

Multi-CAST

Teop

annotation notes

Ulrike Mosel

August 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

Mosel, Ulrike. 2019. Multi-CAST Teop 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The Multi-CAST collection has been archived at the *University of Bamberg*, Germany, and is freely accessible online at multicast.aspra.uni-bamberg.de/.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 Teop annotation notes v1.0 last updated 25 August 2019
This document was typeset by NNS with \LaTeX and the *multicast3* class (v3.2.1).

Contents

1	A very short grammar of Teop	1
1.1	Introduction	1
1.2	Clause structure	3
1.3	Person forms	5
1.4	Zero arguments	5
1.5	Types of phrase	7
1.5.1	The verb complex	8
1.5.2	The noun phrase	9
	References	9
	Appendices	11
A	List of corpus-specific GRAID symbols	11
B	List of abbreviated morphological glosses	12

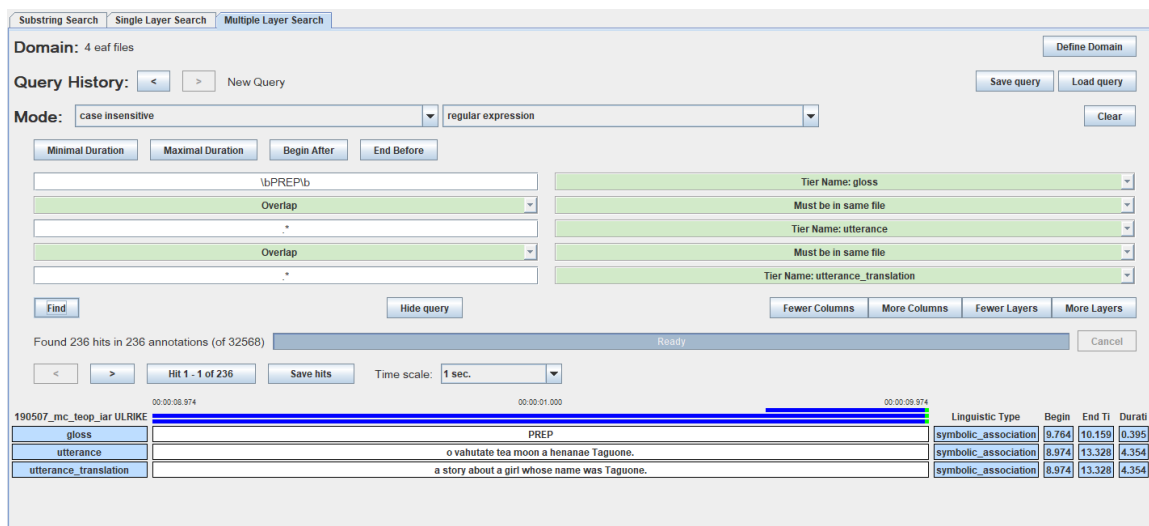
1 A very short grammar of Teop

1.1 Introduction

Teop is an Oceanic, Western Melanesian language, spoken on the north-eastern coast of Bougainville, Papua New Guinea (Lynch et al. 2002: 883). It is a configurational language in which clauses are formed by phrases. The phrases consists of lexical and functional words whose sequential order is fixed so that they form syntactic units with clear borderlines. The only words that inflect are person forms. Although object, action and property words are multifunctional as they can occur as the heads and modifiers of different phrase types, verbs, nouns, and adjectives can be identified as distinct word classes on the basis of distributional criteria (see Mosel 2017).

The following sections start with the structure of clauses (Section 1.2) and the inflection of person forms (Section 1.3), and then deal with zero arguments (Section 1.4) and the structure and function of phrases (Section 1.5).

All sections below frequently refer to the four texts in Multi-CAST corpus and suggest to search (in ELAN or R) by regular expressions for a morphological gloss or a GRAID annotation, or to search simultaneously the utterance and the translation tiers for their context by using with the wild card operator $\langle . * \rangle$. The *Multiple Layer Search* window shown in Figure 1 can be accessed in ELAN



Domain: 4 eaf files

Query History: < > New Query

Mode: case insensitive regular expression

Minimal Duration Maximal Duration Begin After End Before

Search criteria:

- Expression: \bPREP\b
- Overlap: Overlap
- Expression: *
- Overlap: Overlap
- Expression: *
- Overlap: Overlap
- Expression: *

Tier Name: gloss
Must be in same file
Tier Name: utterance
Must be in same file
Tier Name: utterance_translation

Find Hide query Fewer Columns More Columns Fewer Layers More Layers

Found 236 hits in 236 annotations (of 32568) Ready Cancel

Time scale: 1 sec.

