Multi-CAST

Nafsan

annotation notes

Thieberger Nick

Brickell Timothy

January 2021
v1.1

ARC CENTRE OF EXCELLENCE FOR
THE DYNAMICS OF LANGUAGE

Australian Government
Australian Research Council

University of Bamberg

DFG

Multi-CAST *Multilingual Corpus of Annotated Spoken Texts*

Citation for this document

Nick, Thieberger & Timothy, Brickell. 2021. Multi-CAST Nafsan annotation notes. In Haig, Geoffrey & Schnell, Stefan (eds.), *Multi-CAST: Multilingual corpus of annotated spoken texts.* (`multicast.aspra.uni-bamberg.de/#nafsan`) (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The Multi-CAST collection has been archived at the *University of Bamberg*, Germany, and is freely accessible online at `multicast.aspra.uni-bamberg.de/`.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 Nafsan annotation notes v1.1 last updated 27 January 2021
This document was typeset by NNS with XeLaTeX and the *multicast3* class (v3.2.3).

Contents

1 Introduction

The following comprises selected notes on the GRAID (Haig & Schnell 2014) and RefIND (Schiborr et al. 2018) annotations of Nafsan. It corresponds 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.

1.1 Language and speakers

Nafsan is an Oceanic language spoken by aprroximately 5 000 people on the island of Efate in central Vanuatu. It was formerly known as 'South Efate', which is also the name the language is referred to in Thieberger's (2006) descriptive grammar.

1.2 Typological profile

Nafsan is a mostly isolating-analytic language. Bound morphology is restricted to person indexes, namely subject-indexing proclitics that occupy the first slot in the verb complex (VC) and possessor suffixes that attach to possessed nouns. Constituent order in the clause is quite fixed, with only a small set of ordering variations involving fixed alternative slots for certain constituents. The canonical order is SVO. Subject and object are essentially encoded by their fixed position relative to the verbal predicate. Adpositions are preposed to their NP complement. Prepositional flagging is used with oblique arguments and different kinds of adjuncts. There is a pre-clausal, left-dislocated position where different types of phrases can occur. Clause combining is mostly paratactic, involving strings of finite clause constructions. However, Nafsan also has a clause-chaining construction that involves sequential non-finite predicates. A similar construction employs a so-called 'echo-subject' proclitic that occupies the same slot as the subject-indexing (and TAM-marking) proclitics.

2 Clause structure and arguments

2.1 Basic verbal clause structure

Nafsan has SVO word order in verbal clauses, and alignment is accusative. Hence, both S and A arguments (which may or may not be expressed) precede the verb complex whose first constituent is an obligatory portmanteau proclitic combining a subject-indexing and a tense, aspect, mood (TAM) morpheme. S and A are thus encoded by their pre-verbal position and cross-referencing by the subject proclitic. Examples (1) and (2) illustrate this.

(1) *malnen ntwam ipam nua nait iskei …*

	malnen	*ntwam*	*i=*	*pam*	*nua*	*nait*	*i=*	*skei*
	as	devil	1S.RS=	eat	fruit	fig_tree	3S.RS=	one
#ac	other	np.d:a	=lv	v:pred	np:p	rn_np	=rn	rn_np

'As the devil ate the fig, …' [mc_nafsan_ntwam_0041]

(2) *me nmatu gaag nen itae weswes wi.*

	me	*nmatu*	*gaag*	*nen*	*i=*	*tae*	*weswes*	*wi*
	but	wife	2S.POS	that	3S.RS=	know	work	good
##ds	other	np.h:s	rn_pro.2:poss	rn	lv=	aux	v:pred	rv

'But your wife here can work well.' [mc_nafsan_nmatu_0011]

Free pronouns with A or S function occupy the exact same slot as do full NPs in these functions, regardless of their person, as can be seen in (3–6).

(3) *me malfanen ga kipe pam nataṁol ilatol.*

	me	*malfane*	*ga*	*ki=*	*pe*	*pam*	*nataṁool*	*i=*	*latol*
	but	now	3S	3S.PS=	PF	eat	person	3S.RS=	eight
##ds	other	other	pro.h:a	=lv	lv	v:pred	np.h:p	=rn	r_np

'But now he has eaten eight men.' [mc_nafsan_ntwam_0045]

(4) *ga iur etan ...*

	ga	*i=*	*ur*	*etan*
	3S	3S.RS=	go_along	down
##	pro.d:s	=lv	v:pred	np:l

'He flies below, …' [mc_nafsan_maal_0012]

