

# Multi-CAST

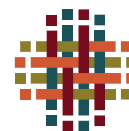
## *Cypriot Greek corpus counts*

---

*Harris Hadjidas*  
University of Bamberg

May 2019  
v2.0

*Maria Vollmer*  
University of Bamberg



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*

Hadjidas, Harris & Vollmer, Maria. 2019. Multi-CAST Cypriot Greek corpus counts. In Haig, Geoffrey & Schnell, Stefan (eds.), *Multi-CAST: Multilingual corpus of annotated spoken texts*. ([multicast.aspra.uni-bamberg.de/](http://multicast.aspra.uni-bamberg.de/)) (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The Multi-CAST collection has been archived at the *University of Bamberg*, Germany, and is freely accessible online at [multicast.aspra.uni-bamberg.de/](http://multicast.aspra.uni-bamberg.de/).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 licensing terms can be reviewed online at [creativecommons.org/licenses/by/4.0/](http://creativecommons.org/licenses/by/4.0/).

*Multi-CAST Cypriot Greek corpus counts* v2.0 last updated 8 May 2019  
This document was typeset by NNS with X<sub>Y</sub>L<sup>A</sup>T<sub>E</sub>X and the *multicast3* class (v3.0.9045).

## Contents

<b>1</b>	<b>Notes on the GRAID counts</b>	_____	<b>1</b>
<b>2</b>	<b>The Cypriot Greek corpus</b>	_____	<b>2</b>
2.1	<i>jitros</i>		3
2.2	<i>minaes</i>		4
2.3	<i>psarin</i>		5



## 1 Notes on the GRAID counts

This document collects tables with frequency counts for combinations of selected GRAID symbols in version 1905 (from May 2019) of the Multi-CAST Cypriot Greek corpus. The tables are intended to offer 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.

The tables in this document can be recreated with the `mc_table` function from *multicastR*, the companion package to Multi-CAST for the statistical computing language R. Please refer to the package documentation (`?multicastR`, `?mc_table`) for more information.

## 2 The Cypriot Greek corpus

GRAID	<:a>	<:s>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:poss>	<:pred>	<:other>	<i>totals</i>
<∅ .1>	116	55	0	1	0	0	0	0	0	0	172
<∅ .2>	116	75	0	0	0	0	0	0	0	1	192
<∅ .h>	182	173	0	5	0	0	0	0	0	1	361
<∅ .d>	7	6	0	1	0	0	0	0	0	0	14
<∅>	13	16	0	71	0	0	0	0	2	0	102
<pro .1>	13	9	0	10	13	20	0	100	0	3	168
<pro .2>	5	9	0	15	13	21	0	45	0	3	111
<pro .h>	2	5	0	54	20	57	1	57	1	6	203
<pro .d>	0	2	0	0	0	0	0	1	0	0	3
<pro>	0	2	0	59	1	1	0	0	0	1	64
<np .h>	28	50	0	38	17	19	0	28	11	5	196
<np .d>	1	14	0	2	0	0	0	0	0	20	37
<np>	9	24	0	219	44	50	15	14	18	82	475
<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	8	0	23	3	26	19	0	41	0	121
<i>totals</i>	493	448	0	498	111	194	35	245	73	122	
<##>											431
<#>											640
<i>totals</i>											1071

**Table 1** Summarized GRAID counts for the entire Cypriot Greek corpus.

2.1 *jitros*

GRAID	<:a>	<:s>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:poss>	<:pred>	<:other>	<i>totals</i>
<∅ .1>	27	10	0	0	0	0	0	0	0	0	37
<∅ .2>	32	17	0	0	0	0	0	0	0	0	49
<∅ .h>	52	48	0	0	0	0	0	0	0	0	100
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	9	3	0	16	0	0	0	0	0	0	28
<pro .1>	5	4	0	3	4	3	0	33	0	1	53
<pro .2>	1	3	0	6	1	4	0	9	0	0	24
<pro .h>	0	2	0	19	5	8	1	27	0	0	62
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	0	1	0	5	0	0	0	0	0	0	6
<np .h>	9	16	0	13	3	4	0	14	4	1	64
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	1	2	0	61	11	13	4	6	4	27	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	2	0	12	1	8	5	0	16	0	44
<i>totals</i>	136	108	0	135	25	40	10	89	24	29	
<##>											104
<#>											168
<i>totals</i>											272

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

## 2.2 *minaes*

GRAID	<:a>	<:s>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:poss>	<:pred>	<:other>	<i>totals</i>
<∅ .1>	31	29	0	1	0	0	0	0	0	0	61
<∅ .2>	39	38	0	0	0	0	0	0	0	1	78
<∅ .h>	40	31	0	1	0	0	0	0	0	0	72
<∅ .d>	7	6	0	1	0	0	0	0	0	0	14
<∅>	1	4	0	25	0	0	0	0	0	0	30
<pro .1>	3	4	0	3	7	12	0	29	0	2	60
<pro .2>	3	4	0	1	10	6	0	24	0	1	49
<pro .h>	2	1	0	7	5	13	0	19	0	4	51
<pro .d>	0	2	0	0	0	0	0	1	0	0	3
<pro>	0	1	0	12	1	0	0	0	0	0	14
<np .h>	6	14	0	13	7	10	0	9	0	2	61
<np .d>	1	14	0	2	0	0	0	0	0	20	37
<np>	4	12	0	74	14	15	1	3	3	37	163
<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	1	0	5	0	9	0	0	6	0	22
<i>totals</i>	138	161	0	145	44	65	1	85	9	67	
<##>											161
<#>											198
<i>totals</i>											359

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



2.3 *psarin*

GRAID	<:a>	<:s>	<:ncs>	<:p>	<:obl>	<:g>	<:l>	<:poss>	<:pred>	<:other>	<i>totals</i>
<∅ .1>	58	16	0	0	0	0	0	0	0	0	74
<∅ .2>	45	20	0	0	0	0	0	0	0	0	65
<∅ .h>	90	94	0	4	0	0	0	0	0	1	189
<∅ .d>	0	0	0	0	0	0	0	0	0	0	0
<∅>	3	9	0	30	0	0	0	0	2	0	44
<pro .1>	5	1	0	4	2	5	0	38	0	0	55
<pro .2>	1	2	0	8	2	11	0	12	0	2	38
<pro .h>	0	2	0	28	10	36	0	11	1	2	90
<pro .d>	0	0	0	0	0	0	0	0	0	0	0
<pro>	0	0	0	42	0	1	0	0	0	1	44
<np .h>	13	20	0	12	7	5	0	5	7	2	71
<np .d>	0	0	0	0	0	0	0	0	0	0	0
<np>	4	10	0	84	19	22	10	5	11	18	183
<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	5	0	6	2	9	14	0	19	0	55
<i>totals</i>	219	179	0	218	42	89	24	71	40	26	
<##>											166
<#>											274
<i>totals</i>											440

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

# Multi-CAST

*Multilingual Corpus of Annotated Spoken Texts*



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

