

# Anexos

---

## **Anexo A: Unidades léxicas determinantes**

Se presentan las unidades léxicas identificadas como determinantes y se lista su categoría (artículo definido, cardinal, etc.) marcando con una **x** las funciones adicionales, detectadas y más comunes, que puede cumplir en la oración (ver 3.2.3).

Unidad léxica	Funciones			
	Determinante	Adjetivo	Pronombre	Nombre
1. el	artículo			
2. la	artículo		x	
3. las	artículo		x	
4. lo	artículo		x	
5. los	artículo		x	
6. cientos	cardinal		x	
7. uno, dos, etc.	cardinal		x	
8. miles	cardinal		x	
9. millones	cardinal		x	
10. aquel	demostrativo	X		
11. aquella	demostrativo	X		
12. aquellas	demostrativo	X		
13. aquellos	demostrativo	X		
14. esa	demostrativo	X		
15. esas	demostrativo	X		
16. ese	demostrativo	X		
17. esos	demostrativo	X		
18. esta	demostrativo	X		
19. estas	demostrativo	X		
20. este	demostrativo	X		
21. estos	demostrativo	X		
22. semejante	demostrativo	X		x
23. semejantes	demostrativo	X		x
24. tal	demostrativo			
25. tales	demostrativo			
26. cada	distributivo	X	x	
27. algún	indefinido			

Unidad léxica	Funciones			
	Determinante	Adjetivo	Pronombre	Nombre
28. alguna	indefinido		x	
29. algunas	indefinido		x	
30. algunos	indefinido		x	
31. ambas	indefinido		x	
32. ambos	indefinido		x	
33. bastante	indefinido	X	x	
34. bastantes	indefinido	X	x	
35. cierta	indefinido	X		
36. ciertas	indefinido	X		
37. cierto	indefinido	X		x
38. ciertos	indefinido	X		
39. cualquier	indefinido			
40. cuanta	indefinido	X	x	
41. cuantas	indefinido	X	x	
42. cuantísima	indefinido	X	x	
43. cuantísimas	indefinido	X	x	
44. cuantísimo	indefinido	X	x	
45. cuantísimos	indefinido	X	x	
46. cuanto	indefinido	X	x	x
47. cuantos	indefinido	X	x	
48. demasiada	indefinido	X	x	
49. demasiadas	indefinido	X	x	
50. demasiado	indefinido	X	x	
51. demasiados	indefinido	X	x	
52. más	indefinido	X	x	
53. menos	indefinido	X	x	
54. mucha	indefinido	X	x	
55. muchas	indefinido	X	x	
56. muchísima	indefinido	X	x	
57. muchísimas	indefinido	X	x	
58. muchísimo	indefinido	X	x	
59. muchísimos	indefinido	X	x	
60. mucho	indefinido	X	x	
61. muchos	indefinido	X	x	
62. ningún	indefinido			
63. ninguna	indefinido		x	
64. ningunas	indefinido		x	
65. ningunos	indefinido		x	
66. otra	indefinido		x	
67. otras	indefinido		x	
68. otro	indefinido		x	
69. otros	indefinido		x	
70. poca	indefinido	X	x	

Unidad léxica	Funciones			Nombre
	Determinante	Adjetivo	Pronombre	
71. pocas	indefinido	X	x	
72. poco	indefinido	X	x	
73. pocos	indefinido	X	x	
74. poquísima	indefinido	X	x	
75. poquísimas	indefinido	X	x	
76. poquísimo	indefinido	X	x	
77. poquísimos	indefinido	X	x	
78. tanta	indefinido	X	x	
79. tantas	indefinido	X	x	
80. tantísima	indefinido	X	x	
81. tantísimas	indefinido	X	x	
82. tantísimo	indefinido	X	x	
83. tantísimos	indefinido	X	x	
84. tanto	indefinido	X	x	X
85. tantos	indefinido	X	x	X
86. toda	indefinido	X	x	
87. todas	indefinido	X	x	
88. todo	indefinido	X	x	X
89. todos	indefinido	X	x	
90. un	indefinido			
91. una	indefinido		x	x
92. unas	indefinido		x	x
93. unos	indefinido		x	x
94. varias	indefinido			
95. varios	indefinido			
96. cuánta	interrogativo		x	
97. cuántas	interrogativo		x	
98. cuánto	interrogativo		x	
99. cuántos	interrogativo		x	
100. qué	interrogativo		x	
101. cuya	posesivo			
102. cuyas	posesivo			
103. cuyo	posesivo			X
104. cuyos	posesivo			X
105. mi	posesivo			
106. mis	posesivo			
107. nuestra	posesivo	X		
108. nuestras	posesivo	X		
109. nuestro	posesivo	X		
110. nuestros	posesivo	X		
111. su	posesivo		x	
112. sus	posesivo		x	

## Anexo B: Características de documentos usados en los experimentos

Archivo	Palabra	Adj	Adv	Det	Nom	Verb	Pron	Inter	Conj	Prep	Punt	Abrev	Num	Fecha	Desc	UL
a1	2152	174	122	333	510	370	152	2	186	303	382	0	0	7	0	2541
a2	194	20	7	36	55	22	7	0	7	40	22	0	2	0	0	218
a4	204	21	17	35	45	25	16	0	16	29	18	0	0	0	0	222
a10a	912	102	59	152	199	135	71	0	53	141	171	0	1	0	0	1084
a10b	908	74	58	133	203	172	84	0	78	106	180	0	0	0	0	1088
a11a	812	47	65	114	192	150	76	3	55	110	187	0	0	0	0	999
a11b	645	66	42	111	138	107	45	1	50	85	93	0	0	0	0	738
a12	1355	101	82	234	310	187	122	1	112	206	140	0	0	0	0	1495
a13a	984	78	54	164	238	158	84	2	53	153	163	0	2	0	0	1149
a13b	918	58	49	136	220	160	79	1	66	149	181	0	1	0	0	1100
a13c	471	40	35	68	106	81	38	1	26	76	69	0	1	0	0	541
a14	824	58	40	140	198	136	52	0	68	132	134	0	2	1	0	961
a15a	982	70	53	157	272	154	62	3	71	140	186	0	2	7	0	1177
a15b	552	40	27	83	118	102	45	0	45	92	62	0	3	3	0	620
a15c	496	50	30	77	111	76	33	0	34	85	101	0	0	0	0	597
a18	717	39	42	119	150	121	57	0	63	126	74	0	3	0	0	794
a19	95	8	6	17	23	14	8	0	6	13	11	0	0	0	0	106
a20	501	37	25	85	127	68	28	0	43	88	65	0	0	0	0	566
a21a	1361	159	93	210	314	198	74	0	90	223	216	0	11	1	0	1589
a21b	610	66	41	84	142	92	42	0	45	98	85	0	6	4	0	705
a21c	390	43	24	66	89	54	26	0	24	64	53	0	0	5	0	448
a22a	770	50	44	119	157	160	70	2	57	111	130	0	0	0	0	900
a22b	297	26	16	42	63	58	20	0	30	42	36	0	0	0	0	333
a23a	997	123	59	160	229	131	70	0	65	160	130	0	0	0	0	1127
a23b	1229	111	92	202	255	201	98	0	84	186	163	0	0	0	0	1392
a24	2094	134	84	429	545	243	171	0	161	327	338	0	1	1	0	2434
a25a	1688	150	58	319	418	234	89	4	99	317	230	0	2	13	0	1933
a25b	664	35	38	122	159	108	51	0	42	109	84	0	0	0	0	748
a26a	965	81	53	158	216	165	77	0	76	139	121	0	0	0	0	1086
a26b	869	107	28	194	224	96	38	0	60	122	121	0	0	0	0	990
a26c	845	77	59	129	184	151	62	0	68	115	104	0	0	1	0	950
a27	547	37	26	100	132	75	48	0	41	88	57	0	1	2	0	607
a28a	870	60	43	140	213	146	57	0	73	138	157	0	1	1	0	1029
a28b	821	62	43	134	219	116	51	1	53	142	151	0	1	0	0	973
a28c	242	21	14	44	59	35	12	0	17	40	30	0	0	2	0	274
a29	419	40	36	51	89	67	49	0	43	44	71	0	0	0	0	490
a30a	895	44	75	125	196	152	97	12	79	115	301	0	0	0	0	1196
a30b	161	5	12	23	34	35	13	5	15	19	55	0	0	0	0	216
Suma	30456	2514	1751	5045	7152	4755	2274	38	2254	4673	4872	0	40	48	0	35416
Promedio	801.5	66.2	46.1	132.8	188.2	125.1	59.8	1.0	59.3	123.0	128.2	0.0	1.1	1.3	0.0	932.0
Desvstd	467.3	41.2	25.5	85.7	115.7	70.1	36.1	2.2	36.5	74.1	84.4	0.0	2.1	2.7	0.0	545.7

## **Anexo C: Ejemplo de texto usado para el experimento**

1. Cuando escribo esto la Madre\_Coraje peruana acaba de ser reventada por los senderistas.
2. Veo su foto en los periódicos: una mujer joven, atractiva, probablemente zamba, esto\_es, mestiza de negra e india; oscura de color, en\_fin, como son oscuros todos los habitantes de las villas limeñas, arrabales de miseria en donde se hacinan cientos de miles de personas.
3. Son, en su mayoría, indígenas que bajaron de los Andes huyendo del hambre, del atraso y la tuberculosis; quisieron llegar a la ciudad, pero quedaron varados en las afueras, a una decena de kilómetros, en los sórdidos arenales que rodean Lima, en donde plantaron sus chabolas, precarios tenderetes de cartón y cajones astillados.
4. Aspiraban a más: a mejorar su situación, a ser felices; pero les atrapó la miseria suburbana y subhumana, la ferocidad y la violencia de los arrabales.
5. Alguno consigue escapar de allí, muy de\_tarde\_en\_tarde; pero para la inmensa mayoría no hay ni salida ni retorno.
6. En los alrededores de Lima malviven entre chabolas varios millones de personas; sólo en Villa\_El\_Salvador, el suburbio de la Madre\_Coraje, hay 300.000 habitantes.
7. Hace 14 años visité uno de estos asentamientos, y creo recordar que fue precisamente Villa\_El\_Salvador.
8. Me acompañaba un muchacho que había logrado la proeza de escaparse del barrio; su familia seguía viviendo allí, pero él había conseguido un trabajo y una cama en la ciudad.
9. Yo quería conocer aquello, hacer un reportaje, y él se ofreció a servirme de guía.
10. Fuimos hasta allí, tras un largo trayecto de traqueteantes autobuses, para almorzar con su hermana.
11. La villa había nacido como una excrescencia cancerosa, desordenadamente, allí donde termina el asfalto y la esperanza.
12. Imaginen un pueblo de chabolas de varios cientos de miles de personas: caminas y caminas por los mugrientos arenales y la miseria resulta inacabable, inabarcable.
13. Era media mañana y había bastante gente en las calles, esto\_es, en las veredas sin urbanizar que habían quedado abiertas entre las chozas.
14. Mi guía iba evitando, precisamente, los lugares más poblados: avanzábamos por las callejas solitarias, huyendo de la gente, porque te podían rajar en\_mitad\_de una muchedumbre sin que nadie hiciera nada.
15. La policía no se atrevía a entrar en las villas: eran lugares sin ley, territorios prohibidos.
16. Llegamos al\_fin a la casa de la hermana: un tenderete de latas y cartones de apenas tres metros por tres metros, con un infernillo de gas, una silla de enea, un plástico y una manta en un rincón del suelo, sobre la sucia arena.
17. Allí vivían, no sé cómo, la hermana de mi amigo, su marido y cuatro niños.
18. Comimos un arroz con pollo muy sabroso.

19. A mí la invitada de honor, me hicieron sentar en la única silla.
20. Pienso ahora en María\_Elena\_Moyano, la Madre\_Coraje, y recuerdo aquel almuerzo conmovedor, aquel poblado espeluznante.
21. El liderazgo de María\_Elena nació de aquella miseria y de una increíble voluntad de superación.
22. De la generosidad, de la inteligencia, del tesón.
23. Consiguí, me dicen, organizar el suburbio, y sacarlo de la violencia y el salvajismo.
24. Construyó un marco de dignidad en el que reconocerse: luchó por otorgar una dimensión humana a unas vidas embrutecidas por la miseria.
25. Por todo esto le dieron el premio Príncipe\_de\_Asturias\_de\_la\_Paz de 1987: porque logró ser una persona, aunque nada en su entorno se lo permitiese.
26. Y por todo esto ha sido ahora asesinada.
27. Porque los fanáticos de Sendero\_Luminoso no pueden admitir que haya seres libres.
28. La verdadera heroicidad no es un acto único: el soldado que se inmola para salvar a sus compañeros, el hombre que entra en un edificio en llamas a rescatar a un niño, son sin\_duda personas admirables porque supieron responder a un instante de gloria, de generosidad y de exigencia.
29. Pero ese momento heroico, me parece, es más una especie de fiebre que un talante.
30. Por\_el\_contrario, la verdadera heroicidad se construye calladamente, día a día, sobreponiéndose una y otra vez a circunstancias dolorosas y extremas.
31. Como hizo María\_Elena.
32. Es heroico levantarse todas las mañanas a luchar contra la desesperación y la incultura.
33. Y es heroico sentir miedo ante las amenazas senderistas y seguir actuando año tras año, sin\_embargo, como si el propio miedo no existiese.
34. Pero, al\_fin, el terror atrapó a María\_Elena; fueron a buscarla, y duele imaginar el instante de pánico que debió de sentir ante sus asesinos.
35. ¿Sirve de algo esa muerte salvaje y absurda?
36. ¿Es cierta esa versión cristiana y consoladora que asegura que los sacrificios no son inútiles?
37. Probablemente no; probablemente, en la historia concreta de los suburbios limeños la desaparición de María\_Elena sea tan sólo una catástrofe.
38. Pero sí creo que la vida de Moyano, su entereza hasta el final y su coraje, forma parte del legado de los humanos, del inconsciente colectivo, de la sustancia común que todos somos.
39. Y aun cuando la olvidemos, como hemos olvidado a todos los demás héroes anónimos, es gracias\_a ella, y a gentes como ella, que la humanidad puede perseverar en el sueño de la felicidad y la razón.
40. Porque ellos nos demuestran que el bien también existe.

## Anexo D: Ejemplo de archivo de entrada etiquetado

Cuando cuando CS  
escribo escribir VMIP1S0  
esto este PD0NS000  
la el DA0FS0  
Madre\_Coraje madre\_coraje NP00000  
peruana peruano AQ0FS0  
acaba acabar VMIP3S0  
de de SPS00  
ser ser VSN0000  
reventada reventar VMP00SF  
por por SPS00  
los el DA0MP0  
senderistas senderista NCCP000  
. . Fp  
Veo ver VMIP1S0  
su su DP3CS0  
foto foto NCFS000  
en en SPS00  
los el DA0MP0  
periódicos periódico NCMP000  
: : Fd  
una uno DI0FS0  
mujer mujer NCFS000  
joven joven AQ0CS0  
, , Fc  
atractiva atractivo AQ0FS0  
, , Fc  
probablemente probablemente RG  
zamba zambo AQ0FS0  
, , Fc  
esto\_es esto\_es CC  
, , Fc  
mestiza mestizo AQ0FS0  
de de SPS00  
negra negro AQ0FS0  
e e CC  
india indio AQ0FS0  
; ; Fx  
oscura oscuro AQ0FS0  
de de SPS00  
color color NCMS000  
, , Fc  
en\_fin en\_fin RG  
, , Fc  
como como CS  
son ser VSIP3P0  
oscuros oscuro AQ0MP0  
todos todo DI0MP0  
los el DA0MP0  
habitantes habitante NCCP000  
de de SPS00  
las el DA0FP0  
villas villa NCFP000  
limeñas limeño AQ0FP0  
, , Fc  
arrabales arrabal NCMP000  
de de SPS00  
miseria miseria NCFS000  
en en SPS00  
donde donde PR000000  
se él P0300000  
hacinan hacinar VMIP3P0  
cientos ciento PN0CP000  
de de SPS00  
miles mil PN0CP000  
de de SPS00  
personas persona NCFP000  
. . Fp  
Son ser VSIP3P0  
, , Fc  
en en SPS00  
su su DP3CS0  
mayoría mayoría NCFS000  
, , Fc  
indígenas indígena NCCP000  
que que PR0CN000  
bajaron bajar VMIS3P0  
de de SPS00  
los el DA0MP0  
Andes andes NP00000  
huyendo huir VMG0000  
del del SPCMS  
hambre hambre NCFS000  
, , Fc  
del del SPCMS  
atraso atraso NCMS000  
y y CC  
la el DA0FS0  
tuberculosis tuberculosis NCFN000  
; ; Fx  
quisieron querer VMIS3P0  
llegar llegar VMN0000  
a a SPS00  
la el DA0FS0  
ciudad ciudad NCFS000  
, , Fc  
pero pero CC  
quedaron quedar VMIS3P0  
varados varado AQ0MPP  
en en SPS00  
las el DA0FP0  
afueras afueras NCFP000  
, , Fc  
a a SPS00

