

Multi-CAST Kamang annotation notes

Katherine Walker, George Saad, Antoinette Schapper, Clemens Mayer

August 24, 2026

Contents

1	Introduction	3
2	Language profile	3
3	Referential expressions	3
3.1	Nouns and pronouns	3
3.2	Core arguments	4
3.3	Non-core arguments	5
4	Verbs	6
4.1	Basic verb structure	6
4.2	Person/number prefixes	6
4.3	Applicativising prefixes	8
5	Complex and non-verbal predicates	8
5.1	Cognition and emotion predicates	8
5.2	Serial Verb Constructions	9
5.3	Non-verbal predicates	12
6	Clause types and clause combining	12
6.1	Clause types	13
6.2	Clause combining	13
7	References	13
8	Appendices	15
8.1	List of corpus-specific GRAID symbols	15
8.1.1	Pronouns	15
8.1.2	Person/number prefixes	15
8.1.3	Other form symbols	15
8.1.4	Functions	16
8.2	List of abbreviated morphological glosses	16

1 Introduction

This document provides an outline of Kamang morphosyntax and describes the implementation of the GRAID glossing conventions (Haig and Schnell, 2014) and RefIND annotation scheme (Schiborr et al., 2018). A brief profile of Kamang is given in given in Section 2, and the specific application of GRAID annotations to particular morphosyntactic features are given in Sections 3-6. All grammatical analysis is based on the sketch grammar of Kamang by Schapper (Schapper, 2014). A full grammar is in preparation by the same author.

Some of the texts in this collection contain words and phrases in Indonesian/Malay, the local lingua franca. Malay items integrated in otherwise Kamang clauses are glossed ⟨BI⟩ for ‘Bahasa Indonesia’ and annotated with GRAID as usual. Whole clauses in Malay are not annotated, and sometimes not translated.

2 Language profile

Word order in Kamang is primarily verb-final. Core arguments are freely elided, but when they are present the order is APV or SV. Kamang has few grammatical affixes. On nouns, these are limited to a prefix indexing the possessor. Verbs have a suffix slot for a (non-obligatory) aspectual suffix and up to three prefix slots. The slot closest to the verb may contain one of six different series that index the person and number of the S or P argument (S, A, P are used as specified in Haig and Schnell, 2014 § 2.4). The middle slot may contain one of two different multi-functional prefixes, *mi-* and *wo-*. The left-most slot is for ‘self-benefactive’ (SBEN) prefixes. Kamang makes use of serialisation to express various aspectual categories, for instance, and to add arguments such as a theme argument in ditransitives (Kamang has no true trivalent verbs [see Klamer and Schapper, 2012], perhaps with the exception of *wolom* ‘to name’; see Section 3.2).

3 Referential expressions

3.1 Nouns and pronouns

Kamang has various sets of pronouns: an additional tag followed by an underscore before the basic form tag ⟨pro⟩ identifies which pronoun paradigm the form comes from. (Full list of tags in Section 8.1; full paradigms in Schapper, 2014, ch. 5.) Example (1) shows a third person pronoun *gera* from the focus pronoun paradigm with the form tag ⟨fc_pro⟩.

(1) *Dum kiding lami kui yera ilaisak...*

<i>dum</i>	<i>kiding</i>	<i>lami</i>	<i>kui</i>	<i>gera</i>	<i>ilai_ilai</i>
child	small	male	dog	3.CONTR	look
## np.h:dt_a	rn	rn	rn_np.n	fc_pro.h:a	-0_p v:pred
0007			0004	0000	

‘The boy and his dog look...’

[mc_kamang_frog_0005]

Example (1) also shows a pattern in which two expressions correspond to a single argument function in the same clause. When, like in (1), two adjacent co-referential expressions occur clause-initially, the ‘dislocated topic’ tag ⟨dt⟩ (with a function extension, e.g. ⟨dt_a⟩) is used for the first one. The second one is annotated as usual with the basic function tag only, i.e. ⟨a⟩.

Where the adjacent co-referential expressions are not clause-initial, the ⟨appos⟩ tag is used for the second expression (the ordering is an arbitrary convention). In (2), two adjacent expressions of P are not

clause-initial, because they are preceded by the A argument. The first P argument *Takailubui* is tagged ⟨p⟩, while the second *geifu* ‘they together’ is tagged ⟨appos⟩.

(2) *gafaa Takailubui geifu mauuha...*

=bo	<i>gafaa</i>	<i>Takailubui</i>	<i>geifu</i>	<i>mauu_mauu</i>	<i>-ih</i>
=REL	3.INDP	Takailubui	they_together	war	-PURP
#cc:p	=other	pl_q_pro.h:a	pl_np.h:p	qg_pro.h:appos	-pl_0_h_p v:pred
0005					0005

‘(They think) that if they alone make war with Takailubui...’ [mc_kamang_war_0006]

Kamang also has three demonstratives that can be used either adnominally (tagged ⟨rn⟩ for ‘right of noun’) and pronominally (tagged ⟨dem_pro⟩, see e.g. (12)). Noun phrases may also be marked for definiteness and specificity with clitics =*ak* ‘DEF’ and =*a* ‘SPEC’, both tagged ⟨=rn⟩ as in (4), for example. See Schapper (2014, ch. 4.3) for details.

