UNIVERSITY OF MURCIA

International Journal of English Studies

IJES

http://revistas.um.es/ijes

Testing Delta on the *Disputed Federalist Papers*

ANTONIO MIRANDA-GARCÍA & JAVIER CALLE-MARTÍN
Universidad de Málaga

Received: 9 June 2011 / Accepted: 27 September 2011

ABSTRACT
The Federalist Papers stand out as an excellent proving ground in the field of authorship attribution, being nowadays considered a breaking issue in literary detection. The crucial point of the *Federalist Papers* is the set of the *Disputed Papers*, twelve articles traditionally attributed either to Alexander Hamilton or James Madison. This authorial obscurity, together with the existence of undisputed samples, surely explains the proliferation of studies trying to spot the hand responsible for the *Disputed Papers*, particularly throughout the second half of the 20th century, both with traditional and non-traditional approaches. Since the publication of Mosteller and Wallace's masterpiece, there has been a consensus as to consider them exclusively Madisonian (Mosteller & Wallace 1963: 300; 1964: 16). Notwithstanding this incessant activity on the *Federalist Papers* as a test probe for authorial purposes, the use of Burrows' Delta is still deemed a desideratum in the field, a technique proposing that the salient features which characterize an author's style can be obtained from the hierarchy of the most common function words (Burrows 2002: 267-87; 2003: 5-32). The present paper then proposes the testing of Burrows' model in a twofold version: a) modified Delta; and b) simplified Delta. The results come to corroborate the lexical differences between Hamilton and Madison, a fact allowing us to validate the hypothesis of the Madisonian composition of the *Disputed Papers*, exception being made of Paper 55.

KEYWORDS:
Authorship Attribution, Lexical richness, Standard Delta, *Federalist Papers*, Simplified Delta, Stylometry, Zipf *Z*.

RESUMEN
Escritos durante los años 1787 y 1788 por James Madison, Alexander Hamilton y John Jay, *The Federalist Papers* contienen un total de 85 artículos periodísticos con un fin propagandístico a favor de la Constitución americana. Dado el carácter anónimo de los denominados *Disputed Papers*, 12 artículos escritos por Alexander Hamilton o James Madison, éstos han sido tradicionalmente fuente primaria fundamental para los estudios de atribución de autoría, especialmente desde la segunda mitad del siglo XX, utilizando enfoques tanto tradicionales como no tradicionales. A raíz de la publicación de la obra seminal de Mosteller y Wallace, los *Disputed Papers* se han considerado desde entonces responsabilidad de James Madison (1963: 300; 1964: 16). A pesar de la incesante actividad investigadora en torno a los *Federalist Papers* como piedra de toque para los estudios de autoría, éstos no han sido aún analizados desde la perspectiva de Delta, una técnica acuñada inicialmente por John Burrows que propone la caracterización de los rasgos estilísticos de un determinado autor en función de la jerarquía de las palabras funcionales de mayor frecuencia de aparición (Burrows, 2002: 267-87; 2003: 5-32). El presente estudio, por tanto, aplica el modelo de Burrows a *The Federalist Papers* desde una doble perspectiva, con el uso de Delta tanto en su variante modificada como simplificada. Los resultados validan la hipótesis de partida apuntando a una probable autoría de James Madison de los *Disputed Papers*, con la única excepción del artículo 55.

PALABRAS CLAVE:
Atribución de autoría, Delta, Delta simplificado, Estilometría, *Federalist Papers*, Riqueza léxica, Zipf Z.

Address for correspondence: Antonio Miranda-García. Departamento de Filología Inglesa, Francesa y Alemana, Campus de *Teatinos s/n* 29071, Málaga, Spain. Tel: 00 34 951952317; E-mail: amiranda@uma.es

Print ISSN: 1578-7044; Online ISSN: 1989-6131

1. INTRODUCTION

The present contribution has been conceived with the following objectives, i.e. the authorship attribution of the *Disputed Federalist Papers* and the testing of Burrows' Delta methodology on them. The *Federalist Papers* consist of 85 anonymous journalistic contributions written as propaganda to indoctrinate the New York citizens in favour of the American Constitution. Of these, 51 papers are by Alexander Hamilton, 14 by James Madison, 5 by John Jay whilst 3 are the result of the joint contribution of Hamilton and Madison. However, the moot point is the set of the *Disputed Papers*, 12 pieces claimed either by Hamilton or by Madison, thereby becoming a recurrent topic in the field of authorship attribution.

The relevant literature is abundant, both from historical and authorial perspectives (Adair, 1944a: 98-122; 1944b: 235-64). Mosteller and Wallace have reviewed the historical approaches to the *Federalist Papers* in a monograph that has eventually become a masterpiece for reference in the field (Mosteller and Wallace, 1964: 2-6). A number of seminal papers have been thenceforth published by scholars and statisticians attempting to find the Holy Grail with new statistical models, many of which come to validate Mosteller and Wallace's conclusions about the Madisonian composition of the *Disputed Papers* (Mosteller and Wallace, 1963: 306; 1964: 263-65), though some doubts are cast on Paper 55, deemed more Hamiltonian (Mosteller and Wallace, 1963: 286; Tweedie, Singh and Holmes 1996: 6). Most of these studies are based, for convenience and opportunity, on lexical variables on the assumption that the words are the most accessible items for an objective analysis with the least human manipulation. Accordingly, they have been approached from different perspectives, i.e. word length, sentence length, verb-adjective ratio, punctuation style, type-token ratio, function words, etc. (Miranda-García and Calle-Martín, 2008: 155-57).

The *Federalist Papers* have not escaped the attention of the new approaches stemming from computer-assisted research, i.e. artificial neural networks (Kjell, 1994: 119-24; Tweedie, Singh and Holmes 1996: 1-10), supportive vector machines (Diederich et al., 2003: 109-23; Fung 2003: 1-8), genetic algorithms (Holmes and Forsyth, 1995: 111-27) or linear discriminant analysis (Baayen et al., 2002: 1-7), to mention some of the most innovative. However, as far as we have been able to investigate and notwithstanding Diederich's words on the topic (Diederich et al., 2003: 111), the application of Burrows' Delta to the *Federalist Papers* is still a desideratum.

Delta was first presented at the 2001 ACH-ACLLC in New York (Burrows, 2003) and a saga of publications has thereafter proliferated assuming that the lexical salient features that characterize an author's style can be obtained from the rating of a hierarchy of the most common words (particularly function words), if compared with those in a model corpus. Like other methods in computational stylistics, Delta seeks to assess these numerical differences in word frequency as an authorship test designed "to complement *principal component analysis* (p.c.a) […] and to consolidate it in the role for which it is best suited, in the middle stages of

Print ISSN: 1578-7044; Online ISSN: 1989-6131

the game" (Burrows, 2003: 7). In this fashion, Delta comes to remedy the drawback that *cluster analysis* (Craig, 1999: 103-13) presents in open games, i.e. "the detailed evidence remains opaque and two rather similar specimens can turn away from each other in early iterations and end up much more widely separated than they should" (Burrows, 2003: 5).

Since its inception, Delta has also been tested to evaluate its potential and reliability. Even though there is a general consensus as to its effectiveness, a set of modifications and improvements has been proposed after due experimentation (van Dalen-Oskam and van Zundert, 2007: 1-18; Diederich et al., 2003: 110; Hoover, 2004a: 453-75; 2004b: 477-95; Miranda-García and Calle-Martín, 2007: 49-66). In this line, the present paper measures the potential of Burrows' technique as applied to the *Federalist Papers* with a twofold version: (a) modified Delta; and (b) simplified Delta, thus filling a void in the relevant literature. The singularity does not come from the results, known beforehand thanks to the previous insights into the topic, both from traditional and non-traditional approaches, but from the methodology adopted. Therefore, our immediate objective is the re-attribution of the *Disputed Papers* to Madison with the use of the simplest statistics of Delta. The analysis has been accomplished from a lexical standpoint, taking the raw word as the unit. Burrows' Delta has been accordingly adopted as the rationale, summarized in "add them up and average them out" (Burrows, 2003: 11), with the difference that no predetermined distinction is made between content and function words[1].

