Appendix A: Evaluation Tables for all Dataset.

Table 5.1. Results of name matching in **English** Paris 60.

No	String.		Compared Algorithms								Proposed Algorithms	
			LD		DLD		MDLD		N-DIST		E-N-DIST	
	Source	Target	Dist	Sim %	Dist	Sim %	Dist	Sim %	Dist	Sim %	Dist	Sim %
1	absence	absense	1	0.86	1	0.86	1	0.86	1	0.86	1.00	0.86
2	acceptable	acceptible	1	0.90	1	0.90	1	0.90	1	0.90	1.00	0.90
3	accidentally	accidentaly	1	0.92	1	0.92	1	0.92	1	0.92	0.50	0.96
4	achieve	acheive	2	0.71	1	0.86	1	0.86	2	0.71	1.00	0.86
5	acknowledge	aknowledge	1	0.91	1	0.91	1	0.91	1.5	0.86	1.00	0.91
6	acquire	aquire	1	0.86	1	0.86	1	0.86	1.5	0.79	1.00	0.86
7	alcohol	alchohol	1	0.88	1	0.88	1	0.88	1.5	0.81	1.00	0.88
8	algorithm	algorythm	1	0.89	1	0.89	1	0.89	1	0.89	1.00	0.89
9	almost	allmost	1	0.86	1	0.86	1	0.86	1	0.86	0.50	0.93
10	amateur	amatuer	2	0.71	1	0.86	1	0.86	2	0.71	1.00	0.86
11	amend	ammend	1	0.83	1	0.83	1	0.83	1	0.83	0.50	0.92
12	atheist	athiest	2	0.71	1	0.86	1	0.86	2	0.71	1.00	0.86
13	balance	balence	1	0.86	1	0.86	1	0.86	1	0.86	1.00	0.86
14	because	becuase	2	0.71	1	0.86	1	0.86	2	0.71	1.00	0.86
15	believe	beleive	2	0.71	1	0.86	1	0.86	2	0.71	1.00	0.86
16	burglar	burgler	1	0.86	1	0.86	1	0.86	1	0.86	1.00	0.86
17	calendar	calender	1	0.88	1	0.88	1	0.88	1	0.88	1.00	0.88
18	cemetery	cemetary	1	0.88	1	0.88	1	0.88	1	0.88	1.00	0.88
19	chief	cheif	2	0.60	1	0.80	1	0.80	2	0.60	1.00	0.80
20	colleague	collegue	1	0.89	1	0.89	1	0.89	1.5	0.83	1.00	0.89
21	conscious	concious	1	0.89	1	0.89	1	0.89	1.5	0.83	1.00	0.89
22	deceive	decieve	2	0.71	1	0.86	1	0.86	2	0.71	1.00	0.86
23	definite	definate	1	0.88	1	0.88	1	0.88	1	0.88	1.00	0.88
24	desperate	desparate	1	0.89	1	0.89	1	0.89	1	0.89	1.00	0.89
25	dilemma	dilema	1	0.86	1	0.86	1	0.86	1	0.86	0.50	0.93
26	embarrass	embarass	1	0.89	1	0.89	1	0.89	1	0.89	0.00	1.00
27	existence	existance	1	0.89	1	0.89	1	0.89	1	0.89	1.00	0.89
28	experience	experiance	1	0.90	1	0.90	1	0.90	1	0.90	1.00	0.90
29	fiery	firey	2	0.60	1	0.80	1	0.80	2	0.60	1.00	0.80
30	foreign	foriegn	2	0.71	1	0.86	1	0.86	2	0.71	1.00	0.86
31	friend	freind	2	0.67	1	0.83	1	0.83	2	0.67	1.00	0.83
32	gauge	guage	2	0.60	1	0.80	1	0.80	2	0.60	1.00	0.80
33	guidance	guidence	1	0.88	1	0.88	1	0.88	1	0.88	1.00	0.88
34	harass	harrass	1	0.86	1	0.86	1	0.86	1	0.86	0.00	1.00
35	humorous	humerous	1	0.88	1	0.88	1	0.88	1	0.88	1.00	0.88

