

# Multi-CAST

## *Matukar Panau corpus counts*

*Danielle Barth*

*Kira Davey*

*Maria Matheas*

*November 2023*  
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*

Barth, Danielle & Davey, Kira & Matheas, Maria. 2023. Multi-CAST Matukar Panau corpus counts. In Haig, Geoffrey & Schnell, Stefan (eds.), *Multi-CAST: Multilingual corpus of annotated spoken texts*. ([multicast.aspra.uni-bamberg.de/#matukar](http://multicast.aspra.uni-bamberg.de/#matukar)) (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/](http://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The Multi-CAST collection has been archived at the *University of Bamberg*, Germany, and is freely accessible online at [multicast.aspra.uni-bamberg.de/](http://multicast.aspra.uni-bamberg.de/).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/](http://creativecommons.org/licenses/by/4.0/).

*Multi-CAST Matukar Panau corpus counts* v1.0 last updated 19 November 2023  
This document was typeset by NNS with  $\text{\LaTeX}$  and the *multicast3* class (v3.2.6).

## Contents

<b>1</b>	<b>Notes on the GRAID counts</b>	_____	<b>1</b>
<b>2</b>	<b>The Matukar Panau corpus</b>	_____	<b>2</b>
2.1	<i>bklife</i>		3
2.2	<i>fishing</i>		4
2.3	<i>kadagoi</i>		5
2.4	<i>manub</i>		6
2.5	<i>mariu</i>		7
2.6	<i>niu</i>		8
2.7	<i>ww2</i>		9
2.8	<i>yali</i>		10



## 1 Notes on the GRAID counts

This document collects tables with frequency counts for combinations of selected GRAID symbols in version 2311 (from November 2023) of the Multi-CAST Matukar Panau corpus. Unless a more recent version of this document exists, it also applies to any later versions of the annotations. Note that the tables are intended to offer only cursory impressions of the relative proportions between different types of referring expression. They do not provide exact summaries of the annotations.

Only a small number of basic GRAID symbols are counted:

### *Function symbols*

⟨0⟩	zero
⟨pro⟩	definite pronoun
⟨np⟩	full noun phrase
⟨other⟩	form not further specified

### *Person/Animacy symbols*

⟨.1⟩	first person
⟨.2⟩	second person
⟨.h⟩	third person, human
⟨.d⟩	third person, anthropomorphic
∅	third person, non-human

### *Function symbols*

⟨:s⟩	subject of an intransitive clause
⟨:a⟩	subject of a transitive clause
⟨:ncs⟩	non-canonical subject
⟨:p⟩	direct object
⟨:ob1⟩	oblique argument
⟨:g⟩	goal argument
⟨:l⟩	locational argument
⟨:pred⟩	predicate
⟨:poss⟩	possessive
⟨:other⟩	function not further specified

### *Clause boundary symbols*

⟨##⟩	independent clause
⟨#⟩	other clause

Only basic categories are listed; categories represented by complex symbols with additional specifiers (e.g. ⟨dem\_pro⟩ ‘demonstrative pronoun’) have been subsumed under the more basic category (e.g. ⟨pro⟩ ‘definite pronoun’). Please refer to the annotation notes for this corpus for information on all annotated categories, including those not listed here.

## 2 The Matukar Panau corpus

GRAID	<:s>	<:a>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:pred>	<:poss>	<:other>	<i>totals</i>
<∅ .1>	221	131	0	21	0	8	0	0	0	0	381
<∅ .2>	16	76	0	4	0	3	0	0	0	0	99
<∅ .h>	173	233	0	65	0	17	0	0	0	0	488
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	30	2	0	147	0	0	0	0	0	0	179
<pro .1>	41	17	0	0	3	1	0	0	81	0	143
<pro .2>	13	8	0	0	0	1	0	0	6	0	28
<pro .h>	66	27	0	0	1	0	0	0	34	0	128
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	28	1	0	9	1	0	1	1	3	0	44
<np .h>	107	37	0	55	8	12	0	3	4	5	231
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	124	6	0	246	22	107	97	75	0	62	739
<other .h>	0	0	0	0	0	0	0	0	0	0	0
<other .d>	0	0	0	0	0	0	0	0	0	0	0
<other>	2	3	0	7	0	1	60	20	0	487	580
<i>totals</i>	821	541	0	554	35	150	158	99	128	554	
<##>											0
<#>											1423
<i>totals</i>											1423

