

Multi-CAST

*Tondano
corpus counts*

Timothy Brickell

January 2021
v2.1



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

Brickell, Timothy. 2021. Multi-CAST Tondano corpus counts. In Haig, Geoffrey & Schnell, Stefan (eds.), *Multi-CAST: Multilingual corpus of annotated spoken texts*. (multicast.aspra.uni-bamberg.de/#tondano) (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The Multi-CAST collection has been archived at the *University of Bamberg*, Germany, and is freely accessible online at multicast.aspra.uni-bamberg.de/.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 Tondano corpus counts v2.1 last updated 27 January 2021
This document was typeset by NNS with $\text{X}_{\text{L}}^{\text{A}}\text{T}_{\text{E}}\text{X}$ and the *multicast3* class (v3.2.3).

Contents

1	Notes on the GRAID counts	1
2	The Tondano corpus	2
2.1	<i>gulamera</i>	3
2.2	<i>holiday</i>	4
2.3	<i>kiniar01</i>	5
2.4	<i>kiniar02</i>	6
2.5	<i>kiniar03</i>	7
2.6	<i>mapalus</i>	8
2.7	<i>water</i>	9
2.8	<i>watulaney</i>	10

1 Notes on the GRAID counts

This document collects tables with frequency counts for combinations of selected GRAID symbols in version 2101 (from January 2021) of the Multi-CAST Tondano 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

⟨:a⟩	subject of a transitive clause
⟨:s⟩	subject of an intransitive clause
⟨:ncs⟩	non-canonical subject
⟨:p⟩	direct object
⟨:ob1⟩	oblique argument
⟨:g⟩	goal argument
⟨:l⟩	locational argument
⟨:poss⟩	possessive
⟨:pred⟩	predicate
⟨: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 Tondano corpus

GRAID	<:a>	<:s>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:poss>	<:pred>	<:other>	<i>totals</i>
<∅ .1>	38	36	0	6	0	0	0	0	0	0	80
<∅ .2>	10	4	0	1	0	0	0	0	0	0	15
<∅ .h>	284	13	0	9	0	0	0	0	0	0	306
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	0	70	0	201	0	0	0	0	1	0	272
<pro .1>	31	52	0	13	0	1	0	59	0	0	156
<pro .2>	8	10	0	2	0	0	0	0	0	0	20
<pro .h>	215	83	0	12	2	0	0	19	0	0	331
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	3	62	0	77	0	1	0	37	0	3	183
<np .h>	32	26	0	14	13	2	0	11	9	1	108
<np .d>	1	0	0	0	0	0	0	0	0	0	1
<np>	5	93	0	271	21	51	130	15	65	69	720
<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	0	109	0	109
<i>totals</i>	627	449	0	606	36	55	130	141	184	73	
<##>											913
<#>											172
<i>totals</i>											1085

Table 1 Summarized GRAID counts for the entire Tondano corpus.

2.1 *gulamera*

GRAID	<:a>	<:s>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:poss>	<:pred>	<:other>	<i>totals</i>
<∅ .1>	1	0	0	0	0	0	0	0	0	0	1
<∅ .2>	0	0	0	0	0	0	0	0	0	0	0
<∅ .h>	34	0	0	0	0	0	0	0	0	0	34
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	0	15	0	32	0	0	0	0	0	0	47
<pro .1>	1	0	0	0	0	0	0	0	0	0	1
<pro .2>	0	0	0	0	0	0	0	0	0	0	0
<pro .h>	36	11	0	2	0	0	0	1	0	0	50
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	0	10	0	20	0	0	0	3	0	0	33
<np .h>	8	1	0	0	1	0	0	0	0	0	10
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	0	11	0	26	0	16	7	0	9	7	76
<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	0	10	0	10
<i>totals</i>	80	48	0	80	1	16	7	4	19	7	
<##>											112
<#>											17
<i>totals</i>											129

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

2.2 holiday

GRAID	<:a>	<:s>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:poss>	<:pred>	<:other>	<i>totals</i>
<∅ .1>	15	10	0	3	0	0	0	0	0	0	28
<∅ .2>	0	0	0	0	0	0	0	0	0	0	0
<∅ .h>	10	5	0	3	0	0	0	0	0	0	18
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	0	3	0	8	0	0	0	0	0	0	11
<pro .1>	4	14	0	1	0	0	0	7	0	0	26
<pro .2>	2	1	0	0	0	0	0	0	0	0	3
<pro .h>	4	6	0	1	0	0	0	3	0	0	14
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	0	1	0	2	0	0	0	2	0	0	5
<np .h>	2	4	0	0	1	0	0	2	0	0	9
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	1	7	0	18	4	0	27	0	8	11	76
<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	0	6	0	6
<i>totals</i>	38	51	0	36	5	0	27	14	14	11	
<##>											79
<#>											10
<i>totals</i>											89

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

2.3 *kiniar01*

GRAID	<:a>	<:s>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:poss>	<:pred>	<:other>	<i>totals</i>
<∅ .1>	3	1	0	0	0	0	0	0	0	0	4
<∅ .2>	0	0	0	0	0	0	0	0	0	0	0
<∅ .h>	61	0	0	0	0	0	0	0	0	0	61
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	0	8	0	48	0	0	0	0	1	0	57
<pro .1>	13	0	0	0	0	0	0	0	0	0	13
<pro .2>	3	2	0	1	0	0	0	0	0	0	6
<pro .h>	34	1	0	1	0	0	0	2	0	0	38
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	0	3	0	12	0	0	0	3	0	0	18
<np .h>	7	2	0	2	1	0	0	1	1	0	14
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	0	4	0	48	3	6	2	4	3	2	72
<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	0	8	0	8
<i>totals</i>	121	21	0	112	4	6	2	10	13	2	
<##>											124
<#>											18
<i>totals</i>											142

Table 4 Summarized GRAID counts for the *kiniar01* text.