36	hygiene	hygene	1	0.86	1	0.86	1	0.86	1.5	0.79	1.00	0.86
37	ignorance	ignorence	1	0.89	1	0.89	1	0.89	1	0.89	1.00	0.89
38	independent	independant	1	0.91	1	0.91	1	0.91	1	0.91	1.00	0.91
39	liaison	liason	1	0.86	1	0.86	1	0.86	1.5	0.79	0.50	0.93
40	library	libary	1	0.86	1	0.86	1	0.86	1.5	0.79	1.00	0.86
41	memento	momento	1	0.86	1	0.86	1	0.86	1	0.86	1.00	0.86
42	minuscule	miniscule	1	0.89	1	0.89	1	0.89	1	0.89	1.00	0.89
43	niece	neice	2	0.60	1	0.80	1	0.80	2	0.60	1.00	0.80
44	noticeable	noticable	1	0.90	1	0.90	1	0.90	1.5	0.85	1.00	0.90
45	occurred	occured	1	0.88	1	0.88	1	0.88	1	0.88	0.50	0.94
46	original	orignal	1	0.88	1	0.88	1	0.88	1.5	0.81	0.50	0.94
47	perceive	percieve	2	0.75	1	0.88	1	0.88	2	0.75	1.00	0.88
48	precede	preceed	2	0.71	1	0.86	1	0.86	1.5	0.79	0.00	1.00
49	promise	promiss	1	0.86	1	0.86	1	0.86	0.5	0.93	0.00	1.00
50	really	realy	1	0.83	1	0.83	1	0.83	1	0.83	0.50	0.92
51	receipt	reciept	2	0.71	1	0.86	1	0.86	2	0.71	1.00	0.86
52	relevant	relevent	1	0.88	1	0.88	1	0.88	1	0.88	1.00	0.88
53	religious	religous	1	0.89	1	0.89	1	0.89	1.5	0.83	0.50	0.94
54	secretary	secretery	1	0.89	1	0.89	1	0.89	1	0.89	1.00	0.89
55	similar	similer	1	0.86	1	0.86	1	0.86	1	0.86	1.00	0.86
56	speech	speach	1	0.83	1	0.83	1	0.83	1	0.83	0.50	0.92
57	surprise	suprise	1	0.88	1	0.88	1	0.88	1.5	0.81	1.00	0.88
58	until	untill	1	0.83	1	0.83	1	0.83	1	0.83	0.50	0.92
59	vacuum	vaccuum	1	0.86	1	0.86	1	0.86	1	0.86	0.00	1.00
60	whether	wether	1	0.86	1	0.86	1	0.86	1.5	0.79	1.00	0.86
Average (percentage similarity)				**0.83**		**0.86**		**0.86**		**0.81**		**0.89**

Table 5.2. Results of name matching in Dataset 13 (Paris 115).

No	String.		Compared Algorithms								Proposed Algorithms	
			LD		DLD		MDLD		N-DIST		E-N-DIST	
	Source	Target	Dist	Sim %	Dist	Sim %	Dist	Sim %	Dist	Sim %	Dist	Sim %
1	Achrochaetium	Acrochaetium	1.00	0.92	1.00	0.92	1.00	0.92	1.50	0.88	1.00	0.92
2	Acropaginula	Arcopaginula	2.00	0.83	1.00	0.92	1.00	0.92	2.00	0.83	1.00	0.92
3	Acrostrigma	Acrosterigma	1.00	0.92	1.00	0.92	1.00	0.92	1.50	0.88	1.00	0.92
4	Aetopedes	Aetiopedes	1.00	0.90	1.00	0.90	1.00	0.90	1.50	0.85	1.00	0.90
5	Amesopus	Amesopous	1.00	0.89	1.00	0.89	1.00	0.89	1.50	0.83	0.50	0.94
6	Andriphiloscia	Andricophiloscia	2.00	0.88	2.00	0.88	2.00	0.88	2.50	0.84	2.00	0.88
7	Anodonta	Anodontia	1.00	0.89	1.00	0.89	1.00	0.89	1.50	0.83	0.50	0.94
8	Ascolepis	Ascolepsis	1.00	0.90	1.00	0.90	1.00	0.90	1.50	0.85	1.00	0.90
9	Asteracmaea	Asteracmea	1.00	0.91	1.00	0.91	1.00	0.91	1.50	0.86	1.00	0.91
10	Asteroceras	Astroceras	1.00	0.91	1.00	0.91	1.00	0.91	1.50	0.86	1.00	0.91
11	Atiliosa	Attiliosa	1.00	0.89	1.00	0.89	1.00	0.89	1.00	0.89	0.50	0.94
12	Australophocoena	Australophocaena	1.00	0.94	1.00	0.94	1.00	0.94	1.00	0.94	1.00	0.94
13	Bannerus	Bannereus	1.00	0.89	1.00	0.89	1.00	0.89	1.50	0.83	0.50	0.94
14	Bougainvillii	Bougainvillia	1.00	0.92	1.00	0.92	1.00	0.92	0.50	0.96	0.00	1.00
15	Buffonellodes	Buffonelloides	1.00	0.93	1.00	0.93	1.00	0.93	1.50	0.89	1.00	0.93
16	Centropala	Centrapala	1.00	0.90	1.00	0.90	1.00	0.90	1.00	0.90	1.00	0.90
17	Chloridiella	Chlorodiella	1.00	0.92	1.00	0.92	1.00	0.92	1.00	0.92	1.00	0.92
18	Chlorisquilla	Chorisquilla	1.00	0.92	1.00	0.92	1.00	0.92	1.50	0.88	0.50	0.96
19	Chuculba	Chucullba	1.00	0.89	1.00	0.89	1.00	0.89	1.00	0.89	0.00	1.00
20	Cicatella	Cicatellia	1.00	0.90	1.00	0.90	1.00	0.90	1.50	0.85	1.00	0.90
21	Clathriopsis	Clathropsis	1.00	0.92	1.00	0.92	1.00	0.92	1.50	0.88	1.00	0.92
22	Clavellina	Clavelina	1.00	0.90	1.00	0.90	1.00	0.90	1.00	0.90	0.50	0.95
23	Clavigella	Clavagella	1.00	0.90	1.00	0.90	1.00	0.90	1.00	0.90	1.00	0.90
24	Coccopigyra	Coccopigya	1.00	0.91	1.00	0.91	1.00	0.91	1.50	0.86	1.00	0.91
25	Comasterias	Cosmasterias	1.00	0.92	1.00	0.92	1.00	0.92	1.50	0.88	1.00	0.92
26	Condujius	Condukius	1.00	0.89	1.00	0.89	1.00	0.89	1.00	0.89	1.00	0.89
27	Crassisorophium	Crassicorophium	1.00	0.93	1.00	0.93	1.00	0.93	1.00	0.93	1.00	0.93
28	Cribularia	Cribralaria	2.00	0.82	2.00	0.82	2.00	0.82	2.00	0.82	2.00	0.82
29	Cysticus	Cystiscus	1.00	0.89	1.00	0.89	1.00	0.89	1.50	0.83	1.00	0.89
30	Diplonotus	Diplonotos	1.00	0.90	1.00	0.90	1.00	0.90	1.00	0.90	1.00	0.90
31	Echinophora	Echinophoria	1.00	0.92	1.00	0.92	1.00	0.92	1.50	0.88	1.00	0.92
32	Emarginula	Emarginella	2.00	0.82	2.00	0.82	2.00	0.82	2.00	0.82	1.50	0.86
33	Eogynodiastylus	Eogynodiastylis	1.00	0.93	1.00	0.93	1.00	0.93	1.00	0.93	1.00	0.93
34	Epidrona	Epidirona	1.00	0.89	1.00	0.89	1.00	0.89	1.50	0.83	0.50	0.94
35	Epipidonana	Epipedonana	1.00	0.91	1.00	0.91	2.00	0.82	1.00	0.91	1.00	0.91
36	Erichthonius	Ericthonius	1.00	0.92	1.00	0.92	1.00	0.92	1.50	0.88	1.00	0.92
37	Eryhtrops	Erythrops	2.00	0.78	1.00	0.89	1.00	0.89	2.00	0.78	1.00	0.89
38	Etisius	Etisus	1.00	0.86	1.00	0.86	1.00	0.86	1.50	0.79	0.50	0.93
39	Eulaospira	Eulaeospira	1.00	0.91	1.00	0.91	1.00	0.91	1.50	0.86	1.00	0.91
40	Eurataster	Euretaster	1.00	0.90	1.00	0.90	1.00	0.90	1.00	0.90	1.00	0.90
41	Eusyllus	Eusyllis	1.00	0.88	1.00	0.88	1.00	0.88	1.00	0.88	1.00	0.88
42	Exioperna	Exosiperna	3.00	0.70	3.00	0.70	3.00	0.70	3.00	0.70	1.50	0.85
43	Exocereis	Exocerceis	1.00	0.90	1.00	0.90	1.00	0.90	1.50	0.85	1.00	0.90
44	Falsilunata	Falsilunatia	1.00	0.92	1.00	0.92	1.00	0.92	1.50	0.88	1.00	0.92