## Anexos

una uno DI0FS0  
decena decena NCFS000  
de de SPS00  
kilómetros kilómetro NCMP000  
, , Fc  
en en SPS00  
los el DA0MP0  
sórdidos sórdido AQ0MP0  
arenales arenal NCMP000  
que que PR0CN000  
rodean rodear VMIP3P0  
Lima lima NP00000  
, , Fc  
en en SPS00  
donde donde PR000000  
plantaron plantar VMIS3P0  
sus su DP3CP0  
chabolas chabola NCFP000  
, , Fc  
precarios precario AQ0MP0  
tenderetes tenderete NCMP000  
de de SPS00  
cartón cartón NCMS000  
y y CC  
cajones cajón NCMP000  
astillados astillado AQ0MPP  
. . Fp  
Aspiraban aspirar VMII3P0  
a a SPS00  
más más RG  
: : Fd  
a a SPS00  
mejorar mejorar VMN0000  
su su DP3CS0  
situación situación NCFS000  
, , Fc  
a a SPS00  
ser ser VSN0000  
felices felice AQ0CP0  
; ; Fx  
pero pero CC  
les él PP3CPD00  
atrapó atrapar VMIS3S0  
la el DA0FS0  
miseria miseria NCFS000  
suburbana suburbano AQ0FS0  
y y CC  
subhumana subhumano AQ0FS0  
, , Fc  
la el DA0FS0  
ferocidad ferocidad NCFS000  
y y CC  
la el DA0FS0  
violencia violencia NCFS000  
de de SPS00

los el DA0MP0  
arrabales arrabal NCMP000  
. . Fp  
Alguno alguno PI0MS000  
consigue conseguir VMIP3S0  
escapar escapar VMN0000  
de de SPS00  
allí allí RG  
, , Fc  
muy mucho RG  
de\_tarde\_en\_tarde de\_tarde\_en\_tarde RG  
; ; Fx  
pero pero CC  
para para SPS00  
la el DA0FS0  
inmensa inmenso AQ0FS0  
mayoría mayoría NCFS000  
no no RN  
hay haber VAIP3S0  
ni ni CC  
salida salida NCFS000  
ni ni CC  
retorno retorno NCMS000  
. . Fp  
En en SPS00  
los el DA0MP0  
alrededores alrededor NCMP000  
de de SPS00  
Lima lima NP00000  
malviven malvivir VMIP3P0  
entre entre SPS00  
chabolas chabola NCFP000  
varios varios DI0MP0  
millones millón PN0CP000  
de de SPS00  
personas persona NCFP000  
; ; Fx  
sólo sólo RG  
en en SPS00  
Villa\_El\_Salvador villa\_el\_salvador NP00000  
, , Fc  
el el DA0MS0  
suburbio suburbio NCMS000  
de de SPS00  
la el DA0FS0  
Madre\_Coraje madre\_coraje NP00000  
, , Fc  
hay haber VAIP3S0  
300.000 300.000 Z  
habitantes habitante NCCP000  
. . Fp  
Hace hacer VMIP3S0  
14 14 Z  
años año NCMP000  
visité visitar VMIS1S0

uno uno PI0MS000  
 de de SPS00  
 estos este DD0MP0  
 asentamientos asentamiento NCMP000  
 , , Fc  
 y y CC  
 creo crear VMIP1S0  
 recordar recordar VMN0000  
 que que CS  
 fue ser VSIS3S0  
 precisamente precisamente RG  
 Villa\_El\_Salvador villa\_el\_salvador NP00000  
 . . Fp  
 Me yo PP1CS000  
 acompañaba acompañar VMII3S0  
 un uno DI0MS0  
 muchacho muchacho NCMS000  
 que que PROCN000  
 había haber VAII3S0  
 logrado lograr VMP00SM  
 la el DA0FS0  
 proeza proeza NCFS000  
 de de SPS00  
 escaparse escapar VMN0000  
 del del SPCMS  
 barrio barrio NCMS000  
 ; ; Fx  
 su su DP3CS0  
 familia familia NCFS000  
 seguía seguir VMII3S0  
 viviendo vivir VMG0000  
 allí allí RG  
 , , Fc  
 pero pero CC  
 él él PP3MS000  
 había haber VAII3S0  
 conseguido conseguir VMP00SM  
 un uno DI0MS0  
 trabajo trabajo NCMS000  
 y y CC  
 una uno DI0FS0  
 cama cama NCFS000  
 en en SPS00  
 la el DA0FS0  
 ciudad ciudad NCFS000  
 . . Fp  
 Yo yo PP1CSN00  
 quería querer VMII1S0  
 conocer conocer VMN0000  
 aquello aquel PD0NS000  
 , , Fc  
 hacer hacer VMN0000  
 un uno DI0MS0  
 reportaje reportaje NCMS000  
 , , Fc

y y CC  
 él él PP3MS000  
 se él P0300000  
 ofreció ofrecer VMIS3S0  
 a a SPS00  
 servirme servir VMN0000  
 de de SPS00  
 guía guía NCMS000  
 . . Fp  
 Fuimos ir VMIS1P0  
 hasta hasta SPS00  
 allí allí RG  
 , , Fc  
 tras tras SPS00  
 un uno DI0MS0  
 largo largo AQ0MS0  
 trayecto trayecto NCMS000  
 de de SPS00  
 traqueteantes traqueteante AQ0CP0  
 autobuses autobús NCMP000  
 , , Fc  
 para para SPS00  
 almorzar almorzar VMN0000  
 con con SPS00  
 su su DP3CS0  
 hermana hermana NCFS000  
 . . Fp  
 La el DA0FS0  
 villa villa NCFS000  
 había haber VAII3S0  
 nacido nacer VMP00SM  
 como como CS  
 una uno DI0FS0  
 excrescencia excrescencia NCFS000  
 cancerosa canceroso AQ0FS0  
 , , Fc  
 desordenadamente desordenadamente RG  
 , , Fc  
 allí allí RG  
 donde donde PR000000  
 termina terminar VMIP3S0  
 el el DA0MS0  
 asfalto asfalto NCMS000  
 y y CC  
 la el DA0FS0  
 esperanza esperanza NCFS000  
 . . Fp  
 Imaginen imaginar VMM03P0  
 un uno DI0MS0  
 pueblo pueblo NCMS000  
 de de SPS00  
 chabolas chabola NCFP000  
 de de SPS00  
 varios varios DI0MP0  
 cientos ciento PN0CP000

## Anexos

de de SPS00  
miles mil PNOCP000  
de de SPS00  
personas persona NCFP000  
: : Fd  
caminas caminar VMIP2S0  
y y CC  
caminas caminar VMIP2S0  
por por SPS00  
los el DA0MP0  
mugrientos mugriento AQ0MP0  
arenales arenal NCMP000  
y y CC  
la el DA0FS0  
miseria miseria NCFS000  
resulta resultar VMIP3S0  
inacabable inacabable AQ0CS0  
, , Fc  
inabarcable inabarcable AQ0CS0  
. . Fp  
Era ser VSII3S0  
media medio DN0FS0  
mañana mañana NCFS000  
y y CC  
había haber VAI3S0  
bastante bastante DIOCS0  
gente gente NCFS000  
en en SPS00  
las el DA0FP0  
calles calle NCFP000  
, , Fc  
esto\_es esto\_es CC  
, , Fc  
en en SPS00  
las el DA0FP0  
veredas vereda NCFP000  
sin sin SPS00  
urbanizar urbanizar VMN0000  
que que PROCN000  
habían haber VAI3P0  
quedado quedar VMP00SM  
abiertas abierto AQ0FPP  
entre entre SPS00  
las el DA0FP0  
chozas choza NCFP000  
. . Fp  
Mi mi DP1CSS  
guía guía NCFS000  
iba ir VMII1S0  
evitando evitar VMG0000  
, , Fc  
precisamente precisamente RG  
, , Fc  
los el DA0MP0  
lugares lugar NCMP000

más más RG  
poblados poblado AQ0MPP  
: : Fd  
avanzábamos avanzar VMII1P0  
por por SPS00  
las el DA0FP0  
callejas calleja NCFP000  
solitarias solitario AQ0FP0  
, , Fc  
huyendo huir VMG0000  
de de SPS00  
la el DA0FS0  
gente gente NCFS000  
, , Fc  
porque porque CS  
te tú PP2CS000  
podían poder VMII3P0  
rajar rajar VMN0000  
en\_mitad\_de en\_mitad\_de SPS00  
una uno DIOFS0  
muchedumbre muchedumbre NCFS000  
sin sin SPS00  
que que CS  
nadie nadie PIOCS000  
hiciera hacer VMSI3S0  
nada nada PIOCS000  
. . Fp  
La el DA0FS0  
policía policía NCCS000  
no no RN  
se él P0300000  
atrevia atreverse VMII3S0  
a a SPS00  
entrar entrar VMN0000  
en en SPS00  
las el DA0FP0  
villas villa NCFP000  
: : Fd  
eran ser VSII3P0  
lugares lugar NCMP000  
sin sin SPS00  
ley ley NCFS000  
, , Fc  
territorios territorio NCMP000  
prohibidos prohibido AQ0MPP  
. . Fp  
Llegamos llegar VMIP1P0  
al\_fin al\_fin RG  
a a SPS00  
la el DA0FS0  
casa casa NCFS000  
de de SPS00  
la el DA0FS0  
hermana hermana NCFS000  
: : Fd

un uno DI0MS0  
 tenderete tenderete NCMS000  
 de de SPS00  
 latas lata NCFP000  
 y y CC  
 cartones cartón NCMP000  
 de de SPS00  
 apenas apenas RG  
 tres tres DN0CP0  
 metros metro NCMP000  
 por por SPS00  
 tres tres DN0CP0  
 metros metro NCMP000  
 , , Fc  
 con con SPS00  
 un uno DI0MS0  
 infernillo infernillo NCMS000  
 de de SPS00  
 gas gas NCMS000  
 , , Fc  
 una uno DI0FS0  
 silla silla NCFS000  
 de de SPS00  
 enea enea NCFS000  
 , , Fc  
 un uno DI0MS0  
 plástico plástico NCMS000  
 y y CC  
 una uno DI0FS0  
 manta manta NCFS000  
 en en SPS00  
 un uno DI0MS0  
 rincón rincón NCMS000  
 del del SPCMS  
 suelo suelo NCMS000  
 , , Fc  
 sobre sobre SPS00  
 la el DA0FS0  
 sucia sucio AQ0FS0  
 arena arena NCFS000  
 . . Fp  
 Allí allí RG  
 vivían vivir VMII3P0  
 , , Fc  
 no no RN  
 sé saber VMIP1S0  
 cómo cómo PT000000  
 , , Fc  
 la el DA0FS0  
 hermana hermana NCFS000  
 de de SPS00  
 mi mi DP1CSS  
 amigo amigo NCMS000  
 , , Fc  
 su su DP3CS0

marido marido NCMS000  
 y y CC  
 cuatro cuatro DN0CP0  
 niños niño NCMP000  
 . . Fp  
 Comimos comer VMIS1P0  
 un uno DI0MS0  
 arroz arroz NCMS000  
 con con SPS00  
 pollo pollo NCMS000  
 muy mucho RG  
 sabroso sabroso AQ0MS0  
 . . Fp  
 A a SPS00  
 mí yo PP1CS000  
 la el DA0FS0  
 invitada invitada NCFS000  
 de de SPS00  
 honor honor NCMS000  
 , , Fc  
 me yo PP1CS000  
 hicieron hacer VMIS3P0  
 sentar sentar VMN0000  
 en en SPS00  
 la el DA0FS0  
 única único AQ0FS0  
 silla silla NCFS000  
 . . Fp  
 Pienso pensar VMIP1S0  
 ahora ahora RG  
 en en SPS00  
 María\_Elena\_Moyano maría\_elenamoyano NP00000  
 , , Fc  
 la el DA0FS0  
 Madre\_Coraje madre\_coraje NP00000  
 , , Fc  
 y y CC  
 recuerdo recordar VMIP1S0  
 aquel aquel DD0MS0  
 almuerzo almuerzo NCMS000  
 conmovedor conmovedor AQ0MS0  
 , , Fc  
 aquel aquel DD0MS0  
 poblado poblado NCMS000  
 espeluznante espeluznante AQ0CS0  
 . . Fp  
 El el DA0MS0  
 liderazgo liderazgo NCMS000  
 de de SPS00  
 María\_Elena maría\_elenamoyano NP00000  
 nació nacer VMIS3S0  
 de de SPS00  
 aquella aquel DD0FS0  
 miseria miseria NCFS000  
 y y CC

## Anexos

de de SPS00  
una uno DI0FS0  
increíble increíble AQ0CS0  
voluntad voluntad NCFS000  
de de SPS00  
superación superación NCFS000  
. . Fp  
De de SPS00  
la el DA0FS0  
generosidad generosidad NCFS000  
, , Fc  
de de SPS00  
la el DA0FS0  
inteligencia inteligencia NCFS000  
, , Fc  
del del SPCMS  
tesón tesón NCMS000  
. . Fp  
Consiguió conseguir VMIS3S0  
, , Fc  
me yo PP1CS000  
dicen decir VMIP3P0  
, , Fc  
organizar organizar VMN0000  
el el DA0MS0  
suburbio suburbio NCMS000  
, , Fc  
y y CC  
sacarlo sacar VMN0000  
de de SPS00  
la el DA0FS0  
violencia violencia NCFS000  
y y CC  
el el DA0MS0  
salvajismo salvajismo NCMS000  
. . Fp  
Construyó construir VMIS3S0  
un uno DI0MS0  
marco marco NCMS000  
de de SPS00  
dignidad dignidad NCFS000  
en en SPS00  
el el DA0MS0  
que que PROCN000  
reconocerse reconocer VMN0000  
: : Fd  
luchó luchar VMIS3S0  
por por SPS00  
otorgar otorgar VMN0000  
una uno DI0FS0  
dimensión dimensión NCFS000  
humana humano AQ0FS0  
a a SPS00  
unas uno DI0FP0  
vidas vida NCFP000

embrutecidas embrutecer AQ0FPP  
por por SPS00  
la el DA0FS0  
miseria miseria NCFS000  
. . Fp  
Por por SPS00  
todo todo DI0MS0  
esto este PD0NS000  
le él PP3CSD00  
dieron dar VMIS3P0  
el el DA0MS0  
premio premio NCMS000  
Príncipe\_de\_Asturias\_de\_la\_Paz  
príncipe\_de\_asturias\_de\_la\_paz NP00000  
de de SPS00  
1987 [?:?/?/1987:?.?] W  
: : Fd  
porque porque CS  
logró lograr VMIS3S0  
ser ser VSN0000  
una uno DI0FS0  
persona persona NCFS000  
, , Fc  
aunque aunque CS  
nada nada PI0CS000  
en en SPS00  
su su DP3CS0  
entorno entorno NCMS000  
se él PP3CN000  
lo él PP3MSA00  
permitiese permitir VMSI3S0  
. . Fp  
Y y CC  
por por SPS00  
todo todo DI0MS0  
esto este PD0NS000  
ha haber VAIP3S0  
sido ser VSP00SM  
ahora ahora RG  
asesinada asesinar VMP00SF  
. . Fp  
Porque porque CS  
los el DA0MP0  
fanáticos fanático NCMP000  
de de SPS00  
Sendero\_Luminoso sendero\_luminoso NP00000  
no no RN  
pueden poder VMIP3P0  
admitir admitir VMN0000  
que que CS  
haya haber VASP3S0  
seres ser NCMP000  
libres libre AQ0CP0  
. . Fp  
La el DA0FS0

verdadera verdadero AQ0FS0  
 heroicidad heroicidad NCFS000  
 no no RN  
 es ser VSIP3S0  
 un uno DI0MS0  
 acto acto NCMS000  
 único único AQ0MS0  
 : : Fd  
 el el DA0MS0  
 soldado soldado NCMS000  
 que que PROCN000  
 se él PP3CN000  
 inmola inmolar VMIP3S0  
 para para SPS00  
 salvar salvar VMN0000  
 a a SPS00  
 sus su DP3CP0  
 compañeros compañero NCMP000  
 , , Fc  
 el el DA0MS0  
 hombre hombre NCMS000  
 que que PROCN000  
 entra entrar VMIP3S0  
 en en SPS00  
 un uno DI0MS0  
 edificio edificio NCMS000  
 en en SPS00  
 llamas llama NCFP000  
 a a SPS00  
 rescatar rescatar VMN0000  
 a a SPS00  
 un uno DI0MS0  
 niño niño NCMS000  
 , , Fc  
 son ser VSIP3P0  
 sin\_duda sin\_duda RG  
 personas persona NCFP000  
 admirables admirable AQ0CP0  
 porque porque CS  
 supieron saber VMIS3P0  
 responder responder VMN0000  
 a a SPS00  
 un uno DI0MS0  
 instante instante NCMS000  
 de de SPS00  
 gloria gloria NCFS000  
 , , Fc  
 de de SPS00  
 generosidad generosidad NCFS000  
 y y CC  
 de de SPS00  
 exigencia exigencia NCFS000  
 . . Fp  
 Pero pero CC  
 ese ese DD0MS0

momento momento NCMS000  
 heroico heroico AQ0MS0  
 , , Fc  
 me yo PP1CS000  
 parece parecer VMIP3S0  
 , , Fc  
 es ser VSIP3S0  
 más más RG  
 una uno DI0FS0  
 especie especie NCFS000  
 de de SPS00  
 fiebre fiebre NCFS000  
 que que CS  
 un uno DI0MS0  
 talante talante NCMS000  
 . . Fp  
 Por\_el\_contrario por\_el\_contrario RG  
 , , Fc  
 la el DA0FS0  
 verdadera verdadero AQ0FS0  
 heroicidad heroicidad NCFS000  
 se se P0000000  
 construye construir VMIP3S0  
 calladamente calladamente RG  
 , , Fc  
 día día NCMS000  
 a a SPS00  
 día día NCMS000  
 , , Fc  
 sobreponiéndose sobreponer VMG0000  
 una uno DI0FS0  
 y y CC  
 otra otro DI0FS0  
 vez vez NCFS000  
 a a SPS00  
 circunstancias circunstancia NCFP000  
 dolorosas doloroso AQ0FP0  
 y y CC  
 extremas extremo AQ0FP0  
 . . Fp  
 Como como CS  
 hizo hacer VMIS3S0  
 María\_Elena maría\_elena NP00000  
 . . Fp  
 Es ser VSIP3S0  
 heroico heroico AQ0MS0  
 levantarse levantar VMN0000  
 todas todo DI0FP0  
 las el DA0FP0  
 mañanas mañana NCFP000  
 a a SPS00  
 luchar luchar VMN0000  
 contra contra SPS00  
 la el DA0FS0  
 desesperación desesperación NCFS000