Noun phrases include compounds such as *ili salii* ‘water pipe (lit. water bamboo pipe)’, see (6), or other modifiers, as in *dum kiding lami* ‘boy (lit. child small male)’ in (1). The first element is annotated with function tags, and following elements are tagged ⟨rn⟩, along with form and animacy tags. This includes nouns in coordinated noun phrases: in (1), *kui* ‘dog’ is also tagged ⟨rn_np.n⟩, since it forms a coordinated noun phrase ‘the boy and his dog’.

3.2 Core arguments

The tag ⟨s⟩ is assigned to the single argument of intransitive verbs. This assignment is based primarily on the semantics of the verb, with addition evidence provided by a person/number prefix that indexes the S argument, shown in (3a) (see Section 4.2), though this is not always present for all verbs; see (3b).

(3) a. *kuia yetak.*

<i>kui</i>	=a	<i>ge-</i>	<i>tak_tak2</i>
dog	=SPEC	3.GEN	get_up_and_run
##	np.an:s	=rn	-gen_pro_n_s v:pred
0004		0004	

‘the dog also starts to run.’ [mc_kamang_frog_0046]

b. *Sukuu nok diiak...*

<i>sukuu</i>	<i>nok</i>	<i>dii_dii</i>	=ak
hole	one	lie	=DEF
##	np:s	rn	-0_s v:pred =other
0017		0017	

‘There is a hole...’ [mc_kamang_frog_0031]

The functions ⟨a⟩ and ⟨p⟩ are assigned to the more agent-like and more patient-like arguments of transitive verbs, respectively. The functions are assigned based on the presence of a prefix: if there is a person/number prefix on the verb (which is not always the case; see Section 4.2), this indexes the P argument, never the A argument. This is shown in (4), where the prefix *ga-* ‘3.PAT-’ indexes the P argument ‘tree’.

(4) *Naabo ga bonga woyayok.*

<i>naabo</i>	<i>ga</i>	<i>bong</i>	<i>=a</i>	<i>wo-</i>	<i>ga-</i>	<i>yok_yok</i>
nevertheless	3.AGT	tree	=SPEC	3.LOC	3.PAT	shake
## other	wpro.n:a	np:p	=rn	-lv_wo	-pat_pro_p	v:pred
	0004	0013			0013	

‘Then he shakes the tree.’

[mc_kamang_frog_0036]

We consider a single predicate *lom* ‘say’ to have two P arguments when used in a construction involving the prefix *wo-* (see Section 4.3) that means ‘to name someone something’. As illustrated in (5), ⟨p⟩ is used for the noun *nei* ‘name’, and ⟨p2⟩ for the name itself, in this case *Tuttaa*.

(5) *Ang ganei wolom ee "Tuttaa"*

<i>ang</i>	<i>ga-</i>	<i>nei</i>	<i>wo</i>	<i>lom_lom</i>	<i>ee</i>	<i>Tuttaa</i>
DIST.NWKN	3.INAL	name	3.LOC	say	HESIT	Tuttaa
## ln_dem_pro:poss	-ln_pat_pro_poss	np:p	-pro_p	v:pred	nc	np:p2
0026	0026	0027	0027			

‘It is called Tuttaa.’

[mc_kamang_aribei_0057]

Note that verbs like ‘give’, which typically take two P arguments in other languages, occur in serial verb constructions in Kamang (see Section 5.2).

Difficulties sometimes arise in identifying core arguments due to: 1) lack of gender/number distinction in third-person prefixes and pronouns, 2) absence of independent expression of A and/or P, 3) absence of a person/number indexing prefix on the verb (conditioned by various factors, including lexical preference of the verb, animacy of the P argument and event semantics [see Walker et al., 2024; Walker, 2024]). In transitive clauses, these factors can lead to difficulty in identifying A and P (particularly in scenarios in which both roles are third person). In some clauses, it may be unclear whether the verb is transitive or intransitive. In unclear cases, we rely on our understanding of the discourse context, verb semantics, other uses of the verb or whether it is listed as transitive or intransitive in the lexicon.

One additional clue is the set of agentive pronouns. Though not all A (or agentive S) arguments are encoded with agentive pronouns, they cannot be used for P arguments. Basic pronouns, on the other hand, may encode S, A, or P arguments. An example of an agentive pronoun, tagged ⟨wpro⟩, is shown in (4) above.

3.3 Non-core arguments

Non-core arguments often occur with a postposition, tagged ⟨adp⟩, as shown in (6).

(6) *Ili salii mi gatoota dii.*

<i>ili</i>	<i>salii</i>	<i>mi</i>	<i>ga-</i>	<i>toota_toota</i>	<i>dii_dii</i>
water	bamboo_pipe	IN	3.PAT	lay_crosswise	lie_down
## 0.h:a	np:l	rn_np	adp	0:p	-pat_pro_p
0002	0014		0011	0011	v:pred_svc
					rv_v

‘At the water pipe, he set it down.’