45	Gastrolepida	Gastrolepidia	1.00	0.92	1.00	0.92	1.00	0.92	1.50	0.88	0.50	0.96
46	Glaphrymenia	Glaphyrymenia	1.00	0.92	1.00	0.92	1.00	0.92	1.50	0.88	1.00	0.92
47	Gonodactoideus	Gonodactyloideus	2.00	0.88	2.00	0.88	2.00	0.88	2.50	0.84	1.50	0.91
48	Halydictyon	Halodictyon	1.00	0.91	1.00	0.91	1.00	0.91	1.00	0.91	1.00	0.91
49	Hemipodia	Hemipodus	2.00	0.78	2.00	0.78	2.00	0.78	1.50	0.83	1.50	0.83
50	Histiocidaris	Histocidaris	1.00	0.92	1.00	0.92	1.00	0.92	1.50	0.88	1.00	0.92
51	Homalimetra	Homalometra	1.00	0.91	1.00	0.91	1.00	0.91	1.00	0.91	1.00	0.91
52	Hypereryhtrops	Hypererythrops	2.00	0.86	1.00	0.93	2.00	0.86	2.00	0.86	1.00	0.93
53	Isosecuriflustra	Isoseculiflustra	1.00	0.94	1.00	0.94	1.00	0.94	1.00	0.94	1.00	0.94
54	Laevilittorina	Laevilitorina	1.00	0.93	1.00	0.93	1.00	0.93	1.00	0.93	0.00	1.00
55	Leiopurga	Leiopyrga	1.00	0.89	1.00	0.89	1.00	0.89	1.00	0.89	1.00	0.89
56	Lensisquilla	Lenisquilla	1.00	0.92	1.00	0.92	1.00	0.92	1.50	0.88	0.50	0.96
57	Leognathus	Leongathus	2.00	0.80	1.00	0.90	1.00	0.90	2.00	0.80	1.00	0.90
58	Litogynodiastylus	Litogynodiastylis	1.00	0.94	1.00	0.94	1.00	0.94	1.00	0.94	1.00	0.94
59	Lytocarpa	Lytocarpia	1.00	0.90	1.00	0.90	1.00	0.90	1.50	0.85	1.00	0.90
60	Macroramphosus	Macrorhamphosus	1.00	0.93	1.00	0.93	1.00	0.93	1.50	0.90	0.50	0.97
61	Marplestonia	Maplestonia	1.00	0.92	1.00	0.92	1.00	0.92	1.50	0.88	1.00	0.92
62	Meosarmatium	Neosarmatium	1.00	0.92	1.00	0.92	1.00	0.92	1.50	0.88	1.50	0.88
63	Metalphius	Metalpheus	1.00	0.90	1.00	0.90	1.00	0.90	1.00	0.90	1.00	0.90
64	Montastrea	Montastraea	1.00	0.91	1.00	0.91	1.00	0.91	1.50	0.86	1.00	0.91
65	Musculum	Musculium	1.00	0.89	1.00	0.89	1.00	0.89	1.50	0.83	1.00	0.89
66	Myriozoisis	Myriozotisis	1.00	0.92	1.00	0.92	1.00	0.92	1.50	0.88	0.50	0.96
67	Nanophora	Nanaphora	1.00	0.89	1.00	0.89	1.00	0.89	1.00	0.89	1.00	0.89
68	Naxoides	Naxioides	1.00	0.89	1.00	0.89	1.00	0.89	1.50	0.83	1.00	0.89
69	Neoglypha	Neoglyphea	1.00	0.90	1.00	0.90	1.00	0.90	1.50	0.85	1.00	0.90
70	Nesopontonia	Mesopontonia	1.00	0.92	1.00	0.92	1.00	0.92	1.50	0.88	1.50	0.88
71	Ollophon	Ollaphon	1.00	0.88	1.00	0.88	1.00	0.88	1.00	0.88	1.00	0.88
72	Ophioconus	Ophioconis	1.00	0.90	1.00	0.90	1.00	0.90	1.00	0.90	1.00	0.90
73	Ophiopallus	Ophiopallas	1.00	0.91	1.00	0.91	1.00	0.91	1.00	0.91	1.00	0.91
74	Ophyrotrocha	Ophryotrocha	2.00	0.83	1.00	0.92	1.00	0.92	2.00	0.83	1.00	0.92
75	Palliola	Paliolla	2.00	0.75	2.00	0.75	2.00	0.75	2.00	0.75	0.00	1.00
76	Paraphrinima	Paraphronima	1.00	0.92	1.00	0.92	1.00	0.92	1.00	0.92	1.00	0.92
77	Patinonemertes	Pantinonemertes	1.00	0.93	1.00	0.93	1.00	0.93	1.50	0.90	1.00	0.93
78	Perculator	Peculator	1.00	0.90	1.00	0.90	1.00	0.90	1.50	0.85	1.00	0.90
79	Perusa	Pherusa	1.00	0.86	1.00	0.86	1.00	0.86	1.50	0.79	1.50	0.79
80	Pholae	Pholoe	1.00	0.83	1.00	0.83	1.00	0.83	1.00	0.83	1.00	0.83
81	Phsychroteuthis	Psychroteuthis	1.00	0.93	1.00	0.93	1.00	0.93	1.50	0.90	1.00	0.93
82	Platycheles	Pachycheles	3.00	0.73	3.00	0.73	3.00	0.73	3.00	0.73	2.00	0.82
83	Platyclina	Platyclinia	1.00	0.91	1.00	0.91	1.00	0.91	1.50	0.86	0.50	0.95
84	Posidonemertes	Poseidonemertes	1.00	0.93	1.00	0.93	1.00	0.93	1.50	0.90	1.00	0.93
85	Prophias	Prophlias	1.00	0.89	1.00	0.89	1.00	0.89	1.50	0.83	1.00	0.89
86	Pseudocrptocoeloma	Pseudocryptocoeloma	1.00	0.95	1.00	0.95	1.00	0.95	1.50	0.92	1.00	0.95
87	Pseuechinus	Pseudechinus	1.00	0.92	1.00	0.92	1.00	0.92	1.50	0.88	1.00	0.92
88	Pseuochydorus	Pseudochydorus	1.00	0.93	1.00	0.93	1.00	0.93	1.50	0.89	1.00	0.93
89	Ptoterato	Proterato	1.00	0.89	1.00	0.89	1.00	0.89	1.00	0.89	1.00	0.89
90	Pulojacquesia	Pylojacquesia	1.00	0.92	1.00	0.92	1.00	0.92	1.00	0.92	1.00	0.92
91	Rhinocetes	Rhinoecetes	1.00	0.91	1.00	0.91	1.00	0.91	1.50	0.86	1.00	0.91
92	Rhodocorton	Rhodochorton	1.00	0.92	1.00	0.92	1.00	0.92	1.50	0.88	0.50	0.96
93	Saganafusus	Saginafusus	1.00	0.91	1.00	0.91	2.00	0.82	1.00	0.91	1.00	0.91
94	Samacis	Salmacis	1.00	0.88	1.00	0.88	1.00	0.88	1.50	0.81	1.00	0.88
95	Scruoaria	Scruparia	1.00	0.89	1.00	0.89	1.00	0.89	1.00	0.89	1.00	0.89

