*	*ama*	*=na*	*=y*
PST.INTR-die	=COMP	PL.DEF	father	=3PL.GEN	=SPC
## v:pred	=rv	ln	np.h:s_a	=rn_pro.h:poss	=rn

aydi:	*ina*	*=na*
and	mother	=3SG.GEN
rn	rn	=rn

'His father and mother died.'   [mc_arta_husband_0021]

These realizations are, in accordance with Haig & Schnell (2014), tagged as ⟨=pro⟩ and ⟨np⟩ respectively. The equal sign ⟨=⟩ 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:

(9)   *Saya iggaman na a:yi: ni kanakannaki.*

saya	*iggam-an*	*=na*	*a:yi:*	*ni*	*kanakannak*	*=i*
DEM.DIST	hold-TR	=3SG.GEN	DEM.PROX	GEN.DEF	child	=SPC
## dem_pro:other	v:pred	=pro_a_u	dem_pro:p_u	ln	np.h:bpi_a_u	=rn

'Then the child held this.'   [mc_arta_t0110_0114]

This type of argument encoding appears at a relatively low frequency. If this pattern is observed, the person index is taken as the primary instantiation of the argument, while the function of the independent NP is annotated as ⟨bpi_⟩ 'bound person index' plus the function of the primary person index.

1.3   Voice tags on nominals

In Arta, almost every verb is marked. Morphological roots take various kinds of verbal (or adjectival) affixes to formulate predicates (Table 4). Depending on the possible morphosyntactic behaviours, these affixes fall into two verb classes: dynamic verbs and potentive verbs. As mentioned in Section 1.2, GRAID annotations in Arta include a cross-reference tag on their function about the relevant predicate type. Dynamic verbs, a morphologically unmarked category, are tagged as ⟨_a⟩ (actor voice) or ⟨_u⟩ (undergoer voice); potentive verbs, a morphologically marked category, is tagged either ⟨_ap⟩ (actor voice) or ⟨_up⟩ (undergoer voice). Other predicate categories, which lack voice distinctions, are specified ⟨_np⟩ or ⟨_other⟩, as shown in Table 4.

(10) *Amma mamurab tidi amamiti.*

   amma *mam-purab* *tidi* *ama*   *=mi*   *=ti*
   if   AV-hunt PL.DEF father =1PL.GEN   =SPC
 #ac other v:pred   ln   np.h:s_a =rn_pro.1:poss =rn

   'If our father went hunting, …'   [mc_arta_disubu_0006]

(11) *Sa:biténdid, ngay ti bunbunmi.*

    sa:bit-én *=di*   *=d*   *ngay*   *0*   *ti*
  0   carry-PV =3PL.GEN =POST go   0   OBL.DEF
 ## 0:p_u v:pred =pro.h:a_u =other vother:pred 0.h:s_other ln

   bunbun *=mi*
   house =1PL.GEN
   np:g =rn_pro.1:poss

   '(They) carry it on the shoulder, going (with it) to our house.'
                   [mc_arta_disubu_0060-0061]

(12) *Awantep maka:ngay ta ayta lugar.*

    awan *=tep* *maka:ngay*   *ta* *ayta*   *lugar*
    NEG =ANT POT.AV-go 0   OBL OBL.DIST place
 ##neg other =other v:pred   0.h:a_ap ln dem_pro:p_ap np:g_ap

   'They could not come there yet, (to that) place.'   [mc_arta_disubu_0033-0034]

(13) *Saya napi:piyad i pamiliami.*

   saya *na-pi:piya*   *=d* *i* *pamilia* *=mi*
   then PST.POT.PV-good =COMP DEF family   1PL.GEN
 ## other v:pred   =rv ln np.h:s_up =rn_pro.1:poss

   'Then our families got better.'   [mc_arta_t0601_0046]

(14) *Med-dès i uga:likuy.*

   med-dès *i* *uga:li*   *=ku*   *=y*
   ADJ-bad DEF habit   =1SG.GEN   =SPC
 ## ap_other ln np:s_other =rn_pro.1:poss =rn

   'My habit was bad.'   [mc_arta_t0601_0074]

(15) *Siye: wam kuwartom.*

   siye:,   *wa* *=m,*   *kuwarto* *=m*
   DEM.PROX   PLH =2SG.GEN money =2SG.GEN
 ## dem_pro:s_np other =other   np:pred =rn_pro.2:poss

   'This is yours, your money.'   [mc_arta_t0601_0094]

1.4   Referential expressions

When an independent referential expression is headed by a lexical noun, the noun is preceded by a determiner that inflects for 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 ⟨ln⟩, and specificity markers are, when they appear after a noun, tagged as ⟨rn⟩.