(5) *e, me kineu akano pam nua nait.*

	e	*me*	*kineu*	*a=*	*kano*	*pam*	*nua*	*nait*
	hey	but	1S	1S.RS=	unable	eat	fruit	fig_tree
##ds	other	other	pro.1:a	=lv	aux	v:pred	np:p	rn_np

'Hey, but I cannot eat Nait figs.' [mc_nafsan_ntwam_0040]

(6) *me kineu atap nrogteesawes mau.*

	me	*kineu*	*a=*	*tap*	*nrogteesa*	*-wes*	*mau*
	but	1S	1S.RS=	NEG1	feel_bad	-3PL.OBL	NEG2
##neg	other	pro.1:s	=lv	other	v:pred	-pro.h:obl	other

'But I didn't feel bad about them.' [mc_nafsan_tafra_0010]

Where S or A are left zero we put the zero annotation in this same slot, as illustrated in (7) and (8).

(7) *rapreg nasuṁ gar, ...*

		ra=	*preg*	*nasuṁ*	*gar*
	0	3D.RS=	make	house	3P.POS
##	0.h:a	=lv	v:pred	np:p	rn_pro.h:poss

'They made their house …' [mc_nafsan_ntwam_0002]

(8) *ifla tu msak, ...*

		i=	*fla*	*tu*	*msak*
	0	3S.RS=	CND	stay	sick
##	0.h:s	=lv	lv	v:pred	rv

'If he is sick, …' [mc_nafsan_lisau_0012]

Note that the proclitics that cross-reference S and A are not annotated, given that they are absolutely obligatory and hence predictable from the presence of whatever annotation is found for the respective clause-level slot, be that NP, pronoun, or zero. In other words, the S and A function in Nafsan are never entirely zero, since they are minimally represented by the proclitic,[1] and where the NP or pronoun occur on clause level, this entails multirepresentation of these functions. For P arguments, the realisation patterns are somewhat more intricate: pronominal P arguments either occur as a free form in the same position as a free NP, as in (9), or they are realised as pronominal suffixes on the verb, as in (10).

(9) a. *kin i= wel ag, …*

	kin	*ZERO*	*i=*	*wel*	*ag*
	REL	0	3S.RS=	like	2S
#rc	other	0.d:a	=lv	v:pred	pro.2:p

‘(… it is the spirit) who is like you …’ [mc_nafsan_lelep_0020]

b. *me waak p̃ur iskei ip̃as komam.*

	me	*waak*	*p̃ur*	*i=*	*skei*	*i=*	*p̃as*	*komam*
	but	pig	big	3S.RS=	one	3S.RS=	chase	1P.EX
##	other	np:s	rn	=rn	rn_np	=;v	v:pred	pro.1:p

‘… but a pig chased us.’ [mc_nafsan_ntwam_0020]

(10) a. *Atua ikano mai watgik.*

	atua	*i=*	*kano*	*mai*	*wat*	*-gik*
	god	3S.RS=	unable	come	hit	-2S.O
##	np.d:a	=lv	lv	lv	v:pred	-pro.2:p

‘God cannot hurt you.’ [mc_nafsan_lelep_0021]

b. *waak p̃ur nen imai, kaip̃asir.*

	waak	*p̃ur*	*nen*	*i=*	*mai*			*kai=*	*p̃as-i*	*-r.*
	pig	big	that	3S.RS=	come	0		ES=	chase-TS	-3P.O
##	np:s	rn	rn	=lv	v:pred	##	es_f0:a	=lv	v:pred	-pro.h:p

‘The big pig came and chased them.’ [mc_nafsan_ntwam_0007]

Where a P argument is left zero, the respective annotation in GRAID aligns with the clause-level position, as in (11). Note that this zero annotation here means that there is no form of expression for P, neither on clause level nor within the verb morphology.

(11) *me rupami p̃ulp̃og.*

	me		*ru=*	*pam-i*		*p̃ulp̃og*
	and	0	3P.RS=	eat-TS	0	morning
##	other	0.h:a	=lv	v:pred	0:p	np:other

‘… and they ate (it) in the morning.’ [mc_nafsan_kori_0012]

Overt suffixes can never co-occur with overt free object pronouns or object NPs, which is what

1 Note that this is what is in many descriptive frameworks accounted for by assuming that such cross-indexes are “the actual arguments”. That the proclitics themselves are not annotated in GRAID is merely a practical consideration and should not be confused with any theoretical standpoint that would counter such a descriptive account.

motivates our annotation practice.

