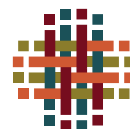


Multi-CAST

*Jinghpaw
corpus counts*

Keita Kurabe

September 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

Kurabe, Keita. 2021. Multi-CAST Jinghpaw corpus counts. In Haig, Geoffrey & Schnell, Stefan (eds.), *Multi-CAST: Multilingual corpus of annotated spoken texts*. (multicast.aspra.uni-bamberg.de/#jinghpaw) (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The Multi-CAST collection has been archived at the *University of Bamberg*, Germany, and is freely accessible online at multicast.aspra.uni-bamberg.de/.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 Jinghpaw corpus counts v1.1 last updated 13 September 2021
This document was typeset by NNS with \LaTeX and the *multicast3* class (v3.2.5).

Contents

1	Notes on the GRAID counts	1
2	The Jinghpaw corpus	2
2.1	<i>chyeju</i>	3
2.2	<i>dwi</i>	4
2.3	<i>galang</i>	5
2.4	<i>ganu</i>	6
2.5	<i>hkaili</i>	7
2.6	<i>hpaji</i>	8
2.7	<i>manau</i>	9
2.8	<i>natga</i>	10
2.9	<i>nchyang</i>	11
2.10	<i>nga</i>	12
2.11	<i>shanngayi</i>	13

1 Notes on the GRAID counts

This document collects tables with frequency counts for combinations of selected GRAID symbols in version 2108 (from August 2021) of the Multi-CAST Jinghpaw 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 Jinghpaw corpus

GRAID	<:s>	<:a>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:pred>	<:poss>	<:other>	<i>totals</i>
<∅ .1>	28	49	0	10	0	0	0	0	0	0	87
<∅ .2>	20	37	0	8	0	0	0	0	0	0	65
<∅ .h>	199	136	0	38	1	0	0	0	1	0	375
<∅ .d>	82	39	0	6	0	0	0	0	0	0	127
<∅>	52	2	0	119	1	8	3	0	0	1	186
<pro .1>	25	36	0	14	1	0	0	0	27	0	103
<pro .2>	18	29	0	8	0	0	1	0	15	0	71
<pro .h>	74	67	0	15	0	0	0	0	37	0	193
<pro .d>	29	16	0	9	0	0	0	0	12	0	66
<pro>	21	1	0	26	17	3	32	3	1	61	165
<np .h>	76	56	0	49	11	2	0	19	15	3	231
<np .d>	28	16	0	6	1	0	0	2	3	1	57
<np>	91	4	0	208	24	68	58	35	3	149	640
<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	14	0	2629	2643
<i>totals</i>	743	488	0	516	56	81	94	73	114	2844	
<##>											1081
<#>											195
<i>totals</i>											1276

Table 1 Summarized GRAID counts for the entire Jinghpaw corpus.

2.1 *chyeju*

GRAID	<:s>	<:a>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:pred>	<:poss>	<:other>	<i>totals</i>
<∅ .1>	1	16	0	3	0	0	0	0	0	0	20
<∅ .2>	2	6	0	3	0	0	0	0	0	0	11
<∅ .h>	3	1	0	0	0	0	0	0	0	0	4
<∅ .d>	30	16	0	2	0	0	0	0	0	0	48
<∅>	4	0	0	23	0	0	1	0	0	0	28
<pro .1>	2	8	0	3	0	0	0	0	3	0	16
<pro .2>	2	6	0	3	0	0	0	0	2	0	13
<pro .h>	0	1	0	0	0	0	0	0	0	0	1
<pro .d>	13	7	0	1	0	0	0	0	0	0	21
<pro>	0	0	0	4	0	1	2	1	0	7	15
<np .h>	0	0	0	0	0	0	0	0	0	0	0
<np .d>	5	3	0	3	0	0	0	0	0	0	11
<np>	10	0	0	21	0	6	5	2	0	6	50
<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	284	285
<i>totals</i>	72	64	0	66	0	7	8	4	5	297	
<##>											120
<#>											23
<i>totals</i>											143

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

2.2 *dwi*

GRAID	<:s>	<:a>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:pred>	<:poss>	<:other>	<i>totals</i>
<∅ .1>	4	10	0	3	0	0	0	0	0	0	17
<∅ .2>	5	12	0	2	0	0	0	0	0	0	19
<∅ .h>	65	63	0	10	0	0	0	0	0	0	138
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	21	0	0	50	1	1	1	0	0	0	74
<pro .1>	2	11	0	2	1	0	0	0	3	0	19
<pro .2>	4	9	0	2	0	0	0	0	3	0	18
<pro .h>	21	27	0	3	0	0	0	0	16	0	67
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	11	1	0	8	2	2	4	1	0	20	49
<np .h>	31	24	0	7	6	1	0	9	3	1	82
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	21	1	0	70	6	19	21	6	3	33	180
<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	2	0	759	761
<i>totals</i>	185	158	0	157	16	23	26	18	28	813	
<##>											317
<#>											41
<i>totals</i>											358