No	Source	Target	Dist	Sim %	Dist	Sim %	Dist	Sim %	Dist	Sim %	Dist	Sim %
96	Scyllaercturella	Scyllarcturella	1.00	0.94	1.00	0.94	1.00	0.94	1.50	0.91	0.50	0.97
97	Sigamba	Sigambra	1.00	0.88	1.00	0.88	1.00	0.88	1.50	0.81	1.00	0.88
98	Simicarcinus	Simocarcinus	1.00	0.92	1.00	0.92	1.00	0.92	1.00	0.92	1.00	0.92
99	Soleniscus	Soleneiscus	1.00	0.91	1.00	0.91	1.00	0.91	1.50	0.86	0.50	0.95
100	Solster	Solaster	1.00	0.88	1.00	0.88	1.00	0.88	1.50	0.81	1.00	0.88
101	Sphaerodorpsis	Sphaerodoropsis	1.00	0.93	1.00	0.93	1.00	0.93	1.50	0.90	0.50	0.97
102	Sternapsis	Sternaspis	2.00	0.80	1.00	0.90	1.00	0.90	2.00	0.80	1.00	0.90
103	Stichpathes	Stichopathes	1.00	0.92	1.00	0.92	1.00	0.92	1.50	0.88	1.00	0.92
104	Stolsterias	Stolasterias	1.00	0.92	1.00	0.92	1.00	0.92	1.50	0.88	1.00	0.92
105	Styllopallene	Stylopallene	1.00	0.92	1.00	0.92	1.00	0.92	1.00	0.92	0.00	1.00
106	Sycyonella	Sicyonella	1.00	0.90	1.00	0.90	1.00	0.90	1.00	0.90	1.00	0.90
107	Syneryhtrops	Synerythrops	2.00	0.83	1.00	0.92	1.00	0.92	2.00	0.83	1.00	0.92
108	Synoicium	Synoicum	1.00	0.89	1.00	0.89	1.00	0.89	1.50	0.83	0.50	0.94
109	Trematrochus	Trematotrochus	2.00	0.86	2.00	0.86	2.00	0.86	2.00	0.86	1.50	0.89
110	Tubastrea	Tubastraea	1.00	0.90	1.00	0.90	1.00	0.90	1.50	0.85	1.00	0.90
111	Tulbergella	Tullbergella	1.00	0.92	1.00	0.92	1.00	0.92	1.00	0.92	0.00	1.00
112	Ventrojassa	Ventojassa	1.00	0.91	1.00	0.91	1.00	0.91	1.50	0.86	0.50	0.95
113	Vermitexa	Vermitexta	1.00	0.90	1.00	0.90	1.00	0.90	1.50	0.85	1.00	0.90
114	Wombalana	Wombalano	1.00	0.89	1.00	0.89	1.00	0.89	0.50	0.94	0.50	0.94
115	Zobraco	Zobracho	1.00	0.88	1.00	0.88	1.00	0.88	1.50	0.81	1.00	0.88
Average (percentage similarity)				0.89		0.90		0.90		0.87		0.91