## Anexos

y y CC  
la el DA0FS0  
incultura incultura NCFS000  
. . Fp  
Y y CC  
es ser VSIP3S0  
heroico heroico AQ0MS0  
sentir sentir VMN0000  
miedo miedo NCMS000  
ante ante SPS00  
las el DA0FP0  
amenazas amenaza NCFP000  
senderistas senderista AQ0CP0  
y y CC  
seguir seguir VMN0000  
actuando actuar VMG0000  
año año NCMS000  
tras tras SPS00  
año año NCMS000  
, , Fc  
sin\_embargo sin\_embargo CC  
, , Fc  
como como CS  
si si CS  
el el DA0MS0  
propio propio AQ0MS0  
miedo miedo NCMS000  
no no RN  
existiese existir VMSI3S0  
. . Fp  
Pero pero CC  
, , Fc  
al\_fin al\_fin RG  
, , Fc  
el el DA0MS0  
terror terror NCMS000  
atrapó atrapar VMIS3S0  
a a SPS00  
María\_Elena maría\_elen NP00000  
; ; Fx  
fueron ser VSIS3P0  
a a SPS00  
buscarla buscar VMN0000  
, , Fc  
y y CC  
duele doler VMIP3S0  
imaginar imaginar VMN0000  
el el DA0MS0  
instante instante NCMS000  
de de SPS00  
pánico pánico NCMS000  
que que PROCN000  
debió deber VMIS3S0  
de de SPS00  
sentir sentir VMN0000

ante ante SPS00  
sus su DP3CP0  
asesinos asesino NCMP000  
. . Fp  
¿ ¿ Fia  
Sirve servir VMIP3S0  
de de SPS00  
algo algo PIOCS000  
esa ese DD0FS0  
muerte muerte NCFS000  
salvaje salvaje AQ0CS0  
y y CC  
absurda absurdo AQ0FS0  
? ? Fit  
¿ ¿ Fia  
Es ser VSIP3S0  
cierta cierto AQ0FS0  
esa ese DD0FS0  
versión versión NCFS000  
cristiana cristiano AQ0FS0  
y y CC  
consoladora consolador AQ0FS0  
que que PROCN000  
asegura asegurar VMIP3S0  
que que CS  
los el DA0MP0  
sacrificios sacrificio NCMP000  
no no RN  
son ser VSIP3P0  
inútiles inútil AQ0CP0  
? ? Fit  
Probablemente probablemente RG  
no no RN  
; ; Fx  
probablemente probablemente RG  
, , Fc  
en en SPS00  
la el DA0FS0  
historia historia NCFS000  
concreta concreto AQ0FS0  
de de SPS00  
los el DA0MP0  
suburbios suburbio NCMP000  
limeños limeño AQ0MP0  
la el DA0FS0  
desaparición desaparición NCFS000  
de de SPS00  
María\_Elena maría\_elen NP00000  
sea ser VSSP3S0  
tan tanto RG  
sólo sólo RG  
una uno DI0FS0  
catástrofe catástrofe NCFS000  
. . Fp  
Pero pero CC

sí sí RG  
 creo crear VMIP1S0  
 que que CS  
 la el DA0FS0  
 vida vida NCFS000  
 de de SPS00  
 Moyano moyano NP00000  
 , , Fc  
 su su DP3CS0  
 entereza entereza NCFS000  
 hasta hasta SPS00  
 el el DA0MS0  
 final final NCMS000  
 y y CC  
 su su DP3CS0  
 coraje coraje NCMS000  
 , , Fc  
 forma formar VMIP3S0  
 parte parte NCFS000  
 del del SPCMS  
 legado legado NCMS000  
 de de SPS00  
 los el DA0MP0  
 humanos humano NCMP000  
 , , Fc  
 del del SPCMS  
 inconsciente inconsciente NCMS000  
 colectivo colectivo AQ0MS0  
 , , Fc  
 de de SPS00  
 la el DA0FS0  
 sustancia sustancia NCFS000  
 común común AQ0CS0  
 que que PROCN000  
 todos todo PI0MP000  
 somos ser VSIP1P0  
 . . Fp  
 Y y CC  
 aun aun RG  
 cuando cuando CS  
 la él PP3FSA00  
 olvidemos olvidar VMSP1P0  
 , , Fc  
 como como CS  
 hemos haber VAIP1P0

olvidado olvidar VMP00SM  
 a a SPS00  
 todos todo DI0MP0  
 los el DA0MP0  
 demás demás DIOCP0  
 héroes héroe NCMP000  
 anónimos anónimo AQ0MP0  
 , , Fc  
 es ser VSIP3S0  
 gracias\_a gracias\_a SPS00  
 ella él PP3FS000  
 , , Fc  
 y y CC  
 a a SPS00  
 gentes gente NCFP000  
 como como CS  
 ella él PP3FS000  
 , , Fc  
 que que CS  
 la el DA0FS0  
 humanidad humanidad NCFS000  
 puede poder VMIP3S0  
 perseverar perseverar VMN0000  
 en en SPS00  
 el el DA0MS0  
 sueño sueño NCMS000  
 de de SPS00  
 la el DA0FS0  
 felicidad felicidad NCFS000  
 y y CC  
 la el DA0FS0  
 razón razón NCFS000  
 . . Fp  
 Porque porque CS  
 ellos él PP3MP000  
 nos yo PP1CP000  
 demuestran demostrar VMIP3P0  
 que que CS  
 el el DA0MS0  
 bien bien NCMS000  
 también también RG  
 existe existir VMIP3S0  
 . . Fp

## Anexo E: Ejemplo de la salida del programa

A continuación se muestra la impresión de salida en la detección de correferencia y anáfora indirecta donde los números indican:

- 0** = Nombre no precedido por determinante (no es expresión referencial)
- 1** = Nombre propio **sin** correferencia
- 2** = Nombre propio **con** correferencia
- 3** = Nombre común **sin** correferencia
- 4** = Nombre común **con** correferencia
- 6** = Nombre común **con** anáfora indirecta

### Marcado de nombres propios

2 Madre_Coraje	1 Andes	2 Lima
2 Lima	2 Villa_El_Salvador	2 Madre_Coraje
2 Villa_El_Salvador	2 María_Elena_Moyano	2 Madre_Coraje
2 María_Elena	1 Príncipe_de_Asturias_de_la_Paz	1 Sendero_Luminoso
2 María_Elena	2 María_Elena	2 María_Elena
2 Moyano		

### Correferencias encontradas entre nombres comunes

un largo <b>trayecto</b> ↔ servirme de <b>guía</b>	más una <b>especie</b> ↔ <b>Exigencia</b>
Imaginen un <b>pueblo</b> ↔ La <b>villa</b> había	el propio <b>miedo</b> ↔ heroico sentir <b>miedo</b>
había bastante <b>gente</b> ↔ de <b>personas</b>	el <b>terror</b> ↔ el propio <b>miedo</b>
una <b>muchedumbre</b> ↔ de la <b>gente</b>	una <b>catástrofe</b> ↔ <b>versión</b> cristiana
mí la <b>invitada</b> ↔ de mi <b>amigo</b>	Hasta el <b>final</b> ↔ una <b>catástrofe</b>
del <b>tesón</b> ↔ increíble <b>voluntad</b>	Su <b>coraje</b> , ↔ una <b>catástrofe</b>
dieron el <b>premio</b> ↔ de la <b>violencia</b>	de la <b>sustancia</b> ↔ de los <b>humanos</b>
el <b>hombre</b> que ↔ haya <b>seres</b> libres	la <b>humanidad</b> ↔ <b>gentes</b> como ella
Pero ese <b>momento</b> ↔ <b>instante</b> de gloria	

**Marcado de nombres comunes**

3 senderistas	3 Foto	3 periódicos	3 Mujer	0 color
3 Habitantes	3 Villas	0 arrabales	0 miseria	0 personas
3 Mayoría	0 Indígenas	3 hambre	3 atraso	3 tuberculosis
3 Ciudad	3 Afueras	3 decena	0 kilómetros	3 arenales
3 Chabolas	0 Tenderetes	0 cartón	0 cajones	3 situación
3 Miseria	3 Ferocidad	3 violencia	3 arrabales	3 mayoría
0 salida	0 Retorno	3 alrededores	0 chabolas	3 personas
3 suburbio	0 Habitantes	0 años	3 asentamientos	3 muchacho
3 proeza	3 Barrio	3 familia	3 trabajo	3 cama
3 ciudad	3 Reportaje	4 guía	4 trayecto	0 autobuses
3 hermana	4 Villa	3 excrescencia	3 asfalto	3 esperanza
4 pueblo	0 Chabolas	4 personas	3 arenales	3 miseria
3 mañana	4 Gente	3 calles	3 veredas	3 chozas
3 Guía	3 Lugares	3 callejas	4 gente	4 muchedumbre
3 policía	3 Villas	0 lugares	0 ley	0 territorios
3 Casa	3 Hermana	3 tenderete	0 latas	0 cartones
3 metros	3 Metros	3 infernillo	0 gas	3 silla
0 Enea	3 Plástico	3 manta	3 rincón	3 suelo
3 Arena	3 Hermana	4 amigo	3 marido	3 niños
3 Arroz	0 Pollo	4 invitada	0 honor	3 silla
3 almuerzo	3 Poblado	3 liderazgo	3 miseria	4 voluntad
0 superación	3 Generosidad	3 inteligencia	4 tesón	3 suburbio
4 violencia	3 Salvajismo	3 marco	0 dignidad	3 dimensión
3 vidas	3 Miseria	4 premio	3 persona	3 entorno
3 fanáticos	4 Seres	3 heroicidad	3 acto	3 soldado
3 compañeros	4 Hombre	3 edificio	0 llamas	3 niño
0 personas	4 Instante	0 gloria	0 generosidad	4 exigencia
4 momento	4 Especie	0 fiebre	3 talante	3 heroicidad
0 día	0 Día	3 vez	0 circunstancias	3 mañanas
3 desesperación	3 Incultura	4 miedo	3 amenazas	0 año
0 año	4 Miedo	4 terror	3 instante	0 pánico
3 asesinos	3 Muerte	4 versión	3 sacrificios	3 historia
3 suburbios	3 desaparición	4 catástrofe	3 vida	3 entereza
4 final	4 Coraje	0 parte	3 legado	4 humanos
3 inconsciente	4 Sustancia	3 héroes	4 gentes	4 humanidad
3 sueño	3 Felicidad	3 razón	3 bien	

**Anáforas indirectas encontradas entre nombres**

Nº UL	Unidades léxicas	Rel.	Unidades léxicas	Nº UL
23	una mujer	↔	Madre_Coraje peruana	5
73	su mayoría	↔	personas,	67
102	las afueras	↔	ciudad,	95
122	sus chabolas	↔	arenales que	113
149	la miseria	↔	tenderetes de	125
195	de personas	↔	mayoría no	176
218	estos asentamientos	↔	suburbio de	202
231	un muchacho	↔	Villa_El_Salvador .	226
243	su familia	↔	muchacho que	231
286	largo trayecto	↔	guía.	277
295	su hermana	↔	guía.	277
319	un pueblo	↔	villa había	298
328	de personas	↔	villa había	298
336	mugrientos arenales	↔	chabolas de	321
339	la miseria	↔	chabolas de	321
351	bastante gente	↔	personas:	328
372	Mi guía	↔	gente en	351
392	la gente	↔	guía iba	372
416	las villas	↔	muchedumbre sin	400
430	la casa	↔	villas :	416
436	un tenderete	↔	casa de	430
487	mi amigo	↔	hermana :	433
529	aquel almuerzo	↔	Comimos un	495
845	los sacrificios	↔	muerte salvaje	828

**Estadísticas**

<b>Nombres propios</b>			<b>Nombres comunes</b>				<b>Otras</b>	
Total	Corref	No corref	Total	Corref	AnaInd	No corref	Pron	Verb
16	10	6	179	31	24	95	53	137

## **Anexo F: Fuentes de programas**

**Sólo se presentan los programas fundamentales que ilustran los algoritmos**

```
/*
  Archivo de encabezado para incluir en programas para el
  procesamiento de archivos etiquetados de TnT.

  Desarrollado por: Raúl Morales Carrasco    20030225
  Modificado por: Raúl Morales Carrasco    20030225
*/
#define nl '\n'
#define MAXBUFFER 1024
#define REN 250
#define COL 100

// Se definen variables globales
extern char *arch_entrada;
extern char *arch_escena;
extern char *arch_sinocom;
extern char *arch_sinopro;
extern char *arch_fanomb;
extern char *arch_faesce;
extern char *arch_estad;
extern char cadban[];
extern char bandera;
extern int conta_ai;
extern int imprime;
extern char buffer[];

/* Se define la clase de Componentes lexicos (tokens) */
struct Inf_lex {
  char *lex; // Componente lexico
  char *lema; // Lema
  char *cat; // Categoria, tipo, etc
  int ind; // indicador de coreferencia, etc
  int ancla; // identificador único para enlace
  int enlace; // enlace a la referencia anterior
};

extern Inf_lex *Dat1;
```

## Anexos

```
struct Dnodo { // El elemento de la lista
    Inf_lex *Elem; // apunta a la Informacion del nodo
    Dnodo *Der; // apunta al siguiente nodo a la derecha
    Dnodo *Izq; // apunta al siguiente nodo a la izquierda
};

extern Dnodo *paso;
extern Dnodo *derecha;
extern Dnodo *izquierda;
extern Dnodo *prev_izq;
extern Dnodo *prev_der;

class Dlista { // La lista de objetos accesados como nodos
    Dnodo *Dactual; // apunta al nodo actual de la lista
    Dnodo *Dinicial; // apunta al nodo inicial de la lista
    Dnodo *Dfinal; // apunta al ultimo nodo insertado
    int Dnodos; // Numero de Nodos en la lista
public:
    // constructor
    Dlista();
    // destructor
    ~Dlista();
    // agrega un elemento a la lista
    void DInserta(Inf_lex *Nuevo);
    // Obtiene el nodo inicial de la lista
    Dnodo * Inicio(){ return Dinicial;};
    // Obtiene el nodo final de la lista
    Dnodo * Fin(){ return Dfinal;};
    // Obtiene el nodo actual de la lista
    Dnodo * Actual(){ return Dactual;};
    // Listado de elementos
    void Dlistado();
};

extern Dlista texto; // Se declara texto como una variable de lista

/*
Funciones para manipulación de estructura Inf_lex
*/
int verifica(char * clave);
void mostrar(Inf_lex *v);
int semejanza(char * cadena1, char *cadena2);
char primera_letra ( char letra );
int coref ( char * cadena1, char * cadena2);
int anaf_ind ( char * cadena1, char * cadena2 );
int cadenas( char *cadena , const char* archivo, char cad[REN][COL]);
```

```

int relaciones( char *cadena , const char* archivo, char cad[REN][COL]);
void convierte_htm();
void imp_cadenas();
char *referencia( Dnodo *inicio);
int lexcom( char *cadena, int *longitud);
char *nom_arch( char *band, int conta, char * term);
int es_vocal( char letra);
void desacentua( char *cadena);
void singular( char *cadena);

```

```

/* Programa principal para el procesamiento de
   archivos etiquetados de Clic-TALP.

```

```

    Desarrollado por: Raúl Morales Carrasco    20020205

```

```

    Modificado por: Raúl Morales Carrasco    20031208

```

```

*/

```

```

#include "..\general\tnt.h"

```

```

#include <exception> // Para manipular errores en new

```

```

#include <iostream.h>

```

```

#include <string.h>

```

```

#include <fstream.h>

```

```

char *arch_entrada;

```

```

char *arch_escena = "escen\?.txt";

```

```

char *arch_sinocom = "sinon\?.txt";

```

```

char *arch_fanomb = "res\fnombre.txt";

```

```

char *arch_faesce = "res\fescena.txt";

```

```

char *arch_estad = "res\estadis.txt";

```

```

char cadban[3], bandera;

```

```

int conta_cd, conta_ci, conta_ai, imprime;

```

```

FILE *archivo, *fprop;

```

```

Dnodo *paso, *derecha, *izquierda, *prev_izq, *prev_der;

```

```

Inf_lex *Dat1;

```

```

char buffer[MAXBUFFER];

```

```

Dlista texto; // Se declara texto como una variable de lista

```

```

int main ( int argc, char* argv[])

```

```

{

```

```

    char *p1, *p2, *p3, *log;

```

```

    int cmp1, cmp2, indice, linea, temp, anclant, k;

```

```

    int cp, ncp, cd, ci, ai, nc, nr, nv, npron, pr;

```

```

    ifstream f_in;

```

```

    // se verifican parámetros de llamada

```

```

    if ( argc < 2 ) {

```

```

        cerr << " Se debe correr asi:" << nl;

```

```

cerr << " anaind archivo imprime bandera conta_cd conta_ci conta_ai" << nl;
cerr << nl << " Siendo:" << nl;
cerr << " archivo = nombre de archivo TnT a procesar"<< nl;
cerr << " imprime = clave de impresion:" << nl;
cerr << "      bit 0 = 1 = seguimiento(log)" << nl;
cerr << "      bit 1 = 2 = archivo texto" << nl;
cerr << "      bit 2 = 4 = archivo htm (default)" << nl;
cerr << "      bit 3 = 8 = falta de info en diccionarios" << nl;
cerr << "      bit 4 = 16 = estadisticas" << nl;
cerr << " bandera = señalador de busqueda hacia atras:" << nl;
cerr << "      V = verbo(default)" << nl;
cerr << "      N = Nombre" << nl;
cerr << "      F = Signo de puntuacion" << nl;
cerr << "cont_cd = contador de correfer. directa (backtraking)" << nl;
cerr << "cont_ci = contador de correfer. indirecta (backtraking)" << nl;
cerr << "cont_ai = contador de anafora indirecta (backtraking)" << nl;
cerr << endl;
return(1);
}
// se reciben parámetros de entrada
arch_entrada = argv[1];
if ( argc < 3 ) imprime = 4; // clave de impresión por omisión htm
else imprime = atoi(argv[2]);
if ( imprime > 31 || imprime < 1 ) {
    cerr << " imprime debe ser mayor que 0 y menor que 32" << imprime << endl;
    return(2);
}
if ( argc < 4 ) {
    bandera = 'F'; // bandera por omisión cualquier "punto"
    cadban[0]= bandera;
    cadban[1]= 'p';
}
else {
    bandera = argv[3][0];
    if (bandera == 'V' || bandera == 'N' || bandera == 'F'){
        cadban[0]= bandera;
        cadban[1]= 0;
    }
    else {
        cerr << "bandera incorrecta " << bandera << endl;
        return(3);
    }
}
if ( argc < 5 ) conta_cd = 9; // contador por omisión igual a nueve
else conta_cd = atoi(argv[4]); // búsquedas hacia atrás
if ( argc < 6 ) conta_ci = 9; // contador por omisión igual a nueve
else conta_ci = atoi(argv[5]); // búsquedas hacia atrás

```

```

if ( argc < 7 ) conta_ai = 9; // contador por omisión igual a nueve
else conta_ai = atoi(argv[6]); // búsquedas hacia atrás

cerr << " Corriendo con archivo " << arch_entrada << " " << imprime
  << " " << bandera << " " << conta_cd
  << " " << conta_ci << " " << conta_ai << endl;
f_in.open(arch_entrada);
if (!f_in ) { // error: no se pudo abrir el archivo de entrada
  cerr << "No se pudo abrir el archivo " << arch_entrada << endl;
  return(5);
}
if (imprime & 1){
  log = nom_arch( cadban, conta_ai, "log");
  archivo = fopen(log,"w+");
  if (!archivo){ // error: no se pudo abrir el archivo de seguimiento
    cerr << "No se pudo abrir el archivo " << archivo << endl;
    return(6);
  } }
/*
Se lee archivo de entrada hasta el fin de archivo
*/
linea = 1;
while ( f_in.getline(buffer,MAXBUFFER, nl) ) {
  p1 = strtok(buffer, " \t"); // Se obtiene el token o lex
  if( (p3 = strtok(NULL, " \t")) == NULL) { // categoria
    cerr << " Linea " << linea << " sin categoria "
      << p1 << nl;
    continue;
  }
  Dat1->lex = new char[strlen(p1)+1];
  strcpy(Dat1->lex, p1);
  Dat1->cat = new char[strlen(p3)+1];
  strcpy(Dat1->cat, p3);
  Dat1->ind = 0;
  Dat1->ancla = linea++;
  Dat1->enlace = NULL;
  texto.DInserta(Dat1); // Se inserta en la lista
} // fin de while
f_in.close ();
/*
Se marca correferencia de nombres propios
*/
cerr << " Correferencia de Nombres propios " << endl;
if (imprime & 1){
  fprintf(archivo," \n Correferencia de Nombres propios \n");
}