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

2.3 *galang*

GRAID	<:s>	<:a>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:pred>	<:poss>	<:other>	<i>totals</i>
<∅ .1>	5	1	0	1	0	0	0	0	0	0	7
<∅ .2>	3	7	0	0	0	0	0	0	0	0	10
<∅ .h>	40	6	0	3	0	0	0	0	0	0	49
<∅ .d>	4	1	0	1	0	0	0	0	0	0	6
<∅>	3	0	0	1	0	0	1	0	0	0	5
<pro .1>	7	3	0	4	0	0	0	0	0	0	14
<pro .2>	4	1	0	1	0	0	1	0	0	0	7
<pro .h>	15	2	0	1	0	0	0	0	3	0	21
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	2	0	0	1	0	0	13	0	0	6	22
<np .h>	5	0	0	1	0	0	0	0	1	0	7
<np .d>	4	1	0	0	0	0	0	1	0	0	6
<np>	8	0	0	8	2	10	14	7	0	15	64
<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	4	0	249	253
<i>totals</i>	100	22	0	22	2	10	29	12	4	270	
<##>											118
<#>											8
<i>totals</i>											126

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

2.4 ganu

GRAID	<:s>	<:a>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:pred>	<:poss>	<:other>	<i>totals</i>
<∅ .1>	0	7	0	0	0	0	0	0	0	0	7
<∅ .2>	1	0	0	1	0	0	0	0	0	0	2
<∅ .h>	10	14	0	4	0	0	0	0	0	0	28
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	3	0	0	8	0	6	0	0	0	1	18
<pro .1>	0	0	0	0	0	0	0	0	0	0	0
<pro .2>	0	0	0	1	0	0	0	0	0	0	1
<pro .h>	10	14	0	5	0	0	0	0	2	0	31
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	0	0	0	0	1	0	0	0	0	7	8
<np .h>	10	9	0	10	1	1	0	3	6	0	40
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	1	0	0	23	4	5	3	4	0	24	64
<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	2	0	148	150
<i>totals</i>	35	44	0	52	6	12	3	9	8	180	
<##>											61
<#>											24
<i>totals</i>											85

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

2.5 *hkaili*

GRAID	<:s>	<:a>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:pred>	<:poss>	<:other>	<i>totals</i>
<∅ .1>	1	5	0	0	0	0	0	0	0	0	6
<∅ .2>	2	8	0	0	0	0	0	0	0	0	10
<∅ .h>	19	12	0	5	1	0	0	0	0	0	37
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	6	1	0	7	0	0	0	0	0	0	14
<pro .1>	4	2	0	2	0	0	0	0	4	0	12
<pro .2>	1	2	0	1	0	0	0	0	0	0	4
<pro .h>	5	10	0	1	0	0	0	0	10	0	26
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	1	0	0	8	5	0	5	0	0	9	28
<np .h>	17	12	0	18	1	0	0	1	0	2	51
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	9	0	0	13	1	3	2	2	0	15	45
<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	4	0	279	283
<i>totals</i>	65	52	0	55	8	3	7	7	14	305	
<##>											101
<#>											22
<i>totals</i>											123

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

2.6 hpaji

GRAID	<:s>	<:a>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:pred>	<:poss>	<: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>	0	0	0	0	0	0	0	0	0	0	0
<∅ .d>	14	9	0	0	0	0	0	0	0	0	23
<∅>	0	0	0	7	0	1	0	0	0	0	8
<pro .1>	0	2	0	0	0	0	0	0	0	0	2
<pro .2>	4	1	0	0	0	0	0	0	5	0	10
<pro .h>	0	0	0	0	0	0	0	0	0	0	0
<pro .d>	5	5	0	2	0	0	0	0	2	0	14
<pro>	0	0	0	4	2	0	0	0	0	1	7
<np .h>	0	0	0	0	0	0	0	0	1	0	1
<np .d>	5	3	0	0	0	0	0	1	1	0	10
<np>	8	0	0	5	1	2	3	0	0	2	21
<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	0	123	123
<i>totals</i>	36	21	0	18	3	3	3	1	9	126	
<##>											50
<#>											7
<i>totals</i>											57

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

2.7 *manau*

GRAID	<:s>	<:a>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:pred>	<:poss>	<:other>	<i>totals</i>
<∅ .1>	3	3	0	0	0	0	0	0	0	0	6
<∅ .2>	0	1	0	1	0	0	0	0	0	0	2
<∅ .h>	1	0	0	0	0	0	0	0	0	0	1
<∅ .d>	16	5	0	3	0	0	0	0	0	0	24
<∅>	2	0	0	6	0	0	0	0	0	0	8
<pro .1>	3	1	0	1	0	0	0	0	2	0	7
<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>	4	1	0	4	0	0	0	0	2	0	11
<pro>	1	0	0	0	0	0	0	0	0	1	2
<np .h>	0	0	0	0	0	0	0	0	0	0	0
<np .d>	4	6	0	1	1	0	0	0	1	0	13
<np>	3	0	0	9	2	3	2	4	0	14	37
<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	0	85	85
<i>totals</i>	37	17	0	25	3	3	2	4	5	100	
<##>											40
<#>											15
<i>totals</i>											55