This study has been accordingly organised into four different sections. The second deals with the methodological procedure followed; the third accounts for modified Delta and its results when applied to the *Federalist Papers*; the fourth section presents the simplified version of Delta followed by an evaluation of the results; our conclusions are summarized in section 5.

2. METHODOLOGY

In light of these premises, the main task is to compare each paper in the test corpus against the members of the main corpus (authorial corpus), which contains articles by (a) Hamilton, (b) Madison, and (c) Jay along with Hamilton and Madison's *Joint Papers*. For the purpose of compiling both corpora for the experiment, the 85 *Federalist Papers* are downloaded from the Gutenberg Project website[2], and saved as .txt individual files which are stored in folders *H* (for Hamilton), *M* (for Madison), *JAY*, *D* (for *Disputed Papers*), and *JOI* (for *Joint Papers*)[3].

Next, all the files are processed with *WordSmith Tools 5.0* (Scott, 2009) to generate a frequency word-list[4], from which the 100 most common words are retrieved in terms of their representativeness (accounting for 59% of the words in the main corpus). Thus, we have considered Burrows' estimation that a 150-word list is unusually long and that the 60 most common words would suffice for the purpose (Burrows, 2003: 11). Table 1 reproduces the

Print ISSN: 1578-7044; Online ISSN: 1989-6131

frequency hierarchy of the 100 most common words, either function or content words (the latter in italics), which are in a ratio 3:1, the first content word *states* being ranked as the 28th. The last word in the hierarchy is *first* with 1.19 occurrences every thousand words. This rank sharply contrasts with Burrows' hierarchy wherein at least the 50 topmost common items are all function words. Although previous experiments have been carried out with function words (Ellegärd 1962; Miranda-García and Calle-Martín 2005: 49-66; Mosteller and Wallace, 1963: 275-309; 1964: 16-91), the present insight has been accomplished without this restriction.

Rank	*Word*	*Hits*	*Rank*	*Word*	*Hits*	*Rank*	*Word*	*Hits*	*Rank*	*Word*	*Hits*
1	the	17613	26	an	922	51	there	456	76	were	296
2	of	11703	27	they	921	52	*constitution*	454	77	*great*	289
3	to	7013	28	*states*	841	53	these	454	78	*authority*	286
4	and	5043	29	*government*	825	54	must	444	79	had	284
5	in	4403	30	may	804	55	who	431	80	*public*	282
6	a	3959	31	been	788	56	such	412	81	some	273
7	be	3810	32	*state*	779	57	so	400	82	ought	269
9	it	2524	34	but	685	59	upon	385	84	shall	255
10	is	2168	35	other	644	60	I	367	85	*powers*	254
11	which	2048	36	its	642	61	his	365	86	*time*	248
12	by	1735	37	*people*	607	62	*union*	361	87	between	246
13	as	1706	38	*power*	603	63	should	355	88	*executive*	244
14	this	1392	39	has	590	64	every	350	89	each	239
15	would	1274	40	more	571	65	*same*	343	90	could	238
16	have	1252	41	at	569	66	*national*	340	91	*united*	238
17	will	1245	42	if	561	67	against	328	92	*men*	237
18	or	1228	43	than	545	68	was	324	93	what	231
19	for	1226	44	them	544	69	might	322	94	less	220
20	not	1201	45	one	543	70	*federal*	320	95	*part*	226
21	their	1090	46	any	542	71	*new*	313	96	*body*	224
22	with	1040	47	no	496	72	under	308	97	*members*	224
23	from	1039	48	those	481	73	our	307	98	*number*	221
24	are	1019	49	can	465	74	into	299	99	us	221
25	on	939	50	we	459	75	only	297	100	first	220

Table 1. Frequency hierarchy for the Federalist corpus

The third step is the adoption of some preliminary measures as for the range and composition of the corpora. The main corpus must comprise Papers by *H* and *M* since they are the disputing authors, though those by *JAY* and *JOI* are also included for comparison and illustrative purposes. The test corpus, in turn, must necessarily include the *Disputed Papers*. However, in our first experiment the test corpus also contains some Papers by *H* and *M* to check whether they are assigned to their right counterparts in the main corpus. This modification requires that no paper can be included in both corpora. To accomplish this and to minimize the effect of text-length dependency, the *H* corpus is randomly partitioned into three blocks of a similar length to *M* corpus (N=41005), called for short *H-I* (N=37162), *H-II* (N= 38672) and *H-III* (N= 35658), and their lexical richness is measured with Zipf's *Z* (Tweedie and Baayen, 1998: 331) just in case remarkable deviations could affect the final result.[5] After

Print ISSN: 1578-7044; Online ISSN: 1989-6131

discarding this possibility, the main corpus comprises *H-II* elected as a representative partition for *H*, *H-IV* (*N*= 11335) constituted by three Papers selected from *H-I* (*H8*, *H28*, and *H68)* and three from *H-III (H21*, *H67*, and *H85)*, *M-III* (*N*= 22290) formed with eight *M* Papers (*M10*, *M14*, *M37*, *M41*, *M42*, *M43*, *M47*, and *M48*), along with *JOI* and *JAY*. The test corpus, on the other hand, comprises three *Papers* from *H-I* (*H22*, *H34* and *H80)*, three from *H-III* (*H21*, *H67*, and *H85*), six from *M* (*M38*, *M39*, *M40*, *M44*, *M45*, and *M46*), and the twelve *Disputed*, as summarised in Table 2.

Main corpus (authorial corpus)	*Test corpus*
	Modified Delta
H-II, H-IV (H8, H28, H68, H21, H67, H85) *M-III (M10, M14, M37, M41, M42, M43, M47, M48), JOI, JAY*	*H22, H34, H80, H21, H67, H85* *M38, M39, M40, M44, M45, M46* *D49, D50, D51, D52, D53, D54, D55, D56, D57, D58, D62, D63*
	Simplified Delta
H, M, JAY, JOI	*D49, D50, D51, D52, D53, D54, D55, D56, D57, D58, D62, D63*

Table 2. Composition of the corpora

In the fourth stage, the individual files are the input for *WordSmith Tools 5.0* to obtain their corresponding frequency lists. A specific programming routine facilitates the arrangement of the items in accordance with the frequency hierarchy for the *Federalist Corpus*. The values for *H*, *M*, *JAY*, *JOI*, and *D49* (*Disputed 49*) in terms of their respective ranking order, the occurrences and the score (the percentage of hits with respect to *N*, i.e., the number of running words or tokens, reduced to three figures after the decimal point) are also shown in Table 2. The ranking slightly changes if *the*, which is the most common word, is left

Federalist Corpus		*H* (*N* = 111494)			*M* (*N* = 41005)			*JAY* (*N* = 8430)			*JOI* (*N* = 5658)			*D49* (*N* = 1656)		
Rank	*Word*	*Rank*	*Hits*	*%*	*Rank*	*Hits*	*%*	*Rank*	*Hits*	*%*	*Rank*	*Hits*	*%*	*Rank*	*Hits*	*%*
1	*the*	1	10300	9.238	1	4115	10.035	1	526	6.240	1	579	10.233	1	178	10.749
2	*of*	2	7214	6.470	2	2451	5.977	3	369	4.377	2	339	5.992	2	101	6.099
3	*to*	3	4537	4.069	3	1322	3.224	4	293	3.476	4	154	2.722	3	58	3.502
4	*and*	5	2663	2.388	4	1210	2.951	2	408	4.840	3	215	3.800	5	42	2.536
5	*in*	4	2778	2.492	5	864	2.107	5	164	1.945	5	121	2.139	6	34	2.053
6	*a*	6	2470	2.215	6	819	1.997	10	100	1.186	6	92	1.626	7	33	1.993
7	*be*	7	2285	2.049	7	806	1.966	6	160	1.898	24	27	.477	4	47	2.838
8	*that*	8	1682	1.509	8	573	1.397	7	150	1.779	13	43	.760	10	21	1.268
9	*it*	9	1523	1.366	9	521	1.271	8	138	1.637	8	57	1.007	8	26	1.570
10	*is*	10	1299	1.165	10	498	1.214	20	57	.676	14	41	.725	12	18	1.087
11	*which*	11	1224	1.098	12	451	1.100	21	56	.664	9	56	.990	11	20	1.208
12	*by*	15	849	.761	11	473	1.154	14	82	.973	7	74	1.308	17	15	.906
13	*as*	12	955	.857	13	393	.958	9	102	1.210	19	36	.636	16	15	.906
14	*this*	14	905	.812	19	263	.641	33	38	.451	20	36	.636	47	5	.302
15	*would*	13	926	.831	32	171	.417	16	68	.807	64	11	.194	9	22	1.329
16	*have*	16	775	.695	16	283	.690	23	53	.629	25	26	.460	41	5	.302
17	*will*	19	692	.621	15	284	.693	15	73	.866	125	6	.106	563	1	.060

