

ÍNDIX

AGRAÏMENTS	vii
RESUM	ix
CAPÍTOL 1: INTRODUCCIÓ	1
1.1 Antecedents i motivació	1
1.2 Objectiu i parts de la tesi	2
CAPÍTOL 2: LITERATURA I ESTADÍSTICA	3
2.1 Estilometria i autoria	4
2.2 Caracterització de l'estil literari	6
2.2.1 Llargada de paraula i frase	6
2.2.2 Vocabulari i paraules eina	9
2.2.3 Distribució de les parts de l'oració	10
2.2.4 Diversitat i Riquesa de Vocabulari	11
2.2.5 Distribucions de Vocabulari	12
2.2.6 Dependència del llenguatge	13
2.2.7 Tècniques Multivariants i de Classificació	14
CAPÍTOL 3: LA QÜESTIÓ DE L'AUTORIA DEL TIRANT LO BLANC1.....	7
3.1 Hipòtesis a favor de la doble autoria Martorell/Galba	18
3.2 Hipòtesis a favor de l'autoria única de Martorell	21
3.3 Joan Roís de Corella i el Tirant	23
CAPÍTOL 4: DESCRIPCIÓ DE LA BASE DE DADES.....	25
4.1 Quantificació de l'estil del Tirant : passar del text a xifres	26
4.1.1 Quantificació de l'estil	26
4.1.2 El Tirant en Dades	26
4.2 Subdivisió del Tirant en trossos de text	27
4.2.1 Subdivisió del Tirant en capítols	27
4.2.2 Subdivisió en blocs de 1000 paraules consecutives	28
4.3 Descripció de la base de dades	28
4.3.1 Unitats d'estadística textual analitzades	29
4.3.1.1 Llargada de bloc	29
4.3.1.2 Llargada de Paraula	30
4.3.1.3 Llargada de frase i de capítol	30

4.3.1.4	Freqüència d'ús de lletres.....	31
4.3.1.5	Freqüència d'ús de paraules més freqüents.....	31
4.3.1.6	Mesures de Riquesa i Diversitat de vocabulari.....	31
4.3.2	Distribució de Vocabulari.....	32
4.3.3	Freqüència d'ús de totes les paraules.....	32
CAPÍTOL 5: EINES D'ANÀLISI EXPLORATÒRIA DE DADES		35
5.1	Gràfics de Control.....	35
5.2	Gràfics de Control per dades univariants.....	36
5.2.1	Gràfics de Control per variables contínues.....	37
5.2.2	Gràfics de Control per dades categòriques.....	38
5.2.3	Gràfics Cusum.....	40
5.3	Gràfics de Control per dades Multivariants.....	41
5.3.1	Gràfics de control T2 per a dades Normals.....	41
5.3.2	Gràfics de Control per dades Multinomials.....	42
5.4	Anàlisi de Correspondències.....	43
CAPÍTOL 6: EL PROBLEMA DEL PUNT DE CANVI		49
6.1	El problema del punt de canvi.....	49
6.1.1	Seqüència de Normals.....	50
6.1.2	Seqüència de binomials.....	51
6.1.3	Seqüència de multivariants.....	51
6.1.4	Més d'un punt de canvi.....	52
6.2	Punt de canvi per seqüència de Normals.....	52
6.2.1	Comparació de mitjanes.....	53
6.2.2	Estimació del punt de canvi via regressió lineal.....	54
6.3	Punt de canvi per seqüència de binomials.....	56
6.3.1	Comparació de dues proporcions.....	56
6.3.2	Estimació del punt de canvi via regressió logística.....	58
6.4	Punt de canvi en seqüència de multinomials.....	59
6.4.1	Estimació del punt de canvi per regressió politòmica.....	60
6.4.2	Aproximació univariant al cas multinomial.....	62
6.5	Estimació de més d'un punt de canvi.....	63
6.5.1	Comparació de més de dues mitjanes o proporcions.....	63
6.5.2	Estimació de més d'un punt de canvi per seqüència de binomials.....	64
6.6	Conclusions.....	66
CAPÍTOL 7: ANÀLISI CLUSTER EN UNA POBLACIÓ MULTINOMIAL		67
7.1	Mesures de dissimilaritat entre objectes o variables.....	68
7.2	Tècniques Cluster basades en particions.....	69
7.3	Tècniques Cluster Jeràrquiques.....	70
7.4	Cluster no jeràrquic per files d'una taula de contingència.....	71
7.4.1	Algorisme basat en la Deviança.....	72
7.4.2	Algorisme basat en la distància χ^2	73
7.4.3	Relació dels algorismes amb l'anàlisi de correspondències i discussió de resultats.....	77