190507_mc_teop_iar_ULRIKE	Linguistic Type	Begin	End	Duration
gloss	PREP	9.764	10.159	0.395
utterance	o vahutate tea moon a henanae Taguone.	8.974	13.328	4.354
utterance_translation	a story about a girl whose name was Taguone.	8.974	13.328	4.354

utterance_id	utterance	grammatical_word	gloss	graid	utterance_translation
teop_iar_0004	o vahutate tea moon a henanae Taguone.	o vahutate te =a moon #	ART3.SG story PREP =ART2.SG girl #	ln np:p rn_adp =rn rn_np.h:poss ##	a story about a girl whose name was Taguone.

Figure 1 Multiple Layer Search for a gloss, an utterance, and a translation in ELAN (top), and the first result of a search for $\langle \text{PREP} \rangle$, i.e. prepositions, in the Multi-CAST Teop corpus (bottom).

via *Search* → *Structured Search Multiple eaf...* → *Multiple Layer Search*. For a general guide to searches with regular expressions in ELAN, see Mosel (2015).

A further source of information on the meaning and construction of functional and content words is *A multifunctional Teop-English dictionary* (MTED, Mosel 2019),¹ which facilitates single or combined searches for Teop headwords, parts of speech and their constructions, meanings, morphemes, morphological glossings, and semantic domains. A search for ditransitive constructions as in Figure 2, for example, yields 81 entries with morphological analyses, glosses, and examples.

Showing 1 to 81 of 81 entries (filtered from 6,488 total entries)

full entry	headword	part of speech	meaning description
<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="vd.const"/>	<input type="text" value="Search"/>
more	<i>aviavisi vabero ki</i>	vd.constr	serve someone more (food)
more	<i>avisi ki</i>	vd.constr	serve some food for someone
more	<i>boha me</i>	vd.constr	say something to somebody
more	<i>boha mi</i>	vd.constr	say something to somebody; say something about somebody
more	<i>busu kahi</i>	vd.constr	remove something from something and throw it away
more	<i>busu ki</i>	vd.constr	throw something to someone; distribute something to someone

Part of speech	vd.constr
MWE structure	vt - adv - prep
morphology	avi-avisi va-bero ki
gloss	RED-serve.food ADVR-much DAT
serve someone more (food)	
semantic domain: transfer food & drink	
Example 464:	
<i>Ean toro aviavisi vabero kanaa sa mata muu!</i>	
'You must serve me more taro pieces!'	
Source: Mom 01R 090	

Figure 2 Ditransitive constructions in the MTED, found by searching for <vd.constr> among parts of speech (top), and the first entry found by the search (bottom).

¹ dictionaria.clld.org/contributions/teop#twords

<i>1st position</i>	<i>2nd position</i>	<i>3rd position</i>	<i>4th position</i>
subject	verbal predicate		
subject	verbal predicate	object	
subject	verbal predicate	primary object	secondary object
object	verbal predicate	subject	
primary object	verbal predicate	subject	secondary object
secondary object	verbal predicate	subject	primary object

Table 1 The variable order of NP arguments in verbal clauses.

<i>1st position</i>	<i>2nd position</i>	<i>3rd position</i>	<i>4th position</i>
subject 1 st /2 nd pers. pron.	verbal predicate	primary object ARTX	secondary object OBJ.ARTX
subject 1 st /2 nd pers. pron.	verbal predicate	primary object 1 st /2 nd pers. pron.	secondary object OBJ.ARTX

Table 2 Ditransitive constructions.

1.2 Clause structure

Clauses with verbal predicates (i.e. predicates formed by a verb complex) maximally contain three arguments: a subject, a primary object, and a secondary object. The predicative verb complex holds the second position in the clause, while the first position may be occupied by any topical argument, see Table 1. If all arguments are formed by NPs, the arguments printed in blue in Table 1 are marked by a basic article, the others are preceded by an object article.