Table 5.3. Results of name matching in Portuguese (pairs 120).

No	String.		Compared Algorithm								Proposed Algorithm	
			LD		N-DIST		DLD		MDLD		E-N-DIST	
	Source	Target	Dist	Sim %	Dist	Sim %	Dist	Sim %	Dist	Sim %	Dist	Sim %
1	acerca	àcerca	1	0.83	1.5	0.75	1	0.83	1	0.83	1.2	0.80
2	açoriano	açoreano	1	0.88	1	0.88	1	0.88	1	0.88	0.4	0.95
3	alcoolémia	alcoolemia	1	0.90	1	0.90	1	0.90	1	0.90	0.4	0.96
4	ameixial	ameixeal	1	0.88	1	0.88	1	0.88	1	0.88	0.4	0.95
5	antárctico	antártico	1	0.90	1.5	0.85	1	0.90	1	0.90	1.2	0.88
6	antepor	antepôr	1	0.86	1	0.86	1	0.86	1	0.86	0.4	0.94
7	árctico	artico	2	0.71	3	0.57	2	0.71	2	0.71	2.4	0.66
8	artífice	artífece	1	0.88	1	0.88	1	0.88	1	0.88	0.4	0.95
9	bainha	baínha	1	0.83	1	0.83	1	0.83	1	0.83	0.4	0.93
10	bebé	bébé	1	0.75	1	0.75	1	0.75	1	0.75	0.4	0.90
11	bege	beje	1	0.75	1	0.75	1	0.75	1	0.75	0.4	0.90
12	bênção	benção	1	0.83	1	0.83	1	0.83	1	0.83	0.4	0.93
13	benefcência	benefciência	1	0.92	1.5	0.88	1	0.92	1	0.92	1.2	0.90
14	biopsia	biópsia	1	0.86	1	0.86	1	0.86	1	0.86	0.4	0.94
15	burburinho	borborinho	2	0.80	2	0.80	2	0.80	2	0.80	0.8	0.92
16	caiem	caem	1	0.80	1.5	0.70	1	0.80	1	0.80	1.2	0.76
17	calvície	calvíce	1	0.88	1.5	0.81	1	0.88	1	0.88	1.2	0.85
18	camoniano	camoneano	1	0.89	1	0.89	1	0.89	1	0.89	0.4	0.96

