

Tabelle1

Word layer	Number of clusters	Tag set	Cosine similarity	Rounded	Standard Deviation	Variance	Median
Lemma	250	Mixed	0.6747289057262376	0,674	0.1843759	0.03399446	0.664648
Lemma	250	Nom	0.6704259902407466	0,67	0.1835163	0.03367823	0.6471398
Token	250	Nom	0.6611438631736772	0,661	0.1984968	0.03940099	0.6355575
Token	250	Mixed	0.6579997372448713	0,657	0.2265156	0.0513093	0.5923313
Lemma	250	Adj	0.6396767677190237	0,639	0.1678194	0.02816336	0.6286488
Lemma	250	Verb	0.637427734374412	0,637	0.2427922	0.05894804	0.5878782
Token	250	Verb	0.63645505516621	0,636	0.2410656	0.05811264	0.5740713
Token	250	Adj	0.6184052207873473	0,618	0.174669	0.03050928	0.607096
Token	250	Other	0.5971165784022904	0,597	0.2265156	0.0513093	0.5923313
Lemma	250	Other	0.5967887469416676	0,596	0.225804	0.05098747	0.5853017
Token	100	Nom	0.5428885572692168	0,542	0.2114618	0.04471611	0.5022774
Lemma	100	Nom	0.5428605249107482	0,542	0.2111284	0.04457519	0.4962749
Lemma	100	Mixed	0.5422300762512144	0,542	0.2075966	0.04309637	0.4973985
Token	100	Mixed	0.5404054811750055	0,54	0.206554	0.04266455	0.4953801
Lemma	100	Verb	0.5391255908572484	0,539	0.2402116	0.05770161	0.4641851
Token	100	Verb	0.5338250765998395	0,533	0.2317562	0.05371091	0.4513787
Lemma	100	Adj	0.5205757974826017	0,52	0.1636844	0.02679257	0.5087038
Token	100	Adj	0.4961986778398667	0,496	0.1649072	0.0271944	0.4654908
Lemma	100	Other	0.48575336427799015	0,485	0.1802125	0.03247653	0.4671078
Token	100	Other	0.47823932793478136	0,478	0.1918088	0.03679061	0.4578859

Sorted by: cosine similarity → standard deviation → variance

Stand: 02. Juni 2017