The syntactic functions nominal subject, primary object, and secondary object are distinguished by the constituent order in combination with two kinds of article, the basic and the object article, in the following way:

- ◆ Irrespective of their position, subject NPs always take the basic article (ART). To find examples, search the GRAID tier for the regular expression $\langle np.*?:[as] \$ \rangle$.
- ◆ Topical NPs always precede the verbal predicate and take the basic article (ART) irrespective of their syntactic function.
- ◆ Non-topical primary object NPs only take the basic article (ART) if the subject of the clause is a first or a second person pronoun, as in (1), otherwise they take the object article (OBJ.ART). For more examples, search the GRAID tier for $\langle pro.[12]:a \$ \rangle$.

(1) *are ma nahu a guu vai bona tahii.*

```

a-re          ma    nahu a    guu vai bona    tahii
1PL.IN-CONSEC come  cook ART2.SG pig  DEM6 OBJ.ART.SG saltwater
##ds p1_pro.1:a  v:pred rv_v ln    np:p rn    ln    np:p2
‘[He said,] “... so that we can cook this pig with saltwater.”’
[mc_teop_mat_0069]
```

- ◆ Non-topical secondary objects only take the basic article (ART) if the subject and the primary object refer to the speech act participants, otherwise they take the object article (OBJ.ART). Thirteen instances of non-human secondary objects can be found by searching the corpus for the regular expression $\langle np:p2 \rangle$, of which only one shows a secondary object with a basic article:

Figure 3 Multiple Layer Search for secondary objects in ELAN.

(2) *ena a pasi vaasusu avameam a meha taba.*

e-naa *pasi* *vaasusu* *a=* *v=* *ame-am*
 H-1SG.PRON TAM4 teach OBJM= IMM= 2PL.OBJM-2PL.PRON
 ##ds pro.1:a lv v:pred =lv =rv pl_pro.2:p

a *meha* *taba*
 ART2.SG other2.SG thing
 ln ln np:p2

‘[He said,] “I’ll teach you something else.”’

[mc_teop_sii_0046]

Clauses are not obligatorily introduced by a topical argument. In two kinds of context topical arguments can be missing: (i) in presentational constructions (Lambrecht 1994: 177–181), and (ii) with zero anaphora. Presentational constructions are also typically found at the beginning of stories, as exemplified in (3).

(3) *teitei roho a peha vaan.*

tei-tei *roho* *a* *peha* *vaan*
 RED-stay before ART2.SG INDEF2.SG village
 ## v:pred rv ln ln np:s

‘There was once a village.’

[mc_teop_mat_0001]

The term zero anaphora means that an argument position that is required by the valency of the predicate is left empty by the speaker because they assume that the hearer gets sufficient information from the context to identify the implied referent. To find, for instance, 90 examples with zero transitive subjects, search the GRAID tier for $\langle \emptyset . h : a \rangle$.

Clauses with predicates that are not formed by verb complexes can be found by searching the GRAID tier with the the four types of annotations explained in Table 3.

<i>regular expression</i>	<i>matched form of predicate</i>
<code>np.*?:pred</code>	full noun phrase
<code>#cc.*?:pred</code>	complement clause
<code>[_^]other:pred</code>	(i) the negative existential word <i>ahiki</i> ‘non exist, there is not’ (ii) interrogative adverb <i>havee</i> ‘where’ (iii) the interrogative determiner <i>pomae</i> ‘whose, from where’ (iv) the demonstrative <i>eiei</i> ‘this here’
<code>vother:pred</code>	verb complex without pre-head position for tense-aspect-mood marking, which only occurs in complement clauses introduced by <i>tea</i> COMP1

Table 3 Clauses with non-verbal predicates

1.3 Person forms

The Teop person forms can be classified into pronouns and cross-referencing agreement markers (also called indexing markers, cf. Malchukov et al. 2010: 8; Mosel 2010: 488). The class of pronouns comprises five subclasses:

1. Heavy pronouns are marked by the prefix *e-* and function as left-dislocated topics, pre-verbal topical subjects, and in a few contexts as object pronouns.
2. Basic pronouns function as post-verbal non-topical subjects, objects, possessors in adnominal possessive constructions, and as complements of prepositions. The 1SG and 1PL.EX pronouns can also function as topical subjects.
3. Speech act participant object pronouns are marked by a prefix that is formally related to the clitic object marker within the verb complex.
4. Fourth person object pronouns are used in transitive clauses with a third person subject and in ditransitive clauses with a third person subject and/or the third person primary object.
5. Consecutive person forms are subject pronouns that are prefixed to the consecutive conjunction *re* ‘then, so that’.