```

```

derecha = texto.Inicio();
while( derecha != NULL) { // de inicio hacia final
  if( (derecha->Elem->cat[0] == 'N') && // Nombre propio
      (derecha->Elem->cat[1] == 'P') && // sin marcar
      (derecha->Elem->ind == 0) ) {
    cmp2 = strlen(derecha->Elem->lex);
    derecha->Elem->ind = 1; // Se marca la primera ocurrencia
    anclant = derecha->Elem->ancla; // se almacena referencia numerica
    // se buscan correferencias hasta el final
    paso = derecha->Der; // apunta al siguiente nodo
    while( paso->Der != NULL ) {
      if( (paso->Elem->cat[0] == 'N') && // nombre propio
          (paso->Elem->cat[1] == 'P') && // sin marcar
          (paso->Elem->ind == 0) ) {
        cmp1 = strlen(paso->Elem->lex);
        p2 = NULL;
        if ( cmp1 > cmp2)p2 = strstr(paso->Elem->lex, derecha->Elem->lex);
        else p2 = strstr(derecha->Elem->lex,paso->Elem->lex);
        if (p2 != NULL){ //si se encuentra se marcan como correferencia
          // impresion de resultado en log file
          if (imprime & 1){
            fprintf(archivo,"%d %s ",paso->Elem->ancla,
                    paso->Elem->lex);
            fprintf(archivo," >CD--> ");
            fprintf(archivo," %s %d\n",derecha->Elem->lex,anclant);
          }
          paso->Elem->ind = 2;
          paso->Elem->enlace = anclant;
          anclant = paso->Elem->ancla;
        } }
        paso = paso->Der; // apunta al siguiente nodo
      } }
    derecha = derecha->Der; // apunta al siguiente nodo
  }
}
/*
Se marcan los nombres precedidos por determinativos
*/

```

```

cerr << " Marcado de Nombres comunes" << endl;
derecha = texto.Inicio();
while( derecha != NULL) { // de izquierda a derecha
  temp = strcmp(derecha->Elem->cat, "SPCMS"); // preposicion contraida?
  // localiza determinativo o preposicion contraida(al, del)
  if( (derecha->Elem->cat[0] == 'D') || (temp == 0) ) {
    while( derecha != NULL ) { // se busca nombre comun
      if( (derecha->Elem->cat[0] == 'N') &&
          (derecha->Elem->cat[1] == 'C') ) {

```

```

        derecha->Elem->ind = 3; // y se marca como candidato
        break;
    }
    // el ciclo se interrumpe si hay ...
    else if( derecha->Elem->cat[0]== 'V') break; // Verbo
    else if( derecha->Elem->cat[0]== 'F') break; // Puntuacion
    else if( derecha->Elem->cat[0]== 'R') break; // Adverbio
    else if( derecha->Elem->cat[0]== 'P') break; // Pronombre
    else if( derecha->Elem->cat[0]== 'S') break; // prepoSición
    else if( derecha->Elem->cat[0]== 'C') break; // o Conjuncion
        else derecha = derecha->Der; // de otra forma continúa
    } }
    derecha = derecha->Der; // apunta al siguiente nodo
}
/*
Se marcan las relaciones predefinidas por la relación
"de" o "del" en las secuencias:
D->[A]->N->de->[A]->N
D->[A]->N->del->[A]->N
*/
cerr << " Marcado de relaciones predefinidas [de]" << endl;
if (imprime & 1){
    fprintf(archivo,"\n Relaciones predefinidas \n");
}
derecha = texto.Inicio();
while( derecha != NULL) { // de izquierda a derecha
    if( (derecha->Elem->cat[0] == 'N') && // nombre común?
        (derecha->Elem->cat[1] == 'C') ) {
        prev_der = derecha; // es primer nombre y prepara enlace
        derecha = derecha->Der; // avanza al siguiente token
        if( (derecha->Elem->lex[0] == 'd') && // cadena igual a "de"?
            (derecha->Elem->lex[1] == 'e') ) {
            derecha = derecha->Der; // avanza al siguiente token
            if( derecha->Elem->cat[0] == 'N') { // seguido de nombre?
                prev_der->Elem->ind = 7; // relación preestablecida
                prev_der->Elem->enlace = derecha->Elem->ancla;
                // impresión de resultado en log file
                if (imprime & 1){
                    fprintf(archivo,"%d %s",prev_der->Elem->ancla,
                        prev_der->Elem->lex);
                    fprintf(archivo," >pr--> ");
                    fprintf(archivo,"%s %d\n",derecha->Elem->lex,
                        derecha->Elem->ancla);
                } } } }
                derecha = derecha->Der; // apunta al siguiente nodo
            }
        }
    }
}

```

```

/*
Resolución de correferencia directa
*/
// Se localiza posición de la segunda oración completa sea
// simple, compuesta o compleja (hasta primer punto y seguido)
derecha = texto.Inicio();
while (derecha != NULL ) { // de inicio hacia final
    if( derecha->Elem->cat[0] == 'F' &&
        derecha->Elem->cat[1] == 'p') break;
    else derecha = derecha->Der; // apunta al siguiente nodo
};
cerr << " Correferencias directas de nombres comunes " << endl;
if (imprime & 1){
    fprintf(archivo, "\n Correferencias directas de nombres comunes \n");
}
prev_der = derecha;
while( derecha != texto.Fin() ) { // hasta el fin de lista
    if( (derecha->Elem->cat[0] == 'N') && // nombre
        (derecha->Elem->cat[1] == 'C') && // comun
        (derecha->Elem->ind == 3) ) { // marcado como candidato
        // Localiza posición antes de esta oración (del punto y seguido)
        izquierda = derecha->Izq;
        while (izquierda != NULL ) { // de derecha a izquierda
            if( izquierda->Elem->cat[0] == 'F') break;
            else izquierda = izquierda->Izq;
        };
        // Se recorre de derecha a izquierda hasta encontrar relaciones
        indice = 0;
        prev_izq = izquierda;
        while ( izquierda != NULL && indice < conta_cd ) {
            // se valida la bandera y el conta_cd
            if ( bandera == 'F' ) {
                if ( izquierda->Elem->cat[0] == bandera &&
                    izquierda->Elem->cat[1] == 'p')indice++;
            }
            else if ( izquierda->Elem->cat[0] == bandera)indice++;
            if ( izquierda->Elem->cat[0] == 'N' &&
                izquierda->Elem->cat[1] == 'C') {
                // se verifica correferencia directa entre cadenas
                k = strcmp(izquierda->Elem->lex, derecha->Elem->lex);
                if( k == 0 ) { // se marca correferencia
                    derecha->Elem->ind = 4; // Directa
                    derecha->Elem->enlace = izquierda->Elem->ancla;
                    indice = conta_cd;
                    // impresión de resultado en log file
                    if (imprime & 1){

```

```

        fprintf(archivo,"%d %s %s",derecha->Elem->ancla,
                prev_der->Elem->lex, derecha->Elem->lex);
        fprintf(archivo," >cd--> ");
        fprintf(archivo,"%s %s %d\n",izquierda->Elem->lex,
                prev_izq->Elem->lex, izquierda->Elem->ancla);
    }
    break;
} }
prev_izq = izquierda;
izquierda = izquierda->Izq; // apunta al siguiente nodo
} // termina while de derecha a izquierda
} // fin de if marcado
prev_der = derecha;
derecha = derecha->Der; // apunta al siguiente nodo
} // termina while correferencia
/*
Resolución de correferencia indirecta
*/
// Se localiza posición después de la primera oración completa sea
// simple, compuesta o compleja (hasta primer punto y seguido)
derecha = texto.Inicio();
while (derecha != NULL ) { // de inicio hacia final
    if( derecha->Elem->cat[0] == 'F' &&
        derecha->Elem->cat[1] == 'p') break;
    else derecha = derecha->Der; // apunta al siguiente nodo
};
// Se verifica relacion con nombres comunes del diccionario
// para cada componente léxico marcado
cerr << " Correferencias indirectas de nombres comunes " << endl;
if (imprime & 1){
    fprintf(archivo,"\n Correferencias indirectas de nombres comunes \n");
}
prev_der = derecha;
while( derecha != texto.Fin() ) { // hasta el fin de lista
    if( (derecha->Elem->cat[0] == 'N') && // nombre
        (derecha->Elem->cat[1] == 'C') && // comun
        (derecha->Elem->ind == 3 ) ) { // marcado como candidato
        // Localiza posición antes de esta oración (del punto y seguido)
        izquierda = derecha->Izq;
        while (izquierda != NULL ) { // de derecha a izquierda
            if( izquierda->Elem->cat[0] == 'F') break;
            else izquierda = izquierda->Izq;
        };
        // Se recorre de derecha a izquierda hasta encontrar relaciones
        indice = 0;
        prev_izq = izquierda;

```

```

while ( izquierda != NULL && indice < conta_ci ) {
// se valida la bandera y limite del conta_ci
if ( bandera == 'F' ) {
    if ( izquierda->Elem->cat[0] == bandera &&
        izquierda->Elem->cat[1] == 'p')indice++;
    }
else if ( izquierda->Elem->cat[0] == bandera)indice++;
    if ( izquierda->Elem->cat[0] == 'N' &&
        izquierda->Elem->cat[1] == 'C') {
// se verifica correferencia indirecta
k = coref(derecha->Elem->lex, izquierda->Elem->lex);
if( k > 0 ) { // se marca correferencia
    derecha->Elem->ind = 5; // Indirecta
    derecha->Elem->enlace = izquierda->Elem->ancla;
    indice = conta_ci;
// impresión de resultado en log file
if (imprime & 1){
    fprintf(archivo,"%d %s %s",derecha->Elem->ancla,
        prev_der->Elem->lex, derecha->Elem->lex);
    fprintf(archivo," >ci--> ");
    fprintf(archivo,"%s %s %d\n",izquierda->Elem->lex,
        prev_izq->Elem->lex, izquierda->Elem->ancla);
    }
    break;
} }
    prev_izq = izquierda;
    izquierda = izquierda->Izq; // apunta al siguiente nodo
} // termina while de derecha a izquierda
} // fin de if marcado
prev_der = derecha;
derecha = derecha->Der; // apunta al siguiente nodo
} // termina while correferencia indirecta
/*
Resolución de anafora indirecta
*/
// Se localiza posición después de la primera oración completa sea
// simple, compuesta o compleja (hasta primer punto y seguido)
derecha = texto.Inicio();
while (derecha != NULL ) { // de inicio hacia final
    if( derecha->Elem->cat[0] == 'F' &&
        derecha->Elem->cat[1] == 'p') break;
    else derecha = derecha->Der; // apunta al siguiente nodo
};
// Se verifica relacion con nombres comunes del diccionario
// para cada componente léxico marcado
cerr << " marcado de anafora indirecta " << endl;

```

```

if (imprime & 1){
    fprintf(archivo, "\n Anáfora indirecta \n");
}
prev_der = derecha;
while( derecha != texto.Fin() ) { // hasta el fin de lista
    if( (derecha->Elem->cat[0] == 'N') && // nombre comun
        (derecha->Elem->cat[1] == 'C') && // marcado
        (derecha->Elem->ind == 3) ) { // como candidato
        // Se localiza posición antes de esta oración (del punto y seguido)
        izquierda = derecha->Izq;
        while( izquierda != NULL ) { // de derecha a izquierda
            if( izquierda->Elem->cat[0] == 'F') break;
            else izquierda = izquierda->Izq;
        };
        // Se recorre de derecha a izquierda hasta encontrar relaciones
        indice = 0;
        prev_izq = izquierda;
        while ( izquierda != NULL && indice < conta_ai ) {
            // se valida la bandera y el conta_ai
            if ( bandera == 'F' ) {
                if ( izquierda->Elem->cat[0] == bandera &&
                    izquierda->Elem->cat[1] == 'p')indice++;
            }
            else if ( izquierda->Elem->cat[0] == bandera)indice++;
            if ( izquierda->Elem->cat[0] == 'N' &&
                izquierda->Elem->cat[1] == 'C') {
                // se verifica relacion anaforica
                k = anaf_ind(derecha->Elem->lex, izquierda->Elem->lex);
                if( k > 0 ) { // se marcan los coreferentes por anaind
                    derecha->Elem->ind = 6; // anafora indirecta
                    derecha->Elem->enlace = izquierda->Elem->ancla;
                    indice = conta_ai;
                    if (imprime & 1){
                        fprintf(archivo, "%d %s %s", derecha->Elem->ancla,
                            prev_der->Elem->lex, derecha->Elem->lex);
                        fprintf(archivo, " >ai--> ");
                        fprintf(archivo, "%s %s %d\n", izquierda->Elem->lex,
                            prev_izq->Elem->lex, izquierda->Elem->ancla);
                    }
                    break;
                } }
            prev_izq = izquierda;
            izquierda = izquierda->Izq; // apunta al siguiente nodo
        } // termina while de derecha a izquierda
    } // fin de if marcado
    prev_der = derecha;
}

```

```

    derecha = derecha->Der; // apunta al siguiente nodo
} // termina while anafora indirecta

// Se imprimen líneas de nombres y cuentan ocurrencias
cerr << " Impresion de marcado de nombres " << endl;
if (imprime & 1){
    fprintf(archivo,"\n Impresion de marcado de nombres \n");
}
cp = ncp = cd = ci = ai = nc = nr = nv = npron = pr = 0;
derecha = texto.Inicio();
while ( derecha != NULL ) { // de izquierda a derecha
    if ( derecha->Elem->cat[0] == 'N' ) {
        if ( derecha->Elem->ind == 0 ) nr++; // no determinante
        else if ( derecha->Elem->ind == 1 ) ncp++; // no corref. propio
        else if ( derecha->Elem->ind == 2 ) cp++; // corref. nombre propio
        else if ( derecha->Elem->ind == 3 ) nc++; // determinante no corref.
        else if ( derecha->Elem->ind == 4 ) cd++; // corref. directa
        else if ( derecha->Elem->ind == 5 ) ci++; // corref. indirecta
        else if ( derecha->Elem->ind == 6 ) ai++; // anafora indirecta
        else if ( derecha->Elem->ind == 7 ) pr++; // rel. predefinida
        if (imprime & 1){
            fprintf(archivo,"%d %s %d",derecha->Elem->ind,
                derecha->Elem->lex, derecha->Elem->ancla);
            if ( derecha->Elem->enlace ){
                if ( derecha->Elem->ind == 2 )fprintf(archivo," >CP--> ");
                else if ( derecha->Elem->ind == 4 )fprintf(archivo," >cd--> ");
                else if ( derecha->Elem->ind == 5 )fprintf(archivo," >ci--> ");
                else if ( derecha->Elem->ind == 6 )fprintf(archivo," >ai--> ");
                else if ( derecha->Elem->ind == 7 )fprintf(archivo," >pr--> ");
                fprintf(archivo,"%d", derecha->Elem->enlace);
            }
            fprintf(archivo,"\n");
        }
    }
    else if ( derecha->Elem->cat[0] == 'V' )nv++;
    else if ( derecha->Elem->cat[0] == 'P' )npron++;
    derecha = derecha->Der; // apunta al siguiente nodo
}
if (imprime & 1)fclose(archivo);
// Tiene alguna falla preferible archivo log
if (imprime & 2) {
    cerr << " Imprime archivo de cadenas " << endl;
    imp_cadenas();
}
if (imprime & 4) {
    cerr << " Imprime archivo html " << endl;

```

```

    convierte_htm();
}
// imprime archivo de estadísticas de corridas
if(imprime & 16) {
    cerr << " Imprime archivo de estadística " << endl;
    archivo = fopen(arch_estad,"a+");
    if (!archivo ) {
        cerr << "No se pudo abrir archivo " << archivo << endl;
        return(4);
    }
    fprintf( archivo,"archivo\t\t ban\t vent\t cd\t ci\t ai\t rel\t nrel\t");
    fprintf( archivo,"nref\t com\t CP\t NCP\t Prop\t Verb\t Pre\t Pron\n");
    fprintf(archivo,"%s\t %c\t %d\t %d\t %d\t %d\t %d\t %d\t",
        arch_entrada, bandera, conta_ai, cd, ci, ai,(cd+ci+ai), nc);
    fprintf(archivo," %d\t %d\t %d\t %d\t %d\t %d\t %d\t %d\n",
        nr, (cd+ci+ai+nc+nr), cp, ncp, (cp+ncp), nv, pr, npron);
    fclose(archivo);
}
cerr << " Terminado OK" << endl;
return(0);
}

```

/\*

**Función para verificar la correferencia** entre sinónimos  
p1 es el posible referente y p2 es el posible referido

Desarrollado por: Raúl Morales Carrasco      20020604  
Modificado por: Raúl Morales Carrasco      20020604

\*/

```

int coref ( char *p1, char *p2 )
{
    FILE* fout;
    char cad1[REN][COL];
    char cad2[REN][COL];
    int i, j, k, l, semeja;
    // se inicializan variables de cadenas
    for ( k = 0; k < REN; k++ )
        for ( l = 0; l < COL; l++) cad1[k][l] = 0;
    // se abre el diccionario correspondiente
    arch_sinocom[6] = primera_letra(p1[0]);
    i = cadenas ( p1, arch_sinocom, cad1 );
    if ( i == 0 ) { // no hay entrada en diccionario
        if (imprime & 8) {
            fout = fopen(arch_fanomb,"a+");
            fprintf( fout, "%s \t %s\n",p1, arch_entrada);
            fclose(fout);
        }
    }
}

```

```

    }
    strcpy (cad1[0], p1); // unicamente la entrada
}
// se inicializan variables de cadenas
for ( k = 0; k < REN; k++ )
    for ( l = 0; l < COL; l++) cad2[k][l] = 0;
arch_sinocom[6] = primera_letra(p2[0]);
j = cadenas ( p2, arch_sinocom, cad2 );
if ( j == 0 ) { // no hay entrada en diccionario
    if (imprime & 8){
        fout = fopen( arch_fanom, "a+");
        fprintf( fout, "%s \t %s\n", p2, arch_entrada);
        fclose(fout);
    }
    strcpy (cad2[0], p2); // unicamente la entrada
}
// Se comparan todas las cadenas de cad1 vs cad2
for ( k = 0; k < i; k++ ) {
    cout << cad1[k];
    for ( l = 0; l < j; l++) {
        cout << " " << cad2[l];
        semeja = strcmp(cad1[k], cad2[l]);
        if ( semeja == 0 ) {
            cout << endl << " cd_OK" << endl;
            return(1);
        }
    }
    cout << endl;
}
return(0);
}
/*

```

### **Función para verificar la anáfora indirecta**

p1 es el posible referente y p2 es el posible referido

Desarrollado por: Raúl Morales Carrasco 20020604

Modificado por: Raúl Morales Carrasco 20020604

\*/

```

int anaf_ind ( char * p1, char * p2 )
{
    FILE* fout;
    char cad1[REN][COL];
    char cad2[REN][COL];
    int i, j, k, l, semeja;
    // se inicializa variable de cadenas
    for ( k = 0; k < REN; k++ )
        for ( l = 0; l < COL; l++) cad1[k][l] = 0;