Print ISSN: 1578-7044; Online ISSN: 1989-6131

Federalist Corpus		*H* (*N* = 111494)			*M* (*N* = 41005)			*JAY* (*N* = 8430)			*JOI* (*N* = 5658)			*D49* (*N* = 1656)		
Rank	*Word*	*Rank*	*Hits*	*%*	*Rank*	*Hits*	*%*	*Rank*	*Hits*	*%*	*Rank*	*Hits*	*%*	*Rank*	*Hits*	*%*
18	*or*	17	746	.669	22	243	.593	11	90	1.068	28	21	.371	25	10	.604
19	*for*	18	707	.634	17	272	.663	19	57	.676	22	30	.530	23	11	.664
20	*not*	20	689	.618	18	270	.658	17	65	.771	50	14	.247	19	13	.785
21	*their*	25	570	.511	21	247	.602	13	83	.985	10	54	.954	18	14	.845
22	*with*	23	603	.541	24	225	.549	18	60	.712	15	41	.725	24	11	.664
23	*from*	22	618	.554	27	202	.493	24	50	.593	12	44	.778	39	5	.302
24	*are*	24	570	.511	25	215	.524	22	53	.629	21	32	.566	15	15	.906
25	*on*	41	371	.333	14	313	.763	30	44	.522	16	39	.689	14	16	.966

Table 3. Reduced wordlist for Hamilton, Madison, JAY, JOI and Disputed 49

out. The words *of, to, and, in* are the following ones, but in a jumble order. The difference becomes greater with *will*, which is the 17th in the whole corpus, the 19th in *H*, the 15th in *M* and *JAY*, but the 125th in JOI and 563rd in *D49*. Likewise, *on* ranks 25th, 41st, 14th, 30th, 16th and 14th in the samples, respectively. Note that a difference in their sequencing affects the rate, which may result in a helpful clue to distinguish between two authors.

The fifth task which implies the actual computation for Delta involves the accomplishment of the following operations: (a) *score* each word in its own sub-corpus or Paper; (b) *z-score* each word (difference of the rate for each word to the mean of the main corpus, μ**,** divided by the standard deviation, σ) to represent the divergences from the mean of the main corpus yielding comparable figures for all the words in a hierarchy; (c) *find the difference* of the *z-score* for each word in the texts being compared as each one in the testing corpus is, in succession, measured against the ones constituting the main corpus; (d) *add* the absolute values of the differences and *average* them (dividing by 100) to obtain the Delta-score, which is formally defined by Burrows as "the mean of the absolute differences between the z-scores for a set of word-variables in an authorial text-group and the z-scores for the same set of word-variables in a target set" (2003: 13).

The figures in the score columns indicate the percentage that the occurrences of this word represent within its own corpus. The *score* for *the* is 9.622 in *H-II,* 9.610 in *M-III,* 8.290 *in H22* and 8.145 in *M38* whilst its mean and standard deviation in the main corpus are 9.175 and 1.475, respectively. Their respective *z-scores* (.303, .294, -.600, and -.698) result from subtracting the mean of the main corpus to their scores, and from dividing the difference by the standard deviation. The two rightmost columns hold the absolute values of the difference between the *z-scores*: .903 and .992 when *H22* and *M38*, respectively, are compared against *H-II* and *M-III*. The absolute differences are added and averaged out at the bottom. Thus, the resulting values are .996 and 1.326, respectively, as shown in Table 4.

Print ISSN: 1578-7044; Online ISSN: 1989-6131

	H-II	*M-III*	*Main corpus*		*H22*	*M38*	*H-II*	*M-III*	*H22*	*M38*	*H22_H-II*	*M38_M-III*
	scores		*μ*	*σ*	*scores*		*z-scores*				*difference*	
the	9.622	9.610	9.175	1.475	8.290	8.145	.303	.294	-.600	-.698	.903	.992
of	6.690	6.002	5.970	0.868	6.827	5.711	.829	.0376	.988	-.299	.159	.336
to	4.016	3.140	3.523	0.639	4.131	3.517	.770	-.599	.950	-.011	.180	.589
and	2.472	3.060	3.303	0.889	2.295	2.825	-.935	-.274	-1.134	-.538	.199	.263
in	2.410	2.032	2.151	0.219	2.467	1.863	1.181	-.543	1.440	-1.313	.260	.770
a	2.170	2.086	1.953	0.460	2.754	2.765	.469	.288	1.741	1.766	1.271	1.478
be	1.999	1.880	1.792	0.684	1.520	1.533	.303	.127	-.398	-.380	.699	.507
The data for the remaining items spread down in the complete list												
us	.109	.076	0.116	0.088	.115	.180	-.089	-.455	-.019	.724	.069	1.178
first	.098	.103	0.102	0.016	.143	.180	-.234	0.077	2.621	4.954	2.855	4.877
∑											99.576	132.622
Δ											0.996	1.326

Table 4. Reduced work-sheet for the calculation of Delta

The procedure in our second experiment is the same but the z-scoring is skipped on the assumption that the difference of the scores in the frequency hierarchy may provide remarkable differences once added and averaged out, in opposition to Burrows' opinion that "it must be derived from z-scores and not from original text-percentages" (Burrows, 2003: 13). This simplified version of Delta was inspired by Hoover, who anticipated that "the results are slightly less accurate" (Hoover, 2004b: 480).

3. MODIFIED DELTA: RESULTS AND ANALYSIS

The computations are carried out in an Excel spreadsheet and the results summarised in Table 5, wherein the first row contains the names of the members of the main corpus and the first column the names of the 24 Papers tested. Thus, each Paper in the test corpus is compared against those in the main corpus and the value for standard Delta is shown in the intersection: .996 and 2.202 in the first row indicate that *H-II* and *JOI* are the most and least unlike H22; contrariwise, when read horizontally, .996 in the first row and 2.386 in the 14th row indicate that *H22* and *d50* are, respectively, the least and most unlike *H-II*.

From the values of *Δ*, the corresponding *μ* and *σ* are calculated to obtain *Δz* (the difference of each *Δ* to *μ*, divided by *σ*). The Papers are then arranged in terms of increasing *Δ*. For space, the results of only twenty are shown in Table 6. In the particular case of *H22* (the three leftmost columns), *H-II* is ranked as the least unlike, followed by *H-IV*, *M-III*, and *JOI* is the most unlike, preceded by *JAY*. Likewise, in the study measuring the likeliness to *H80, M-III* is considered the least unlike, followed by *H-II*, *H-IV*, while *JAY* is the most unlike, preceded by *JOI*.

The results indicate that in the six studies comparing *H Papers* (*H22, H33, H80, H21, H67,* and *H85*) with the main corpus, *H-II* is ranked as the least unlike five times whereas *M-III* is thus ranked once (*H80*), as plotted in Figure 1. As expected, the two bottommost positions, which indicate the greatest unlikeness, are occupied by *JOI* (five times) and *JAY* (*once*).