The indexing person forms form three classes:

1. The object markers are part of the verb complex and cross-reference the object of transitive clauses and the primary object on ditransitive clauses.
2. The imperfective aspect markers are inflecting portemanteau morphs denoting the imperfective aspect and cross-referencing the subject or the object of the clause;
3. The possessive markers are used to form adnominal inalienable possessive constructions. They are affixed to the possessee, i.e. the noun forming the head of the NP, but are sometimes written as a separate word when they combine with a clitic article or pronoun.

There are no occurrences of reflexive pronoun constructions in the Multi-CAST Teop corpus, but examples can be found in the MTED when simultaneously searching for the `<pron. constr>` in the “part-of-speech” field and for `<self>` or `<selfes>` in the “meaning description” field. These pronoun constructions are formed by the object article and a heavy or basic pronoun.

1.4 Zero arguments

A referent may not be phonetically represented, but nevertheless be retrievable from the context because its context contains a grammatical construction with an empty argument position and provides sufficient information for the hearer to understand which referent is meant. There are

	<i>heavy pronouns</i>	<i>basic pronouns</i>	<i>object pronouns</i>	<i>object markers</i>	<i>consecutive subj. pro.</i>	<i>imperative markers</i>	<i>possessive markers</i>
1SG	<i>ena</i>	<i>na</i>	<i>a-na</i>	<i>a=, ama(=)</i>	—	<i>nom</i>	—
2SG	<i>ea</i>	<i>an</i>	<i>vu-an</i>	<i>vu=, u=, =u</i>	—	<i>nom</i>	<i>m-</i>
3SG	<i>eove, eve</i>	<i>eve, e</i>	<i>eve, ee, =e, =i</i>	—	<i>o-</i>	<i>nana, =na</i>	<i>na=</i>
1PL.IN	<i>eara</i>	<i>ara</i>	<i>ara-ara</i>	<i>ara, ama</i>	<i>a-</i>	<i>rara, =ra</i>	<i>ra-</i>
1PL.EX	<i>enam</i>	<i>nam</i>	<i>a-nam</i>	<i>ama</i>	—	<i>nom</i>	<i>-ma-</i>
2PL	<i>eam</i>	<i>am</i>	<i>ame-am</i>	<i>ame(=)</i>	—	<i>nom</i>	<i>me-</i>
3PL	<i>eori</i>	<i>ori</i>	<i>ori</i>	<i>=ri, ri=</i>	<i>o-</i>	<i>rori, ri=, =ri</i>	<i>ri=</i>
4SG1	—	—	<i>bona, bari</i>	—	—	—	—
4SG2	—	—	<i>bari</i>	—	—	—	—

Table 4 Person forms in Teop.

three types of zero in Teop, the canonical anaphoric zero, the associative zero, and the situational zero:

1. A canonical anaphoric zero shares the same referent with some antecedent nominal or pronominal form. Examples for zero transitive subjects, for example, can be found by searching the GRAID tier for $\langle \emptyset . * ? : a \rangle$.
2. The reference of the associative zero is not given by an antecedent coreferent pronoun or NP, but understood by the association of a verb or some other word with this referent, see (4).
3. The reference of a situational zero is given by the speech situation; second person zero arguments in imperatives, for instance, can be found with the regular expression $\langle \emptyset . 2 \rangle$. Since imperative clauses can also be formed with subject pronouns in Teop, the assumption of zero subjects in imperative clauses is justified.

(4) a. *o bai na vua nana.*

o bai na vua nana
 ART3.SG mango_tree TAM2 bear_fruit 3SG.IPFV
 ## 1n np:s lv v:pred rv-pro_s
 ‘The mango tree was bearing fruit.’

- b. *mepaa taneo tea tasu, ...*
- | | | | |
|----------------|------------|--------------|-------------------------|
| <i>me=</i> | <i>paa</i> | <i>taneo</i> | <i>tea</i> |
| and4= | 3SG.ZERO | TAM3 | start |
| ## =conj_other | 0.h:a | lv | v:pred #cc:p conj_other |
-
- | | |
|-------------|-----------|
| <i>tasu</i> | |
| throw | 4SG1.ZERO |
| vother:pred | 0:p |
- 'And (she) started to throw (a stick) (at the mango).'
- [mc_teop_mat_0017-0018]

Example (4) is taken from a story about a woman who wants to eat a mango fruit. When she finds a tree bearing fruit, she throws a stick at the fruit. The verb *vua* 'bear fruit' implies that there is a fruit, but grammatically it is not a referential nominal antecedent. The verb *tasu* 'throw something at somebody or something' is ditransitive (see the MTED entries on *vua*₂ and *tasu*).