```

```

// se abre el diccionario correspondiente
arch_escena[6] = primera_letra(p1[0]);
i = relaciones(p1, arch_escena, cad1 );
if ( i == 0 ) { // no hay entrada en diccionario
    if (imprime & 8){
        fout = fopen( arch_faesce,"a+");
        fprintf( fout, "%s \t %s\n",p1, arch_entrada);
        fclose(fout);
    }
    strcpy (cad1[0], p1);
}
for ( k = 0; k < REN; k++ )
    for ( l = 0; l < COL; l++) cad2[k][l] = 0;
arch_escena[6] = primera_letra(p2[0]);
j = relaciones ( p2, arch_escena, cad2 );
if ( j == 0 ) { // no hay entrada en diccionario
    if (imprime & 8){
        fout = fopen(arch_faesce,"a+");
        fprintf( fout, "%s \t %s\n", p2, arch_entrada);
        fclose(fout);
    }
    strcpy (cad2[0], p2);
}
// Se comparan todas las cadenas de p1 vs p2
for ( k = 0; k < i; k++ ) {
    cout << cad1[k];
    for ( l = 0; l < j; l++ ) {
        cout << " " << cad2[l];
        semeja = strcmp(cad1[k], cad2[l]);
        if ( semeja == 0 ) {
            cout << endl << " ai_OK" << endl;
            return(1);
        }
    }
    cout << endl;
}
cout << endl << " ai_NO" << endl;
return(0);
}

```

## Anexo G: Resultados para determinar el tamaño de ventana

En este anexo se presentan ejemplos de tablas de resultados que permitieron analizar y determinar el tamaño de la ventana de búsqueda hacia atrás. En las columnas se utilizaron las siguientes abreviaturas:

- ban** = tipo de **bandera** a evaluar
- vent** = tamaño de **ventana** utilizado
- cd** = correferencia **directa**
- ci** = correferencia **indirecta**
- ai** = anáfora **indirecta**
- det** = **detectados** (suma de cd, ci, y ai)
- ndet** = **no detectados** (nuevas referencias o falla del método)
- nr** = **no referenciados** (nombres no precedidos por determinante)
- com** = total de nombres **comunes** marcados en el texto.
- CP** = Correferencia entre nombres **Propios**
- NCP** = **No** Correferencia entre nombres **Propios**
- Prop** = total de nombres **Propios** marcados en el texto
- Verb** = total de **Verbos** marcados en el texto
- Pron** = total de **Pronombres** marcados en el texto

Resultados del archivo Instituto Oriente													
		Nombres comunes						Nom. Propios					
<b>ban</b>	<b>vent</b>	<b>cd</b>	<b>ci</b>	<b>ai</b>	<b>cor</b>	<b>det</b>	<b>nr</b>	<b>com</b>	<b>CP</b>	<b>NCP</b>	<b>Prop</b>	<b>Verb</b>	<b>Pron</b>
F	2	3	1	2	6	45	21	72	8	34	42	32	9
F	3	3	3	3	9	42	21	72	8	34	42	32	9
F	4	3	3	4	10	41	21	72	8	34	42	32	9
F	5	7	4	5	16	35	21	72	8	34	42	32	9
F	6	8	4	5	17	34	21	72	8	34	42	32	9
F	7	8	4	5	17	34	21	72	8	34	42	32	9
F	8	8	4	6	18	33	21	72	8	34	42	32	9
F	9	9	5	6	20	31	21	72	8	34	42	32	9
F	10	9	9	4	22	29	21	72	8	34	42	32	9
F	11	9	10	4	23	28	21	72	8	34	42	32	9
F	12	9	11	3	23	28	21	72	8	34	42	32	9
F	13	10	11	3	24	27	21	72	8	34	42	32	9
F	14	10	11	3	24	27	21	72	8	34	42	32	9
F	15	11	11	3	25	26	21	72	8	34	42	32	9
F	16	11	11	3	25	26	21	72	8	34	42	32	9
F	17	11	11	3	25	26	21	72	8	34	42	32	9
F	18	11	11	3	25	26	21	72	8	34	42	32	9
F	19	11	11	3	25	26	21	72	8	34	42	32	9
F	20	12	13	3	28	23	21	72	8	34	42	32	9

Resultados del archivo Instituto Oriente													
		Nombres comunes						Nom. Propios					
ban	vent	cd	ci	ai	cor	det	nr	com	CP	NCP	Prop	Verb	Pron
F	20	12	13	3	28	23	21	72	8	34	42	32	9
F	21	12	13	3	28	23	21	72	8	34	42	32	9
F	22	12	13	3	28	23	21	72	8	34	42	32	9
F	23	12	13	3	28	23	21	72	8	34	42	32	9
F	24	12	13	3	28	23	21	72	8	34	42	32	9
F	25	12	13	3	28	23	21	72	8	34	42	32	9
N	2	3	0	1	4	47	21	72	8	34	42	32	9
N	3	4	1	1	6	45	21	72	8	34	42	32	9
N	4	4	2	1	7	44	21	72	8	34	42	32	9
N	5	5	2	4	11	40	21	72	8	34	42	32	9
N	6	7	2	4	13	38	21	72	8	34	42	32	9
N	7	7	2	5	14	37	21	72	8	34	42	32	9
N	8	7	2	7	16	35	21	72	8	34	42	32	9
N	9	7	2	7	16	35	21	72	8	34	42	32	9
N	10	7	3	7	17	34	21	72	8	34	42	32	9
N	11	7	3	7	17	34	21	72	8	34	42	32	9
N	12	7	5	7	19	32	21	72	8	34	42	32	9
N	13	7	6	6	19	32	21	72	8	34	42	32	9
N	14	7	9	3	19	32	21	72	8	34	42	32	9
N	15	7	9	3	19	32	21	72	8	34	42	32	9
N	16	7	9	3	19	32	21	72	8	34	42	32	9
N	17	8	9	3	20	31	21	72	8	34	42	32	9
N	18	8	10	3	21	30	21	72	8	34	42	32	9
N	19	8	10	3	21	30	21	72	8	34	42	32	9
N	20	10	10	4	24	27	21	72	8	34	42	32	9
N	21	11	11	3	25	26	21	72	8	34	42	32	9
N	22	11	12	3	26	25	21	72	8	34	42	32	9
N	23	12	12	3	27	24	21	72	8	34	42	32	9
N	24	12	12	3	27	24	21	72	8	34	42	32	9
N	25	12	13	3	28	23	21	72	8	34	42	32	9
P	2	3	1	2	6	45	21	72	8	34	42	32	9
P	3	3	3	3	9	42	21	72	8	34	42	32	9
P	4	7	4	6	17	34	21	72	8	34	42	32	9
P	5	7	4	6	17	34	21	72	8	34	42	32	9
P	6	9	5	5	19	32	21	72	8	34	42	32	9
P	7	9	9	5	23	28	21	72	8	34	42	32	9
P	8	9	11	3	23	28	21	72	8	34	42	32	9
P	9	9	11	3	23	28	21	72	8	34	42	32	9
P	10	10	11	3	24	27	21	72	8	34	42	32	9
P	11	11	11	3	25	26	21	72	8	34	42	32	9
P	12	11	11	3	25	26	21	72	8	34	42	32	9
P	13	12	13	3	28	23	21	72	8	34	42	32	9

Resultados del archivo Instituto Oriente													
		Nombres comunes						Nom. Propios					
ban	vent	cd	ci	ai	cor	det	nr	com	CP	NCP	Prop	Verb	Pron
P	14	12	13	3	28	23	21	72	8	34	42	32	9
P	15	12	13	3	28	23	21	72	8	34	42	32	9
P	16	12	13	3	28	23	21	72	8	34	42	32	9
P	17	12	13	3	28	23	21	72	8	34	42	32	9
P	18	13	13	3	29	22	21	72	8	34	42	32	9
P	19	13	14	3	30	21	21	72	8	34	42	32	9
P	20	13	14	3	30	21	21	72	8	34	42	32	9
P	21	13	14	3	30	21	21	72	8	34	42	32	9
P	22	13	14	3	30	21	21	72	8	34	42	32	9
P	23	13	14	3	30	21	21	72	8	34	42	32	9
P	24	13	14	3	30	21	21	72	8	34	42	32	9
P	25	13	14	3	30	21	21	72	8	34	42	32	9
V	2	5	1	2	8	43	21	72	8	34	42	32	9
V	3	8	4	3	15	36	21	72	8	34	42	32	9
V	4	8	7	4	19	32	21	72	8	34	42	32	9
V	5	8	9	4	21	30	21	72	8	34	42	32	9
V	6	8	9	5	22	29	21	72	8	34	42	32	9
V	7	10	10	5	25	26	21	72	8	34	42	32	9
V	8	10	10	5	25	26	21	72	8	34	42	32	9
V	9	11	10	5	26	25	21	72	8	34	42	32	9
V	10	11	12	3	26	25	21	72	8	34	42	32	9
V	11	11	13	3	27	24	21	72	8	34	42	32	9
V	12	12	13	3	28	23	21	72	8	34	42	32	9
V	13	12	13	3	28	23	21	72	8	34	42	32	9
V	14	12	13	3	28	23	21	72	8	34	42	32	9
V	15	12	13	3	28	23	21	72	8	34	42	32	9
V	16	12	13	3	28	23	21	72	8	34	42	32	9
V	17	13	13	3	29	22	21	72	8	34	42	32	9
V	18	13	13	3	29	22	21	72	8	34	42	32	9
V	19	13	13	3	29	22	21	72	8	34	42	32	9
V	20	13	13	3	29	22	21	72	8	34	42	32	9
V	21	13	13	3	29	22	21	72	8	34	42	32	9
V	22	13	14	3	30	21	21	72	8	34	42	32	9
V	23	13	14	3	30	21	21	72	8	34	42	32	9
V	24	13	14	3	30	21	21	72	8	34	42	32	9
V	25	13	14	3	30	21	21	72	8	34	42	32	9

<b>Resultados del archivo El Cerebro</b>														
	Nombres comunes							Nom. Propios						
	ban	vent	cd	ci	ai	cor	det	nr	com	CP	NCP	Prop	Verb	Pron
F	2	3	6	1	10	77	33	120	0	4	4	4	92	48
F	3	5	8	6	19	68	33	120	0	4	4	4	92	48
F	4	8	13	12	33	54	33	120	0	4	4	4	92	48
F	5	8	14	12	34	53	33	120	0	4	4	4	92	48
F	6	10	15	10	35	52	33	120	0	4	4	4	92	48
F	7	10	16	12	38	49	33	120	0	4	4	4	92	48
F	8	10	16	17	43	44	33	120	0	4	4	4	92	48
F	9	10	17	18	45	42	33	120	0	4	4	4	92	48
F	10	10	18	19	47	40	33	120	0	4	4	4	92	48
F	11	10	19	18	47	40	33	120	0	4	4	4	92	48
F	12	10	19	23	52	35	33	120	0	4	4	4	92	48
F	13	11	23	22	56	31	33	120	0	4	4	4	92	48
F	14	11	24	21	56	31	33	120	0	4	4	4	92	48
F	15	11	24	21	56	31	33	120	0	4	4	4	92	48
F	16	11	26	19	56	31	33	120	0	4	4	4	92	48
F	17	11	27	19	57	30	33	120	0	4	4	4	92	48
F	18	11	27	19	57	30	33	120	0	4	4	4	92	48
F	19	11	27	19	57	30	33	120	0	4	4	4	92	48
F	20	11	27	20	58	29	33	120	0	4	4	4	92	48
F	20	11	27	20	58	29	33	120	0	4	4	4	92	48
F	21	11	28	20	59	28	33	120	0	4	4	4	92	48
F	22	12	28	19	59	28	33	120	0	4	4	4	92	48
F	23	12	28	21	61	26	33	120	0	4	4	4	92	48
F	24	12	28	21	61	26	33	120	0	4	4	4	92	48
F	25	12	28	22	62	25	33	120	0	4	4	4	92	48
N	2	4	10	3	17	70	33	120	0	4	4	4	92	48
N	3	5	10	7	22	65	33	120	0	4	4	4	92	48
N	4	6	10	12	28	59	33	120	0	4	4	4	92	48
N	5	6	11	16	33	54	33	120	0	4	4	4	92	48
N	6	7	15	15	37	50	33	120	0	4	4	4	92	48
N	7	7	16	17	40	47	33	120	0	4	4	4	92	48
N	8	8	16	19	43	44	33	120	0	4	4	4	92	48
N	9	8	16	22	46	41	33	120	0	4	4	4	92	48
N	10	10	20	22	52	35	33	120	0	4	4	4	92	48
N	11	11	21	21	53	34	33	120	0	4	4	4	92	48
N	12	11	21	21	53	34	33	120	0	4	4	4	92	48
N	13	11	21	21	53	34	33	120	0	4	4	4	92	48
N	14	11	23	20	54	33	33	120	0	4	4	4	92	48
N	15	11	23	21	55	32	33	120	0	4	4	4	92	48
N	16	11	25	20	56	31	33	120	0	4	4	4	92	48
N	17	11	25	20	56	31	33	120	0	4	4	4	92	48

<b>Resultados del archivo El Cerebro</b>														
	Nombres comunes							Nom. Propios						
	<b>ban</b>	<b>vent</b>	<b>cd</b>	<b>ci</b>	<b>ai</b>	<b>cor</b>	<b>det</b>	<b>nr</b>	<b>com</b>	<b>CP</b>	<b>NCP</b>	<b>Prop</b>	<b>Verb</b>	<b>Pron</b>
N	18	11	26	20	57	30	33	120	0	4	4	92	48	
N	19	11	26	20	57	30	33	120	0	4	4	92	48	
N	20	11	26	20	57	30	33	120	0	4	4	92	48	
N	21	11	26	20	57	30	33	120	0	4	4	92	48	
N	22	11	26	21	58	29	33	120	0	4	4	92	48	
N	23	11	26	23	60	27	33	120	0	4	4	92	48	
N	24	11	26	23	60	27	33	120	0	4	4	92	48	
N	25	11	27	23	61	26	33	120	0	4	4	92	48	
P	2	8	13	9	30	57	33	120	0	4	4	92	48	
P	3	10	15	16	41	46	33	120	0	4	4	92	48	
P	4	10	19	20	49	38	33	120	0	4	4	92	48	
P	5	11	22	20	53	34	33	120	0	4	4	92	48	
P	6	11	26	20	57	30	33	120	0	4	4	92	48	
P	7	11	28	20	59	28	33	120	0	4	4	92	48	
P	8	12	28	23	63	24	33	120	0	4	4	92	48	
P	9	12	28	23	63	24	33	120	0	4	4	92	48	
P	10	12	29	22	63	24	33	120	0	4	4	92	48	
P	11	12	31	21	64	23	33	120	0	4	4	92	48	
P	12	12	31	22	65	22	33	120	0	4	4	92	48	
P	13	12	34	19	65	22	33	120	0	4	4	92	48	
P	14	12	36	18	66	21	33	120	0	4	4	92	48	
P	15	12	36	18	66	21	33	120	0	4	4	92	48	
P	16	12	36	18	66	21	33	120	0	4	4	92	48	
P	17	12	36	18	66	21	33	120	0	4	4	92	48	
P	18	12	36	18	66	21	33	120	0	4	4	92	48	
P	19	12	36	18	66	21	33	120	0	4	4	92	48	
P	20	12	36	18	66	21	33	120	0	4	4	92	48	
P	21	12	36	18	66	21	33	120	0	4	4	92	48	
P	22	12	36	18	66	21	33	120	0	4	4	92	48	
P	23	12	36	18	66	21	33	120	0	4	4	92	48	
P	24	12	36	18	66	21	33	120	0	4	4	92	48	
P	25	12	36	18	66	21	33	120	0	4	4	92	48	
V	2	6	9	6	21	66	33	120	0	4	4	92	48	
V	3	6	10	7	23	64	33	120	0	4	4	92	48	
V	4	7	14	10	31	56	33	120	0	4	4	92	48	
V	5	7	15	12	34	53	33	120	0	4	4	92	48	
V	6	8	17	13	38	49	33	120	0	4	4	92	48	
V	7	9	18	14	41	46	33	120	0	4	4	92	48	
V	8	10	18	15	43	44	33	120	0	4	4	92	48	
V	9	11	19	19	49	38	33	120	0	4	4	92	48	
V	10	11	19	19	49	38	33	120	0	4	4	92	48	
V	11	11	22	19	52	35	33	120	0	4	4	92	48	

<b>Resultados del archivo El Cerebro</b>													
Nombres comunes								Nom. Propios					
<b>ban</b>	<b>vent</b>	<b>cd</b>	<b>ci</b>	<b>ai</b>	<b>cor</b>	<b>det</b>	<b>nr</b>	<b>com</b>	<b>CP</b>	<b>NCP</b>	<b>Prop</b>	<b>Verb</b>	<b>Pron</b>
V	12	11	23	19	53	34	33	120	0	4	4	92	48
V	13	11	24	19	54	33	33	120	0	4	4	92	48
V	14	11	24	21	56	31	33	120	0	4	4	92	48
V	15	11	24	21	56	31	33	120	0	4	4	92	48
V	16	11	26	19	56	31	33	120	0	4	4	92	48
V	17	11	27	19	57	30	33	120	0	4	4	92	48
V	18	11	28	21	60	27	33	120	0	4	4	92	48
V	19	11	28	22	61	26	33	120	0	4	4	92	48
V	20	11	28	23	62	25	33	120	0	4	4	92	48
V	21	12	28	22	62	25	33	120	0	4	4	92	48
V	22	12	28	22	62	25	33	120	0	4	4	92	48
V	23	12	28	23	63	24	33	120	0	4	4	92	48
V	24	12	28	23	63	24	33	120	0	4	4	92	48
V	25	12	28	23	63	24	33	120	0	4	4	92	48

## **Anexo H: Archivos de texto libre utilizados para la evaluación**

En este anexo se presentan listados de los archivos de texto libre en formato texto de windows, tal como fueron utilizados para su evaluación.

El primero es un archivo de correo recibido por e-mail titulado "Contra la guerra"

Contra la guerra.

En todo el mundo 10000000, dicen, hemos desfilado contra la guerra.

Participé durante un trecho en la de Barcelona.

Era impresionante.

Dicen que éramos 1300000 personas clamando por la paz.

El problema es por qué no sucede lo mismo ante cualquier guerra o pisoteo de los derechos humanos.

No se trata sólo de esta guerra, sino de muchas que se están librando por todo el mundo.

Además creo que debería quedar claro que el marchar por la paz no significa que estemos a favor de Sadam, pues "gracias" a él su pueblo se está hundiendo.

¿Vivimos realmente en un mundo de buenos y malos?

¿Están los malos legitimados para permitir que, en su nombre, mueran inocentes?

¿Está tan claro lo que es el bien y el mal?

Hemos elaborado los derechos humanos y parece que no sirven para nada, pues en todas partes se siguen atropellando.

Quizá llegue el momento en que muchos millones nos manifestemos en contra de este atropello, que a menudo sucede a nuestro lado.

No hace falta alejarse mucho para observarlo.

Como dice la sabiduría popular "es fácil ver la paja en el ojo del vecino, sin darse cuenta de la viga en nuestro propio ojo".

Un abrazo a todos.

Maite.

El segundo es un archivo de correo recibido por e-mail del Instituto Oriente en Puebla, Pue. (mis hijos estudian en esta institución).

Instituto Oriente.