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

2.8 *natga*

GRAID	<:s>	<:a>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:pred>	<:poss>	<:other>	<i>totals</i>
<∅ .1>	8	1	0	2	0	0	0	0	0	0	11
<∅ .2>	0	1	0	0	0	0	0	0	0	0	1
<∅ .h>	23	6	0	5	0	0	0	0	1	0	35
<∅ .d>	2	5	0	0	0	0	0	0	0	0	7
<∅>	3	0	0	1	0	0	0	0	0	0	4
<pro .1>	3	1	0	1	0	0	0	0	1	0	6
<pro .2>	2	0	0	0	0	0	0	0	1	0	3
<pro .h>	6	1	0	2	0	0	0	0	3	0	12
<pro .d>	2	0	0	0	0	0	0	0	0	0	2
<pro>	0	0	0	0	1	0	1	0	0	0	2
<np .h>	5	2	0	3	1	0	0	4	3	0	18
<np .d>	1	1	0	0	0	0	0	0	0	0	2
<np>	3	0	0	4	1	8	0	3	0	13	32
<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	150	151
<i>totals</i>	58	18	0	18	3	8	1	8	9	163	
<##>											54
<#>											22
<i>totals</i>											76

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

2.9 *nchyang*

GRAID	<:s>	<:a>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:pred>	<:poss>	<:other>	<i>totals</i>
<∅ .1>	3	2	0	0	0	0	0	0	0	0	5
<∅ .2>	0	0	0	1	0	0	0	0	0	0	1
<∅ .h>	21	11	0	6	0	0	0	0	0	0	38
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	0	0	0	2	0	0	0	0	0	0	2
<pro .1>	2	2	0	1	0	0	0	0	1	0	6
<pro .2>	0	0	0	0	0	0	0	0	0	0	0
<pro .h>	12	5	0	1	0	0	0	0	1	0	19
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	0	0	0	0	0	0	1	0	0	5	6
<np .h>	6	1	0	4	1	0	0	1	0	0	13
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	9	2	0	10	1	3	0	1	0	10	36
<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	0	182	182
<i>totals</i>	53	23	0	25	2	3	1	2	2	197	
<##>											64
<#>											13
<i>totals</i>											77

Table 10 Summarized GRAID counts for the *nchyang* text.

2.10 nga

GRAID	<:s>	<:a>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:pred>	<:poss>	<:other>	<i>totals</i>
<∅ .1>	1	3	0	1	0	0	0	0	0	0	5
<∅ .2>	1	2	0	0	0	0	0	0	0	0	3
<∅ .h>	17	23	0	3	0	0	0	0	0	0	43
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	1	1	0	12	0	0	0	0	0	0	14
<pro .1>	0	6	0	0	0	0	0	0	12	0	18
<pro .2>	1	8	0	0	0	0	0	0	3	0	12
<pro .h>	5	7	0	2	0	0	0	0	2	0	16
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	4	0	0	1	6	0	5	0	1	0	17
<np .h>	1	6	0	6	0	0	0	1	1	0	15
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	3	1	0	40	5	3	3	5	0	7	67
<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	0	199	199
<i>totals</i>	34	57	0	65	11	3	8	6	19	206	
<##>											82
<#>											13
<i>totals</i>											95

Table 11 Summarized GRAID counts for the *nga* text.

2.11 shanngayi

GRAID	<:s>	<:a>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:pred>	<:poss>	<:other>	<i>totals</i>
<∅ .1>	2	0	0	0	0	0	0	0	0	0	2
<∅ .2>	6	0	0	0	0	0	0	0	0	0	6
<∅ .h>	0	0	0	2	0	0	0	0	0	0	2
<∅ .d>	16	3	0	0	0	0	0	0	0	0	19
<∅>	9	0	0	2	0	0	0	0	0	0	11
<pro .1>	2	0	0	0	0	0	0	0	1	0	3
<pro .2>	0	2	0	0	0	0	0	0	1	0	3
<pro .h>	0	0	0	0	0	0	0	0	0	0	0
<pro .d>	5	3	0	2	0	0	0	0	8	0	18
<pro>	2	0	0	0	0	0	1	1	0	5	9
<np .h>	1	2	0	0	1	0	0	0	0	0	4
<np .d>	9	2	0	2	0	0	0	0	1	1	15
<np>	16	0	0	5	1	6	5	1	0	10	44
<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	0	171	171
<i>totals</i>	68	12	0	13	2	6	6	2	11	187	
<##>											74
<#>											7
<i>totals</i>											81

Table 12 Summarized GRAID counts for the *shanngayi* text.

Multi-CAST

Multilingual Corpus of Annotated Spoken Texts



multicast.aspra.uni-bamberg.de/