(16)   *Atti:tep i gilangani ta ayta Danak.*

atti:	*=tep*	*i*	*gilangan*	*=i*	*ta*	*ayta*	*Danak*
exist	=ANT	DEF	male	=SPC	OBL	there	Danak
## other:predex	=other	ln	np.h:s_other	=rn	ln	ln	np:s

'The man was still there in Danak.'   [mc_arta_udulan_0011]

When a specificity marker appears within a noun phrase, it occupies the slot immediately after the first lexical element. For example, when a modifier appears before a head noun, the specificity marker longer follows the noun but is encliticized to the modifier:

(17)   *Tidi tallip=i a buka:gan, awantid nakapanga:dal.*

tidi	*tallip*	*=i*	*a*	*buka:gan,*	*awan*	*=tid*	*naka-panga:dal*
PL.DEF	two	=SPC	LIG	woman	NEG	=3PL	PST.POT-learn
## ln	ln	=ln	ln	np.h:dt_s	other	=pro.h:s_ap	v:pred

'As for the two women, they were not able to go to school.'   [mc_arta_t0110_0046]

Some optional elements may modify a head noun with the intervening connective *a* (LIG, ligature), as shown by the numeral quantifier in ((17)). 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 cases in which the adjectives, quantifiers, and/or demonstratives modify the head nouns. 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.

(18)   *Basta inan'anu:sanmitéddi ay ka:mani a to:luda.*

basta	*in-an-'anu:s-an*	*=mi*	*=téddi*	*ay*	*ka:man*	*=i*	*a*	*to:luda*
just/even	PST-RDP-tolerate-LV	=1PL.GEN	=only	TOP	big	=SPC	LIG	tent
## other	v:pred	=pro.1:a_u	=other	ln	ln	=ln	ln	np:p_u

'We put up with just a big tent [instead of our houses].'   [mc_arta_typhoon_0008]

(19)   *Umangayde:tènti, man na ne:but attanan a meddès a uga:li.*

um-angay	*=de:*	*=tèn*	*=ti,*
INTR-go	=COMP	=1SG	=DEM.PROX.OBL
## v:pred	=other	=pro.1:s_a	=dem_pro:g_a

	man	*na*	*ne:but*	*attanan*	*a*	*med-dès*	*a*	*uga:li*
0	as.if	GEN	PST.POT.lose	all	LIG	ADJ-bad	LIG	habit
#ac 0.1:a_up	other	other	v:pred	ln	ln	ln	ln	ln

'After I came here, it seems that (I) have lost every bad habit.'   [mc_arta_t0601_0055]

(20) *Ènsi:na di:sanna i gissay a lingo aynina a me'a:du a baggat.*

	ènsi:na	di:san	=na	i	gissa	=y	a	lingo	aynina	a	me'a:du
	so.that	reach	=3SG.GEN	DEF	one	=SPC	LIG	week	DEM.MED	LIG	ADJ-plenty
##	other	v:pred	=rn	ln	ln	=ln	ln	np:p_u	ln		ln	ln

a	baggat
LIG	rice
ln	np:a_u

'So that that large amount of rice would last for one week.' [mc_arta_t0601_0055]

Note that the same constructional template [modifier *a* head] or [head *a* modifier] is employed both for adjective modifications and relative clauses; that is, 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 demostratives are treated as ⟨ln⟩ or ⟨rn⟩, while adnominal modifications by means of verbs (or more precisely verb-headed clauses) are treated as relative clauses and annotated hence annotated as ⟨#rc⟩, which are touched upon in Section 1.5.

1.5 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 *ka:huy* 'cassava' and immediately followed by a relative clause *a nimulamula=mi* 'that we planted':

(21) *Ayde:yi ka:huy a nimulamulami.*

	ayde:yi	ka:huy		a	n-i-mula-mula	=mi
	and	cassava	0	LIG	PST-CV-RDP-plant	=1PL.GEN
##	other	np:dt_s	#rc	f0:p_u	other	v:pred	=pro.1:a_u

'And the cassava, that we planted.' [mc_arta_typhoon_0022]

Although the transitive verb *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 inferrable, but because the construction does not allow overt appearance of the argument. In standard GRAID, the zero realization caused by grammatical suppressions is not specified (by ⟨0⟩); in the Arta corpus, such cases are instead glossed as ⟨f0⟩, indicating that the argument is a forced zero within a relative clause.

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).

Kimoto, Yukinori. 2014. A preliminary report on the grammar of Arta. *Proceedings of the Thirty-Eighth Annual Meeting of The Kansai Linguistic Society (KLS34)*. 85–96.

Kimoto, Yukinori. 2017a. *A grammar of Arta: A Philippine Negrito language*. Kyoto: Kyoto University Ph.D. dissertation.

Kimoto, Yukinori. 2017b. Vowel length, mora, and diachrony: The case for Arta, a Philippine Negrito language. In Liao, Hsiu-chuan (ed.), *Issues in Austronesian historical linguistics*, 1–22. Honolulu: University of Hawai'i Press.