19	campeão	campião	1	0.86	1	0.86	1	0.86	1	0.86	0.4	0.94
20	chiita	xiita	2	0.67	2.5	0.58	2	0.67	2	0.67	2.2	0.63
21	comboio	combóio	1	0.86	1	0.86	1	0.86	1	0.86	0.4	0.94
22	compor	compôr	1	0.83	1	0.83	1	0.83	1	0.83	0.4	0.93
23	comummente	comumente	1	0.90	1	0.90	1	0.90	1	0.90	1	0.90
24	constituia	constituía	1	0.90	1	0.90	1	0.90	1	0.90	0.4	0.96
25	constituiu	constituíu	1	0.90	1	0.90	1	0.90	1	0.90	0.4	0.96
26	cor	côr	1	0.67	1	0.67	1	0.67	1	0.67	0.4	0.87
27	crânio	crâneo	1	0.83	1	0.83	1	0.83	1	0.83	0.4	0.93
28	definição	defenição	1	0.89	1	0.89	1	0.89	1	0.89	0.4	0.96
29	definido	defenido	1	0.88	1	0.88	1	0.88	1	0.88	0.4	0.95
30	definir	defenir	1	0.86	1	0.86	1	0.86	1	0.86	0.4	0.94
31	desequilíbrio	desequilibrio	1	0.92	1	0.92	1	0.92	1	0.92	0.4	0.97
32	despretensioso	despretencioso	1	0.93	1	0.93	1	0.93	1	0.93	0.4	0.97
33	dignatários	dignitários	1	0.91	1	0.91	1	0.91	1	0.91	0.4	0.96
34	dispender	despender	1	0.89	1	0.89	1	0.89	1	0.89	0.4	0.96
35	dispêndio	dispendio	1	0.89	1	0.89	1	0.89	1	0.89	0.4	0.96
36	ecrã	ecran	2	0.60	1.5	0.70	2	0.60	2	0.60	1.2	0.76
37	emirados	emiratos	1	0.88	1	0.88	1	0.88	1	0.88	0.4	0.95
38	esotérico	isotérico	1	0.89	1.5	0.83	1	0.89	1	0.89	1.2	0.87
39	esquisito	esquesito	1	0.89	1	0.89	1	0.89	1	0.89	0.4	0.96
40	estratego	estratega	1	0.89	0.5	0.94	1	0.89	1	0.89	0.2	0.98
41	feminino	femenino	1	0.88	1	0.88	1	0.88	1	0.88	0.4	0.95
42	feminismo	femininismo	2	0.82	2	0.82	2	0.82	2	0.82	2	0.82
43	fôr	for	1	0.67	1	0.67	1	0.67	1	0.67	0.4	0.87
44	gineceu	geneceu	1	0.86	1	0.86	1	0.86	1	0.86	0.4	0.94
45	gorjeta	gorgeta	1	0.86	1	0.86	1	0.86	1	0.86	0.4	0.94
46	granjear	grangear	1	0.88	1	0.88	1	0.88	1	0.88	0.4	0.95
47	guisar	guizar	1	0.83	1	0.83	1	0.83	1	0.83	0.4	0.93
48	halariedade	hilaridade	2	0.82	2.5	0.77	2	0.82	2	0.82	1.6	0.85
49	hectare	hectar	1	0.86	1	0.86	1	0.86	1	0.86	1	0.86
50	hiroshima	hiroxima	2	0.78	2	0.78	2	0.78	2	0.78	1.4	0.84
51	ilacção	elação	2	0.71	3	0.57	2	0.71	2	0.71	2.4	0.66
52	indispensável	indespensável	1	0.92	1	0.92	1	0.92	1	0.92	0.4	0.97
53	inflacção	inflação	1	0.89	1.5	0.83	1	0.89	1	0.89	1.2	0.87
54	interveio	interviu	2	0.78	2	0.78	2	0.78	2	0.78	1.4	0.84
55	intervindo	intervido	1	0.90	1.5	0.85	1	0.90	1	0.90	1.2	0.88
56	invocar	evocar	2	0.71	2.5	0.64	2	0.71	2	0.71	2.2	0.69
57	ípsilon	ipslon	2	0.71	3	0.57	2	0.71	2	0.71	2.4	0.66
58	irisar	irizar	1	0.83	1	0.83	1	0.83	1	0.83	0.4	0.93
59	irupção	irrupção	1	0.88	1	0.88	1	0.88	1	0.88	1	0.88
60	jeropiga	geropiga	1	0.88	1.5	0.81	1	0.88	1	0.88	1.2	0.85
61	juiz	juíz	1	0.75	1	0.75	1	0.75	1	0.75	0.4	0.90