2.2 Three-participant and benefactive constructions

Nafsan three-participant constructions exhibit a direct-indirect and a double-object pattern. Crucially, in the direct-indirect pattern, it is the theme argument that is encoded as a kind of oblique argument, a prepositional phrase headed by *ki*, as in (12), that Thieberger (2006: 189) terms "Theme/Instrument". This construction is relatively rare in the Multi-CAST corpus of Nafsan, and there is only one example with zero objects expressing the recipient, as in (12). Its function is captured as ⟨:obl⟩ in GRAID, taking note of the prepositional marking.

(12) *me rutua kin*

	me		*ru=*	*tua*		*ki*	*-n*
	and	0	3P.RS=	give-TS	0	PREP	-3S.O
##	other	0.h:a	=lv	v:pred	0.d:p	adp	-pro.h:obl

'… and they gave him her.' [mc_nafsan_litog_0024]

In the alternative construction, the theme is expressed by a plain NP that follows the NP expressing the recipient, without a preposition *ki*. This unmarked NP is considered to be a secondary object in GRAID annotations, hence applying the function gloss ⟨:p2⟩, as shown in (13) and (14):

(13) *Kineu katuok nmatu neu, …*

	kineu	*katuok*	*nmatu*	*neu*
	1S	1S.IRS=	give-V	-2S.O
##	pro.1:a	=lv	v:pred	-pro.2:p

'I will give you my wife.' [mc_nafsan_nmatu_0014]

(14) *Kafo tuok nalkes.*

		ka=	*fo*	*tu-o*	*-k*	*nalkes*
	0	1S.IRS=	PSP.IR	give-TS	-2S.O	herbs
##	0.1:a	=lv	lv	v:pred	-pro.2:p	np:p2

'I will give you herbs.' [mc_nafsan_maal_0004]

Hence, the difference in function glossing here takes note of the presence versus absence of prepositional flagging of the theme argument expression.

The same three-participant construction is used to express acts of communication, where information can be seen as being transferred from one party to the other, and the same annotation practice is applied here, as shown in (15) and (16):

(15) *… kefo nrik mam ki naɸet nafsan*

		ke=	*fo*	*nrik*	*-mam*	*ki*	*naɸet*	*nafsan*
	0	3S.IRS=	PSP.IR	tell	-1PL	PREP	meaning	story
#cc	pro.h:a	=lv	lv	v:pred	-pro.1:p	adp	np:obl	rn

'… he would tell us the meaning of that story, …' [mc_nafsan_tafra_0011]

(16) *... kin maarik nen inrikin kin*

	kin	*maarik*	*nen*	*i=*	*nrik-i*	*-n*	*ki*	*-n*
	REL	husband	that	3S.RS=	tell-TS	-3S.O	PREP	-3S.O
#rc_rn	other	np.h:a	rn	=lv	v:pred	-pro:p	adp	-pro:obl

'(I didn't understand) what the man said to it.' [mc_nafsan_nmatu_0022]

A further type of three-participant construction in Nafsan involves the expression of a beneficiary by special kind of pronoun that occupies a preverbal position within the verb complex, as shown in (17), where it refers to the person for whom the mats are made. The form receives a specifier ⟨lv_⟩, thus noting its VC-internal position. In these examples the function gloss of the benefactive pronoun is ⟨:obl⟩.

(17) *malen kin ruga pregptaki m̃iit, ...*

	malen	*kin*		*ru=*	*ga*	*pregpta-ki*	*m̃iit*
	then	SUB	0	3P.RS=	3S.BEN	make_good-TR	mat
##	other	other	0.d:a	=lv	lv_pro.h:obl	v:pred	np:p

'Then they prepared mats for her.' [mc_nafsan_litog_0018]

The benefactive construction can also be used to express recipients or addressees in communicative events, as in (18). In these cases we gloss the function as ⟨:g⟩ rather than ⟨:obl⟩ since this is a kind of oblique argument whose semantic role is a goal in the wider sense, hence encompassing recipients and addressees.

(18) *... ineu trausi teetwei*

		i=	*neu*	*traus-i*		*teetwei*
	0	3S.RS=	1S.BEN	tell-TS	0	before
##	0.h:a	=lv	lv_pro.1:g	v:pred	0:p	other

'(the story of Willi Santo,) (he) told (it) to me long ago.' [mc_nafsan_nmatu_0001]

2.3 Echo-subject constructions

Nafsan has a non-finite clause construction that involves a so-called 'echo subject marker'. An echo subject marker occupies the same syntactic slot as a subject-indexing/TAM proclitic, and thus the predicate does not inflect for these categories and is hence non-finite, with the subject argument being suppressed. The use of an echo subject marker entails that the would-be subject of this clause is co-referent with that of the preceding (finite) clause. This can be seen in (19), where the people who go are the same as the ones who took their garden stuff. This construction is treated in GRAID as embedded (with single ⟨#⟩), and there is no regular zero subject. Instead, the annotation ⟨f0⟩ 'forced zero' is used according to general GRAID conventions for non-finite constructions, and it takes a preposed specifier ⟨es_⟩ to identify this as an echo subject.