1.5 Types of phrase

The phrase types are not defined by the word class affiliation of their heads, but by their internal structure as determined by the combination of functional words with lexical words. Consequently, Teop noun phrases and verb complexes are not exclusively headed by nouns and verbs, respectively. The following types of phrase can be distinguished:

1. Verb complexes (VCs) are marked by tense-aspect mood markers, including imperatives that can be marked by the imperfective aspect marker *nom*, e.g. *hagi nom* 'keep dancing!' (see Section 1.5.1)
2. Noun phrases (NPs) are marked by articles (see Section 1.5.2)
3. Locative phrases (LocPs) do not have an article, but otherwise have the same structure as NPs and may combine with the same kind of modifying constructions, e.g. *inu* 'at home, home'. Search the GRAID tier for ⟨np: 1⟩ and look for NPs without an article.
4. Prepositional phrases (PPs) are introduced by one of the nine prepositions, e.g. *te=a inu* 'in/to the house'. Search the morphological glosses tier for ⟨DAT⟩, ⟨for⟩, ⟨PREP⟩.
5. Adjectival phrases (APs) are marked by an article that agrees with the article of the NP they relate to as an attribute or predicate; e.g. *a inu a mataa* 'a good house', *o naono o mataa* 'a good tree'. They function as attributes and predicates. In attributive function they follow the NP.
6. Numeral phrases (NumPs) are like adjectival phrases. They are marked by an article that agrees with the article of the NP they relate to. But in attributive function they precede the NP, e.g. *a buaku a aba* 'two people', *o buaku o naono* 'two trees', while APs follow the NP they modify. In predicative function both NumPs and APs follow.
7. Adverb phrases (AdvPs) are formed by an adverb which may only be modified by another following adverb, e.g. *vamataa rakaha* (ADVR-good very) 'very well'. AdvPs and LocPs have in common that they are not marked by functional words.
8. Linker phrases (LnkP) consist of the linker *vaa* and a PP, LocP, AdvP, or complement clause. Search the morphological glosses tier for ⟨LNK⟩.
9. Possessor phrases (PossP) follow the head of a NP and consist of the possessive marker (POSS) and a NP or pronoun, e.g. *a hena-na-e* (ART name-3SG.POSS-3SG.PRON) 'her name'. Search the morphological glosses tier for ⟨POSS⟩.
10. Specifier phrases (SpecP) precede the NP and are introduced by an article that agrees with the article of the head. The head of the SpecP may fuse with the article of the fol-

<i>phrase type</i>	<i>as predicate</i>	<i>as argument or adjunct</i>	<i>as modifier of NP</i>	<i>as modifier in VC</i>
VC	<v:pred>	–	–	–
NP	<np(.h):pred>	<np(.h):s>, <np(.h):a>, <np(.h):p>, <np(.h):p2>	<np(.h):appos>	–
LocP	+	<np:l>	+	–
PP	+	+	<rn_np(.h):poss>	–
AdjP	<ap_other:pred>	–	<rn>	–
NumP	+	–	+	–
AdvP	+	+	–	–
LnkP	+	–	<rn>	+
PossP	–	–	<rn_np(.h):poss>	–
SpecP	–	–	+	–

Table 5 Syntactic functions of phrases and their GRAID annotations.

<i>type</i>	<i>gloss</i>
tense-aspect-mood marker	TAM, must
negation	NEG
adverb	all

Table 6 Pre-head constituents.

<i>type</i>	<i>gloss</i>
tense-aspect-mood marker	IPFV, IMM
negation	NEG
object marker	OBJM
incorporated preposition	DAT, GOAL, like, with
applicative particle	APPL
adverb	CONT, ADVR-, hither, thither, very
nouns, verbs	(lexemes)

Table 7 Post-head constituents.

lowing NP, e.g. *a meha inu* ‘the other/another house’, *o meho naono* ‘the other/another tree’. Search the morphological glosses tier for <other2> and <other3>.