62	lampião	lampeão	1	0.86	1	0.86	1	0.86	1	0.86	0.4	0.94
63	lêem	lêm	1	0.75	1.5	0.63	1	0.75	1	0.75	1.2	0.70
64	linguista	linguísta	1	0.89	1	0.89	1	0.89	1	0.89	0.4	0.96
65	lisonjear	lisongear	1	0.89	1	0.89	1	0.89	1	0.89	0.4	0.96
66	logótipo	logotipo	1	0.88	1	0.88	1	0.88	1	0.88	0.4	0.95
67	maciço	massiço	2	0.71	2	0.71	2	0.71	2	0.71	1.4	0.80
68	majestade	magestade	1	0.89	1	0.89	1	0.89	1	0.89	0.4	0.96
69	manjerico	mangerico	1	0.89	1	0.89	1	0.89	1	0.89	0.4	0.96
70	manjerona	mangerona	1	0.89	1	0.89	1	0.89	1	0.89	0.4	0.96
71	meteorologia	metereologia	2	0.83	2	0.83	2	0.83	2	0.83	1.4	0.88
72	miscigenação	miscegenação	1	0.92	1	0.92	1	0.92	1	0.92	0.4	0.97
73	nonagésimo	nonagessimo	2	0.82	2	0.82	2	0.82	2	0.82	1.4	0.87
74	oceânia	oceania	1	0.86	1	0.86	1	0.86	1	0.86	0.4	0.94
75	oficina	ofecina	1	0.86	1	0.86	1	0.86	1	0.86	0.4	0.94
76	opróbrio	opróbio	1	0.88	1.5	0.81	1	0.88	1	0.88	1.2	0.85
77	organograma	organigrama	1	0.91	1	0.91	1	0.91	1	0.91	0.4	0.96
78	paralisar	paralizar	1	0.89	1	0.89	1	0.89	1	0.89	0.4	0.96
79	perserverança	preseverança	3	0.77	3.5	0.73	2	0.85	2	0.85	1.6	0.88
80	persuasão	persuação	1	0.89	1	0.89	1	0.89	1	0.89	0.4	0.96
81	pirinéus	pirenéus	1	0.88	1	0.88	1	0.88	1	0.88	0.4	0.95
82	pretensioso	pretencioso	1	0.91	1	0.91	1	0.91	1	0.91	0.4	0.96
83	privilégio	previlégio	1	0.90	1	0.90	1	0.90	1	0.90	0.4	0.96
84	quadricromia	quadricomia	1	0.92	1.5	0.88	1	0.92	1	0.92	1.2	0.90
85	quadruplicado	quadriplicado	1	0.92	1	0.92	1	0.92	1	0.92	0.4	0.97
86	quasímodo	quasimodo	1	0.89	1	0.89	1	0.89	1	0.89	0.4	0.96
87	quilo	kilo	2	0.60	2.5	0.50	2	0.60	2	0.60	2.2	0.56
88	quilograma	kilograma	2	0.80	2.5	0.75	2	0.80	2	0.80	2.2	0.78
89	quilómetro	kilómetro	2	0.80	2.5	0.75	2	0.80	2	0.80	2.2	0.78
90	quis	quiz	1	0.75	0.5	0.88	1	0.75	1	0.75	0.2	0.95
91	rainha	raínha	1	0.83	1	0.83	1	0.83	1	0.83	0.4	0.93
92	raiz	raíz	1	0.75	1	0.75	1	0.75	1	0.75	0.4	0.90
93	raul	raúl	1	0.75	1	0.75	1	0.75	1	0.75	0.4	0.90
94	rectaguarda	retaguarda	1	0.91	1.5	0.86	1	0.91	1	0.91	1.2	0.89
95	rédea	rédia	1	0.80	1	0.80	1	0.80	1	0.80	0.4	0.92
96	regurgitar	regurjitar	1	0.90	1	0.90	1	0.90	1	0.90	0.4	0.96
97	rejeitar	regeitar	1	0.88	1	0.88	1	0.88	1	0.88	0.4	0.95
98	requeiro	requero	1	0.88	1.5	0.81	1	0.88	1	0.88	1.2	0.85
99	réstia	réstea	1	0.83	1	0.83	1	0.83	1	0.83	0.4	0.93
100	rubrica	rúbrica	1	0.86	1	0.86	1	0.86	1	0.86	0.4	0.94
101	saem	saiem	1	0.80	1.5	0.70	1	0.80	1	0.80	1.2	0.76
102	saloiice	saloice	1	0.88	1	0.88	1	0.88	1	0.88	1	0.88
103	sarjeta	sargeta	1	0.86	1	0.86	1	0.86	1	0.86	0.4	0.94
104	semear	semiar	1	0.83	1	0.83	1	0.83	1	0.83	0.4	0.93

105	suíça	suiça	1	0.80	1	0.80	1	0.80	1	0.80	0.4	0.92
106	supor	supôr	1	0.80	1	0.80	1	0.80	1	0.80	0.4	0.92
107	trânsfuga	transfuga	1	0.89	1	0.89	1	0.89	1	0.89	0.4	0.96
108	transpôr	transpor	1	0.88	1	0.88	1	0.88	1	0.88	0.4	0.95
109	urano	úrano	1	0.80	1.5	0.70	1	0.80	1	0.80	1.2	0.76
110	ventoinha	ventoínha	1	0.89	1	0.89	1	0.89	1	0.89	0.4	0.96
111	verosímil	verosímel	1	0.89	1	0.89	1	0.89	1	0.89	0.4	0.96
112	vigilante	vegilante	1	0.89	1	0.89	1	0.89	1	0.89	0.4	0.96
113	vôo	voo	1	0.67	1	0.67	1	0.67	1	0.67	0.4	0.87
114	vultuoso	vultoso	1	0.88	1.5	0.81	1	0.88	1	0.88	1.2	0.85
115	xadrez	xadrês	2	0.67	1.5	0.75	2	0.67	2	0.67	1.2	0.80
116	xamã	chamã	2	0.60	2.5	0.50	2	0.60	2	0.60	2.2	0.56
117	xelindró	xilindró	1	0.88	1	0.88	1	0.88	1	0.88	0.4	0.95
118	zângão	zangão	1	0.83	1	0.83	1	0.83	1	0.83	0.4	0.93
119	zepelin	zeppelin	1	0.88	1	0.88	1	0.88	1	0.88	1	0.88
120	zoo	zoô	1	0.67	0.5	0.83	1	0.67	1	0.67	0.2	0.93
	Average (percentage similarity)			**0.84**		**0.82**		**0.84**		**0.84**		**0.89**