[mc_kamang_saiko_0021]

Possession is usually expressed in Kamang by indexing the possessor with the patientive prefix (for inalienable possession) or genitive prefix (for alienable possession) on the possessed noun. Some nouns may also form possessive constructions without prefixes, in which case the possessor is obligatory. In

constructions with prefixes, the possessor may optionally be expressed by a preposed NP and/or pronoun; see e.g. (5) for an example with a pronominal possessor. Note that, where the possessor is expressed by the prefix only, as in example (17), we do not add a zero independent expression of the possessor.

A possessive pronoun series also exists, tagged ⟨*poss_pro*⟩ when used pronominally (like English ‘mine’; Schapper, 2014:318). The third person form with the specific clitic *geen=a* ‘3.POSS=SPEC’ is occasionally used as a (non-referential) discourse marker, in which case it is tagged ⟨*other*⟩.

4 Verbs

4.1 Basic verb structure

Kamang verbs are generally clause-final and may have prefixes, suffixes or neither. Most morphology is annotated separately from the verb root, using an initial hyphen ⟨-⟩ to indicate affix status. The hyphen occurs at the beginning of the tag, regardless of whether it is a suffix or a prefix. Person/number prefixes receive special tagging and are discussed in Section 4.2. Other prefixes and suffixes are tagged ⟨-lv⟩ and ⟨-rv⟩, respectively. Applicativising prefixes have an additional tag; see Section 4.3. Examples of verb suffixes tagged ⟨-rv⟩ include the purposive suffix *-ih* in (2) and the imperfective suffix *-si* in (9).

Some verb lexemes have more than one stem; for instance, a handful of verbs have perfective and imperfective stems such as *sua* ‘arrive.PFV’ vs *sue* or *see* ‘arrive.IPFV’. In order to identify the lexeme to which a verb stem belongs, verbs in the grammatical word tier have the structure ‘verb_lexeme’, as in *sua_see* ‘arrive.PFV’, where *sua* is the form that is uttered, and *SEE* the arbitrarily chosen lexeme. For homophonous verb lexemes, a number is used to differentiate them; for instance, **tak_tak1* refers to *tak* ‘see’, while *tak_tak2* refers to *tak* ‘get up and run’. The asterisk in, e.g., **tak_tak1* indicates that this lexical item is obligatorily prefixed.

4.2 Person/number prefixes

Kamang has six prefix series that index the person/number of the S or P argument. A number of different conditions determine the use of a particular prefix series or whether the prefix slot is empty. The Multi-CAST Kamang corpus was created to explore these factors for the study reported in Walker et al. 2024.¹ The prefix series are indicated with an additional tag before the initial form tag ⟨*pro*⟩. Absence of a form tag indicates the most common prefix paradigm, locative, while the other series are tagged as shown in Table 1.² The structure of person/number prefix tags is given in (7).

- (7) Structure of person/number prefix tag
 -[series]_pro_[animacy]_[function]
 Eg.: ⟨-pat_pro_h_s⟩
 ‘prefix in the patientive paradigm; third person pronoun; human; single participant of intransitive clause’

¹The results reported in Walker et al., 2024 are based on a previous version of the annotated corpus, cited as ‘Schapper et al. in prep’. There are minor differences: in the current published version, the animacy tag for non-human animates is ⟨.n⟩ (replacing ⟨.an⟩); text names are in the format [mc_kamang_NAME_UTTERANCE.ID] (replacing [kamang_NAME]); and various tags have been updated to conform to GRAID standards.

²Here we follow the series designations locative, patientive, genitive, dative and self-benefactive given in Schapper (2014). The uses of the series go beyond what these labels suggest, and other publications reflect this in different nomenclature: Roman numerals (Fedden et al., 2013; Fedden et al., 2014) or theme vowel (Walker et al., 2024).

Table 1: Kamang person/number prefix series and associated GRAID tags

Series	GRAID tag
LOCATIVE	unmarked; i.e. ⟨-pro_⟩
PATIENTIVE	⟨-pat_⟩
GENITIVE	⟨-gen_⟩
DATIVE	⟨-dat_⟩
DIRECTIVE	⟨-dir_⟩
ASSITIVE/COMITATIVE	⟨-com_⟩
SELF-BENEFACTIVE	⟨-ben_⟩

Like independent arguments, we also add a zero person/number prefix when the verb is unmarked: examples of ‘zero’ prefixes can be seen in (1), (2) and (3b). This is with the exception of positional/posture/directional/aspect ‘particles’ in SVCs, which never bear prefixes, and therefore are not annotated with zero prefixes either (see Section 5.2). The decision to add ‘zero’ prefixes in GRAID was to facilitate analysis of prefixing rates for our project, and does not reflect any theoretical claims about unprefixing verbs.