With the exception of VCs and PossPs, all types of phrase have more than one syntactic function. In Table 5, the presence of a GRAID annotation indicates that the specific phrase–function combination is found in the Multi-CAST Teop corpus, a plus sign indicates that it is found only in the larger *Teop Language Corpus*, and a minus sign that there are no instances of its occurrence in either corpus.

1.5.1 The verb complex

The VC consists of a head formed by a verb, adjective, or noun, which in independent clauses are accompanied by tense-aspect-mood particles. The head of the VC is annotated with <v:pred>, irrespective of its word-class affiliation. In addition the VC may contain incorporated nouns, serial verbs, and adverbs. In the GRAID annotations VC constituents left of the head are indicated by

The screenshot shows the 'Multiple Layer Search' window. At the top, there are tabs for 'Substring Search', 'Single Layer Search', and 'Multiple Layer Search'. Below the tabs, the 'Domain' is set to '4 eaf files'. The 'Query History' section includes navigation buttons and a 'New Query' button. The 'Mode' is set to 'case insensitive' and 'regular expression'. There are buttons for 'Minimal Duration', 'Maximal Duration', 'Begin After', and 'End Before'. The search criteria are defined by three rows of input fields and dropdown menus. The first row has 'lv' in the input field and 'Tier Name: graid' in the dropdown. The second row has 'Overlap' in the dropdown and 'Must be in same file' in the dropdown. The third row has '*' in the input field and 'Tier Name: utterance' in the dropdown. The fourth row has 'Overlap' in the dropdown and 'Must be in same file' in the dropdown. The fifth row has '*' in the input field and 'Tier Name: utterance_translation' in the dropdown. At the bottom, there are buttons for 'Find', 'Hide query', 'Fewer Columns', 'More Columns', 'Fewer Layers', and 'More Layers'. The status bar at the bottom indicates 'Found 572 hits in 572 annotations (of 32624)' and a 'Ready' status.

Figure 4 Multiple Layer Search for pre-head constituents in VCs.

articles	with 1 st class nouns		with 2 nd class nouns		with 3 rd class nouns	
	singular	plural	singular	plural	singular	plural
specific basic article	<i>e</i>	<i>ere</i>	<i>a</i>	<i>o</i>	<i>o</i>	<i>a</i>
	ART1.SG	ART1.PL	ART2.SG	ART2.PL	ART3.SG	ART3.PL
specific object article	<i>bone,</i> <i>bene</i>	<i>bere,</i> <i>benere</i>	<i>bona</i>	<i>bono</i>	<i>bono</i>	<i>bona</i>
	OBJ.ART1.SG	OBJ.ART1.PL	OBJ.ART2.SG	OBJ.ART2.PL	OBJ.ART3.SG	OBJ.ART3.PL
non-specific article	—	—	<i>ta</i>	(?)	<i>to</i>	<i>ta</i>
partitive article	—	—	<i>sa</i>	<i>sa</i>	<i>sa</i>	<i>sa</i>

Table 8 Articles in Teop.

⟨lv⟩ and VC constituents right of the head by ⟨rv⟩. Figure 4 shows how one can find VCs with constituents left of the head and their contexts with translations. Table 6 lists the pre-head and Table 7 the post-head constituents with their respective glosses.

1.5.2 The noun phrase

Noun phrases are marked by articles (Mosel 2014), see Table 8. In GRAID the head of a NP is annotated as ⟨np⟩ plus an abbreviation of the function of the NP, see Table 5.