Print ISSN: 1578-7044; Online ISSN: 1989-6131

	H-II	*H-IV*	*M-III*	*JAY*	*JOI*
H22	**0.996**	1.104	1.259	1.676	**2.202**
H34	1.907	2.010	2.154	2.429	2.390
H80	2.301	2.313	2.203	2.823	2.486
H21	1.414	1.539	1.644	2.003	1.933
H67	1.762	1.905	1.910	2.354	2.150
H85	1.414	1.579	1.559	2.149	2.062
M38	1.317	1.428	1.326	1.592	1.693
M39	1.876	1.957	1.697	2.065	1.889
M40	1.615	1.631	1.533	1.828	1.754
M44	1.858	1.842	1.702	2.438	2.074
M45	1.730	1.884	1.637	1.864	1.880
M46	1.702	1.770	1.681	1.729	1.928
D49	1.587	1.702	1.399	1.860	1.944
D50	**2.386**	2.350	2.342	2.558	2.336
D51	1.924	1.924	1.734	2.176	2.195
D52	1.774	1.773	1.760	2.202	2.266
D53	1.622	1.811	1.485	1.999	1.995
D54	1.832	1.992	1.744	2.172	2.057
D55	1.954	2.116	2.051	2.216	2.394
D56	1.760	1.910	1.737	1.951	1.910
D57	1.571	1.658	1.469	1.576	1.882
D58	1.673	1.681	1.511	2.045	1.985
D62	1.565	1.786	1.448	1.878	1.979
D63	1.249	1.362	1.250	1.765	1.689

Table 5. Values for modified Delta (**Δ)**

Similarly, when *M38, M39, M40, M44, M45,* and *M46* are studied, *M-III* is ranked as the least unlike five times whilst *H-II* is thus considered once (*M38*), as plotted in Figure 1. *JOI* and *JAY* are ranked as the most unlike.

Finally, in the Disputed Papers, *M-III* is top-ranked eleven times and only *H-II* is taken as the least unlike when compared to *D55*, as plotted in Figure 2. As in the case of *M Papers*, *JOI* and *JAY* are the most unlike.

In summary, 21 out of 24 papers are correctly attributed, which means that the experiment is successful in 87.5% of the cases, confirming the validity of Delta for authorial attribution. If the assignment of *D55* to *H* is taken as valid in line with the findings of other researchers (Mosteller and Wallace, 1963: 306; Tweedie, Singh and Holmes, 1996: 6), the rate of success will rise to 91.66%. However, a less successful rate (41.67%) is obtained when the members of the main corpus are the individual papers in *H-IV (H8, H28, H68, H21, H67, H85)* and *M-III (M10, M14, M37, M41, M42, M43, M47, M48, M58*), but it rises to 62.5% when they are grouped in three's (one in two's): *H-1 (H8, H28, H68), H-2 (H21, H67, H85), M-1(M10, M14, M37), M-2 (M41, M42, M43),* and *M-3 (M47, M48).* This points to the fact that the larger the text-length of the members in the main corpus, the greater the success of Delta in relating plausible likeness.

Print ISSN: 1578-7044; Online ISSN: 1989-6131

H22	*Δ*	*Δ z*	*H34*	*Δ*	*Δ z*	*H80*	*Δ*	*Δ z*	*H21*	*Δ*	*Δ z*	*H67*	*Δ*	*Δ z*
H-II	0.996	-0.913	*H-II*	1.907	-1.182	*M-III*	2.203	-0.909	*H-II*	1.414	-1.156	*H-II*	1.762	-1.084
H-IV	1.104	-0.695	*H-IV*	2.010	-0.732	*H-II*	2.301	-0.508	*H-IV*	1.539	-0.661	*H-IV*	1.905	-0.473
M-III	1.259	-0.380	*M-III*	2.154	-0.107	*H-IV*	2.313	-0.457	*M-III*	1.644	-0.247	*M-III*	1.910	-0.453
JAY	1.676	0.462	*JOI*	2.390	0.925	*JOI*	2.486	0.247	*JOI*	1.933	0.895	*JOI*	2.150	0.571
JOI	2.202	1.525	*JAY*	2.429	1.095	*JAY*	2.823	1.627	*JAY*	2.003	1.169	*JAY*	2.354	1.439
μ	1.447		*μ*	2.178		*μ*	2.425		*μ*	1.707		*μ*	2.016	
σ	0.495		*σ*	0.229		*σ*	0.245		*σ*	0.254		*σ*	0.235	
M38	*Δ*	*Δ z*	*M39*	*Δ*	*Δ z*	*M40*	*Δ*	*Δ z*	*M44*	*Δ*	*Δ z*	*M45*	*Δ*	*Δ z*
H-II	1.317	-.929	*M-III*	1.697	-1.488	*M-III*	1.533	-1.187	*M-III*	1.702	-.978	*M-III*	1.637	-1.465
M-III	1.326	-.871	*H-II*	1.876	-.154	*H-II*	1.615	-.483	*H-IV*	1.842	-.490	*H-II*	1.730	-.620
H-IV	1.428	-.261	*JOI*	1.889	-.056	*H-IV*	1.631	-.347	*H-II*	1.858	-.434	*JAY*	1.864	.585
JAY	1.592	.725	*H-IV*	1.957	.450	*JOI*	1.754	.693	*JOI*	2.074	.317	*JOI*	1.880	.732
JOI	1.693	1.336	*JAY*	2.065	1.248	*JAY*	1.828	1.325	*JAY*	2.438	1.586	*H-IV*	1.884	.768
μ	1.471		*μ*	1.897		*μ*	1.672		*μ*	1.983		*μ*	1.799	
σ	.166		*σ*	.135		*σ*	.118		*σ*	.287		*σ*	.110	
D49	*Δ*	*Δ z*	*D50*	*Δ*	*Δ z*	*D51*	*Δ*	*Δ z*	*D52*	*Δ*	*Δ z*	*D53*	*Δ*	*Δ z*
M-III	1.399	-1.378	*M-III*	2.342	-.558	*M-III*	1.734	-1.322	*M-III*	1.760	-.762	*M-III*	1.485	-1.307
H-II	1.587	-.514	*H-II*	2.386	-.091	*H-II*	1.924	-.343	*H-IV*	1.773	-.712	*H-II*	1.622	-.704
H-IV	1.702	.016	*JOI*	2.336	-.627	*H-IV*	1.924	-.342	*H-II*	1.774	-.708	*H-IV*	1.811	.124
JAY	1.860	.744	*H-IV*	2.350	-.474	*JAY*	2.176	.954	*JAY*	2.202	.966	*JOI*	1.995	.935
JOI	1.944	1.132	*JAY*	2.558	1.750	*JOI*	2.195	1.053	*JOI*	2.266	1.216	*JAY*	1.999	.952
μ	1.698		*μ*	2.394		*μ*	1.991		*μ*	1.955		*μ*	1.783	
σ	.217		*σ*	.093		*σ*	.194		*σ*	.255		*σ*	.228	
D54	*Δ*	*Δ z*	*D55*	*Δ*	*Δ z*	*D56*	*Δ*	*Δ z*	*D57*	*Δ*	*Δ z*	*D58*	*Δ*	*Δ z*
M-III	1.744	-1.252	*H-II*	1.954	-1.143	*M-III*	1.737	-1.193	*M-III*	1.469	-1.042	*M-III*	1.511	-1.046
H-II	1.832	-.740	*M-III*	2.051	-.564	*H-II*	1.760	-.957	*H-II*	1.571	-.387	*H-II*	1.673	-.220
H-IV	1.992	.188	*H-IV*	2.116	-.181	*H-IV*	1.910	.576	*JAY*	1.576	-.356	*JOI*	1.673	-.220
JOI	2.057	.567	*JAY*	2.216	.415	*JOI*	1.910	.575	*H-IV*	1.658	.170	*H-IV*	1.681	-.180
JAY	2.172	1.236	*JOI*	2.394	1.473	*JAY*	1.951	.999	*JOI*	1.882	1.615	*JAY*	2.045	1.666
μ	1.960		*μ*	2.146		*μ*	1.854		*μ*	1.631		*μ*	1.717	
σ	.172		*σ*	.168		*σ*	.098		*σ*	.155		*σ*	.197	