(19) *raslat sernale ni talm̃aat gar kaipa.*

```
   0       ra=      slat      sernale      ni  talm̃at  gar
   0       3D.RS=   take      everything   of  garden  3P.POS
## 0.h:a   =lv      v:pred    np:p         rn  rn_np   rn_pro.h:poss

                kai=  pa
   0            ES=   go
## es_f0.h:s   =lv   v:pred
```

'They carried everything for their garden and they went.' [mc_nafsan_ntwam_0004]

This construction is not very frequent in the current Nafsan corpus in Multi-CAST, comprising only 15 instances.

2.4 Non-verbal and copular clauses

In Nafsan, NPs can function as the predicate of non-verbal clauses, as shown in (20) where the predicate of a non-verbal classificatory clause is the NP 'idol'. The subject here is zero. Such predicates are glossed ⟨np:pred⟩.

(20) *... na natap ...*

```
        na              natap
        COMP  0         idol
#cc:p   other 0.d:s     np:pred
```

'(I find) that (it is) an idol.' [mc_nafsan_lelep_0019]

More typically, nominal predicates occur in copular clauses, following onto the copula *pi*. Copular clauses are verbal clauses, and the copular predicate is inflected for subject/TAM values. The post-copula NP bears the actual predicative function and is glossed as the one in a non-verbal clause without copula; the copula receives the gloss ⟨cop⟩, as shown in (21).

(21) *waak nen ipi waak nmatu.*

```
    waak  nen   i=       pi   waak      nmatu
    pig   that  3S.RS=   be   pig       woman
##  np:s  rn    =lv      cop  np:pred   rn_np
```

'That pig was a female pig.' [mc_nafsan_nmatu_0007]

2.5 Presentational constructions

Somewhat unusual for an Oceanic language, Nafsan has a verb 'have'. Example (22) illustrates its use in a canonical transitive clause expressing possession, and the GRAID annotation is that of a normal transitive clause:

(22) *me ipiatlak kori sees iskei.*

```
    me          i=       piatlak     kori  sees   i=       skei
    and   0     3S.RS=   have       dog   small  3S.RS=   one
##  other 0.h:a =lv      v:predex    np:p  rn     =rn      rn_np
```

'And (she) had a small dog.' [mc_nafsan_kori_0007]

The same verb is also used in what is treated as a presentational construction 'there is'. Structurally, a presentational construction is still transitive, but it lacks a referential subject. Hence, it contains a preverbal proclitic and a post-verbal NP, but only the latter is referential. In GRAID annotations, we treat the construction as intransitive-presentational (similar to the (Southwestern) German or French presentational constructions with verbs meaning 'have'), so that only the post-verbal NP is treated as an argument, receiving the function gloss ⟨:s⟩; the predicate has the function ⟨:predex⟩, see (23).

(23) *me ipiatlak nataṁool nen ...*

	me	*i=*	*piatlak*	*nataṁool*		*nen*
	but	3S.RS=	have	person		REL
##	other	=lv	v:predex	np.h:s	#rc_rn	other

'But are there people who (give them…)' [mc_nafsan_lisau_0007]

3 Complex sentences

In Nafsan, complement clauses are introduced by one or more complementizers. The predicate of a complement clause is finite, taking the usual inflectional marking for subject number/person and TAM values, as can be seen in (24). The complement clause fulfils the function of a P argument, hence the glossing ⟨#cc:p⟩ is applied.

(24) *kitli na kefo sak mai.*

		ki=	*til-i*		*na*		*ke=*	*fo*	*sak*	*mai*
	0	3S.PS=	say-TS		COMP	0	3S.IRS=	PSP:IR	land	come
##	0.d:a	=lv	v:pred	#cc:p	other	0.d:s	=lv	lv	v:pred	rv

'He said (that) he would land.' [mc_nafsan_kori_0026]

In contrast to object NPs, complement clauses in P function can be cross-referenced by an overt object suffix on the verb, as can be seen in (25) and (26). In this case, we treat the suffix as the object and leave the function of the complement clause unannotated. Essentially this means that the correlation between suffix and complement clause is left open.[2]