Puebla, Pue.

Febrero de 2003.

Comparto contigo los puntos de vista de los jesuitas sobre la propuesta de guerra del Presidente de los Estados Unidos de América, contra Irak.

José Amado Fernández Ruiz S.J.

PROVINCIA MEXICANA DE LA COMPAÑÍA DE JESUS.

"LA GUERRA ES UNA DERROTA DE LA HUMANIDAD".

AL PUEBLO Y AL GOBIERNO DE MÉXICO:

Fieles a nuestro compromiso de ser "servidores de la misión de Cristo", queremos reafirmar nuestra oposición a la guerra y nuestra determinación de trabajar a favor de una paz anclada firmemente en la justicia. Deseamos exponer las razones principales que se oponen a una guerra contra Irak.

La 'doctrina' de la guerra preventiva no está de acuerdo con la doctrina y el derecho de la ONU, ni es moralmente sostenible.

(1) La aplicación de esta doctrina abriría las puertas a una guerra infinita, a 'una guerra sin fin'.

(2) En lugar de traer una paz duradera, una guerra contra Irak aumentaría las tensiones entre musulmanes, cristianos y judíos.

(3) Los masivos gastos militares están en contraste con el interés por promover el desarrollo sostenible para todos.

(4) Que los líderes de unos pocos países industrializados hayan tomado decisiones que afectan a todos los pueblos, afecta gravemente el derecho internacional y debilita los organismos multilaterales tan pacientemente contruidos por la humanidad entera.

(5) La experiencia nos ha enseñado que los pobres son siempre las víctimas principales de la violencia y de la guerra.

Estas son las razones por las cuales nuestros esfuerzos a favor de la paz adquieren una apremiante urgencia y por las cuales respaldamos los esfuerzos de las organizaciones sociales en contra de la guerra y la actual política del Gobierno de México en el seno del Consejo de Seguridad de la ONU.

El Papa Juan Pablo II proclamó enfáticamente que la "guerra nunca es una simple fatalidad, es siempre una derrota de la humanidad".

Provincia Mexicana de la Compañía de Jesús  
Juan Luis Orozco Hernández, S.J., Provincial

El tercero es un archivo de correo recibido por e-mail del Club de efectividad que normalmente envía artículos de reflexión, motivación y superación personal; su URL es: <http://www.efectividad.net/club>

La luz del cerebro.

Comparar el cerebro con una galaxia es, sin duda, una analogía modesta. Todos los seres humanos llevamos con nosotros cerca de 1600 gramos de tejidos sin darle mucha importancia y, lo que es más, cada cerebro humano es capaz de realizar más interconexiones perfectamente configuradas que átomos hay en el universo.

Si este número teórico pudiera escribirse, sería un "1" seguido de cerca de diez millones y medio de kilómetros de ceros.

"El cerebro humano es un telar encantado donde millones de deslumbrantes lanzaderas tejen un dibujo que se desvanece en el aire, un dibujo siempre rebosante de un significado nunca perdurable. Es como si la Vía Láctea ejecutara una danza cósmica."

- Sir Charles Sherrington -

El cerebro humano no es, físicamente, muy sorprendente.

Aquellos que lo han visto, no lo describen como una visión particularmente extraordinaria.

Sin embargo, su potencial es fascinante, pero a pesar de ello, lo hacemos funcionar muy por debajo de su capacidad, incluso a veces lo maltratamos.

Es comprensible que un concertista de piano, o un artesano, valoren sus manos por encima de todo; que un pintor aprecie sus ojos; que un atleta se preocupe de sus piernas... Pero las manos son inútiles sin el cerebro, tanto como el piano sin ejecutante.

El potencial del cerebro ha sido largamente subestimado, precisamente por su omnipresencia.

El cerebro está presente en todo lo que hacemos, en cada cosa que nos sucede. Es quizás por ser una "constante", que lo anulamos de la ecuación y así, notamos lo particular de cada experiencia, pasando por alto aquello sin lo cual nada es posible para nosotros.

En cada cabeza se esconde una central eléctrica formidable, un órgano compacto y eficiente cuya capacidad -cuanto más sabemos de él- parecería expandirse hacia el infinito.

"Si la complejidad del cerebro pudiera de alguna manera transformarse en algo visible, de forma tal que se manifestara más claramente a nuestros sentidos, el mundo biológico sería un campo de luz comparado con el mundo físico. El sol, con sus grandes erupciones, se apagaría hasta quedar reducido a la pálida simplicidad de un ramo de rosas; una lombriz sería un faro; un perro, una ciudad de luz, y los seres humanos parecerían soles resplandecientes de complejidad, enceguecedores estallidos de sentido mutuo atravesando la triste noche del mundo físico que los separa. Podríamos herirnos los ojos mutuamente. Mirad las cabezas aureoladas de vuestros raros y complejos semejantes. ¿No es así?"

- John Rader Platt -

Cada cosa que hacemos y experimentamos, desde jugar al tenis hasta pagar las cuentas, tiene como base este complejo sistema bio-eléctrico.

Bien mirado, no es tan desconcertante como parece.  
Sabemos que los ojos no ven por sí mismos, son simplemente lentes.  
Los oídos no oyen por sí mismos; son -por así decirlo- micrófonos.

Cuando miramos un partido de fútbol por televisión, no vemos a los jugadores, sino representaciones electrónicas de ellos en la pantalla.  
Entre el gato vivo que usted mira y la imagen del gato que se forma en su cerebro, hay una serie de procesos neuro-fisiológicos que separan la imagen que usted recibe, de la misma manera que una serie de procesos electrónicos separan el partido de fútbol, ahí donde se está jugando, y lo lleva a su pantalla.

Nuestro cerebro es, casi literalmente, todo.  
Podemos ofrecerle más y, como contrapartida, él también nos dará más.  
Es nuestra arma secreta y silenciosa.  
Si comenzamos a usar más su poder, veremos una luz que herirá y asombrará a nuestros ojos.

## Anexo I: Resultados de las corridas con texto libre

En este anexo se presentan los listados, de los tipos de archivo, de resultados obtenidos por el programa, en una corrida con tipo de bandera “punto” y tamaño de ventana nueve para el archivo reportado “Contra la guerra”. Se presentan como los genera el programa; en ellos se utilizan la siguiente simbología y tipos de abreviatura:

**(NNN)** = número de unidad léxica (token) consecutiva en el archivo

**\*** = marca la unidad léxica enlazada

**<- abr -** tipo de enlace donde abr puede ser

**cd** = correferencia **d**irecta

**ci** = correferencia **i**ndirecta

**ai** = anáfora **i**ndirecta

El primer archivo es un listado de unidades léxicas consecutivas marcando la ocurrencia o no de correferencia directa e indirecta, y la de anáfora indirecta; su terminación es “cad” por ejemplo res\contra.cad

```
( 1) Contra
( 2) la
( 3) * guerra <-cd-(17) guerra
( 4) .
( 5) En
( 6) todo
( 7) el
( 8) * mundo <-cd-(78) mundo
( 9) 10000000
(10) ,
(11) dicen
(12) ,
(13) hemos
(14) desfilado
(15) contra
(16) la
(17) * guerra <-cd-(52) guerra
(18) .
(19) Participé
(20) durante
(21) un
(22) trecho
(23) en
(24) la
(25) de
(26) Barcelona
(27) .
(28) Era
(29) impresionante
(30) .
(31) Dicen
(32) que
(33) éramos
(34) 1300000
(35) personas
(36) clamando
(37) por
(38) la
(39) * paz <-cd-(91) paz
(40) .
(41) El
(42) problema
(43) es
(44) por
(45) qué
(46) no
(47) sucede
(48) lo
(49) mismo
(50) ante
(51) cualquier
(52) * guerra <-cd-(66) guerra
(53) o
(54) pisoteo
(55) de
(56) los
```

( 57) * derechos <-cd-(156)	( 112) .
derechos	( 113) ¿
( 58) humanos	( 114) Vivimos
( 59) .	( 115) realmente
( 60) No	( 116) en
( 61) se	( 117) un
( 62) trata	( 118) mundo
( 63) sólo	( 119) de
( 64) de	( 120) buenos
( 65) esta	( 121) y
( 66) guerra	( 122) * malos <-cd-(127) malos
( 67) ,	( 123) ?
( 68) sino	( 124) ¿
( 69) de	( 125) Están
( 70) muchas	( 126) los
( 71) que	( 127) malos
( 72) se	( 128) legitimados
( 73) están	( 129) para
( 74) librando	( 130) permitir
( 75) por	( 131) que
( 76) todo	( 132) ,
( 77) el	( 133) en
( 78) mundo	( 134) su
( 79) .	( 135) nombre
( 80) Además	( 136) ,
( 81) creo	( 137) mueran
( 82) que	( 138) inocentes
( 83) debería	( 139) ?
( 84) quedar	( 140) ¿
( 85) claro	( 141) Está
( 86) que	( 142) tan
( 87) el	( 143) claro
( 88) marchar	( 144) lo
( 89) por	( 145) que
( 90) la	( 146) es
( 91) paz	( 147) el
( 92) no	( 148) * bien <-ai-(169) partes
( 93) significa	( 149) y
( 94) que	( 150) el
( 95) estemos	( 151) mal
( 96) a	( 152) ?
( 97) favor	( 153) Hemos
( 98) de	( 154) elaborado
( 99) Sadam	( 155) los
( 100) ,	( 156) derechos
( 101) pues	( 157) humanos
( 102) "	( 158) y
( 103) gracias	( 159) parece
( 104) "	( 160) que
( 105) a	( 161) no
( 106) él	( 162) sirven
( 107) su	( 163) para
( 108) pueblo	( 164) nada
( 109) se	( 165) ,
( 110) está	( 166) pues
( 111) hundiendo	( 167) en

## Anexos

( 168)	todas	( 205)	.
( 169)	partes	( 206)	Como
( 170)	se	( 207)	dice
( 171)	siguen	( 208)	la
( 172)	atropellando	( 209)	sabiduría
( 173)	.	( 210)	popular
( 174)	Quizá	( 211)	"
( 175)	llegue	( 212)	es
( 176)	el	( 213)	fácil
( 177)	momento	( 214)	ver
( 178)	en	( 215)	la
( 179)	que	( 216)	paja
( 180)	muchos	( 217)	en
( 181)	millones	( 218)	el
( 182)	nos	( 219)	* ojo <-cd-(232) ojo
( 183)	manifestemos	( 220)	del
( 184)	en	( 221)	vecino
( 185)	contra	( 222)	,
( 186)	de	( 223)	sin
( 187)	este	( 224)	darse
( 188)	atropello	( 225)	cuenta
( 189)	,	( 226)	de
( 190)	que	( 227)	la
( 191)	a	( 228)	viga
( 192)	menudo	( 229)	en
( 193)	sucede	( 230)	nuestro
( 194)	a	( 231)	propio
( 195)	nuestro	( 232)	ojo
( 196)	lado	( 233)	"
( 197)	.	( 234)	.
( 198)	No	( 235)	Un
( 199)	hace	( 236)	abrazo
( 200)	falta	( 237)	a
( 201)	alejarse	( 238)	todos
( 202)	mucho	( 239)	.
( 203)	para	( 240)	Maite
( 204)	observarlo	( 241)	

El segundo archivo es un resumen estadístico de la corrida con nombre estadis.txt

archivo	ban	vent	cd	ci	ai	cor	det	nr	com	CP	NCP	Prop	Verb	Pron
contra	P	9	6	6	2	14	14	9	37	0	2	2	44	14

El tercer archivo es un archivo en formato htm para permitir una mejor navegación en el seguimiento de resultados, con extensión “htm”, por ejemplo res\contra.htm. En este archivo se pueden observar subíndices numéricos entre parentesis que marcan las unidades léxicas (tokens); si el subíndice está subrayado indica que la unidad léxica es una coreferencia o anáfora indirecta. La información adicional se puede obtener con un navegador y colocando el puntero del ratón sobre el elemento (ver anexo J).

Contra la [guerra](#) <sup>(3)</sup>.

En todo el [mundo](#) <sup>(8)</sup>10000000 , dicen , hemos desfilado contra la [guerra](#) <sup>(17)</sup>.

Participé durante un [trecho](#) <sup>(22)</sup>en la de [Barcelona](#) <sup>(26)</sup>.

Era impresionante .

Dicen que éramos 1300000 [personas](#) <sup>(35)</sup>clamando por la [paz](#) <sup>(39)</sup>.

El [problema](#) <sup>(42)</sup>es por qué no sucede lo mismo ante cualquier [guerra](#) <sup>(52)</sup>o [pisoteo](#) <sup>(54)</sup>de los [derechos](#) <sup>(57)</sup>humanos .

No se trata sólo de esta [guerra](#) <sup>(66)</sup>, sino de muchas que se están librando por todo el [mundo](#) <sup>(78)</sup>.

Además creo que debería quedar claro que el marchar por la [paz](#) <sup>(91)</sup>no significa que estemos a [favor](#) <sup>(97)</sup>de [Sadam](#) <sup>(99)</sup>, pues " [gracias](#) <sup>(103)</sup>" a él su [pueblo](#) <sup>(108)</sup>se está hundiendo .

¿ Vivimos realmente en un [mundo](#) <sup>(118)</sup>de [buenos](#) <sup>(120)</sup>y [malos](#) <sup>(122)</sup>? ¿ Están los [malos](#) <sup>(127)</sup>legitimados para permitir que , en su [nombre](#) <sup>(135)</sup>, mueran [inocentes](#) <sup>(138)</sup>? ¿ Está tan claro lo que es el [bien](#) <sup>(148)</sup>y el [mal](#) <sup>(151)</sup>? Hemos elaborado los [derechos](#) <sup>(156)</sup>humanos y parece que no sirven para nada , pues en todas [partes](#) <sup>(169)</sup>se siguen atropellando .

Quizá llegue el [momento](#) <sup>(177)</sup>en que muchos millones nos manifestemos en contra de este [atropello](#) <sup>(188)</sup>, que a menudo sucede a nuestro [lado](#) <sup>(196)</sup>.

No hace falta alejarse mucho para observarlo .

Como dice la [sabiduría](#) <sup>(209)</sup>popular " es fácil ver la [paja](#) <sup>(216)</sup>en el [ojo](#) <sup>(219)</sup>del [vecino](#) <sup>(221)</sup>, sin darse [cuenta](#) <sup>(225)</sup>de la [viga](#) <sup>(228)</sup>en nuestro propio [ojo](#) <sup>(232)</sup>" .

Un [abrazo](#) <sup>(236)</sup>a todos .

[Maite](#) <sup>(240)</sup>.

Finalmente se genera un archivo con extensión “log”, por ejemplo res\contra.log para poder hacer un seguimiento exhaustivo de la corrida donde se tienen todas las comparaciones.



## Anexo j: Tablas de análisis de resultados de anáfora

En todas estas tablas las abreviaturas significan:

vent = tamaño de la ventana

det = expresión referencial = det + nombre comun

Vp = Verdadera (validado manual)

Fp = Falso (validado manual)

Ob = Obtenido por el programa

F = métrica F (F measure)

P = Precisión

R = Especificidad (recall)

E = Exactitud de acuerdo (agreement)

comb = corrida con mejor ventana

Real = de acuerdo al experto

archivo	vent	Correferencia						no_corref			Total				Metricas															
		directa			indirecta			indirecta			indirecta				anafora indirecta				Correferencia directa				Correferencia indirecta				no corref. Indirecta			
		Vf	Fp	Ob	Vf	Fp	Ob	Vp	Fp	Ob	Vf	Fp	Ob	det	F	P	R	E	F	P	R	E	F	P	R	E	F	P	R	E
A12a	1	1	0	1	3	2	5	2	1	3	5	3	8	67	0.42	0.63	0.31	0.79	0.40	1.00	0.25	0.96	0.43	0.60	0.33	0.88	0.40	0.67	0.29	0.91
A12a	2	2	0	2	8	7	15	5	4	9	13	11	24	67	0.65	0.54	0.81	0.79	0.67	1.00	0.50	0.97	0.67	0.53	0.89	0.88	0.63	0.56	0.71	0.91
A12a	3	4	1	5	9	11	20	5	6	11	14	17	31	67	0.60	0.45	0.88	0.72	0.89	0.80	1.00	0.99	0.62	0.45	1.00	0.84	0.56	0.45	0.71	0.88
A12a	4	4	2	6	9	12	21	4	6	10	13	18	31	67	0.55	0.42	0.81	0.69	0.80	0.67	1.00	0.97	0.60	0.43	1.00	0.82	0.47	0.40	0.57	0.87
A12a	5	4	3	7	9	12	21	4	9	13	13	21	34	67	0.52	0.38	0.81	0.64	0.73	0.57	1.00	0.96	0.60	0.43	1.00	0.82	0.40	0.31	0.57	0.82
A12a	6	4	3	7	9	13	22	4	8	12	13	21	34	67	0.52	0.38	0.81	0.64	0.73	0.57	1.00	0.96	0.58	0.41	1.00	0.81	0.42	0.33	0.57	0.84
A12a	7	4	3	7	9	14	23	4	8	12	13	22	35	67	0.51	0.37	0.81	0.63	0.73	0.57	1.00	0.96	0.56	0.39	1.00	0.79	0.42	0.33	0.57	0.84
A12a	8	4	4	8	9	14	23	3	8	11	12	22	34	67	0.48	0.35	0.75	0.61	0.67	0.50	1.00	0.94	0.56	0.39	1.00	0.79	0.33	0.27	0.43	0.82
A12a	9	4	4	8	9	15	24	3	8	11	12	23	35	67	0.47	0.34	0.75	0.60	0.67	0.50	1.00	0.94	0.55	0.38	1.00	0.78	0.33	0.27	0.43	0.82
A12a	10	4	5	9	8	15	23	3	9	12	11	24	35	67	0.43	0.31	0.69	0.57	0.62	0.44	1.00	0.93	0.50	0.35	0.89	0.76	0.32	0.25	0.43	0.81
A12a	11	4	5	9	9	15	24	3	9	12	12	24	36	67	0.46	0.33	0.75	0.58	0.62	0.44	1.00	0.93	0.55	0.38	1.00	0.78	0.32	0.25	0.43	0.81
A12a	12	4	5	9	9	17	26	2	8	10	11	25	36	67	0.42	0.31	0.69	0.55	0.62	0.44	1.00	0.93	0.51	0.35	1.00	0.75	0.24	0.20	0.29	0.81
A12a	13	4	5	9	9	17	26	2	9	11	11	26	37	67	0.42	0.30	0.69	0.54	0.62	0.44	1.00	0.93	0.51	0.35	1.00	0.75	0.22	0.18	0.29	0.79
A12a	14	4	5	9	9	17	26	2	9	11	11	26	37	67	0.42	0.30	0.69	0.54	0.62	0.44	1.00	0.93	0.51	0.35	1.00	0.75	0.22	0.18	0.29	0.79
A12a	15	4	5	9	9	17	26	2	9	11	11	26	37	67	0.42	0.30	0.69	0.54	0.62	0.44	1.00	0.93	0.51	0.35	1.00	0.75	0.22	0.18	0.29	0.79
A12a	16	4	5	9	9	17	26	2	9	11	11	26	37	67	0.42	0.30	0.69	0.54	0.62	0.44	1.00	0.93	0.51	0.35	1.00	0.75	0.22	0.18	0.29	0.79
A12a	17	4	5	9	9	17	26	2	9	11	11	26	37	67	0.42	0.30	0.69	0.54	0.62	0.44	1.00	0.93	0.51	0.35	1.00	0.75	0.22	0.18	0.29	0.79
A12a	18	4	5	9	9	17	26	2	9	11	11	26	37	67	0.42	0.30	0.69	0.54	0.62	0.44	1.00	0.93	0.51	0.35	1.00	0.75	0.22	0.18	0.29	0.79
A12a	19	4	5	9	9	17	26	2	9	11	11	26	37	67	0.42	0.30	0.69	0.54	0.62	0.44	1.00	0.93	0.51	0.35	1.00	0.75	0.22	0.18	0.29	0.79
A12a	20	4	5	9	9	17	26	2	9	11	11	26	37	67	0.42	0.30	0.69	0.54	0.62	0.44	1.00	0.93	0.51	0.35	1.00	0.75	0.22	0.18	0.29	0.79
a12a	comb	4	1	5	8	7	20	5	4	9	13	11	24	67	0.65	0.54	0.81	0.79	0.89	0.80	1.00	0.99	0.55	0.40	0.89	0.81	0.63	0.56	0.71	0.91
	Real	4			9			7			16																			