The prefix series listed in Table 1 all index S or P arguments on the verb, with the exception of the self-benefactive paradigm. The referent of this prefix paradigm ‘is semantically an agentive participant who acts in its own interest and syntactically fills either S or A role of a verb’ (Schapper, 2014: 331). In this use, it is tagged with the syntactic function goal, e.g. ⟨-ben_pro_g⟩, and co-refers to the S or A argument in the same clause. In such cases, the ⟨p⟩ slot is available and is tagged as usual for a person/number P prefix, see (8). A ‘zero’ P prefix is added if the slot is empty.

(8) *Ye’gatak.*

	<i>ge’-</i>	<i>ga-</i>	<i>*tak_tak1</i>
	3.SBEN-	3.PAT-	see
##	0.h:a	0.n:p	-ben_pro_h_g -pat_pro_n_p v:pred
	0007	0004	0007 0004

‘He looked at him.’

[mc_kamang_frog_0019]

However, for a small number of verbs, the self-benefactive paradigm does not index S or A. For *baa* ‘say’, shown in (9), the prefix indexes the addressee, which is semantically like a ⟨g⟩ (goal), but syntactically like ⟨p⟩ since it is indexed with a prefix like other ⟨p⟩ arguments. For verbs like *baa* ‘say’, the function is given as ⟨p_g⟩.

(9) *...gal ge’baasi*

<i>=bo</i>	<i>gal</i>	<i>ge’-</i>	<i>baa_baa1</i>	<i>-si</i>
=LNK	3	3.SBEN	say_to	IPFV
#ac	=other	pro.h:a	0.h:p	-ben_pro_h_p_g v:pred -rv
	0006	0001	0001	

‘(...in order to) say to her’

[mc_kamang_crocodile_0018]

4.3 Applicativising prefixes

The prefixes *mi-* and *wo-* are tagged ⟨-lv_mi⟩ and ⟨-lv_wo⟩, respectively. They have various functions (see Schapper, 2014: 328–331), which are not differentiated in the tags. Schapper glosses *mi-* ‘in’ and *wo-* ‘at’; for expediency, the GRAID corpus uses the glosses ‘APPL’ for *mi-* and ‘3.LOC’ for *wo-*, regardless of function. One of the functions is as an applicative, adding a P argument to an otherwise intransitive verb. The original S argument is tagged ⟨a⟩ and the applicative argument is tagged ⟨p⟩. An example is given in (10): the verb *silang* ‘descend’ is intransitive, but with the addition of *mi-* the clause has both an A and a P argument (the latter included as a zero prefix).

- (10) ...*nok nal misilang*.

<i>nok</i>	<i>nal</i>	<i>mi-</i>	<i>silang_sila</i> (ng)
one	1SG	APPL	descend

##ds np.h:a pro.1:p -lv_mi -0_1_p v:pred
0029 0006 0006

‘...one will follow in my footsteps (lit. descend from me).’ [mc_kamang_crocodile_0021]

Other functions do not involve changing the argument structure of the verb. For example, in (11), the intransitive verb *wai* ‘return’ has the prefix *mi-*, but remains intransitive, reflected in the function tag ⟨:s⟩ for the (omitted) independent argument, and ⟨_s⟩ for the person/number indexing prefix. The function of *mi-* in this example is likely to ‘signal that the event denoted by a verb occurs ‘back’, ‘in return’” (Schapper, 2014: 328).

- (11) “...*misiwaa isi*.”

=nte	<i>mi-</i>	<i>si-</i>	<i>wai_wai</i>	-si
=but	APPL	1PL.INCL.PAT	return	IPFV

##ds =other pl_0.1:s -lv_mi -pl_pat_pro_1_s v:pred -rv
0004 0004

“...and then we will come back.” [mc_kamang_aribei_0008]

The prefix *wo-* is sometimes ambiguous between *wo-* ‘in’ and the person/number index from the locative series *wo-* ‘3.LOC’. Given our limited understanding of Kamang prefixing behaviour, we simply consider all instances of *wo-* to be person/number indexes except when it is followed by a person/number prefix from a different series (e.g. in (4)). Note that the sequence **wo-wo-* ‘in-3.LOC’ is not permitted.

5 Complex and non-verbal predicates

5.1 Cognition and emotion predicates

In Kamang, we consider cognitive and emotion predicates to be transitive, with both A and P arguments. The stimulus is coded as ⟨:a⟩, since it shares morphosyntactic properties with A arguments in canonical transitive clauses: it tends to occur clause-initially and is not indexed on the verb. The nature of the P argument varies. In the cognition predicate *king wo* ‘like’ (lit. ‘thought exist’, annotated as a kind of serial verb construction; see Section 5.2) in (12), the semantic experiencer ‘1PL.EXCL’ is tagged ⟨:p⟩. Like canonical P arguments, it is indexed on the verb with a person/number prefix. Note that in this example it occurs clause-initially before the A argument, rather than in the standard word order APV. In (12), the demonstrative *angak* ‘it’ is the stimulus argument, tagged ⟨:a⟩.