Appendix B:Evaluation F- MEASURE Tables for all Dataset

Table 5.9. F1-scores of different algorithms,, thresholds and similarity calculation methods (Dataset 6 English pairs 60)

	Algorithms \ thresholds	65	70	75	80	85	90
1	(LD)	0.966	0.957	0.868	0.857	0.812	0.182
2	(DLD)	1.000	1.000	1.000	1.000	0.919	0.182
3	(N-DIST)	0.966	0.957	0.879	0.800	0.667	0.154
4	MDLD	1.000	1.000	1.000	1.000	0.919	0.182
5	(E-N-DIST)	1.000	1.000	1.000	1.000	0.957	0.519

Table 5.10. F1-scores of different algorithms,, thresholds and similarity calculation methods (Dataset 4 'Arabic name (600 pairs))

	Algorithms \ thresholds	65	70	75	80	85	90
1	(LD)	0.956	0.937	0.928	0.899	0.250	0.250
2	(DLD)	0.978	0.959	0.951	0.900	0.250	0.250
3	(N-DIST)	0.941	0.931	0.776	0.423	0.096	0.096
4	MDLD	0.978	0.959	0.951	0.900	0.250	0.250
5	(E-N-DIST)	0.996	0.992	0.962	0.881	0.552	0.551

Table 5.11. F1-scores of different algorithms,, thresholds and similarity calculation methods (Dataset 12 'GRIN genera' (189 pairs)).

	Algorithms \ thresholds	65	70	75	80	85	90
1	(LD)	0.997	0.995	0.989	0.970	0.901	0.705
2	(DLD)	0.997	0.997	0.992	0.976	0.923	0.732
3	(N-DIST)	0.995	0.989	0.984	0.953	0.882	0.470
4	MDLD	0.997	0.997	0.992	0.976	0.920	0.727
5	(E-N-DIST)	1.000	1.000	1.000	0.987	0.953	0.792

Table 5.12. F1-scores of different algorithms,, thresholds and similarity calculation methods (Dataset 9 'CAAB' (641 pairs)).

	Algorithms \ thresholds	65	70	75	80	85	90
1	(LD)	1.000	1.000	1.000	0.998	0.988	0.942
2	(DLD)	1.000	1.000	1.000	0.998	0.988	0.947
3	(N-DIST)	1.000	1.000	0.998	0.998	0.987	0.955
4	MDLD	1.000	1.000	1.000	0.998	0.991	0.946
5	(E-N-DIST)	1.000	1.000	1.000	1.000	0.998	0.986

Table 5.13. F1-scores of different algorithms,, thresholds and similarity calculation methods (Dataset 14 'CAABWEB Genera' (853 pairs)).

	Algorithms \ thresholds	65	70	75	80	85	90
1	(LD)	0.998	0.995	0.986	0.954	0.853	0.673
2	(DLD)	0.998	0.996	0.992	0.970	0.886	0.710
3	(N-DIST)	0.994	0.985	0.974	0.911	0.783	0.419
4	MDLD	0.999	0.997	0.992	0.969	0.887	0.711
5	(E-N-DIST)	0.999	0.998	0.995	0.981	0.931	0.786

Table 5.14. F1-scores of different algorithms,, thresholds and similarity calculation methods (Dataset 10 'Dalcin name pairs' (171 pairs)

	Algorithms \ thresholds	65	70	75	80	85	90
1	(LD)	1.000	1.000	1.000	1.000	1.000	0.982
2	(DLD)	1.000	1.000	1.000	1.000	1.000	0.988
3	(N-DIST)	1.000	1.000	1.000	1.000	1.000	0.985
4	MDLD	1.000	1.000	1.000	1.000	1.000	0.982
5	(E-N-DIST)	1.000	1.000	1.000	1.000	1.000	0.982

Table 5.15. F1-scores of different algorithms,, thresholds and similarity calculation methods (Dataset 13 'CAAB Genera' (115 pairs)

	Algorithms \ thresholds	65	70	75	80	85	90
1	(LD)	1.000	1.000	0.991	0.978	0.940	0.777
2	(DLD)	1.000	1.000	0.991	0.982	0.969	0.814
3	(N-DIST)	1.000	1.000	0.991	0.973	0.850	0.456
4	MDLD	1.000	1.000	0.991	0.982	0.959	0.796
5	(E-N-DIST)	1.000	1.000	1.000	0.996	0.978	0.867

Table 5.16. F1-scores of different algorithms,, thresholds and similarity calculation methods (Dataset 11 'CAABWEB' (2047 pairs).

	Algorithms \ thresholds	65	70	75	80	85	90
1	(LD)	1.000	1.000	1.000	0.996	0.976	0.898
2	(DLD)	1.000	1.000	1.000	0.996	0.981	0.913
3	(N-DIST)	1.000	1.000	0.998	0.993	0.966	0.877
4	MDLD	1.000	1.000	1.000	0.996	0.980	0.910
5	(E-N-DIST)	1.000	1.000	1.000	0.996	0.980	0.910