Kimoto, Yukinori. 2018. Operationalizing Philippine-type syntax for GRAID system: Clause structure, case marking, and verb class in Arta. *Asian and African Languages and Linguistics* 12. 17–35. (`http://hdl.handle.net/10108/91147`).

Kimoto, Yukinori. 2019. Multi-CAST Arta. In Haig, Geoffrey & Schnell, Stefan (eds.), *Multi-CAST*: *Multilingual Corpus of Annotated Spoken Texts*. (`https://multicast.aspra.uni-bamberg.de/#arta`) (Accessed 2019-07-05).

Reid, Lawrence A. 1989. Arta, another Philippine Negrito language. *Oceanic Linguistics* 28(1). 47–74.

Appendices

A   Person forms, determiners, and demonstratives

The following tables show the paradigms 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, the labels for "absolutive" and "singular" are omitted for the sake of simplicity.

person	*topical*	*absolutive*	*genitive*	*oblique*
1SG	*tèn*	*=tèn*	*=ku*	*dèn*
1PL	*tami*	*=ami*	*=mi*	*dami*
2SG	*taw*	*=a, =taw*	*=mu*	*daw*
2PL	*tam*	*=am*	*=muyu*	*dam*
1+2SG	*tita*	*=ita*	*=ta*	*dita*
1+2PL	*titam*	*=itam*	*=tam*	*ditam*
3SG	*siya*	*=siya*	*=na*	*dya*
3PL	*tidi~tidu*	*=tid*	*=di*	*did*

Table A.1   Person forms. 1+2SG/PL forms are inclusive, 1PL is exclusive.

			ABS	*GEN*	*OBL*
indefinite				*na*	*ta*
definite	singular	common	*i*	*ni*	*ti*
		personal	*ti*	*ni*	*ni*
	plural		*tidi*	*didi*	*didi*

Table A.2   Determines. Personal determiners are used with proper names.

		TOP	*ABS*	*GEN/ERG*	*OBL*
proximal	SG	*si:yèy*	*a:yi:, =i*	*ni/na a:yi:/ayni, =ni*	*ti/ta a:yi:, =ti*
	PL	*satidi:*	*(ay)tidi a:yi:*	*(ay)didi a:yi:*	*(ay)didi a:yi:*
medial	SG	*sayna*	*a:yina, =ina*	*ni/na ayna, =nina*	*ti/ta ayna, =tina*
	PL	*satidi:na*	*(ay)tidi:na*	*(ay)didi:na*	*(ay)didi:na*
distal	SG	*saya*	*a:ya:, =ya:*	*ni/na a:ya:*	*ti/ta a:ya:, =ta*
	PL	*satiddya:*	*(ay)tiddya*	*(ay)didi a:ya:*	*(ay)didi a:ya:,*

Table A.3   Demonstratives.

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

Form symbols and specifiers

⟨f0⟩   structurally suppressed argument slot of a predicate

⟨dem_pro⟩   demonstrative pronoun
⟨pn_np⟩   proper name
⟨ap_other⟩   adjective phrase as predicate
⟨dem_other⟩   demonstrative adverb

Function symbols and specifiers

⟨:bpi⟩   bound person index; combines with specifiers reflecting the function of the primary argument slot (e.g. ⟨:bpi_a_a⟩)

⟨_a⟩   actor voice of a dynamic verb, attaches to core argument functions
⟨_ap⟩   actor voice of a potentive verb, attaches to core argument functions
⟨_ds⟩   subject of a verb of speech, attaches to ⟨:s⟩ and ⟨:ds⟩
⟨_np⟩   nominal predicate, attaches to core argument functions
⟨_other⟩   other types predicate (incl. existentials), attaches to core argument functions
⟨_stv⟩   subject of a stative verb, attaches to core argument functions
⟨_u⟩   undergoer voice of a dynamic verb, attaches to core argument functions
⟨_up⟩   undergoer voice of a potentive verb, attaches to core argument functions

Other symbols

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

C   List of abbreviated morphological glosses

1	first person		NMZ	nominalizer
1+2	first-second person		OBL	oblique case
2	second person		PL	plural
3	third person		POT	potentive verb
ABS	absolutive case		POST	posterior aspect ('already')
ADJ	adjectivizer		PROG	progressive aspect
AV	actor voice		PROX	proximal demonstrative
CAUS	causative		PSN	personal article
COMP	complementizer		PST	past tense
CV	conveyance voice		PV	patient voice
DEF	definite		RDP	reduplication
DEM	demonstrative		SG	singular
DIST	distal demonstrative		SPC	nominal specificier
GEN	genitive case		STV	stative verb
LIG	ligature (linker)		TOP	topic
LV	locative voice		UV	undergoer voice
MED	medial demonstrative			
NEG	negation		NC	not classifiable

Multi-CAST

Multilingual Corpus of Annotated Spoken Texts

multicast.aspra.uni-bamberg.de/