**Table 1** Summarized GRAID counts for the entire Matukar Panau corpus.

2.1 *bklife*

GRAID	<:s>	<:a>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:pred>	<:poss>	<:other>	<i>totals</i>
<∅ .1>	35	14	0	14	0	0	0	0	0	0	63
<∅ .2>	3	6	0	0	0	0	0	0	0	0	9
<∅ .h>	30	31	0	9	0	0	0	0	0	0	70
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	2	0	0	9	0	0	0	0	0	0	11
<pro .1>	7	2	0	0	0	0	0	0	23	0	32
<pro .2>	0	0	0	0	0	0	0	0	2	0	2
<pro .h>	9	2	0	0	0	0	0	0	3	0	14
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	0	0	0	0	0	0	0	0	0	0	0
<np .h>	13	10	0	10	1	0	0	0	2	1	37
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	20	0	0	25	0	17	11	4	0	5	82
<other .h>	0	0	0	0	0	0	0	0	0	0	0
<other .d>	0	0	0	0	0	0	0	0	0	0	0
<other>	1	0	0	0	0	0	18	3	0	63	85
<i>totals</i>	120	65	0	67	1	17	29	7	30	69	
<##>											0
<#>											188
<i>totals</i>											188

Table 2 Summarized GRAID counts for the *bklife* text.

## 2.2 fishing

GRAID	<:s>	<:a>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:pred>	<:poss>	<:other>	<i>totals</i>
<∅ .1>	42	31	0	1	0	0	0	0	0	0	74
<∅ .2>	1	0	0	0	0	1	0	0	0	0	2
<∅ .h>	5	4	0	1	0	2	0	0	0	0	12
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	3	0	0	15	0	0	0	0	0	0	18
<pro .1>	10	2	0	0	0	0	0	0	10	0	22
<pro .2>	2	0	0	0	0	0	0	0	0	0	2
<pro .h>	2	1	0	0	0	0	0	0	0	0	3
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	1	0	0	0	0	0	0	0	0	0	1
<np .h>	5	0	0	1	1	1	0	0	0	0	8
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	6	1	0	21	1	5	7	6	0	12	59
<other .h>	0	0	0	0	0	0	0	0	0	0	0
<other .d>	0	0	0	0	0	0	0	0	0	0	0
<other>	0	0	0	1	0	0	2	0	0	31	34
<i>totals</i>	77	39	0	40	2	9	9	6	10	43	
<##>											0
<#>											117
<i>totals</i>											117

**Table 3** Summarized GRAID counts for the *fishing* text.



2.3 *kadagoi*

GRAID	<:s>	<:a>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:pred>	<:poss>	<:other>	<i>totals</i>
<∅ .1>	18	18	0	2	0	0	0	0	0	0	38
<∅ .2>	0	0	0	0	0	0	0	0	0	0	0
<∅ .h>	10	7	0	4	0	1	0	0	0	0	22
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	3	0	0	5	0	0	0	0	0	0	8
<pro .1>	2	3	0	0	0	0	0	0	15	0	20
<pro .2>	0	0	0	0	0	0	0	0	0	0	0
<pro .h>	0	0	0	0	0	0	0	0	0	0	0
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	0	0	0	1	0	0	0	0	0	0	1
<np .h>	5	2	0	4	1	0	0	1	0	0	13
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	4	0	0	17	2	5	8	2	0	0	38
<other .h>	0	0	0	0	0	0	0	0	0	0	0
<other .d>	0	0	0	0	0	0	0	0	0	0	0
<other>	0	0	0	0	0	0	0	1	0	8	9
<i>totals</i>	42	30	0	33	3	6	8	4	15	8	
<##>											0
<#>											73
<i>totals</i>											73

Table 4 Summarized GRAID counts for the *kadagoi* text.

2.4 *manub*

GRAID	<:s>	<:a>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:pred>	<:poss>	<:other>	<i>totals</i>
<∅ .1>	7	2	0	0	0	0	0	0	0	0	9
<∅ .2>	0	0	0	0	0	0	0	0	0	0	0
<∅ .h>	25	23	0	2	0	1	0	0	0	0	51
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	1	0	0	3	0	0	0	0	0	0	4
<pro .1>	9	2	0	0	0	0	0	0	3	0	14
<pro .2>	5	0	0	0	0	0	0	0	1	0	6
<pro .h>	24	7	0	0	0	0	0	0	2	0	33
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	3	0	0	0	0	0	0	0	0	0	3
<np .h>	16	3	0	6	1	0	0	0	0	0	26
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	3	0	0	25	0	3	15	7	0	3	56
<other .h>	0	0	0	0	0	0	0	0	0	0	0
<other .d>	0	0	0	0	0	0	0	0	0	0	0
<other>	0	0	0	0	0	1	14	2	0	64	81
<i>totals</i>	93	37	0	36	1	5	29	9	6	67	
<##>											0
<#>											134
<i>totals</i>											134