Table 6. Ranking of least unlikeness (standard **Δ)**

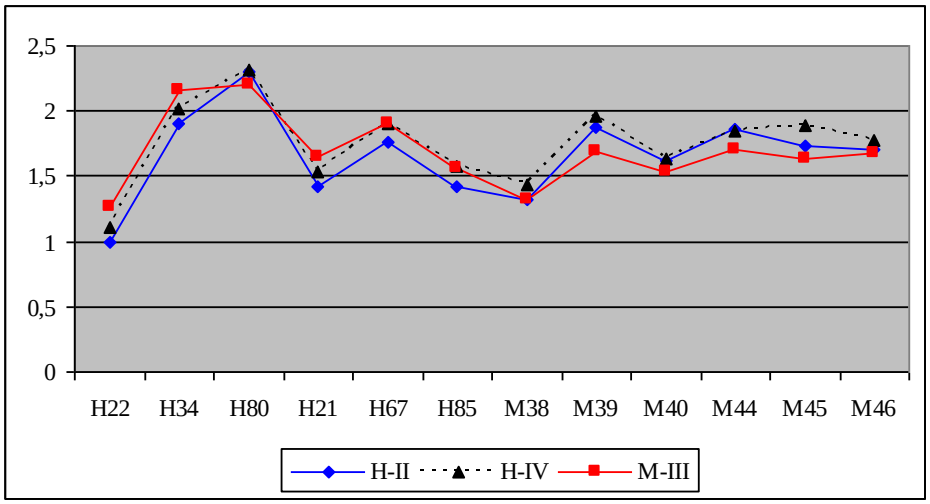

Figure 1. Modified Delta (H and M Papers)

Print ISSN: 1578-7044; Online ISSN: 1989-6131

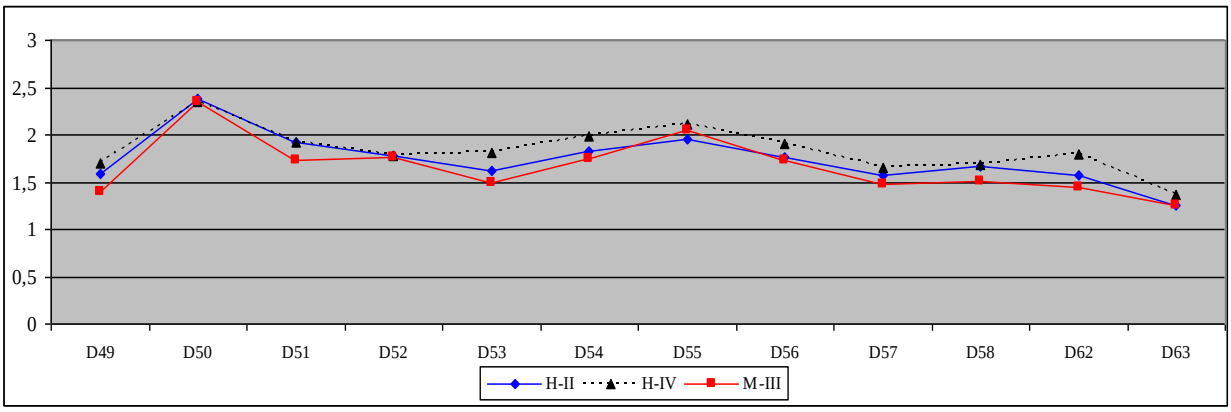

Figure 2. Modified Delta (Disputed Papers)

4. SIMPLIFIED DELTA: RESULTS AND ANALYSIS

As indicated in the section of methodology, in this experiment *H*, *M*, *JOY* and *JAY* constitute the main corpus whereas the test corpus is formed by the *Disputed* only, the frequency hierarchy being kept intact. In addition, the *z-scoring* step is skipped.

	H	***M***	***D49***	***D50***	***H-D49***	***H-D50***	***M-D49***	***M-D50***
the	9.238	10.035	10.749	8.976	1.511	.263	.713	1.060
of	6.470	5.977	6.099	5.258	.371	1.212	.122	.719
to	4.069	3.224	3.502	2.448	.567	1.621	.278	.776
and	2.388	2.951	2.536	2.992	.148	.603	.415	.041
in	2.492	2.107	2.053	2.539	.438	.047	.054	.431
a	2.215	1.997	1.993	1.269	.223	.946	.005	.728
be	2.049	1.966	2.838	2.811	.789	.761	.873	.845
that	1.509	1.397	1.268	1.088	.240	.421	.129	.309
it	1.366	1.271	1.570	1.451	.204	.085	.299	.180
is	1.165	1.214	1.087	.544	.078	.621	.128	.671
which	1.098	1.100	1.208	.816	.110	.282	.108	.284
by	.761	1.154	.906	.997	.144	.236	.248	.156
as	.857	.958	.906	.997	.049	.141	.053	.039
this	.812	.641	.302	.635	.510	.177	.339	.007
would	.831	.417	1.329	.997	.498	.167	.911	.580
have	.695	.690	.302	1.088	.393	.393	.388	.398
will	.621	.693	.060	.635	.164	.014	.632	.058
or	.669	.593	.604	.725	.065	.056	.011	.133
for	.634	.663	.664	.544	.030	.090	.001	.119
not	.618	.658	.785	.907	.167	.289	.127	.248
their	.511	.602	.845	1.179	.334	.667	.243	.576
with	.541	.549	.664	.453	.123	.088	.116	.095
from	.554	.493	.302	.725	.252	.171	.191	.233
are	.511	.524	.906	.363	.395	.149	.381	.162
on	.333	.763	.966	.997	.633	.665	.203	.234
The data for the remaining items spread down in the complete sheet								
us	.132	.078	.060	.000	.071	.132	0,018	.078
first	.104	.129	.060	.272	.044	.168	0,069	.143
∑					**18.379**	**24.484**	**15.736**	**23.406**
μ					**.184**	**.245**	**.157**	**.234**

Table 7. Reduced work-sheet for simplified Delta

Print ISSN: 1578-7044; Online ISSN: 1989-6131

In this line, the rates for each word in the frequency hierarchy are first calculated for all the files in the main and in the test corpus, as shown in the four leftmost columns of Table 7, holding an abbreviated list. The rates for *the* in *H*, *M*, *D49*, and *D50* are 9.238, 10.035, 10.749, and 8.976, respectively. Then, the difference is calculated between the rates for each word in the two archives being compared. The column *H-D49* contains the absolute value for the difference between *H* and *D49*, and the same holds for the columns to the right. The absolute difference for *the* is 1.511, .263, .713, and 1.060 when *H* and then *M* are compared with *D49* and *D50*. The same procedure is repeated with the values for all the files in two's. Next, the absolute values for the differences are added and averaged out as shown in Table 8. The simplified Delta is .184 for *H-D49*, .245 for *H-D50*, .157 for *M-D49*, and .234 for *M-D50*.

It may be tentatively concluded that the smaller the difference with respect to *H*/*M*, the more likely that such a disputed paper can be considered more Hamiltonian or Madisonian, respectively, and vice versa. Accordingly, the results point to the direction that both *D49* and *D50* can be taken as more Madisonian on account of the least difference found (.157<.184; .234<.245), confirming thus previous attributions in this line. A summary of the results is shown in the left side of Table 8, where the values against *JAY* and *JOI* have been added for illustration.