(12) ...*nila angak niking wo*

=sih	nil	=a	ang	=ak	ni-	king_king	wo_wo
STHAT	1PL.EXCL	SPEC	DIST.NKWN	DEF	1PL.EXCL.GEN	thought	exist
#ac	=other	pl_pro.1:p	=rn	dem_pro:a	=rn	-pl_gen_pro_1_p	v:pred_svc rv_vother
	0056		0053		0056		

‘...we like it’

[mc_kamang_childhood_0044]

In other such predicates, the P argument is a body part belonging to the experiencer. In (13), the predicate *lak (wo)* ‘know (exist)’ requires the inalienably possessed P argument *-eng* ‘eye’, tagged ⟨:p⟩. The experiencer (the ‘knower’) is coded as a possessor, which is expressed as a possessor prefix as well as an independent argument (twice in example (13), once as displaced topic and once preceding the possessed noun ‘eye’). The body parts in these constructions are not considered referential (i.e., they receive no refIND index).

(13) ...*na yelaka nenglak wore*

na	ye-	lak	=a	na	n-	*eng
1SG	3.AL	language	SPEC	1SG	-1SG.INAL	eye
wpro.1:dt_poss	-ln_gen_pro_poss	np:a	=rn	ln_wpro.1:poss	-ln_pat_pro_1_poss	np:p
0001	0009	0004		0001	0001	
	lak_lak	wo_wo	-re			
	know	exist	AUX			
	-0_p	v:pred_svc	rv_v	-rv_aux		

‘...(so that) I would know their language.’

[mc_kamang_childhood_0005]

5.2 Serial Verb Constructions

Schapper describes several types of serial verb constructions (SVCs) in Kamang (Schapper, 2014, ch. 8). For the purposes of GRAID annotations, there are two options for annotation, depending on whether the argument structure of the main verb is maintained or matched by the other verbal element, or an argument is added to the basic argument structure of the main verb. The former are treated as a single clause; the latter are treated biclausally.

Types of SVCs in which the argument structure of the main verb is maintained or matched are listed in Table 2. All types have in common the fact that the main verb is tagged ⟨v:pred_svc⟩; how the second element is tagged varies, as shown in the table. Examples of aspectual and parallel serialisations are provided below.

Example (14) shows an aspectual SVC.³ The transitive verb *baila* ‘buy’ is the main verb, which is tagged ⟨v:pred_svc⟩ to indicate that it is part of an SVC. It is followed by *lai* ‘finish’, which denotes completive aspect and is tagged as an auxiliary to the right of the verb: ⟨rv_aux⟩. Only the main verb can have a prefix; that is, no ‘zero’ prefix is added to *lai* as the second element in a serial verb construction. This is the case for any element that is tagged ⟨aux⟩.

³The gloss ‘x’ in (14) is used for inaudible elements or repairs (e.g., false starts).

Table 2: SVC types in Kamang treated as a single clause (for full description, see Schapper, 2014, ch. 8)

SVC type	non-main verbs	tag for non-main verb	example
Aspectual	<i>lai</i> ‘finish’ = completive aspect <i>dii</i> ‘lie (posture)’ = continuative aspect <i>koo</i> ‘stay, live’ = progressive aspect	$\langle rv_aux \rangle$, $\langle lv_aux \rangle$	(14)
Resultative	intransitive verb encoding result of previous verb: of S if intransitive, of P if transitive	$\langle rv_v \rangle$	(6): <i>dii</i> ‘lie down’ = final position of P of main verb
Adverbial	monovalent verbs expressing manner in which event is carried out, e.g. <i>de’tang</i> ‘limp’, <i>likka</i> ‘firm’, etc.	$\langle lv_v \rangle$	-
Motion	motion, direction verb, e.g. <i>silang</i> ‘descend’, <i>yaa</i> ‘go until’, <i>sue</i> ‘arrive’, etc.	$\langle rv_v \rangle$, $\langle lv_v \rangle$	-
Parallel	no restrictions	$\langle rv_v \rangle$	(15)

(14) *Nuaanana gebaila lai...*

nuaanana ge ge ge- baila_baila lai_lai1
 possessions x x 3.GEN buy finish
 ## pl_0.h:a np:p nc nc -gen_pro_p v:pred_svc rv_aux
 0005 0010 0010

‘When they had bought (and sold) their things...’

[mc_kamang_aribei_0045]

Another type of SVC that is annotated as a single clause involves two adjacent verbs with the same argument structure; we term these ‘parallel’ SVCs. The verbs may have similar meanings and form a lexicalised unit, such as *ah wei* ‘eat and drink’, shown in (15). Or the verbs may have different meanings but the same argument structure and form a single unit that we annotate as an SVC in the same way as (15). That is, the first verb in the SVC is tagged $\langle v:pred_svc \rangle$ and the second as $\langle rv_v \rangle$. Both verbs may bear prefixes – no special tags differentiate the prefix of a ‘main’ verb from the other verb in such constructions. The fact that the first and not the second verb is tagged as the predicate is arbitrary.