**Table 5** Summarized GRAID counts for the *manub* text.

## 2.5 mariu

GRAID	<:s>	<:a>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:pred>	<:poss>	<:other>	<i>totals</i>
<∅ .1>	3	9	0	0	0	0	0	0	0	0	12
<∅ .2>	3	42	0	2	0	0	0	0	0	0	47
<∅ .h>	22	43	0	6	0	1	0	0	0	0	72
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	10	1	0	51	0	0	0	0	0	0	62
<pro .1>	0	1	0	0	0	0	0	0	6	0	7
<pro .2>	1	0	0	0	0	0	0	0	0	0	1
<pro .h>	11	12	0	0	0	0	0	0	3	0	26
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	18	0	0	2	1	0	0	0	0	0	21
<np .h>	12	3	0	6	0	1	0	0	0	1	23
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	38	5	0	54	8	15	15	22	0	8	165
<other .h>	0	0	0	0	0	0	0	0	0	0	0
<other .d>	0	0	0	0	0	0	0	0	0	0	0
<other>	0	3	0	1	0	0	6	8	0	92	110
<i>totals</i>	118	119	0	122	9	17	21	30	9	101	
<##>											0
<#>											258
<i>totals</i>											258

Table 6 Summarized GRAID counts for the *mariu* text.

2.6 *niu*

GRAID	<:s>	<:a>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:pred>	<:poss>	<:other>	<i>totals</i>
<∅ .1>	1	4	0	0	0	0	0	0	0	0	5
<∅ .2>	5	17	0	0	0	1	0	0	0	0	23
<∅ .h>	33	45	0	3	0	8	0	0	0	0	89
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	2	0	0	36	0	0	0	0	0	0	38
<pro .1>	0	1	0	0	0	0	0	0	1	0	2
<pro .2>	1	1	0	0	0	0	0	0	2	0	4
<pro .h>	3	1	0	0	0	0	0	0	0	0	4
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	1	1	0	2	0	0	0	1	0	0	5
<np .h>	10	5	0	2	0	2	0	0	0	0	19
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	20	0	0	34	0	13	6	12	0	10	95
<other .h>	0	0	0	0	0	0	0	0	0	0	0
<other .d>	0	0	0	0	0	0	0	0	0	0	0
<other>	1	0	0	0	0	0	8	2	0	64	75
<i>totals</i>	77	75	0	77	0	24	14	15	3	74	
<##>											0
<#>											161
<i>totals</i>											161

**Table 7** Summarized GRAID counts for the *niu* text.

## 2.7 ww2

GRAID	<:s>	<:a>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:pred>	<:poss>	<:other>	<i>totals</i>
<∅ .1>	59	20	0	1	0	2	0	0	0	0	82
<∅ .2>	1	1	0	1	0	0	0	0	0	0	3
<∅ .h>	6	28	0	5	0	0	0	0	0	0	39
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	6	1	0	13	0	0	0	0	0	0	20
<pro .1>	5	2	0	0	1	0	0	0	11	0	19
<pro .2>	0	0	0	0	0	0	0	0	0	0	0
<pro .h>	6	0	0	0	0	0	0	0	1	0	7
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	1	0	0	3	0	0	0	0	0	0	4
<np .h>	19	5	0	2	2	4	0	1	1	0	34
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	19	0	0	37	5	20	27	2	0	5	115
<other .h>	0	0	0	0	0	0	0	0	0	0	0
<other .d>	0	0	0	0	0	0	0	0	0	0	0
<other>	0	0	0	0	0	0	11	1	0	66	78
<i>totals</i>	122	57	0	62	8	26	38	4	13	71	
<##>											0
<#>											190
<i>totals</i>											190

Table 8 Summarized GRAID counts for the ww2 text.

2.8 *yali*

GRAID	<:s>	<:a>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:pred>	<:poss>	<:other>	<i>totals</i>
<∅ .1>	56	33	0	3	0	6	0	0	0	0	98
<∅ .2>	3	10	0	1	0	1	0	0	0	0	15
<∅ .h>	42	52	0	35	0	4	0	0	0	0	133
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	3	0	0	15	0	0	0	0	0	0	18
<pro .1>	8	4	0	0	2	1	0	0	12	0	27
<pro .2>	4	7	0	0	0	1	0	0	1	0	13
<pro .h>	11	4	0	0	1	0	0	0	25	0	41
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	4	0	0	1	0	0	1	0	3	0	9
<np .h>	27	9	0	24	2	4	0	1	1	3	71
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	14	0	0	33	6	29	8	20	0	19	129
<other .h>	0	0	0	0	0	0	0	0	0	0	0
<other .d>	0	0	0	0	0	0	0	0	0	0	0
<other>	0	0	0	5	0	0	1	3	0	99	108
<i>totals</i>	172	119	0	117	11	46	10	24	42	121	
<##>											0
<#>											302
<i>totals</i>											302

Table 9 Summarized GRAID counts for the *yali* text.



# Multi-CAST

*Multilingual Corpus of Annotated Spoken Texts*



[multicast.aspra.uni-bamberg.de/](http://multicast.aspra.uni-bamberg.de/)