



Multilingual Corpus of
Annotated Spoken Texts

English

— corpus counts —

June 2016

Nils Norman Schiborr
University of Bamberg

cite this document as:

Schiborr, Nils N. 2016.
Multi-CAST English corpus counts.
In Haig, Geoffrey & Schnell, Stefan (eds.),
Multi-CAST (Multilingual Corpus of Annotated Spoken Texts).
(handle) (date accessed.)

file v1.0



Language Archive
Cologne

Notes on the GRAID counts

The following tables are intended to provide a top-down view of the data in the Multi-CAST English corpus. As such, they do not include all GRAID categories, and certain subcategories of symbols have been consolidated for the sake of simplicity (e.g. the dt column sums counts for <dt>, <dt.a>, <dt.a>, etc.). For full detail, please refer to the **_counts-raw.csv* files provided with each corpus text.

For information regarding the interpretation of the symbols and their combinations, please consult the *Multi-CAST English annotation notes* and the *GRAID Manual 7.0* (Haig & Schnell 2014).

Contents

Table 1	Counts for all English texts	2
Table 2	Counts for the kent01 text	3
Table 3	Counts for the kent02 text	4

References

Haig, Geoffrey & Schnell, Stefan. 2014. *Annotations using GRAID (Grammatical Relations and Animacy in Discourse): Introduction and guidelines for annotators (version 7.0)*. (<https://lac.uni-koeln.de/en/multicast/>) (accessed 2015-12-30.)

All English texts

(2,245 clause units)

	s	s_ds	a	p	p2	obl	g	l	poss	dt	pred	other	total
0.1	21	1	74	0	0	0	0	0	0	0	0	0	96
0.2	3	0	27	0	0	0	0	0	0	0	0	0	30
0.h	17	9	70	1	0	0	0	0	0	0	0	0	97
0	24	0	10	38	1	0	1	0	0	0	38	1	113
pro.1	230	59	363	43	0	20	16	1	111	1	1	3	848
pro.2	48	1	102	8	0	6	9	0	20	0	0	0	194
pro.h	130	106	216	35	0	16	16	0	59	2	1	0	581
pro	227	0	35	346	2	46	22	8	22	2	19	17	746
=wpro.h	0	0	0	8	0	6	2	0	0	0	0	0	16
=wpro	0	0	0	49	0	2	4	0	0	0	0	1	56
dem_pro.h	3	0	0	0	0	0	0	0	0	0	0	0	3
dem_pro	83	0	9	28	0	4	3	0	0	2	3	11	143
imp_pro.2	16	0	51	2	0	0	1	0	0	0	0	0	70
imp_pro.h	0	0	1	0	0	0	0	0	0	0	0	0	1
imp_pro	1	0	0	6	2	0	2	0	0	0	0	0	11
refl_pro.1	0	0	0	1	0	1	2	0	0	0	0	2	6
refl_pro.h	0	0	0	0	0	1	0	0	0	0	0	0	1
np.h	69	11	62	37	1	28	20	2	11	26	36	4	307
np	78	0	18	488	37	149	131	161	9	27	126	167	1391
pn_np.h	3	1	5	4	0	5	4	0	4	7	5	0	38
pn_np	7	0	1	3	1	12	38	26	1	3	4	2	98
other	10	0	2	17	1	1	64	59	0	0	281	1676	2111
total	970	188	1046	1114	45	297	335	257	237	70	514	1884	6957

Table 1. Combined GRAID counts for all English texts.

English *kent01*

(622 clause units)

	s	s_ds	a	p	p2	obl	g	l	poss	dt	pred	other	total
0.1	6	0	19	0	0	0	0	0	0	0	0	0	25
0.2	2	0	9	0	0	0	0	0	0	0	0	0	11
0.h	3	1	20	0	0	0	0	0	0	0	0	0	24
0	8	0	3	6	1	0	1	0	0	0	10	1	30
pro.1	41	12	79	11	0	4	4	1	31	0	0	1	184
pro.2	11	0	24	0	0	0	0	0	1	0	0	0	36
pro.h	33	25	74	14	0	5	5	0	31	1	0	0	188
pro	69	0	9	95	2	6	9	3	5	1	5	5	209
=wpro.h	0	0	0	4	0	1	2	0	0	0	0	0	7
=wpro	0	0	0	17	0	1	3	0	0	0	0	0	21
dem_pro.h	1	0	0	0	0	0	0	0	0	0	0	0	1
dem_pro	16	0	0	8	0	0	1	0	0	1	0	4	30
imp_pro.2	4	0	14	0	0	0	0	0	0	0	0	0	18
imp_pro.h	0	0	1	0	0	0	0	0	0	0	0	0	1
imp_pro	1	0	0	6	2	0	2	0	0	0	0	0	11
refl_pro.1	0	0	0	1	0	0	0	0	0	0	0	0	1
refl_pro.h	0	0	0	0	0	0	0	0	0	0	0	0	0
np.h	18	1	18	15	0	8	7	0	6	7	11	2	93
np	31	0	10	138	4	43	50	59	4	2	26	40	407
pn_np.h	1	0	1	2	0	3	2	0	1	3	2	0	15
pn_np	2	0	0	0	0	8	17	15	1	0	1	1	45
other	1	0	1	5	0	0	13	17	0	0	93	444	574
total	248	39	282	322	9	79	116	95	80	15	148	498	1931

Table 2. Combined GRAID counts for the *kent01* text.

English *kent02*

(1,623 clause units)

	s	s_ds	a	p	p2	obl	g	l	poss	dt	pred	other	total
0.1	15	1	55	0	0	0	0	0	0	0	0	0	71
0.2	1	0	18	0	0	0	0	0	0	0	0	0	19
0.h	14	8	50	1	0	0	0	0	0	0	0	0	73
0	16	0	7	32	0	0	0	0	0	0	28	0	83
pro.1	189	47	284	32	0	16	12	0	80	1	1	2	664
pro.2	37	1	78	8	0	6	9	0	19	0	0	0	158
pro.h	97	81	142	21	0	11	11	0	28	1	1	0	393
pro	158	0	26	251	0	40	13	5	17	1	14	12	537
=wpro.h	0	0	0	4	0	5	0	0	0	0	0	0	9
=wpro	0	0	0	32	0	1	1	0	0	0	0	1	35
dem_pro.h	2	0	0	0	0	0	0	0	0	0	0	0	2
dem_pro	67	0	9	20	0	4	2	0	0	1	3	7	113
imp_pro.2	12	0	37	2	0	0	1	0	0	0	0	0	52
imp_pro.h	0	0	0	0	0	0	0	0	0	0	0	0	0
imp_pro	0	0	0	0	0	0	0	0	0	0	0	0	0
refl_pro.1	0	0	0	0	0	1	2	0	0	0	0	2	5
refl_pro.h	0	0	0	0	0	1	0	0	0	0	0	0	1
np.h	51	10	44	22	1	20	13	2	5	19	25	2	214
np	47	0	8	350	33	106	81	102	5	25	100	127	984
pn_np.h	2	1	4	2	0	2	2	0	3	4	3	0	23
pn_np	5	0	1	3	1	4	21	11	0	3	3	1	53
other	9	0	1	12	1	1	51	42	0	0	188	1232	1537
total	722	149	764	792	36	218	219	162	157	55	366	1386	5026

Table 3. Combined GRAID counts for the *kent02* text.