CAPÍTOL 8: LLARGADA DE PARAULA.....	79
8.1 Introducció.....	79
8.2 Estudi de la llargada de paraula per capítol.....	80
8.2.1 Relació de la distribució de llargades amb Ni.....	82
8.2.2 Gràfics de Control per la llargada de paraula.....	84
8.2.3 Anàlisi de correspondències per la llargada de paraula.....	86
8.2.4 Estimació del punt de canvi.....	88
8.2.4.1 Punt de canvi en seqüències Normals.....	89
8.2.4.2 Punt de canvi en seqüències binomials.....	90
8.2.4.3 Punt de canvi en seqüència multinomial; model politòmic.....	92
8.2.4.4 Punt de canvi en seqüència multinomial; combinació de binomials.....	93
8.2.4.5 Conclusions.....	95
8.2.5 Anàlisi cluster de les files d'una taula de contingència.....	96
8.2.5.1 Anàlisi cluster basada en la distància χ^2	96
8.2.5.2 Anàlisi cluster basada en la deviança.....	99
8.3 Estudi de la llargada de paraula per blocs.....	100
8.3.1 Notació i forma de les dades.....	100
8.3.2 Anàlisi descriptiva de les dades.....	102
8.3.3 Gràfics de Control per la llargada de paraula.....	102
8.3.4 Anàlisi de correspondències per la llargada de paraula.....	106
8.3.5 Estimació del punt de canvi.....	108
8.3.5.1 Punt de canvi en seqüències Normals.....	108
8.3.5.2 Punt de canvi en les seqüències de binomials.....	110
8.3.5.3 Punt de canvi en seqüència de multinomial; model politòmic.....	112
8.3.5.4 Punt de canvi en seqüència multinomial; combinació de binomials.....	113
8.3.5.5 Conclusions.....	115
8.3.6 Anàlisi cluster de les files d'una taula de contingència.....	116
8.3.6.1 Anàlisi cluster basada en la distància χ^2	116
8.3.6.2 Anàlisi cluster basada en la deviança.....	118
Annex 8.1: Assignació de Capítols a Clusters per llargada de paraula.....	121
Annex 8.1: Assignació de Blocs a Clusters per llargada de paraula.....	123
CAPÍTOL 9: LLARGADA DE FRASE.....	125
9.1 Estudi de la llargada de frase per Capítol.....	125
9.1.1 Notació i forma de les dades.....	126
9.2 Estudi de la relació de la llargada de frase amb Ni.....	128
9.3 Anàlisi descriptiva de les dades.....	129
9.3.1 Gràfics de Control per la Llargada de Frase.....	130
9.3.2 Anàlisi de Correspondències.....	133
9.4 Estimació del punt de canvi.....	136
9.4.1 Punt de canvi en seqüències Normals.....	136
9.4.2 Punt de canvi en seqüència multinomial.....	138
9.5 Anàlisi cluster del les files de la taula de contingència.....	140
9.5.1 Anàlisi cluster basada en la distància χ^2	140
9.5.2 Anàlisi cluster basada en la deviança.....	141
9.6 Conclusions.....	142