Anexos

archivo	vent	Correferencia						no_corref			Total				Metricas															
		directa			indirecta			indirecta			indirecta				anafora indirecta				Correferencia directa				Correferencia indirecta				no corref. Indirecta			
		Vf	Fp	Ob	Vf	Fp	Ob	Vp	Fp	Ob	Vf	Fp	Ob	det	F	P	R	E	F	P	R	E	F	P	R	E	F	P	R	E
A12b	1	2	0	2	3	3	6	0	1	1	3	4	7	77	0.26	0.43	0.19	0.78	0.21	1.00	0.12	0.81	0.30	0.50	0.21	0.82	0.00	0.00	0.00	0.96
A12b	2	10	0	10	8	9	17	2	6	8	10	15	25	77	0.49	0.40	0.63	0.73	0.74	1.00	0.59	0.91	0.52	0.47	0.57	0.81	0.40	0.25	1.00	0.92
A12b	3	11	0	11	13	11	24	2	7	9	15	18	33	77	0.61	0.45	0.94	0.75	0.79	1.00	0.65	0.92	0.68	0.54	0.93	0.84	0.36	0.22	1.00	0.91
A12b	4	11	0	11	14	12	26	2	9	11	16	21	37	77	0.60	0.43	1.00	0.73	0.79	1.00	0.65	0.92	0.70	0.54	1.00	0.84	0.31	0.18	1.00	0.88
A12b	5	13	0	13	14	12	26	2	12	14	16	24	40	77	0.57	0.40	1.00	0.69	0.87	1.00	0.76	0.95	0.70	0.54	1.00	0.84	0.25	0.14	1.00	0.84
A12b	6	14	0	14	14	13	27	3	11	14	17	24	41	77	0.60	0.41	1.06	0.70	0.90	1.00	0.82	0.96	0.68	0.52	1.00	0.83	0.38	0.21	1.50	0.87
A12b	7	14	1	15	14	13	27	2	11	13	16	24	40	77	0.57	0.40	1.00	0.69	0.88	0.93	0.82	0.95	0.68	0.52	1.00	0.83	0.27	0.15	1.00	0.86
A12b	8	14	1	15	14	14	28	2	10	12	16	24	40	77	0.57	0.40	1.00	0.69	0.88	0.93	0.82	0.95	0.67	0.50	1.00	0.82	0.29	0.17	1.00	0.87
A12b	9	14	1	15	14	14	28	2	10	12	16	24	40	77	0.57	0.40	1.00	0.69	0.88	0.93	0.82	0.95	0.67	0.50	1.00	0.82	0.29	0.17	1.00	0.87
A12b	10	17	1	18	14	15	29	2	9	11	16	24	40	77	0.57	0.40	1.00	0.69	0.97	0.94	1.00	0.99	0.65	0.48	1.00	0.81	0.31	0.18	1.00	0.88
A12b	11	17	2	19	14	16	30	2	9	11	16	26	42	77	0.55	0.38	1.00	0.66	0.94	0.89	1.00	0.97	0.64	0.47	1.00	0.79	0.31	0.18	1.00	0.88
A12b	12	17	2	19	14	16	30	2	11	13	16	27	43	77	0.54	0.37	1.00	0.65	0.94	0.89	1.00	0.97	0.64	0.47	1.00	0.79	0.27	0.15	1.00	0.86
A12b	13	17	2	19	14	16	30	2	11	13	16	27	43	77	0.54	0.37	1.00	0.65	0.94	0.89	1.00	0.97	0.64	0.47	1.00	0.79	0.27	0.15	1.00	0.86
A12b	14	17	2	19	14	16	30	2	12	14	16	28	44	77	0.53	0.36	1.00	0.64	0.94	0.89	1.00	0.97	0.64	0.47	1.00	0.79	0.25	0.14	1.00	0.84
A12b	15	17	2	19	14	16	30	2	12	14	16	28	44	77	0.53	0.36	1.00	0.64	0.94	0.89	1.00	0.97	0.64	0.47	1.00	0.79	0.25	0.14	1.00	0.84
A12b	16	17	2	19	14	17	31	2	11	13	16	28	44	77	0.53	0.36	1.00	0.64	0.94	0.89	1.00	0.97	0.62	0.45	1.00	0.78	0.27	0.15	1.00	0.86
A12b	17	17	2	19	14	17	31	2	12	14	16	29	45	77	0.52	0.36	1.00	0.62	0.94	0.89	1.00	0.97	0.62	0.45	1.00	0.78	0.25	0.14	1.00	0.84
A12b	18	17	2	19	14	17	31	2	12	14	16	29	45	77	0.52	0.36	1.00	0.62	0.94	0.89	1.00	0.97	0.62	0.45	1.00	0.78	0.25	0.14	1.00	0.84
A12b	19	17	2	19	14	17	31	2	12	14	16	29	45	77	0.52	0.36	1.00	0.62	0.94	0.89	1.00	0.97	0.62	0.45	1.00	0.78	0.25	0.14	1.00	0.84
A12b	20	17	2	19	14	17	31	2	12	14	16	29	45	77	0.52	0.36	1.00	0.62	0.94	0.89	1.00	0.97	0.62	0.45	1.00	0.78	0.25	0.14	1.00	0.84
A12b	comb	17	1	18	13	10	23	1	7	8	14	17	31	77	0.60	0.45	0.88	0.75	0.97	0.94	1.00	0.99	0.70	0.57	0.93	0.86	0.20	0.13	0.50	0.90
	Real	17			14			2			16																			

archivo	vent	Correferencia						no_corref			Total				Metricas															
		directa			indirecta			indirecta			indirecta				anafora indirecta				Correferencia directa				Correferencia indirecta				no correfer. Indirecta			
		Vf	Fp	Ob	Vf	Fp	Ob	Vp	Fp	Ob	Vf	Fp	Ob	det	F	P	R	E	F	P	R	E	F	P	R	E	F	P	R	E
A12c	1	4	0	4	1	0	1	2	0	2	3	0	3	39	0.46	1.00	0.30	0.82	0.53	1.00	0.36	0.82	0.33	1.00	0.20	0.90	0.57	1.00	0.40	0.92
A12c	2	8	0	8	1	1	2	3	4	7	4	5	9	39	0.42	0.44	0.40	0.72	0.84	1.00	0.73	0.92	0.29	0.50	0.20	0.87	0.50	0.43	0.60	0.85
A12c	3	10	0	10	3	2	5	2	3	5	5	5	10	39	0.50	0.50	0.50	0.74	0.95	1.00	0.91	0.97	0.60	0.60	0.60	0.90	0.40	0.40	0.40	0.85
A12c	4	10	0	10	4	2	6	3	3	6	7	5	12	39	0.64	0.58	0.70	0.79	0.95	1.00	0.91	0.97	0.73	0.67	0.80	0.92	0.55	0.50	0.60	0.87
A12c	5	10	0	10	5	2	7	3	3	6	8	5	13	39	0.70	0.62	0.80	0.82	0.95	1.00	0.91	0.97	0.83	0.71	1.00	0.95	0.55	0.50	0.60	0.87
A12c	6	10	0	10	5	2	7	3	3	6	8	5	13	39	0.70	0.62	0.80	0.82	0.95	1.00	0.91	0.97	0.83	0.71	1.00	0.95	0.55	0.50	0.60	0.87
A12c	7	10	0	10	5	2	7	3	3	6	8	5	13	39	0.70	0.62	0.80	0.82	0.95	1.00	0.91	0.97	0.83	0.71	1.00	0.95	0.55	0.50	0.60	0.87
A12c	8	10	0	10	5	2	7	3	4	7	8	6	14	39	0.67	0.57	0.80	0.79	0.95	1.00	0.91	0.97	0.83	0.71	1.00	0.95	0.50	0.43	0.60	0.85
A12c	9	11	0	11	5	1	6	3	4	7	8	5	13	39	0.70	0.62	0.80	0.82	1.00	1.00	1.00	1.00	0.91	0.83	1.00	0.97	0.50	0.43	0.60	0.85
A12c	10	11	0	11	5	1	6	3	4	7	8	5	13	39	0.70	0.62	0.80	0.82	1.00	1.00	1.00	1.00	0.91	0.83	1.00	0.97	0.50	0.43	0.60	0.85
A12c	11	11	0	11	5	1	6	3	4	7	8	5	13	39	0.70	0.62	0.80	0.82	1.00	1.00	1.00	1.00	0.91	0.83	1.00	0.97	0.50	0.43	0.60	0.85
A12c	12	11	0	11	5	1	6	3	4	7	8	5	13	39	0.70	0.62	0.80	0.82	1.00	1.00	1.00	1.00	0.91	0.83	1.00	0.97	0.50	0.43	0.60	0.85
A12c	13	11	0	11	5	1	6	3	4	7	8	5	13	39	0.70	0.62	0.80	0.82	1.00	1.00	1.00	1.00	0.91	0.83	1.00	0.97	0.50	0.43	0.60	0.85
A12c	14	11	0	11	5	1	6	3	4	7	8	5	13	39	0.70	0.62	0.80	0.82	1.00	1.00	1.00	1.00	0.91	0.83	1.00	0.97	0.50	0.43	0.60	0.85
A12c	15	11	0	11	5	1	6	3	4	7	8	5	13	39	0.70	0.62	0.80	0.82	1.00	1.00	1.00	1.00	0.91	0.83	1.00	0.97	0.50	0.43	0.60	0.85
A12c	16	11	0	11	5	1	6	3	4	7	8	5	13	39	0.70	0.62	0.80	0.82	1.00	1.00	1.00	1.00	0.91	0.83	1.00	0.97	0.50	0.43	0.60	0.85
A12c	17	11	0	11	5	1	6	3	4	7	8	5	13	39	0.70	0.62	0.80	0.82	1.00	1.00	1.00	1.00	0.91	0.83	1.00	0.97	0.50	0.43	0.60	0.85
A12c	18	11	0	11	5	1	6	3	4	7	8	5	13	39	0.70	0.62	0.80	0.82	1.00	1.00	1.00	1.00	0.91	0.83	1.00	0.97	0.50	0.43	0.60	0.85
A12c	19	11	0	11	5	1	6	3	4	7	8	5	13	39	0.70	0.62	0.80	0.82	1.00	1.00	1.00	1.00	0.91	0.83	1.00	0.97	0.50	0.43	0.60	0.85
A12c	20	11	0	11	5	1	6	3	4	7	8	5	13	39	0.70	0.62	0.80	0.82	1.00	1.00	1.00	1.00	0.91	0.83	1.00	0.97	0.50	0.43	0.60	0.85
A12c	comb	11	0	11	5	1	6	3	3	6	8	4	12	39	0.73	0.67	0.80	0.85	1.00	1.00	1.00	1.00	0.91	0.83	1.00	0.97	0.55	0.50	0.60	0.87
	Real	11			5			5			10																			

Anexos

archivo	vent	Correferencia						no_corref			Total				Metricas															
		directa			indirecta			indirecta			indirecta				anafora indirecta				Correferencia directa				Correferencia indirecta				no correfer. Indirecta			
		Vf	Fp	Ob	Vf	Fp	Ob	Vp	Fp	Ob	Vf	Fp	Ob	det	F	P	R	E	F	P	R	E	F	P	R	E	F	P	R	E
A14	1	10	0	10	2	4	6	2	5	7	4	9	13	118	0.25	0.31	0.21	0.80	0.56	1.00	0.38	0.86	0.21	0.33	0.15	0.87	0.31	0.29	0.33	0.92
A14	2	12	0	12	11	9	20	4	8	12	15	17	32	118	0.59	0.47	0.79	0.82	0.63	1.00	0.46	0.88	0.67	0.55	0.85	0.91	0.44	0.33	0.67	0.92
A14	3	15	1	16	11	8	19	5	15	20	16	23	39	118	0.55	0.41	0.84	0.78	0.71	0.94	0.58	0.90	0.69	0.58	0.85	0.92	0.38	0.25	0.83	0.86
A14	4	17	1	18	12	12	24	4	18	22	16	30	46	118	0.49	0.35	0.84	0.72	0.77	0.94	0.65	0.92	0.65	0.50	0.92	0.89	0.29	0.18	0.67	0.83
A14	5	17	1	18	13	12	25	5	19	24	18	31	49	118	0.53	0.37	0.95	0.73	0.77	0.94	0.65	0.92	0.68	0.52	1.00	0.90	0.33	0.21	0.83	0.83
A14	6	19	1	20	13	14	27	5	21	26	18	35	53	118	0.50	0.34	0.95	0.69	0.83	0.95	0.73	0.93	0.65	0.48	1.00	0.88	0.31	0.19	0.83	0.81
A14	7	20	1	21	13	14	27	6	25	31	19	39	58	118	0.49	0.33	1.00	0.67	0.85	0.95	0.77	0.94	0.65	0.48	1.00	0.88	0.32	0.19	1.00	0.79
A14	8	20	1	21	13	17	30	4	26	30	17	43	60	118	0.43	0.28	0.89	0.62	0.85	0.95	0.77	0.94	0.60	0.43	1.00	0.86	0.22	0.13	0.67	0.76
A14	9	21	1	22	13	20	33	4	25	29	17	45	62	118	0.42	0.27	0.89	0.60	0.88	0.95	0.81	0.95	0.57	0.39	1.00	0.83	0.23	0.14	0.67	0.77
A14	10	23	1	24	13	23	36	4	23	27	17	46	63	118	0.41	0.27	0.89	0.59	0.92	0.96	0.88	0.97	0.53	0.36	1.00	0.81	0.24	0.15	0.67	0.79
A14	11	23	1	24	13	23	36	4	23	27	17	46	63	118	0.41	0.27	0.89	0.59	0.92	0.96	0.88	0.97	0.53	0.36	1.00	0.81	0.24	0.15	0.67	0.79
A14	12	23	1	24	13	26	39	4	22	26	17	48	65	118	0.40	0.26	0.89	0.58	0.92	0.96	0.88	0.97	0.50	0.33	1.00	0.78	0.25	0.15	0.67	0.80
A14	13	23	1	24	13	28	41	4	21	25	17	49	66	118	0.40	0.26	0.89	0.57	0.92	0.96	0.88	0.97	0.48	0.32	1.00	0.76	0.26	0.16	0.67	0.81
A14	14	24	1	25	13	28	41	4	24	28	17	52	69	118	0.39	0.25	0.89	0.54	0.94	0.96	0.92	0.97	0.48	0.32	1.00	0.76	0.24	0.14	0.67	0.78
A14	15	24	1	25	13	29	42	4	23	27	17	52	69	118	0.39	0.25	0.89	0.54	0.94	0.96	0.92	0.97	0.47	0.31	1.00	0.75	0.24	0.15	0.67	0.79
A14	16	24	1	25	13	30	43	4	23	27	17	53	70	118	0.38	0.24	0.89	0.53	0.94	0.96	0.92	0.97	0.46	0.30	1.00	0.75	0.24	0.15	0.67	0.79
A14	17	24	1	25	13	30	43	4	23	27	17	53	70	118	0.38	0.24	0.89	0.53	0.94	0.96	0.92	0.97	0.46	0.30	1.00	0.75	0.24	0.15	0.67	0.79
A14	18	25	1	26	13	29	42	4	24	28	17	53	70	118	0.38	0.24	0.89	0.53	0.96	0.96	0.96	0.98	0.47	0.31	1.00	0.75	0.24	0.14	0.67	0.78
A14	19	25	1	26	13	30	43	4	23	27	17	53	70	118	0.38	0.24	0.89	0.53	0.96	0.96	0.96	0.98	0.46	0.30	1.00	0.75	0.24	0.15	0.67	0.79
A14	20	26	1	27	13	31	44	4	22	26	17	53	70	118	0.38	0.24	0.89	0.53	0.98	0.96	1.00	0.99	0.46	0.30	1.00	0.74	0.25	0.15	0.67	0.80
A14	comb	26	1	27	11	8	19	3	7	10	14	15	29	118	0.58	0.48	0.74	0.83	0.98	0.96	1.00	0.99	0.69	0.58	0.85	0.92	0.38	0.30	0.50	0.92
	Real	26			13			6			19																			

