Effect sizes of features in LLM writing

Paired Cohen's d relative to human chunk 2

Feature	GPT			Llama 3 Instruct		Llama 3 Base		
	GPT-4o	GPT-4o Mini	chunk_1	70B	8B	8B	70B	Importance
Present participial clauses	1.38	1.37	−0.02	0.56	0.44	−0.03	0.01	1,766.6
'That' clauses as subject	0.77	1.03	0.02	0.33	0.35	−0.22	−0.20	1,408.3
Mean word length	2.24	1.98	0.21	0.45	0.14	−0.19	0.01	1,112.3
Adverbs	−0.53	−0.43	−0.19	−0.77	−0.77	0.04	0.05	921.8
Nominalizations	1.23	1.18	0.00	0.59	0.49	−0.15	−0.12	805.5
Phrasal co-ordination	0.81	0.44	0.10	0.53	0.61	−0.05	−0.02	766.6
Clausal co-ordination	−0.59	−0.54	−0.17	0.36	0.47	0.21	0.18	728.1
Other nouns	0.10	−0.10	0.32	−0.24	−0.38	−0.09	−0.18	726.1
Downtoners	0.11	0.33	−0.01	−0.29	−0.26	−0.19	−0.18	711.6
Demonstratives	0.32	0.39	−0.06	−0.21	−0.24	−0.25	−0.21	676.7
Attributive adjectives	1.09	0.91	0.18	0.10	−0.01	−0.41	−0.37	657.2
Agentless passives	−0.56	−0.59	−0.01	−0.13	−0.04	0.01	−0.02	647.3
Past participial postnominal	0.55	0.65	0.01	−0.13	−0.13	0.13	0.14	640.1
Infinitives	−0.30	−0.24	−0.06	0.47	0.51	0.22	0.16	605.1
Prepositional phrases	0.68	0.77	0.15	0.02	−0.01	−0.46	−0.09	578.5
Emphatics	−0.39	−0.31	−0.09	−0.02	−0.03	−0.23	−0.26	552.3
Existential 'there'	−0.19	−0.27	−0.06	−0.39	−0.37	0.03	0.05	541.1
Possibility modals	0.04	0.38	−0.07	0.13	0.08	−0.01	−0.01	522.5
Past tense	−0.29	−0.41	−0.21	−0.27	−0.12	0.19	0.16	511.5
Other adverbial subordinators	0.21	0.41	−0.06	0.07	0.13	−0.18	−0.11	505.6

Effect sizes of features in LLM writing

Paired Cohen's d relative to human chunk 2

| Feature | GPT | | | Llama 3 Instruct | | Llama 3 Base | | |
	GPT-4o	GPT-4o Mini	chunk_1	70B	8B	8B	70B	Importance
'That' verb complements	−0.35	−0.23	−0.12	0.08	0.22	0.26	0.22	499.9
Contractions	−0.50	−0.49	−0.10	0.38	0.35	0.29	0.24	499.6
'Be' as main verb	−0.83	−0.94	−0.08	0.00	0.11	0.09	0.02	486.0
Present participial postnominal	0.52	0.65	0.02	0.10	0.10	0.01	0.06	456.0
Public verbs	−0.37	−0.48	−0.16	−0.34	−0.34	0.13	0.10	454.6
Predicative adjectives	−0.11	−0.18	−0.07	0.28	0.32	0.24	0.23	448.8
Private verbs	−0.26	−0.17	−0.20	0.20	0.29	0.28	0.27	442.1
Analytic negation	−0.51	−0.35	−0.18	−0.27	−0.23	0.09	0.05	441.2
Perfect aspect	−0.43	−0.42	0.05	0.09	0.16	−0.03	−0.07	438.7
Third-person pronouns	−0.13	−0.14	−0.20	0.05	0.10	0.24	0.23	437.8
Suasive verbs	0.12	−0.02	−0.08	−0.07	−0.07	0.16	0.10	434.8
Pro-verb 'do'	−0.62	−0.63	−0.08	−0.33	−0.30	0.06	0.04	425.0
Present tense	−0.50	−0.40	−0.05	0.00	−0.01	0.28	0.27	421.1
Pronoun 'it'	−0.11	−0.17	−0.07	0.06	0.06	0.06	0.04	406.6
Gerunds	0.35	0.38	0.02	0.15	0.11	−0.09	−0.09	406.2
Amplifiers	−0.38	−0.25	−0.04	−0.38	−0.33	−0.07	−0.09	404.9
First-person pronouns	−0.62	−0.32	−0.07	0.12	0.15	0.35	0.30	390.9
Split auxiliaries	−0.20	−0.08	0.00	0.13	0.17	−0.01	−0.06	386.0
Because	−0.50	−0.50	−0.11	−0.35	−0.34	0.08	0.02	377.2
Time adverbials	−0.19	−0.21	0.00	−0.26	−0.26	0.07	0.06	357.6
Place adverbials	0.33	0.35	0.00	−0.01	0.01	0.11	0.17	353.7