	1-100				**TOP 1-5**				**TOP 1-25**		**TOP 1-50**		**TOP 1-75**	
	H	*M*	*JAY*	*JOY*	*H*	*M*	*JAY*	*JOY*	*H*	*M*	*H*	*M*	*H*	*M*
D49	.184	.157	.274	.239	.030	.016	.087	.028	.088	.070	.133	.113	.161	.137
D50	.245	.234	.288	.261	.037	.030	.071	.035	.102	.091	.159	.152	.203	.191
D51	.219	.187	.311	.269	.037	.026	.101	.027	.107	.084	.159	.134	.193	.163
D52	.173	.154	.275	.255	.031	.025	.083	.041	.071	.058	.106	.094	.148	.126
D53	.168	.135	.259	.237	.025	.015	.064	.031	.078	.060	.115	.088	.143	.106
D54	.198	.162	.302	.245	.037	.026	.100	.036	.093	.064	.139	.106	.170	.134
D55	**.178**	**.195**	**.261**	**.277**	**.022**	**.030**	**.075**	**.046**	**.065**	**.077**	**.112**	**.128**	**.146**	**.160**
D56	.198	.192	.284	.240	.038	.033	.075	.033	.096	.092	.136	.136	.172	.165
D57	.177	.147	.234	.231	.023	.021	.086	.036	.087	.062	.129	.104	.160	.126
D58	.170	.135	.279	.279	.034	.022	.091	.091	.082	.061	.112	.084	.145	.107
D62	.176	.151	.255	.253	.032	.024	.056	.040	.084	.065	.124	.096	.152	.123
D63	.144	.127	.254	.220	.026	.019	.082	.028	.061	.040	.099	.085	.127	.107

Table 8. (Full and truncated) simplified Delta for the Disputed Papers

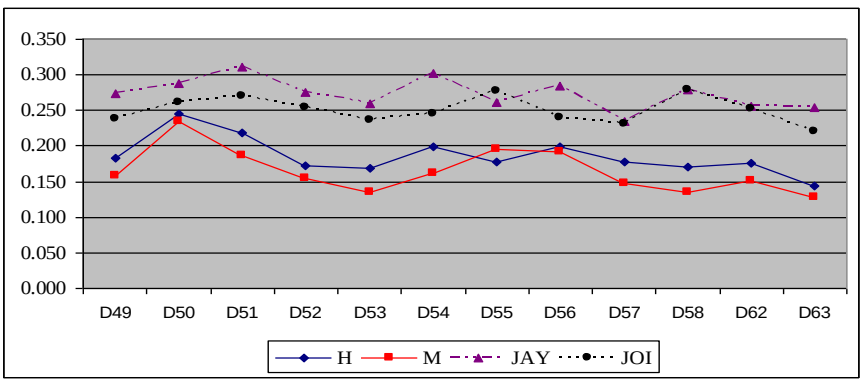

Figure 3. Simplified Delta in the Disputed Papers (100 items)

Print ISSN: 1578-7044; Online ISSN: 1989-6131

The analysis of the data allows us to conclude that the differences, except in the case of *D55*, are smaller with respect to *M* than to *H*. In addition, the latter are smaller than those with respect to *JAY* or the *JOI*, as plotted in Figure 3 below (100 items). These results represent a successful clustering of 91.63%.

In view of these figures, one can tentatively state that the *Disputed Papers* are less Hamiltonian than Madisonian, and that *JAY* is the most unlike (if compared to *H* or to *M*). In fact, the values for *JOI* are between *H* and *JAY*, something reasonable on account of their joint provenance. However, *D55* seems to diverge from the other *Disputed Papers* inasmuch as it turns out to be less unlike *H* than *M* and more similar to *JAY* than to *JOI*. Against this evidence, grounded reasons must be sought to establish an irrefutable argument, if any, to justify this change, which has been the common tendency in all the studies of this paper. Therefore, further investigation is needed to explain this textual anomaly or otherwise to attribute the authorship to Hamilton.

To eliminate any fortuitous concurrency of favourable data, the same experiment has been replicated relying only on the values for the topmost 5, 25, 50 and 75 words (Table 8). The results agree with those for the whole set of words, as plotted in Figure 4 (only the 5 topmost words).

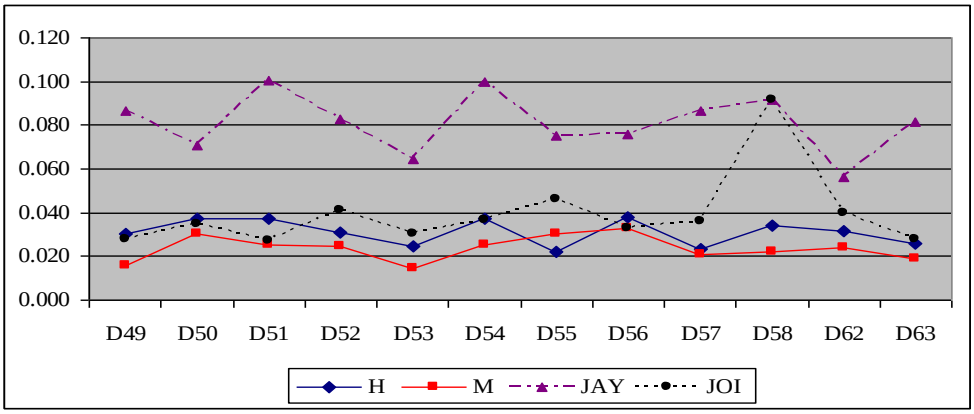

Figure 4. Simplified Delta in the Disputed Papers (5 items)

Likewise, it seems appropriate to replicate the experiment again by disregarding the content words from the analysis. For this purpose, the following content words have been ruled out: *states, government, state, people, power, constitution, union, same, national, federal, new, great, authority, public, general, powers, time, executive, united, men, part, body, members, number*. The results, however, do not diverge from those obtained from the whole set as, in general terms, the values for *M* are lower than those for *H*, exception being made of *D55*. Likewise, the values for *JAY* are also greater than those for *JOI* which, in turn, exceed those for *H* or *M*, as expected.

Print ISSN: 1578-7044; Online ISSN: 1989-6131

Figure 5 shows a great similarity to Figure 2, as the lines plotting each author keeps the same relative position (*M*, *H*, *JOI*, and *JAY*), and those representing *H* and *JAY* can be taken as coincident.

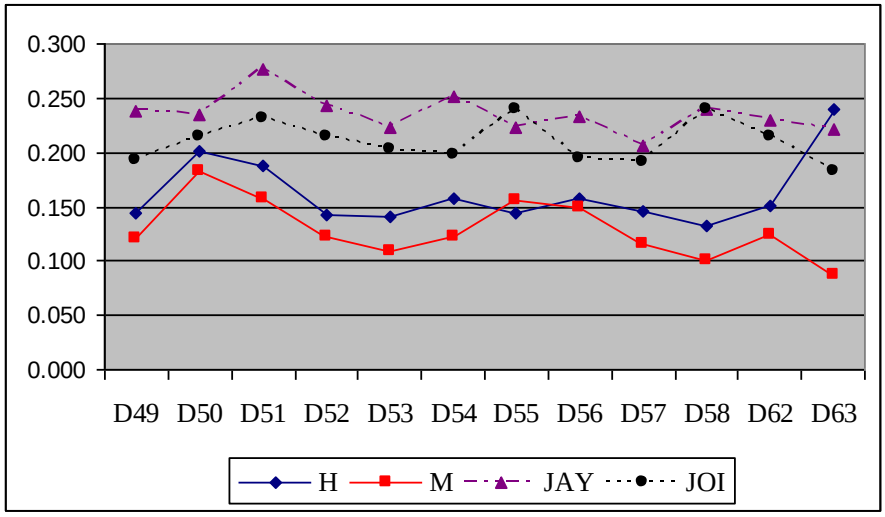

Figure 5. Simplified Delta for the Disputed Papers (function words)

5. CONCLUSIONS

The main conclusion after the application of Delta to the *Disputed Federalist Papers* is that their attribution does not differ from other studies in the field using different methodologies or approaches. In light of the results derived from modified and simplified Delta, it has been proved that the *Disputed Papers* can be re-assigned to Madison, with the exception of *D55*, more Hamiltonian. However, the evaluation of *H80* and *M38* as more Madisonian and more Hamiltonian, respectively, has been held fast throughout the various Delta experiments (modified version), despite the text-size of the sub-groups under scrutiny. It is fair to mention here that no previous experimentation had been done to cluster homologous Hamilton's and Madison's papers for it involved an added difficulty, though Jay's writings are also considered for control and illustration of the differences.