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).
- Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus, and the mental representation of discourse referents*. Cambridge: Cambridge University Press.
- Lynch, John & Ross, Malcolm & Crowley, Terry (eds.). 2002. *The Oceanic languages*. Vol. 1. London: Curzon.
- Malchukov, Andrej & Haspelmath, Martin & Comrie, Bernard (eds.). 2010. *Studies in ditransitive constructions: A comparative handbook*. Berlin: Mouton de Gruyter.
- Mosel, Ulrike. 2010. The fourth person in Teop. In Bowden, John & Himmelmann, Nikolaus P. & Ross, Malcolm D. (eds.), *A journey through Austronesian and Papuan linguistic and cultural space: Papers in honour of Andrew K. Pawley*, vol. 615, 391–404. Canberra: Australian National University.
- Mosel, Ulrike. 2014. Type shifts and noun class changes under determination in Teop. In Gerland, Doris & Horn, Christian & Latrouite, Anja & Ortmann, Albert (eds.), *Meaning and grammar of nouns and verbs*, vol. 1, 49–76. Düsseldorf: dup.
- Mosel, Ulrike. 2015. *Searches in ELAN with regular expressions*. (https://tla.mpi.nl/wp-content/uploads/2011/12/Searches_in_ELAN_with_regular_expressions.pdf) (Accessed 2019-08-19).
- Mosel, Ulrike. 2017. Teop — An Oceanic language with multifunctional verbs, nouns and adjectives. *Studies in Language* 41(2). 255–293. (<https://doi.org/10.1075/sl.42.2.02mos>).
- Mosel, Ulrike. 2019. A multifunctional Teop-English dictionary. *Dictionaria* 4(1-6488). (<https://doi.org/10.5281/zenodo.3257580>). (<https://dictionaria.clld.org/contributions/teop>) (Accessed 2019-08-19).
- Mosel, Ulrike & Schnell, Stefan. 2015. Multi-CAST Teop. In Haig, Geoffrey & Schnell, Stefan (eds.), *Multi-CAST: Multilingual corpus of annotated spoken texts*. (<https://multicast.aspra.uni-bamberg.de/#teop>) (Accessed 2019-03-08).
- Schiborr, Nils N. & Schnell, Stefan & Thiele, Hanna. 2018. *RefIND — Referent Indexing in Natural-language Discourse: Annotation guidelines (v1.1)*. Bamberg / Melbourne: University of Bamberg. (<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 *Teop* corpus. Please refer to the *GRAID manual* (Haig & Schnell 2014: 54–55) for an inventory of basic GRAID symbols.

Form symbols and specifiers

⟨dem_pro⟩	demonstrative pronoun
⟨rel_pro⟩	relative pronoun
⟨pn_np⟩	proper name
⟨ap_other⟩	adjective phrase as predicate (⟨:pred⟩)
⟨conj_other⟩	conjunction
⟨dem_other⟩	adverbial demonstrative (locative, temporal)
⟨intrg_other⟩	interrogative phrase
⟨pl_⟩	<i>specifier</i> : plural form, attaches to ⟨np⟩, ⟨pro⟩, and ⟨θ⟩

Function symbols and specifiers

⟨:s_ds⟩	subject of a verb of speech
---------	-----------------------------

Clause boundary symbols

⟨#rc_rn⟩	relative clause as a clausal subconstituent of a NP, to the right of the head
----------	---

Subconstituent symbols

⟨_adj⟩	attributive adjective; attaches to ⟨rn⟩
⟨_num⟩	attributive numeral; attaches to ⟨ln⟩

Other symbols

⟨rv-pro⟩	bound verbal cross-index; reflects properties of another argument by combining with certain person/animacy symbols (⟨_1⟩, ⟨_2⟩, ⟨_h⟩, and ⟨_d⟩) and function symbols (⟨_s⟩, ⟨_s_ds⟩, ⟨_a⟩, and ⟨_p⟩), e.g. ⟨lv-pro_h_s⟩; should not be conflated with corresponding nominal forms
⟨rv-pl_pro⟩	same as above, but plural
⟨nc_⟩	<i>specifier</i> : marks form glosses with RefIND indices in segments otherwise not considered (i.e. those marked ⟨#nc⟩)