archivo	vent	Correferencia						no_corref			Total			Metricas																
		directa			indirecta			indirecta			indirecta			anafora indirecta				Correferencia directa				Correferencia indirecta				no corref. Indirecta				
		Vf	Fp	Ob	Vf	Fp	Ob	Vp	Fp	Ob	Vf	Fp	Ob	det	F	P	R	E	F	P	R	E	F	P	R	E	F	P	R	E
cerebro	1	4	0	4	1	1	2	4	1	5	5	2	7	66	0.59	0.71	0.50	0.89	0.38	1.00	0.24	0.80	0.29	0.50	0.20	0.92	0.80	0.80	0.80	0.97
cerebro	2	9	0	9	2	5	7	4	8	12	6	13	19	66	0.41	0.32	0.60	0.74	0.69	1.00	0.53	0.88	0.33	0.29	0.40	0.88	0.47	0.33	0.80	0.86
cerebro	3	11	0	11	3	6	9	4	12	16	7	18	25	66	0.40	0.28	0.70	0.68	0.79	1.00	0.65	0.91	0.43	0.33	0.60	0.88	0.38	0.25	0.80	0.80
cerebro	4	12	0	12	4	7	11	5	14	19	9	21	30	66	0.45	0.30	0.90	0.67	0.83	1.00	0.71	0.92	0.50	0.36	0.80	0.88	0.42	0.26	1.00	0.79
cerebro	5	13	0	13	4	11	15	4	14	18	8	25	33	66	0.37	0.24	0.80	0.59	0.87	1.00	0.76	0.94	0.40	0.27	0.80	0.82	0.35	0.22	0.80	0.77
cerebro	6	13	0	13	4	13	17	4	14	18	8	27	35	66	0.36	0.23	0.80	0.56	0.87	1.00	0.76	0.94	0.36	0.24	0.80	0.79	0.35	0.22	0.80	0.77
cerebro	7	13	0	13	5	14	19	3	14	17	8	28	36	66	0.35	0.22	0.80	0.55	0.87	1.00	0.76	0.94	0.42	0.26	1.00	0.79	0.27	0.18	0.60	0.76
cerebro	8	15	1	16	5	11	16	3	16	19	8	27	35	66	0.36	0.23	0.80	0.56	0.91	0.94	0.88	0.95	0.48	0.31	1.00	0.83	0.25	0.16	0.60	0.73
cerebro	9	15	1	16	5	12	17	3	16	19	8	28	36	66	0.35	0.22	0.80	0.55	0.91	0.94	0.88	0.95	0.45	0.29	1.00	0.82	0.25	0.16	0.60	0.73
cerebro	10	16	2	18	5	11	16	3	15	18	8	26	34	66	0.36	0.24	0.80	0.58	0.91	0.89	0.94	0.95	0.48	0.31	1.00	0.83	0.26	0.17	0.60	0.74
cerebro	11	16	2	18	5	12	17	3	14	17	8	26	34	66	0.36	0.24	0.80	0.58	0.91	0.89	0.94	0.95	0.45	0.29	1.00	0.82	0.27	0.18	0.60	0.76
cerebro	12	16	2	18	5	12	17	3	14	17	8	26	34	66	0.36	0.24	0.80	0.58	0.91	0.89	0.94	0.95	0.45	0.29	1.00	0.82	0.27	0.18	0.60	0.76
cerebro	13	16	3	19	5	14	19	2	12	14	7	26	33	66	0.33	0.21	0.70	0.56	0.89	0.84	0.94	0.94	0.42	0.26	1.00	0.79	0.21	0.14	0.40	0.77
cerebro	14	17	3	20	5	14	19	2	12	14	7	26	33	66	0.33	0.21	0.70	0.56	0.92	0.85	1.00	0.95	0.42	0.26	1.00	0.79	0.21	0.14	0.40	0.77
cerebro	15	17	3	20	5	15	20	2	11	13	7	26	33	66	0.33	0.21	0.70	0.56	0.92	0.85	1.00	0.95	0.40	0.25	1.00	0.77	0.22	0.15	0.40	0.79
cerebro	16	17	3	20	5	15	20	2	11	13	7	26	33	66	0.33	0.21	0.70	0.56	0.92	0.85	1.00	0.95	0.40	0.25	1.00	0.77	0.22	0.15	0.40	0.79
cerebro	17	17	3	20	5	15	20	2	11	13	7	26	33	66	0.33	0.21	0.70	0.56	0.92	0.85	1.00	0.95	0.40	0.25	1.00	0.77	0.22	0.15	0.40	0.79
cerebro	18	17	3	20	5	15	20	2	11	13	7	26	33	66	0.33	0.21	0.70	0.56	0.92	0.85	1.00	0.95	0.40	0.25	1.00	0.77	0.22	0.15	0.40	0.79
cerebro	19	17	3	20	5	15	20	2	11	13	7	26	33	66	0.33	0.21	0.70	0.56	0.92	0.85	1.00	0.95	0.40	0.25	1.00	0.77	0.22	0.15	0.40	0.79
cerebro	20	17	3	20	5	15	20	2	11	13	7	26	33	66	0.33	0.21	0.70	0.56	0.92	0.85	1.00	0.95	0.40	0.25	1.00	0.77	0.22	0.15	0.40	0.79
cerebro	comb	17	3	20	4	5	9	3	0	3	7	5	12	66	0.64	0.58	0.70	0.88	0.92	0.85	1.00	0.95	0.57	0.44	0.80	0.91	0.75	1.00	0.60	0.97
	Real	17			5			5			10																			

Anexos

archivo	vent	Correferencia						no_corref			Total				Metricas															
		directa			indirecta			indirecta			indirecta				anafora indirecta				Correferencia directa				Correferencia indirecta				no corref. Indirecta			
		Vf	Fp	Ob	Vf	Fp	Ob	Vp	Fp	Ob	Vf	Fp	Ob	det	F	P	R	E	F	P	R	E	F	P	R	E	F	P	R	E
io	1	7	0	7	1	1	2	1	1	2	2	2	4	56	0.40	0.50	0.33	0.89	0.47	1.00	0.30	0.71	0.67	0.50	1.00	0.98	0.29	0.50	0.20	0.91
io	2	12	0	12	1	2	3	1	3	4	2	5	7	56	0.31	0.29	0.33	0.84	0.69	1.00	0.52	0.80	0.50	0.33	1.00	0.96	0.22	0.25	0.20	0.88
io	3	12	0	12	1	3	4	2	3	5	3	6	9	56	0.40	0.33	0.50	0.84	0.69	1.00	0.52	0.80	0.40	0.25	1.00	0.95	0.40	0.40	0.40	0.89
io	4	16	0	16	1	3	4	4	5	9	5	8	13	56	0.53	0.38	0.83	0.84	0.82	1.00	0.70	0.88	0.40	0.25	1.00	0.95	0.57	0.44	0.80	0.89
io	5	16	0	16	1	5	6	4	3	7	5	8	13	56	0.53	0.38	0.83	0.84	0.82	1.00	0.70	0.88	0.29	0.17	1.00	0.91	0.67	0.57	0.80	0.93
io	6	17	0	17	1	6	7	3	3	6	4	9	13	56	0.42	0.31	0.67	0.80	0.85	1.00	0.74	0.89	0.25	0.14	1.00	0.89	0.55	0.50	0.60	0.91
io	7	17	0	17	1	10	11	3	2	5	4	12	16	56	0.36	0.25	0.67	0.75	0.85	1.00	0.74	0.89	0.17	0.09	1.00	0.82	0.60	0.60	0.60	0.93
io	8	17	0	17	1	10	11	3	3	6	4	13	17	56	0.35	0.24	0.67	0.73	0.85	1.00	0.74	0.89	0.17	0.09	1.00	0.82	0.55	0.50	0.60	0.91
io	9	18	0	18	1	9	10	3	3	6	4	12	16	56	0.36	0.25	0.67	0.75	0.88	1.00	0.78	0.91	0.18	0.10	1.00	0.84	0.55	0.50	0.60	0.91
io	10	19	0	19	1	9	10	3	3	6	4	12	16	56	0.36	0.25	0.67	0.75	0.90	1.00	0.83	0.93	0.18	0.10	1.00	0.84	0.55	0.50	0.60	0.91
io	11	20	0	20	1	9	10	3	4	7	4	13	17	56	0.35	0.24	0.67	0.73	0.93	1.00	0.87	0.95	0.18	0.10	1.00	0.84	0.50	0.43	0.60	0.89
io	12	20	0	20	1	9	10	4	4	8	5	13	18	56	0.42	0.28	0.83	0.75	0.93	1.00	0.87	0.95	0.18	0.10	1.00	0.84	0.62	0.50	0.80	0.91
io	13	21	0	21	1	10	11	4	4	8	5	14	19	56	0.40	0.26	0.83	0.73	0.95	1.00	0.91	0.96	0.17	0.09	1.00	0.82	0.62	0.50	0.80	0.91
io	14	21	0	21	1	10	11	4	4	8	5	14	19	56	0.40	0.26	0.83	0.73	0.95	1.00	0.91	0.96	0.17	0.09	1.00	0.82	0.62	0.50	0.80	0.91
io	15	21	0	21	1	10	11	4	4	8	5	14	19	56	0.40	0.26	0.83	0.73	0.95	1.00	0.91	0.96	0.17	0.09	1.00	0.82	0.62	0.50	0.80	0.91
io	16	22	0	22	0	10	10	4	4	8	4	14	18	56	0.33	0.22	0.67	0.71	0.98	1.00	0.96	0.98	0.00	0.00	0.00	0.80	0.62	0.50	0.80	0.91
io	17	22	0	22	0	10	10	4	4	8	4	14	18	56	0.33	0.22	0.67	0.71	0.98	1.00	0.96	0.98	0.00	0.00	0.00	0.80	0.62	0.50	0.80	0.91
io	18	23	0	23	0	10	10	4	3	7	4	13	17	56	0.35	0.24	0.67	0.73	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.80	0.67	0.57	0.80	0.93
io	19	23	0	23	0	12	12	4	4	8	4	16	20	56	0.31	0.20	0.67	0.68	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.77	0.62	0.50	0.80	0.91
io	20	23	0	23	0	13	13	4	4	8	4	17	21	56	0.30	0.19	0.67	0.66	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.75	0.62	0.50	0.80	0.91
io	comb	23	0	23	0	1	1	3	5	8	3	6	9	56	0.40	0.33	0.50	0.84	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.96	0.46	0.38	0.60	0.88
	Real	23	0		1			5		0	6			0																

## **Anexo k: Cómo correr el programa de demostración e interpretar su archivo de salida**

Una versión de demostración del software desarrollado en esta tesis se anexa a la misma en forma de un CD-ROM. El programa corre dentro de la ventana de DOS en el sistema Windows. En el disco aparece el archivo `leame.txt` que contiene las instrucciones para su uso; aquí se repite lo más importante de ese archivo y se da la información para la interpretación de los resultados de la corrida del programa.

Para utilizar la demostración es necesario hacer los siguientes pasos:

1. Copiar todos los archivos del subdirectorio `demo` del CD-ROM a un subdirectorio en el disco duro de su PC, verificando que **no** queden con el atributo de sólo lectura (read only) por ejemplo, `C:\demo` (porque el sistema necesita crear y almacenar en archivos al desarrollar su proceso).

2. Abrir una ventana MS-DOS.

3. Ir al directorio donde colocó los archivos del CD, usando el comando `cd`, por ejemplo:

```
C:\> cd demo
```

4. El disco ya incluye textos de ejemplo, `demo\f_txt\a12c.txt`. Para correr el programa usando este archivo como entrada, ejecutar:

```
C:\demo> demo a12c
```

5. Si quiere procesar su propio archivo, por ejemplo, `su_archivo.txt`:

- a. Colocarlo en el subdirectorio `f_txt` para, que como en nuestro ejemplo, aparezca como

```
C:\demo\f_txt\su_archivo.txt.
```

- b. Ejecutar:

```
C:\demo> demo su_archivo
```

nótese que el nombre aparece sin la extensión del archivo (`.txt`).

6. Observar que el programa termine con el mensaje “Terminado OK”.

7. Ver el archivo de salida en cualquier visualizador de Internet actualizado, de preferencia en Internet Explorer de version mayor a 5.0. El archivo de salida aparece en el directorio `demo\res` y su nombre se compone del nombre del archivo original más un código que refleja los parámetros del programa. Para visualizar el archivo, se puede abrir de la ventana del visualizador o bien ejecutar desde la línea de comandos:

```
C:\demo> start res\a12c_F_9.htm
```

o en su caso

```
C:\demo> start res\su_archivo_F_9.htm
```

A continuación se explica cómo se interpretan los mensajes que el programa muestra durante su ejecución y más abajo, cómo se interpreta el archivo `.htm` de salida.

### ***Seguimiento de corrida del programa***

Al correr el programa podrá observarse la corrida en la pantalla, véase la figura 23. Los pasos indicados en la figura son los siguientes:

1. Procesando archivo. Este mensaje es de retroalimentación
2. Convierte el texto a tokens. Se transforma el archivo a una unidad léxica por línea. El resultado puede apreciarse en el archivo `temp\su_archivo.tt`
- 3 al 18. Convierte tokens a etiquetado. En la línea tres se llama al etiquetador TnT que imprime las acciones que está realizando, además de información estadística. El resultado puede apreciarse en el archivo `temp\su_archivo.tts`
19. Limpia comentarios de etiquetado. En este paso se elimina información de TnT como la identificación y estadística para obtener un archivo con etiquetado simple. El resultado puede apreciarse en el archivo `f_tnt\su_archivo.tts`
20. convierte tnt a htm. En este paso se llama al programa de detección de anáfora indirecta que despliega mensajes de acuerdo a las tareas que va realizando y termina en un tiempo máximo de dos minutos aproximadamente (con un archivo no mayor de 45 KB). El resultado puede apreciarse en el archivo `res\su_archivo_F_9.htm` (que se muestra en la figura 24). Se puede apreciar que el nombre del archivo se altera con los argumentos por omisión (la bandera y el tamaño de la ventana) que utiliza el programa `demo`. En este caso `F` = signo de punto y el número `9` indica el ancho de la ventana en número de puntos a encontrar en la búsqueda hacia atrás.

21 al 27 son pasos que internamente ejecuta el algoritmo. En el renglón 21 aparece el nombre del archivo interno derivado del nombre del archivo de la entrada, junto con las opciones de corrida: el valor de imprime (ver tabla 5), la letra que identifica el tipo de bandera y el valor del tamaño de la ventana según (ver capítulo 6). Después se dan mensajes al activarse diferentes módulos del programa. Finalmente, se da mensaje de la terminación normal del programa.

```

E:\Return-to-D-Gelbukh-Doc\Tesis\Raul\Version final\demo>demo io
1. Procesando archivo io
2. convierte texto a tokens
3. convierte tokens a etiquetado
4. TNT: Trigrams'n'Tags - Statistical Trigram Tagging - Version 2.2
5. (C) 1993 - 2000 Thorsten Brants, thorsten@coli.uni-sb.de
6. Reading n-grams talp.123 ..... (268 uni-, 5885 bi-,
7. 27273 trigrams)
8. Reading lexicon talp.lex ..... (17362 entries)
9. Building suffix trie ..... (46683 lowercase, 10867 uppercase)
10. Estimating lambdas ..... (done)
11. lambda1 = 1.238826e-01   lambda2 = 3.840267e-01   lambda3 = 4.920907e-01
12. lam_bi1 = 1.627426e-01   lam_bi2 = 8.372574e-01
13. suffix theta = 5.083906e-02
14. Setup: real 00:00:01.76
15. Tagging (381 tokens)
16. 79 (20.73%) unknown tokens
17.     1 recognized as cardinals/ordinals
18. avg. 5.70 tags/token, 1.38 tags/known token
19. Tagging: real 00:00:00.14 (2557 tok/sec)
20. limpia comentarios de etiquetado
21. convierte tnt a htm
22. Corriendo con archivo f_tnt\io.tts 4 F 9
23. Correferencia de Nombres propios
24. Marcado de Nombres comunes
25. Correferencias nombres comunes
26. Anafora indirecta
27. Impresion de marcado de nombres
    Imprime archivo html
    Terminado OK
E:\Return-to-D-Gelbukh-Doc\Tesis\Raul\Version final\demo>

```

Figura 23 Corrida del programa

### ***Interpretación del archivo de salida***

El archivo de salida, mostrado en la figura 24, es del tipo de hipertexto (.htm) y muestra el texto de entrada con el siguiente marcado agregado: los nombres (propios y comunes) los nombres comunes aparecen en azul; las unidades léxicas (*tokens*) se identifican con números secuenciales que aparecen como subíndices entre paréntesis (sólo se muestran los números para los nombres, aunque se cuentan todas las palabras).

La aportación principal del presente trabajo es la identificación de las relaciones de correferencia y de anáfora indirecta entre palabras, las que forman cadenas de palabras una refiriéndose a otra. Estas cadenas encontradas por el programa se muestran a través de las referencias cruzadas en la pantalla de salida. Así, si el subíndice está subrayado indica que la unidad léxica es una correferencia o anáfora indirecta, según la información adicional que se puede obtener con un navegador colocando el puntero del ratón sobre el subíndice de la palabra; en este momento se desplegará en una ventanita amarilla la abreviatura del tipo de correferencia o anáfora indirecta (véase anexo I) seguido del número de palabra con la que existe dicha relación. Al hacer clic con el ratón en ese subíndice de la palabra, el apuntador se desplazará al elemento referenciado. De esta manera se queda demostrada la habilidad del programa a encontrar las relaciones de correferencia y anáfora indirecta en un texto libre (no preparado) en español.

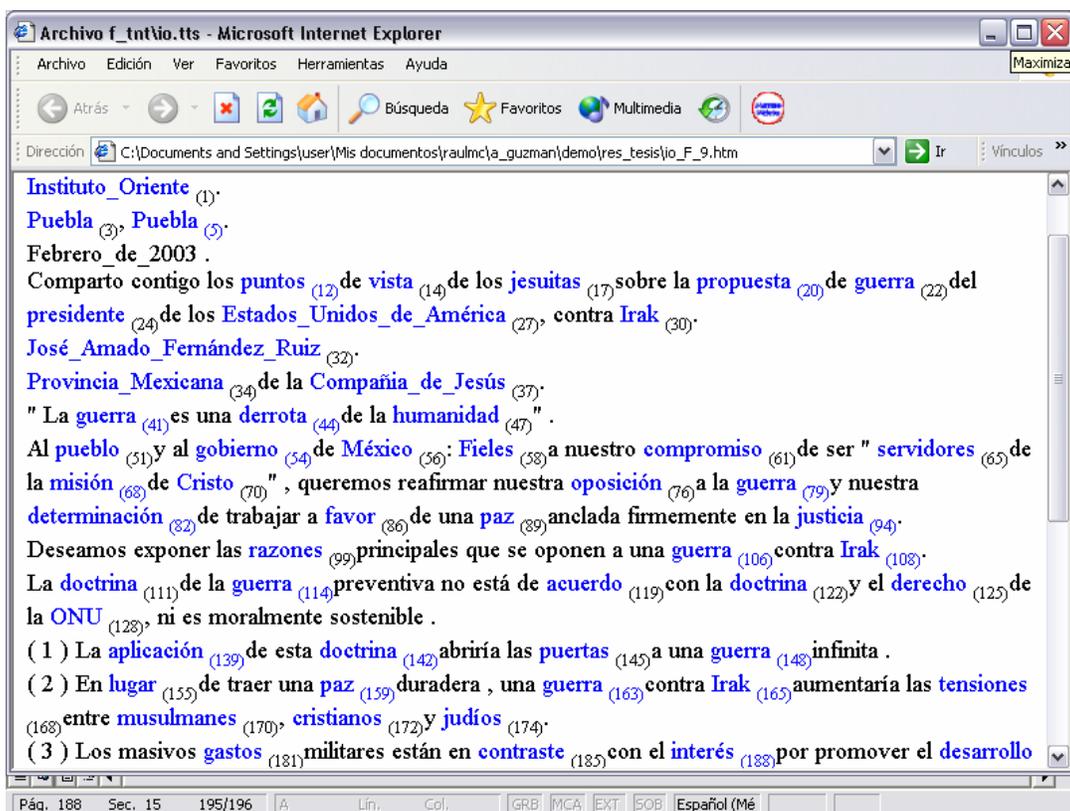


Figura 24 Archivo de salida del programa