

Sphinx4 baseline performance with selected Open Yale Courses lectures (<http://oyc.yale.edu>)

Language model: HUB4 trigram (HUB4_trigram_lm)

Acoustic model: US English HUB4 (hub4opensrc.cd_continuous_8gau)

File	Length (s)	Ref Words	Sentences	Runtime	Realtime ratio (runtime to length)	Output words	Output: original word ratio	Edit distance	Word Error Rate (WER)	Words correct	Word Correct Rate (WCR)	Perplexity (sentence transcript)	Out of vocabulary (OOV) words	OOV %	Unique out of vocabulary (OOV) words	Unique unrecognized words	Unique extraneous words
astr160_21_041707	2 747	6 704	447	6863	2.50	7 312	1.1	2 360	35%	5 185	77%	228	110	1.6%	47	116	422
beng100_05_012908	3 141	7 385	336	6595	2.10	8 166	1.1	2 617	35%	5 780	78%	307	164	2.2%	54	85	486
beng100_09_021208	2 792	6 974	334	6081	2.18	7 706	1.1	2 228	32%	5 675	81%	211	96	1.4%	62	100	415
eeb122_34_042009	2 455	5 795	370	6917	2.82	5 935	1.0	2 312	40%	3 935	68%	331	145	2.5%	97	242	595
engl220_06_092407	3 111	7 350	330	8167	2.63	8 293	1.1	3 185	43%	5 312	72%	535	314	4.3%	143	232	671
engl291_12_022008	2 967	6 201	274	4778	1.61	6 596	1.1	2 015	32%	4 738	76%	379	174	2.8%	96	251	586
engl300_15_030309	3 169	6 701	280	5072	1.60	7 992	1.2	3 162	47%	5 006	75%	274	265	4.0%	116	170	624
hist116_08_020410	2 864	7 902	332	5278	1.84	7 797	1.0	3 131	40%	5 184	66%	309	74	0.9%	54	260	625
hist202_06_092208	2 996	6 643	466	5170	1.73	8 031	1.2	4 059	61%	4 293	65%	252	212	3.2%	131	290	806
phil176_13_022707	2 932	6 603	464	5845	1.99	6 676	1.0	3 230	49%	3 856	58%	475	97	1.5%	31	140	588
plsc114_02_091306	2 734	5 473	249	3479	1.27	5 976	1.1	1 739	32%	4 355	80%	357	103	1.9%	48	124	381
psyc110_05_013107	2 926	7 085	467	5208	1.78	7 156	1.0	2 984	42%	4 661	66%	275	119	1.7%	70	228	510
rlst152_26_042209	2 860	8 196	393	6700	2.34	8 614	1.1	3 379	41%	5 730	70%	286	91	1.1%	58	190	617
Average	2 900	6 847	365	5 858	2.02	7 404	1.1	2 800	41%	4 901	72%	324	151	2.2%	77	187	564