B List of abbreviated morphological glosses

1PL.EX	1 st person plural exclusive
1PL.IN	1 st person plural inclusive
1SG	1 st person singular
2PL	2 nd person plural
2SG	2 nd person singular
3PL	3 rd person plural
3SG	3 rd person singular
4PL.PRON	4 th person plural object pronoun; it follows the verb complex in clauses with a topical subject of a 3 rd person and refers to more than one entity; in the case that the subject is plural, it can also refer to a single entity
4SG.PRON	4 th person singular object pronoun follows the verb complex in clauses with a topical subject of a 3 rd person and refers to a single entity under the condition that the topical subject also refers to a single entity
ADJR	adjectivaliser, suffix deriving adjectives
ADVR	adverbialiser, prefix deriving adverbs
ANA	anaphoric adnominal demonstrative
and1	conjunction coordinating clauses and NPs
and2	conjunction coordinating words and phrases
and3	conjunction coordinating nouns and pronouns that refer to specific persons
and4	conjunction coordinating clauses referring to a sequence of events
ART1–3	basic specific article of NPs headed by a noun of the 1 st (ART1), 2 nd (ART2), and 3 rd (3) class, respectively; with non-topical objects it is only used if the topical subject is a 1 st or 2 nd person
CAUS	causative prefix
COMPL1	conjunction introducing a verbal complement clause without a subject position; it denotes an intended action to be performed by the subject or object referent of the preceding superordinate clause
COMPL2	conjunction introducing complement clauses with a subject position
CONSEC	consecutive marker; the conjunction <i>re</i> ‘then’ indicates that an event was or will be the consequence of a past, present, or future state of affairs; when used with the 1 st person inclusive pronoun it is often translated by ‘let’s’; it holds the second position in the clause (Wackernagel position)
CONT	an adverb that is incorporated in the the verb complex; it indicates that the situation or event denoted by the predicate last for a while
DEM1–6	six different kinds of demonstrative
DEM	diminutive particle preceding the head of NPs
DYAD	a noun of the 2 nd class referring to a group of two or more people who are in a reciprocal kin relationship, e.g. mother and child
EMPH2	the emphatic particle <i>kou</i>
GOAL	the preposition <i>vo</i> ‘to a place’ is mostly incorporated in the verb complex; it governs an adverb, a NP without an article, a prepositional phrase, or a complement clause as its complement, see the gloss like
IMM	immediateness marker indicating that something has just happened or will happen immediately
INDEF2–3	indefinite determiner preceding an NP head of the 2 nd and the 3 rd class, respectively

IPFV	imperfective aspect marker; in the realis aspect marked by <i>na</i> TAM2 it has special forms for the 3 rd person singular, the 1 st person plural inclusive, and the 3 rd person plural and is accordingly glossed as 3SG.IPFV, 1PL.IN.IPFV, or 3PL.IPFV
like	(i) the verb <i>potee</i> ‘be like, be similar to’; (ii) the preposition <i>vo</i> ‘like’, which is mostly incorporated in the verb complex and governs an adverb, a NP without an article, or a prepositional phrase as its complement, see GOAL
LNK	linker, it combines with adverbs and prepositional phrases introduced by the multi-purpose preposition <i>te</i> PREP; the linker construction functions as an attribute and in combination with TAM markers as a predicate
MULT	prefix of verbs indicating that an action, process, or state of being is repeated and involves two or more participants
NSPEC2–3	non-specific article of NPs headed by a noun of the second and third class, respectively
OBJ.ART1–3	a specific object article of NPs with a head of the 1 st , 2 nd , and 3 rd class respectively, which marks a non-topical object in clauses with a topical subject of the 3 rd person, see 4SG.PRON and 4PL.PRON above
OBJM	object marker; a particle – mostly cliticized – that cross-references 1 st and 2 nd person objects and objects referring to a plurality of human beings; the latter kind of object may formally be a singular NP that typically refers to a particular group of people
one2–3	the numeral <i>peha</i> ‘one’; used with NP heads of the 2 nd class (one2); the numeral <i>peho</i> ‘one’, used with NP heads of the 3 rd class (one3)
ORD	prefix deriving ordinal numerals from cardinal numerals
other2–3	the determiner <i>meha</i> ‘other’ preceding a NP head of the 2 nd (other2) and the 3 rd (other3) class, respectively
PART.ART	partitive article, <i>sa</i> ‘some’, used with mass nouns
PLM	plural marker
POSS	a suffix inflecting for person and number attached to the possessee of an inalienable possessive construction and cross-referencing the possessor; an exception is the 1 st person singular possessor that directly follows the possessee
PREP	the multi-purpose preposition <i>te</i>
PRON	pronoun
TAM1	pre-head tense-aspect-mood marker indicating that the event will happen in the future
TAM2	pre-head tense-aspect-mood marker used in affirmative sentences expressing that something happened or existed or, when combined with the imperfect aspect marker, that something usually happens, is happening, or exists
TAM3	pre-head tense-aspect-mood marker indicating a change of situation either in the past or the future
TAM4	pre-head tense-aspect-mood marker indicating that an event which directly or indirectly involves the addressee may or will take place in the future
two2	the numeral <i>bua</i> ‘two’; used with NP heads of the 2 nd class (the corresponding numeral <i>buo</i> two3 used with NP heads of the 3 rd class is not attested in the four texts in the Multi-CAST Teop corpus)
ZERO	an empty argument position which relates to an unambiguously retrievable argument
NC	not classified

Multi-CAST

Multilingual Corpus of Annotated Spoken Texts



multicast.aspra.uni-bamberg.de/