(15) *...gal ge’gaah ge’gaweisi.*

*=sih gal ge’- ga- *ah_ah ge’-*
 =STHAT 3 3.SBEN 3.PAT eat 3.SBEN
 #ac =rv pl_pro.h:s -pl_ben_pro_h_g -pl_pat_pro_h_s v:pred_svc -pl_ben_pro_h_g
 0001 0001 0001 0001
*ga- *wei_wei -si*
 3.PAT wash_up IPFV
 -pl_pat_pro_h_s rv_v -rv
 0001

‘...such that they could eat and drink for themselves.’

[mc_kamang_saiko_0016]

SVCs that are annotated as two clauses – a main and a dependent clause – involve the light verb *me* ‘TAKE’ (lexeme: $\langle_me2\rangle$; the lexeme ‘come’ is $\langle_me1\rangle$). It has a number of uses: it may add an argument

in a serialisation construction with a mono- or bivalent lexical verb (see Schapper, 2014: 340–342), or it may flag an argument that is already part of the lexical semantics of the main verb. We treat *me* as the predicate of a dependent clause, with the lexical verb that follows the predicate of the main clause. This is a practical solution that allows us to keep track of roles and referents; it is not intended to reflect theoretical claims about the status of *me* constructions as biclausal or otherwise.

An example of *me* adding an argument in an SVC with an intransitive verb *yaangme* ‘come down’ is given in (16). The first clause, which contains *me*, is considered the ‘main’ clause, annotated with the main clause tag ⟨##⟩. This clause contains an A (*wati* ‘sun’) and a P (*gebeba* ‘his heat’) argument, followed by the verb *me* ‘TAKE’, which is tagged ⟨vother:pred⟩ to reflect its status as a light verb. It does not bear a prefix. This clause is followed by what we annotate as an ‘adverbial clause’ ⟨#ac⟩, which contains the intransitive lexical verb, tagged ⟨v:pred_svc⟩. This verb bears a prefix, annotated with the special function tag ⟨s_m⟩; that is, an S argument in a *me* construction. The S argument may not be realised independently (it is coreferent with P in the *me* clause, *gebeba* ‘his heat’) and is added as a ‘forced zero’ with the tag ⟨f0_s_m⟩. Again, this tag indicates an S argument in a *me* construction.

(16) *Anna wosol wati gera ge- beba gaima*

anna	wosol	wati	gera	ge-	beba	gaima
then	beginning	sun	3.CONTR	3.AL	hot	all_(of_a_certain_group)
##	other	other	np.d:dt_a	fc_pro.d:a	-ln_gen_pro_d_poss	np:p rn
		0001	0001	0001		0008
<i>me</i>			ge-	<i>yaangme_yaangme</i>		
TAKE			3.GEN	come_down		
vother:pred	#ac	f0:s_m	-gen_pro_s_m	v:pred_svc		
		0008	0008			

‘Then the sun released all his heat...’

[mc_kamang_sun_0007]

Example (17) shows serialisation in which *me* adds an argument to a transitive verb, *ite* ‘put inside’ (here meaning ‘give (a name)’). The similarities to *me* constructions with intransitive verbs are that the main verb is also tagged ⟨v:pred_svc⟩, and the argument in the main clause receives a special tag indicating that it is part of a *me* construction. Here, this tag is ⟨p_m2⟩, used for both the prefix and the independent argument. The independent argument may be realised in *me* constructions with transitive verbs. Differently to intransitive verbs like in (16), transitive verbs in *me* constructions are within the main clause, tagged ⟨##⟩. The main clause contains an A and a P argument. The *me* clause is tagged ⟨#ac⟩, and contains a third argument, tagged with the function ⟨p⟩. Here, it is the recipient of the name. The A argument, coreferent with A in the main clause (first person in (17)), may not be realised, and is annotated as ⟨f0⟩.

(17) *ahbo nerunga nok me yanei wute*

ahbo		ne-		rung	=a	nok	me_me2
then		1SG.AL		child	SPEC	one	TAKE
##	other	0.1:a	#ac	f0.1:a	-ln_gen_pro_1_poss	np.h:p	=other rn vother:pred %
		0001		0001	0001		0041
ya-		nei		wute_ite			
3.INAL		name		put_inside			
-ln_pat_pro_h_poss		np:p_m2	-0_p_m2	v:pred_svc			
0041		0042	0042				

‘Then I gave that name to my son.’

[mc_kamang_childhood_0037]

The third type of *me* construction is annotated as monoclausal, because *me* simply flags an argument present in the underlying argument structure of the main verb in a ‘displaced theme construction’ (Schapper, 2014: 341). In (18), the (unrealised) P argument of *mota* ‘throw implement’ is flagged with *me*, even though the verb is transitive and *me* is not required to add an argument. The P argument in such constructions is tagged ⟨p_m1⟩ and *me* is tagged ⟨lv_vother⟩. The main verb is tagged ⟨v:pred_svc⟩.