(25) *amurin na katrausi tenatrauswen ...*

		a=	*muur-i*	*-n*		*na*		*ka=*	*traus-i*	*te-natrauswen*
	0	1S.RS=	want-TS	3S.O		COMP	0	1S.IRS=	tell-TS	DET-story
##	0.1:a	=lv	v:pred	-pro:p	#cc	other	0.1:a	=lv	v:pred	np:p

'I want to tell the story (of this man Wili Santo).' [mc_nafsan_nmatu_0001]

2 Alternatively one could add a tag to the ⟨cc⟩ gloss, such as ⟨#cc_p⟩, to make this distinction clear.

(26) *... go kinrikin kin na kefreg tenmatun.*

	go		*ki=*	*nrik-i*	*-n*		*kind*	*na*		*ke=*	*freg*
	and	0	3S.PS=	tell-TS	-3S.O		COMP	COMP	0	3S.IRS=	make:IR
##	other	0.h:a	=lv	v:pred	-pro:p	#cc	other	other	0:a	=lv	v:pred

te-namatun
DET-thing
np:p

'… and he said that she should make something.' [mc_nafsan_nmatu_0025]

Where transitive verbs of speaking are used to frame direct reported speech, we follow the same practice, taking the object suffix as the actual P argument and leaving the direct speech without a matrix function gloss (which is general GRAID practice anyway). Examples (27) and (28) illustrate this practice.

(27) *go ntwam nen itok kainrikin kin na, Me...*

	go	*ntwam*	*nen*	*i=*	*tok*		*kai=*	*nri-ki*	*-n*	*kin*	*na*
	and	devil	that	3S.RS=	stay	0	ES=	tell-TS	-3S.O	SUB	SUB
##	other	np.d:s	rn	=lv	v:pred	## es_f0.d:a	=lv	v:pred	-pro:p	#ds	other

me
but
other other

'And the devil (stayed and) said, "Have you (eaten already?)"'
[mc_nafsan_ntwam_0029]

(28) *kitok mrokin na, Nafte kin?*

		ki=	*tok*	*mro-ki*	*-n*		*na*	*nafte*	*kin*
	0	3S.PS=	PROG	think-TR	-3S.O		SUB	what	HES
##	0.h:a	=lv	aux	v:pred	-pro:p	#ds	other	np:pred	other

'… and he thought, "What's that?"' [mc_nafsan_ntwam_0019]

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.). 2019. *Multi-CAST: Multilingual Corpus of Annotated Spoken Texts*. (https://multicast.aspra.uni-bamberg.de/) (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 / University of Melbourne. (https://multicast.aspra.uni-bamberg.de/#annotations) (Accessed 2019-03-08).

Thieberger, Nick. 2006. *A grammar of South Efate: An Oceanic language of Vanuatu*. Honolulu: University of Hawaii Press. (http://hdl.handle.net/11343/31242) (Accessed 2019-07-31).

Thieberger, Nick & Brickell, Timothy. 2019. Multi-CAST Nafsan. In Haig, Geoffrey & Schnell, Stefan (eds.), *Multi-CAST: Multilingual Corpus of Annotated Spoken Texts*. (https://multicast.aspra.uni-bamberg.de/#nafsan) (Accessed 2019-07-31).

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 Nafsan 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
⟨es_f0⟩	echo subject, see Appendix 2.3
⟨dem_pro⟩	demonstrative pronoun
⟨pn_np⟩	proper name

Function symbols and specifiers

⟨:s_ds⟩	subject of a verb of speech

Clause boundary symbols

⟨#rc_rn⟩	relative clause as a subconstituent of a NP

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
ART	article
BEN	benefactive
BI	code-switching to Bislama
CND	conditional
COMP	complementizer
D	dual
DEM	demonstrative
DET	determiner
DP	direct possession
DST	distant
EMPH	emphasis
ES	echo subject
EX	exclusive
EXCL	exclamation
HAB	habitual
HESIT	hesitation
IN	inclusive
IR	irrealis
IRS	irrealis subject
LOC	locative
NEG	negation particle
NEG2	second negation particle
O	object
OBL	oblique
P, PL	plural
PF	perfect
PN	proper name
POS	possessive
PROG	progressive
PS	perfect subject
PSP	prospective
PURP	purpose
REFL	reflexive
REL	relativizer
RS	realis subject
S	subject
SG	singular
SUB	subordinator
TOP	topic marker
TR	transitivizer
TS	transitive suffix
V	epenthetic vowel preceding DP suffixes
NC	not classified

Multi-CAST

Multilingual Corpus of Annotated Spoken Texts

multicast.aspra.uni-bamberg.de/