CAPÍTOL 10: LLARGADA DE CAPÍTOL.....	143
10.1 Descripció de les dades	143
10.2 Estimació del punt de canvi	147
10.3 Conclusió	148
CAPÍTOL 11: FREQUÈNCIA D'ÚS DE LES LLETRES.....	149
11.1 Introducció.....	149
11.2 Notació i forma de les dades	150
11.3 Relació de la proporció de lletres amb Ni	152
11.4 Evolució temporal de les seqüències multinomials.....	153
11.5 Estimació del punt de canvi en seqüències multinomials.....	155
11.5.1 Punt de canvi en seqüències Normals	155
11.5.2 Punt de canvi en seqüència binomial	156
11.5.3 Punt de canvi en les seqüències multinomials; model politòmic	157
11.6 Conclusions	158
CAPÍTOL 12: ÚS DE PARAULES EN EL <i>TIRANT</i>.....	159
12.1 Introducció.....	159
12.2 Estudi de les paraules eina per Capítols	160
12.2.1 Notació i forma de les dades.....	160
12.2.2 Relació de la proporció de paraules eina amb Ni	162
12.2.3 Gràfics de Control per les proporcions de paraules.....	162
12.2.4 Anàlisi de correspondències per l'ús de paraules	168
12.2.5 Estimació del punt de canvi	170
12.2.5.1 Punt de canvi en seqüències Normals	171
12.2.5.2 Punt de canvi en seqüències binomials	172
12.2.5.3 Punt de canvi en seqüències multinomials; model politòmic	176
12.2.5.4 Paraules discriminants.....	177
12.2.5.5 Conclusions.....	178
12.2.6 Anàlisi cluster de les files de la taula de contingència	179
12.2.6.1 Anàlisi cluster basada en la distància χ^2	179
12.2.6.2 Anàlisi cluster basada en la deviança.....	180
12.2.6.3 Validació dels resultats obtinguts en l'anàlisi cluster	180
12.3 Estudi de les paraules eina per blocs	181
12.3.1 Notació i forma de les dades	181
12.3.2 Anàlisi descriptiva univariant de les dades	183
12.3.3 Gràfics de Control per les proporcions de paraules	184
12.3.4 Anàlisi de Correspondències l'ús de paraules.....	190
12.3.5 Estimació del punt de canvi.....	192
12.3.5.1 Punt de canvi en seqüències Normals	192
12.3.5.2 Punt de canvi en seqüències binomials	194
12.3.5.3 Punt de canvi en seqüències multinomials; model politòmic	198
12.3.5.4 Paraules discriminants.....	199
12.3.5.5 Conclusions.....	200
12.3.6 Anàlisi cluster de les files de la taula de contingència	200
12.3.6.1 Anàlisi cluster basada en la distància χ^2	201
12.3.6.2 Anàlisi cluster basada en la deviança.....	202
12.3.6.3 Validació dels resultats obtinguts en l'anàlisi cluster	202

Annex 12.1: Llistat de les 104 paraules més abundants en el <i>Tirant</i>	205
Annex 12.2: Llistat de paraules discriminants	207
Annex 12.3: Assignació de capítols a Clusters per ús de paraules eina	209
CAPÍTOL 13: RIQUESA I DIVERSITAT	217
13.1 Índexs de Diversitat	217
13.2 Estudi dels Índexs per blocs	221
13.2.1 Anàlisi descriptiva univariant de les dades	222
13.2.2 Gràfics de Control univariants pels índexs de diversitat	222
13.2.3 Anàlisi descriptiva multivariant	227
13.2.3.1 Anàlisi en Components Principals	230
13.2.4 Gràfics de Control multivariants	232
13.2.5 Estimació dels punts de canvi	234
13.2.5.1 Estimació d'un punt de canvi	234
13.2.5.2 Estimació de dos punts de canvi	236
13.2.5.3 Conclusions	238
13.2.6 Anàlisi Cluster	239
13.3 Estudi dels Índexs per capítols	241
13.3.1 Dependència dels índexs respecte a la llargada de capítol N_i	241
13.3.2 Evolució temporal de D_i i M_i	243
13.3.3 Estimació del punt de canvi	244
13.3.3.1 Estimació del(s) punt(s) de canvi per a D_i	244
13.3.3.2 Estimació dels punts de canvi per a V_i , V_{1i} , H_{si} i B_i	246
13.3.3.3 Conclusions	247
13.3.4 Anàlisi Cluster	247
Annex 13.1: Assignació de blocs a clusters per riquesa i diversitat	251
Annex 13.1: Assignació de capítols a clusters per riquesa i diversitat	255
CAPÍTOL 14: CONCLUSIONS	259
14.1 Punt de canvi per capítols	259
14.2 Assignació de capítols a un dels dos estils	260
14.2.1 Resum dels resultats de les anàlisi clusters individuals	261
14.2.2 Anàlisi clusters individuals imposant grandària 100	264
14.2.3 Cluster amb llargada de paraula, ús de paraules i diversitat	265
14.3 Conclusions	266
Annex 14.1: Resum de les assignacions de capítols a clusters	267
Annex 14.2: Assignacions de capítols a clusters fixant la grandària dels clusters	277
Annex 14.3: Assignació de capítols mitjançant una anàlisi a clusters per variables contínues	287
CAPÍTOL 15: LÍNIES FUTURES DE RECERCA	291
15.1 Ajust de distribucions de vocabulari	291
15.2 Agrupació de les distribucions de vocabulari	295
15.3 Dependència del llenguatge	296
BIBLIOGRAFIA	297