- (18) *me’ mota...*
- | | | |
|----|---|------------------|
| | <i>me’_me2</i> | <i>mota_mota</i> |
| | TAKE | throw_implement |
| ## | 0.h:a 0:p_m1 lv_vother -0_p_m1 v:pred_svc | |
| | 0008 0019 | 0019 |
- ‘(and it is) thrown away...’ [mc_kamang_barter_0026]

5.3 Non-verbal predicates

In Kamang, words from other classes, such as nouns (19a) and adpositions (19b), can be used as predicates with no additional marking. We tag the noun or adposition with the predicate function – ⟨:pred⟩ – in these instances. If an auxiliary suffix is present, it is tagged ⟨-rn_aux⟩, shown in (19a). A further example can be found below in (20b) where the interrogative *maau* ‘who’ is annotated ⟨other:pred⟩.

- (19) a. *Kollau male fondamante...*
- | | | |
|---------------------------|------------|------------|
| <i>Kollau male fon</i> | <i>-da</i> | <i>-ma</i> |
| Kollau female young_woman | AUX | PFV |
| ## np.h:s np:pred rn_np.h | -rn_aux | -rn |
| 0001 | | |
- ‘When Kollau was a young woman...’ [mc_kamang_crocodile_0002]
- b. *Male dum gatok mi.*
- | | | |
|--|------------|-----------|
| <i>male dum ga-</i> | <i>tok</i> | <i>mi</i> |
| female child 3.INAL- | stomach | IN |
| ## ln_np.h:poss np.h:s -ln_pat_pro_h:poss np:l | adp:pred | |
| 0002 0003 0002 | | |
- ‘A woman was pregnant.’ [mc_kamang_piee_0003]

In (19a), the nominal predicate is a compound; the first noun is tagged as the predicate, the second as ⟨rn⟩ ‘right of noun’. The auxiliary suffix is common, but not obligatory to form nominal predicates. Note that the tag ⟨adp:pred⟩ is a departure from basic GRAID annotation, where the NP in a PP is given the ⟨pred⟩ function tag. Since in Kamang nouns are sometimes elided in adpositional predicates, we avoid the case of having a ⟨0:pred⟩ for the elided noun and instead label the adposition ⟨adp:pred⟩. This also introduces the tags ⟨radp⟩ and ⟨ladp⟩ for ‘right of adposition’ and ‘left of adposition’, respectively.

Though only verbs take person/number prefixes to index S/P arguments, two of the prefix series occur on nouns to index the possessor. Possessor prefixes remain tagged with the function ⟨_poss⟩ in nominal predicates, shown in (19b).

6 Clause types and clause combining

6.1 Clause types

Any of the standard tags for clause types – main or subordinate – can be combined with the <ds> tag to indicate that it is direct speech. Note that direct speech clauses are not considered complements of speech verbs; that is, *lom* ‘say’ is considered intransitive (when no prefix is present; see Section 3.2), and what is said is considered a main clause, shown in the adjacent utterances in (20a-20b).

- (20) a. *male gera gal lom_lom -si*
 female 3.CONTR 3 say IPFV
 ## np.h:dt_s fc_pro.h:dt_s pro.h:s -0_h_s v:pred -rv
 0001 0001 0001 0001
 ‘the woman said:’ [mc_kamang_crocodile_0011]
- b. *“Al maau?”*
 al maau
 2SG who
 ##ds pro.2:s other:pred
 0006
 “‘Who are you?’” [mc_kamang_crocodile_0012]

We use <#nc> ‘not considered’ to tag clauses that cannot be fully annotated – because they are fragments or not understandable, for example.

6.2 Clause combining

Kamang has a number of forms used to combine clauses. These are included in the clause that follows; that is, a clause boundary may occur prior to a clause-combining clitic such as *=nte* ‘but’, shown in (11) above. (For the sake of simplicity, such clitics and their hosts are not given in the first line of examples in this guide.) Table 3 shows a number of such forms and the kinds of clauses they introduce. They are annotated as <=other>.

Table 3: List of clause combining forms

form	gloss	use	example
<i>=bo</i>	REL	‘relativiser’; introduces relative clauses	-
<i>=bo</i>	LNK	‘linker’; links two clauses of various types	(9)
<i>=nte</i>	BUT	combines either two main clauses or main and (adverbial) subordinate clause	(11)
<i>=sa</i>	COMP	‘complementiser’; generally introduces a complement clause	-
<i>=sih</i>	STHAT	‘so that; such that’; often introduces an adverbial clause	(15)

7 References

Klamer, M., & Schapper, A. (2012). ‘Give’ constructions in the Papuan languages of Timor-Alor-Pantar. *Linguistic Discovery*, 10(3), 174–207. <https://doi.org/10.1349/PS1.1537-0852.A.421>