A statistical explanation for the Hamiltonian resemblance of *D55* has been formulated in the sense that it "does not have its share of marker words, no matter who wrote the paper, and the high frequency words produced no information" (Mosteller and Wallace, 1963: 300). We have investigated the aetiology for the failures in *D55*, *H80* and *M38*, but no plausible generalized answer has been found, despite the endless list of activities tested at word level for the purpose: the calculation of *Z*, the comparison of the ratings, the checking of the rank order, the accumulated scoring of the first 100 words in the hierarchy of each paper, the study of the items of the hierarchy which do not occur in each paper, the counting of *hapax*

Print ISSN: 1578-7044; Online ISSN: 1989-6131

legomena and *dislegomena*, etc. It is true that some traces seem to point to the expected direction, for example the ranking order of the first 17 words in each hierarchy or the percentage accumulated by the words in the hierarchy of each paper, but it does not hold valid in all cases. Moreover, if a salient feature were found to be a reliable discriminator to solve this authorship problem, it would hardly become universal for other cases. From this, it follows that the efficiency and robustness of Delta lies in a multiple evaluation of lexical features, as the differences with respect to the model are not blurred by the whole. Then, the cases of *D55*, *H80* and *M38*, with a more Hamiltonian or Madisonian profile, is an unsolved question which encourages us to deepen in their study so as to find a plausible answer by taking into account the syntactic layer or the content analysis of the papers.

The second conclusion is that Hamilton's and Madison's styles, not to mention Jay's, present remarkable differences, notwithstanding their likeness as stated above. They resemble quite alike in terms of the average sentence length (34.55 and 34.59 words respectively for Hamilton and Madison) or in terms of the average standard deviation (19.2 as opposed 20.3) according to Williams and Mosteller's counting in 1941 (Mosteller and Wallace, 1963: 6). However, from a lexical perspective, some divergences are found when their word hierarchies are compared or when the vocabulary richness of individual papers is measured by *Z*. Both the variation in lexical richness (Smith and Kelly, 2002: 412) and the changes in the rating of the most common words can be useful to establish an author's fingerprint and, paradoxically, to signal a natural evolution of style (Malyutov, 2005: 354; Somers and Tweedie, 2003: 412).

An examination of the papers in terms of their lexical richness and of their likeness ranking allows us to discard any one-to-one correspondence between them, in the sense that the more similar the lexical richness of two papers, the less unlike the authors' styles, which may lead to a common authorship. Likewise, the rank of lexical richness does not prove to favour or avert the possibility of becoming the least unlike in the experiments with Delta: *H80* and *M38*, which are ranked in bottom and top positions in terms of lexical richness, are not correctly associated with their homologous sub-groups listed in nearby positions. Conversely, *H8* and *H68* with top and bottom positions in the list of lexical richness frequently appear as less unlike.

The *Federalist Papers* show these differences, mostly if individually treated, even though the chronological span can be dismissed for non-existent, there is no genre change and somewhat related topics are dealt with. However, the range of unlikeness does not reach the threshold required by Burrows to apply Delta successfully in the sense that "any text we care to submit to the test must by definition be 'least unlike' some members of the group" (Burrows, 2003: 15). For this very reason, it is remarkable that the simple rating of a few words can allow the matching of texts by the same author rather accurately.

The third conclusion derives from the application of Delta to the *Federalist Papers*, a time-consuming procedure notwithstanding the invaluable assistance provided by *WordSmith*

Print ISSN: 1578-7044; Online ISSN: 1989-6131

Tools and *MS Excel*. Thanks to the latter, the original proposal lends itself well to the adaptations introduced to improve its efficiency and reliability, and in general terms it deserves a highly favourable assessment as a test for authorship attribution, not only quantitatively (on account of the correct associations produced) but also qualitatively, as it enables to identify the correct sub-group in the main corpus. However, the potential of Delta does not prove to be the same in the various experiments, and the success rate is found to increase from the modified version to the simplified. The initial application of (modified) Delta yields a series of results that do not allow to match homologous papers in the test corpus with those in the main one, as a strong tendency is observed to relate a great number of the tested papers (even those by *Madison* or the *Disputed Papers*) to *H-II* (in the main corpus). This failure is blamed on the minimum text-length that the sub-groups must have so as to convey reliability after realizing that the longer the text-length of the sub-groups in the main corpus, the greater the success in the expected likeness of the associated papers in the test corpus (from 10 correct assignments to just 21 or 22). It is obvious that the text-size increase of the sub-groups implies a reduction in their number (and of candidates), which necessarily leads to a greater success. Therefore, it seems convenient to establish a reliable text-length for the sub-groups constituting the main corpus since Burrows only advises "to set aside shorter texts" than 2,000 words (Burrows, 2003: 21) in reference to the test corpus. Taking into account that the text-length of the members in the main corpus ranges from 29.905 (*H-II*) to 22.290 (*M-III*) through 11.335 (*H-IV*), it seems appropriate to estimate that 17,000 words (the average of *M-III* and *H-IV*) could be a safe threshold for them.

The fourth conclusion, stemming from modified Delta, has to do with the type of word and the size of the frequency hierarchy. Whilst Burrows' prototype employed an *excessively* long list of 150 most common words (function words?), which was then shortened by progressive truncations from the lower end (from 150 to 60 in five stages), our experiments have been carried out successively (a) by using the complete 100 wordlist (most common words, function and content altogether); (b) by relying on function words only; and (c) by truncating the list from below in agreement with Burrows' statement that "the top 40 or so are powerful markers of genre, doing much to delineate the different sorts of texts, the results are no longer reliable" (Burrows, 2003: 24; 28).

Each of the three treatments, however, yields its own results, which cannot be considered as completely homogeneous, as follows: (type a) does not accurately associate *H80*, *M38*, and *D55* as expected, since they are assessed as the least unlike *M-III*, *H-II* and *H-III*, respectively; and (type b) replicates the original results when the 25 bottom words of the complete list are ruled out, but produces new four mismatches if the 50 bottom words are not considered. In summary, the worst results are obtained by eliminating the 50 bottom words from the hierarchy.

Print ISSN: 1578-7044; Online ISSN: 1989-6131

These results confirm that our decision as for the number and type of words (function and content) in the hierarchy is not misleading or groundless. No argument can be made with respect to the number of words that amply fit within the interval 150 to 60, but a warning must be given in the matter of word-type. It is generally agreed that function words "appear to be a fertile source of discriminators, and luckily the high-frequency words are the strongest" (Mosteller & Wallace, 1963: 306), but in the *Federalist Papers* the content words do not seem to produce the expected divergences found in other texts without so many features in common. Therefore, function words are bound to be most reliable unless quite similar texts and/or quite like authors are dealt with.

The fifth conclusion originates from the application of the simplified version of Delta as we seek to find out whether each disputed paper resembles Hamilton's or Madison's style more closely by comparing the rates (instead of the *z*-scores) for the words in the hierarchy one against the other, by adding their differences and averaging them out. This version proves successful in the task of attributing all the *Disputed Papers* to Madison, exception being made of *D55*, which seems more Hamiltonian, on account of the former's lesser simplified Delta, the figures for *JOI* and *JAY* depicting independent trajectories.

The same result is obtained when the wordlist is truncated from the bottom (in four stages of 75, 50, 25 and 5 words), or the content words are discarded from it. This finding can be understood as a proof of test validation and of the reliability of the results as they assign the same author while the wordlist is modified. In addition, that *D55* is considered as the least unlike Hamilton was also a constant in the experiments with modified Delta. Therefore, we can consider this simple modality as more appropriate for dealing with closed game attributions (Hamilton's or Madison's?), although we have also analysed *JAY* and *JOI*.