Effect sizes of features in LLM writing

Paired Cohen's d relative to human chunk 2

| Feature | GPT | | chunk_1 | Llama 3 Instruct | | Llama 3 Base | | Importance |
	GPT-4o	GPT-4o Mini		70B	8B	8B	70B	
Though, although	0.09	−0.13	−0.01	−0.28	−0.32	−0.05	−0.03	348.6
Demonstrative pronouns	−0.56	−0.50	−0.13	−0.26	−0.28	−0.01	−0.03	348.1
Predictive modals	−0.43	−0.28	−0.09	0.04	−0.03	0.11	0.08	340.6
Indefinite pronouns	−0.18	−0.22	−0.04	−0.12	−0.11	0.19	0.20	328.4
Synthetic negation	−0.27	−0.37	−0.02	−0.34	−0.33	−0.02	−0.04	307.0
'Wh-' questions	−0.25	−0.06	−0.02	0.16	0.16	0.11	0.10	306.9
If, unless	−0.30	−0.16	−0.09	−0.27	−0.29	0.09	0.06	305.6
by-passives	−0.17	−0.21	0.03	0.03	0.06	0.04	0.01	296.6
'Seem' and 'appear'	0.27	0.12	−0.01	0.09	0.12	0.00	0.02	295.0
Necessity modals	−0.10	0.02	−0.07	−0.22	−0.22	0.14	0.12	285.7
Conjuncts	−0.03	0.07	−0.06	0.09	−0.02	−0.01	0.02	285.1
Second-person pronouns	−0.49	−0.39	−0.06	−0.20	−0.21	0.11	0.07	274.9
'Wh-' relatives as subject	−0.23	−0.21	0.04	0.01	0.04	0.03	0.01	260.0
Sentence relatives	−0.25	−0.26	0.01	0.06	0.02	0.00	−0.03	259.5
Stranded prepositions	−0.16	−0.16	−0.03	0.05	0.04	−0.04	−0.07	254.8
'Wh-' clauses	−0.24	−0.17	−0.07	−0.14	−0.06	0.11	0.10	246.2
Hedges	−0.16	−0.22	−0.10	−0.13	−0.15	−0.04	−0.02	183.0
Discourse particles	−0.20	−0.20	0.02	−0.04	−0.04	0.14	0.04	182.6
Past participial clauses	0.36	0.33	0.02	0.10	0.11	−0.09	−0.07	175.7
'That' deletion	−0.16	−0.11	−0.06	−0.11	−0.10	0.09	0.11	166.5

Effect sizes of features in LLM writing

Paired Cohen's d relative to human chunk 2

Feature	GPT			Llama 3 Instruct		Llama 3 Base		Importance
	GPT-4o	GPT-4o Mini	chunk_1	70B	8B	8B	70B	
Pied-piping relative clauses	−0.17	−0.16	0.02	−0.21	−0.23	−0.17	−0.17	147.2
'That' clauses as object	−0.17	−0.13	−0.01	−0.15	−0.12	−0.02	−0.03	121.3
'That' adjective complements	−0.16	−0.06	−0.04	0.02	0.05	0.09	0.06	108.7
Split infinitives	−0.02	−0.05	−0.02	0.01	0.02	−0.15	−0.13	94.0
'Wh-' relatives as object	−0.27	−0.25	0.00	−0.26	−0.28	−0.10	−0.10	69.9