2.4 *kiniar02*

GRAID	<:a>	<:s>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:poss>	<:pred>	<:other>	<i>totals</i>
<∅ .1>	0	1	0	0	0	0	0	0	0	0	1
<∅ .2>	0	0	0	0	0	0	0	0	0	0	0
<∅ .h>	70	0	0	0	0	0	0	0	0	0	70
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	0	15	0	54	0	0	0	0	0	0	69
<pro .1>	4	1	0	0	0	0	0	3	0	0	8
<pro .2>	0	1	0	0	0	0	0	0	0	0	1
<pro .h>	63	5	0	2	0	0	0	1	0	0	71
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	1	12	0	16	0	1	0	2	0	1	33
<np .h>	5	1	0	0	0	1	0	5	0	0	12
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	1	13	0	69	0	20	12	0	12	8	135
<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	0	13	0	13
<i>totals</i>	144	49	0	141	0	22	12	11	25	9	
<##>											158
<#>											35
<i>totals</i>											193

Table 5 Summarized GRAID counts for the *kiniar02* text.

2.5 *kiniar03*

GRAID	<:a>	<:s>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:poss>	<:pred>	<:other>	<i>totals</i>
<∅ .1>	1	0	0	0	0	0	0	0	0	0	1
<∅ .2>	1	0	0	0	0	0	0	0	0	0	1
<∅ .h>	37	1	0	0	0	0	0	0	0	0	38
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	0	1	0	25	0	0	0	0	0	0	26
<pro .1>	1	0	0	0	0	0	0	1	0	0	2
<pro .2>	0	1	0	0	0	0	0	0	0	0	1
<pro .h>	35	2	0	0	0	0	0	0	0	0	37
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	0	8	0	15	0	0	0	3	0	0	26
<np .h>	3	0	0	0	0	0	0	1	0	0	4
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	0	8	0	39	0	4	4	5	2	1	63
<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	0	8	0	8
<i>totals</i>	78	21	0	79	0	4	4	10	10	1	
<##>											87
<#>											12
<i>totals</i>											99

Table 6 Summarized GRAID counts for the *kiniar03* text.

2.6 *mapalus*

GRAID	<:a>	<:s>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:poss>	<:pred>	<:other>	<i>totals</i>
<∅ .1>	8	19	0	3	0	0	0	0	0	0	30
<∅ .2>	9	4	0	1	0	0	0	0	0	0	14
<∅ .h>	14	3	0	4	0	0	0	0	0	0	21
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	0	5	0	10	0	0	0	0	0	0	15
<pro .1>	7	24	0	12	0	1	0	6	0	0	50
<pro .2>	3	5	0	1	0	0	0	0	0	0	9
<pro .h>	12	11	0	2	1	0	0	2	0	0	28
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	0	8	0	2	0	0	0	2	0	0	12
<np .h>	3	3	0	2	2	1	0	0	3	1	15
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	2	8	0	16	4	2	23	1	14	13	83
<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	0	9	0	9
<i>totals</i>	58	90	0	53	7	4	23	11	26	14	
<##>											128
<#>											22
<i>totals</i>											150

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

2.7 water

GRAID	<:a>	<:s>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:poss>	<:pred>	<:other>	<i>totals</i>
<θ .1>	0	0	0	0	0	0	0	0	0	0	0
<θ .2>	0	0	0	0	0	0	0	0	0	0	0
<θ .h>	29	0	0	0	0	0	0	0	0	0	29
<θ .d>	0	0	0	0	0	0	0	0	0	0	0
<θ>	0	14	0	11	0	0	0	0	0	0	25
<pro .1>	0	0	0	0	0	0	0	0	0	0	0
<pro .2>	0	0	0	0	0	0	0	0	0	0	0
<pro .h>	5	0	0	0	0	0	0	0	0	0	5
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	2	15	0	7	0	0	0	10	0	2	36
<np .h>	0	0	0	0	0	0	0	0	0	0	0
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	1	14	0	18	3	2	10	3	6	6	63
<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	0	17	0	17
<i>totals</i>	37	43	0	36	3	2	10	13	23	8	
<##>											65
<#>											15
<i>totals</i>											80

Table 8 Summarized GRAID counts for the *water* text.

2.8 *watulaney*

GRAID	<:a>	<:s>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:poss>	<:pred>	<:other>	<i>totals</i>
<∅ .1>	10	5	0	0	0	0	0	0	0	0	15
<∅ .2>	0	0	0	0	0	0	0	0	0	0	0
<∅ .h>	29	4	0	2	0	0	0	0	0	0	35
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	0	9	0	13	0	0	0	0	0	0	22
<pro .1>	1	13	0	0	0	0	0	42	0	0	56
<pro .2>	0	0	0	0	0	0	0	0	0	0	0
<pro .h>	26	47	0	4	1	0	0	10	0	0	88
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	0	5	0	3	0	0	0	12	0	0	20
<np .h>	4	15	0	10	8	0	0	2	5	0	44
<np .d>	1	0	0	0	0	0	0	0	0	0	1
<np>	0	28	0	37	7	1	45	2	11	21	152
<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	0	38	0	38
<i>totals</i>	71	126	0	69	16	1	45	68	54	21	
<##>											160
<#>											43
<i>totals</i>											203

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

Multi-CAST

Multilingual Corpus of Annotated Spoken Texts



multicast.aspra.uni-bamberg.de/