It is a common practice to have the results tested again by other explanatory methods, i.e. p.c.a., to reduce the candidates to 2 or 3, and then apply an ANN or employ rare words for definitive corroboration, but this is not the case here as the results come to coincide with those of previous analyses. Moreover, the preliminary results of an on-going study-case dealing with other texts allow us to state that the method holds whenever the requirements of similar chronology, genre and topic are satisfied.

Despite our bench-work effort on computing the *Federalist Papers* and our constant devotion to the matter of authorship attribution, some flaws can yet be detected in our work. It is our intention to continue the research in this field by introducing the so-called *Antifederalist Papers* into play, both in the main and in the test corpus, on the assumption that a greater textual variety will certainly contribute to upgrade the results as well as by processing the annotated corpus of the *Federalist Papers* in the hope that some valuable findings will arise therefrom.

Print ISSN: 1578-7044; Online ISSN: 1989-6131

ACKNOWLEDGEMENTS

This work was supported by the Spanish Ministry of Science and Technology [grant numbers FFI2008-02336 and FFI2011-26492] and by the Autonomous Government of Andalusia [grant numbers P07-HUM02609 and P11-HUM7597]. These grants are hereby gratefully acknowledged.

NOTES

1. In this line, Burrows affirms that "the most common words of the language have been given more attention than ever before, whether in studies where the words are allowed to choose themselves on no other ground than their relative frequency or else in studies where distinctions between lexical words and function-words are attempted, the former being discarded as too subject-specific" (Burrows, 2003: 7).
2. Available at http://www.gutenberg.org/etext/18, being the only resource available at the outset of this research. The *Federalist Papers* can now be accessed from other on-line resources, i.e. http://www.foundingfathers.info/federalistpapers/fedindex.htm; http://avalon.law.yale.edu/subject_menus/fed.asp; http://thomas.loc.gov/home/histdox/fedpapers.html (accessed 15 October 2009).
3. Following Rudman, the texts have not been manipulated but for eliminating the references and/or the editor's notes, though this does not imply to suppress the editorial and experimental corruption as for insertions, new punctuations, etc. (Rudman 1997: 351-365). In this line, the present study does not accomplish the task of tagging the homographs (Craig, 1999: 105; Burrows, 2003: 11), even when we believe in its benefits for authorship attribution, particularly in the case of inflected languages like Old English and Middle English (Miranda-García and Calle-Martín 2005: 41-45). An annotated corpus of *The Federalist Papers* has been recently compiled and used as the input for an on-going lemma-based research.
4. The wordlist generated by *WordSmith 5.0* includes the word, the number of hits, and the percentage of *N* (the number of running words).
5. Only the values for *JOI* and *JAY* diverge significantly from the remaining, though each in opposing directions. The higher values for JOI come to answer Holmes's and Forsyth's question of "whether collaborative tests are always Richer in vocabulary than texts from separate contributions" (1995: 117).

REFERENCES

Adair, D. (1944a). The Authorship of the Disputed *Federalist Papers*. *The William and Mary Quarterly*, *1*(2), 98-122.

Adair, D. (1944b). The Authorship of the Disputed *Federalist Papers*. Part II. *The William and Mary Quarterly*, *1*(3), 235-64.

Baayen, H., van Halteren, H., Neijt, A. & Tweedie, F. (2002). An Experiment in Authorship Attribution. *6es Journées Internationales d'Analyse Statistique des Données Textuelles*. Retrieved from http://www.cavi.univ-paris3.fr/lexicometrica/jadt/jadt2002/pdf-2002/baayen_vanhalteren_neijt_tweedie.pdf

Burrows, J. F. (2002). 'Delta': a Measure of Stylistic Difference and a Guide to Likely Authorship. *Literary and Linguistic Computing*, *17*(3), 267-287.

Burrows J. F. (2003). Questions of Authorship: Attribution and Beyond. A Lectured Delivered on the Occasion of the Roberto Busa Award ACH-ALLC 2001. *Computers and the Humanities*, *37*(1), 5-32.

Print ISSN: 1578-7044; Online ISSN: 1989-6131

Craig, H. (1999). Authorial Attribution and Computational Stylistics: If You Can Tell Authors Apart, Have You Learned Anything about Them? *Literary and Linguistic Computing, 14*(1), 103-113.

Dalen-Oskam, K. van & van Zundert, J. (2007). Delta for Middle Dutch—Author and Copyist Distinction in Walewein. *Literary and Linguistic Computing, 22*(3), 345-362.

Diederich, J., Kindermann, J., Edda, L. & Gerhard, P. (2003). Authorship Attribution with Support Vector Machines. *Applied Intelligence*, *19*(1-2), 109-123.

Ellegärd, A. (1962). A Statistical Method for Determining Authorship. *The Junius Letters*, 1769-1772. *Gothenburg Studies in English*, *47*, 115.

Fung, G. (2003). The Disputed *Federalist Papers*: SVM Feature Selection via Concave Minimization. *Journal of the ACM*, *5*, 1-8. Retrieved from http://pages.cs.wisc.edu/~gfung/federalist.pdf

Holmes, D. I. & Forsyth, R. S. (1995). The *Federalists* Revisited: New Directions in Authorship Attribution. *Literary and Linguistic Computing*, *10*(2), 111-127.

Hoover, D. L. (2004a). Testing Burrows's Delta. *Literary and Linguistic Computing*, *19*(4), 453-475.

Hoover, D. L. (2004b). Delta Prime. *Literary and Linguistic Computing*, *19*(4), 477-495.

Kjell, B. (1994). Authorship Determination Using Letter-pair Frequency Features with Neural Networks Classifiers. *Literary and Linguistic Computing*, *9*(2), 119-124.

Malyutov, M. B. (2005). Authorship Attribution of Texts: a Review. *Electronic Notes in Discreet Mathematics*, *21*, 353-357.

Miranda-García, A. & Calle-Martín, J. (2005). The Validity of Lemma-based Lexical Richness in Authorship Attribution: a Proposal for the Old English *Gospels*. *ICAME Journal*, *29*, 41-55.

Miranda-García, A. & Calle-Martín, J. (2007). Function Words in Authorship Attribution Studies. *Literary and Linguistic Computing*, *22*(1), 49-66.

Miranda-García, A. & Calle-Martín, J. (2008). A Survey of Non-traditional Authorship Attribution Studies. *Ecdotica*, *5*, 147-168.

Mosteller, F. & Wallace, D. L. (1963). Inference in an Authorship Problem. A Comparative Study of Discrimination Methods Applied to the Authorship of the Disputed *Federalist Papers*. *Journal of the American Statistical Association*, *58*(302), 275-309.

Mosteller, F. & Wallace, D. L. (1964). *Applied Bayesian and Classical Inference: The Case of the Federalist Papers*. New York: Addison-Wesley.

Rudman, J. (1997). The State of Authorship Attribution Studies: Some Problems and Solutions. *Computers and the Humanities*, *31*(4), 351-365.

Rudman, J. (2005). The Non-Traditional Case for the Authorship of the Twelve Disputed *Federalist Papers*: A Monument Built on Sand? *Proceedings of ACH/ALLC*, Victoria. Retrieved from http://mustard.tapor.uvic.ca/cocoon/ach_abstracts/xq/xhtml.xq?id=54

Scott, M. (2009). *WordSmith Tools 5.0*. Oxford: Oxford University Press.

Smith, J. A. & Kelly, C. (2002). Stylistic Constancy and Change across Literary Corpora: Using Measure of Lexical Richness to Date Works. *Computers and the Humanities*, *36*(4), 411-430.

Somers, H, & Tweedie, F. (2003). Authorship Attribution and Pastiche. *Computers and the Humanities*, *37*(4), 407-429.

Tweedie, F. J., Singh, S. & Holmes, D. I. (1996). Neural Networks Applications in Stylometry: The *Federalist Papers*. *Computers and the Humanities*, *30*(1), 1-10.

Tweedie, F. J. & Baayen, R. H. (1998). How Variable May a Constant Be? Measures of Lexical Richness in Perspective. *Computers and the Humanities*, *32*(5), 323-352.

Print ISSN: 1578-7044; Online ISSN: 1989-6131