- Fedden, S., Brown, D., Corbett, G., Holton, G., Klamer, M., Robinson, L. C., & Schapper, A. (2013). Conditions on pronominal marking in the Alor-Pantar languages. *Linguistics*, 51(1), 33–74. <https://doi.org/10.1515/ling-2013-0002>
- Fedden, S., Brown, D., Kratochvíl, F., Robinson, L. C., & Schapper, A. (2014). Variation in pronominal indexing: Lexical stipulation vs. referential properties in Alor-Pantar languages. *Studies in Language*, 38(1), 44–79. <https://doi.org/10.1075/sl.38.1.02fed>.
- Haig, G., & Schnell, S. (2014). *Annotations using GRAID (Grammatical Relations and Animacy)*.
- Schapper, A. (2014). Kamang. In A. Schapper (Ed.), *The Papuan languages of Timor, Alor and Pantar: Volume 1: Sketch Grammars* (pp. 285–350). De Gruyter Mouton.
- Schiborr, N. N., Schnell, S., & Thiele, H. (2018). *RefIND — Referent Indexing in Natural-language Discourse: Annotation guidelines (v1.1)*. University of Bamberg/University of Melbourne.
- Walker, K. (2024). Differential indexing in Kamang: A viewpoint alternation. *Linguistics Vanguard*, 10(s1), 85–96. <https://doi.org/doi:10.1515/lingvan-2023-0052>
- Walker, K., Faghiri, P., & Van Lier, E. (2024). Argument Indexing in Kamang. *Studies in Language*, 48(2), 287–350. <https://doi.org/10.1075/sl.21077.wal>

8 Appendices

8.1 List of corpus-specific GRAID symbols

The following subsections list the non-standard GRAID symbols used in the annotation of the Multi-CAST Kamang corpus. Refer to the GRAID manual (Haig and Schnell, 2014: 54–55) for an inventory of basic GRAID symbols.

8.1.1 Pronouns

Tag	Meaning
⟨wpro⟩	agentive pronouns ('weak pronouns')
⟨refl_pro⟩	reflexive pronoun
⟨poss_pro⟩	possessive pronoun
⟨q_pro⟩	quantifying pronoun 'alone'
⟨qd_pro⟩	quantifying pronoun 'dual'
⟨qa_pro⟩	quantifying pronoun 'all'
⟨qg_pro⟩	quantifying pronoun 'group'
⟨fr_pro⟩	focus pronoun 'restrictive'
⟨fc_pro⟩	focus pronoun 'contrastive'
⟨fo_pro⟩	focus pronoun 'other'

8.1.2 Person/number prefixes

Structure: -[series]_pro_[animacy]_[function]

Tag	Prefix series
unmarked	LOCATIVE
⟨-pat_⟩	PATIENTIVE
⟨-gen_⟩	GENITIVE
⟨-dat_⟩	DATIVE
⟨-dir_⟩	DIRECTIVE
⟨-com_⟩	ASSISTIVE/COMITATIVE

8.1.3 Other form symbols

Tag	Meaning
⟨f0⟩	structurally suppressed argument slot of a predicate
⟨mi, wo⟩	prefixes <i>mi-</i> , <i>wo-</i> with various functions, including as applicatives
⟨.n⟩	animacy value for non-human animals
⟨nc⟩	not considered
⟨other⟩	parts of speech not covered by the basic and Kamang-specific tag set
⟨pl⟩	plural referent (unmarked = singular)
⟨vother⟩	light verb <i>me</i> 'TAKE'

8.1.4 Functions

Tag	Meaning
⟨ p_g ⟩	semantic Goal argument treated syntactically like a P argument (may also be combined with ⟨ _m1 ⟩ or ⟨ _m2 ⟩)
⟨ p_m1 ⟩	P argument of a transitive verb in a monoclausal serialisation with light verb <i>me</i> , where <i>me</i> flags the P argument
⟨ p_m2 ⟩	P argument of a transitive main verb in a serialisation with light verb <i>me</i> , where <i>me</i> adds an additional argument annotated in a subordinate clause
⟨ p2 ⟩	[name] in the construction ‘to name someone/something [name]’
⟨ s_m ⟩	S argument of an intransitive main verb in a serialisation with light verb <i>me</i>
⟨ svc ⟩	element in a serial verb construction

8.2 *List of abbreviated morphological glosses*

_XXX	verb lemma	INDEF	indefinite
1, 2, 3	first, second, third person	INDP	independently; on one’s own
AGT	agentive	INTJ	interjection
AL	alienable	IPFV	imperfective
APPL	applicative	KWN	known
ASSERT	assertive	LNK	linker
ASSOC	associative	LOC	locative
AT	postposition ‘at’	NEG	negation
AUX	auxiliary	NKWN	not known
BI	Malay loan (‘Bahasa Indonesia’)	ORD	ordinal
CLF	classifier	PAT	patientive
CMN	common	PFV	perfective
COMPL	completive	PL	plural
COMP	complementiser	POSS	possessor
CONTR	contrastive	PURP	purposive
CVDUP	duplication of consonant and vowel	RDP	reduplication
DAT	dative	REL	relativiser
DEF	definite	RESTR	restricted
DIST	distal	RFLX	reflexive
DU	dual	SBEN	self-benefactive
EMPH	emphatic	SET	setting; backdrop
EXCL	exclusive	SG	singular
EXCLAM	exclamative	SP	species
FOC	focus	SPEC	specific
GEN	genitive	STHAT	so that; such that
HES	hesitation	TAG	question tag
HUM	human	TAKE	light verb ‘take’
IN	postposition inside; within	TIME	marks time after which an event occurs
INAL	inalienable	TOWARDS	postposition ‘towards’
INCL	inclusive	X	inaudible