

Empirical evaluations of language-based author identification techniques

Carole E. Chaski

Executive Director, Institute for Linguistic Evidence, Inc.

ABSTRACT Recent Court decisions in the United States call for the empirical testing of language-based author identification techniques. This article shows the results of such testing. The tested hypotheses include: syntactic analysis, syntactically classified punctuation, sentential complexity, vocabulary richness, readability, content analysis, spelling errors, punctuation errors, word form errors, and grammatical errors. These hypotheses are tested on a set of documents written by four women who are similar in age, educational level, and dialectal background: two of the women are Euro-American, and two are Afro-American. Each hypothesis is tested separately to determine its ability to differentiate documents from different authors and cluster documents from each author. Hypotheses which quantify linguistic features are tested statistically using the chi-square statistic. Discrimination error rates are calculated. Only two hypotheses successfully differentiate and cluster documents: syntactic analysis and syntactically classified punctuation.

KEYWORDS language-based author identification, document examination, author attribution, questioned document

THE TIMELY NEED FOR EMPIRICAL EVALUATIONS

In the early part of 2000, in the United States District Court, District of New Jersey, Roy Van Wyk was brought to trial for making threatening communications (United States v. Van Wyk, 83 F.Supp.2d 515, D.N.J., 2000). The Government, represented by Assistant US Attorney Charles B. McKenna, proposed that Special Agent James R. Fitzgerald of the Federal Bureau of Investigation be allowed to testify as an expert in forensic stylistics (later text analysis) about the authorship of the threatening letters.¹ The Government argued that Fitzgerald's testimony should be admitted because it relied on McMenamin's peer-reviewed publication (McMenamin 1993), thus meeting at least one of the Daubert criteria which Federal judges must consider when determining the admissibility of evidence. Federal judges consider a variety of guidelines for admitting scientific and technical evidence; the Daubert criteria focus on the empirical reliability of a scientific technique.² Thus, the Government argued that Fitzgerald's technique should be admitted based on the fact that he relied on McMenamin (1993) which demonstrates the empirical reliability deducible from peer review before publication.

Defendant Van Wyk, represented by Assistant Federal Public Defender John H. Yauch, filed a motion *in limine* to exclude the proposed testimony of Agent Fitzgerald. The Defence argued that the ‘proffered expert testimony is subjective, unreliable and lacks measurable standards’ (Van Wyk, 83 F.Supp.2d 515, 521). Thus, the Defence argued that admitting forensic stylistics testimony would violate several other criteria of the Daubert standard of empirical reliability, such as falsifiability of the technique, known error rate, and standard operating procedures for performing the technique (see Note 2 for more discussion of admissibility factors).

The Court did, in fact, recognize the ‘lack of scientific reliability of forensic stylistics’:

Although Fitzgerald employed a particular methodology that may be subject to testing, neither Fitzgerald nor the Government has been able to identify a known rate of error, establish what amount of samples is necessary for an expert to be able to reach a conclusion as to probability of authorship, or pinpoint any meaningful peer review. Additionally, as Defendant argues, there is no universally recognized standard for certifying an individual as an expert in forensic stylistics.

(Van Wyk, 83 F.Supp.2d 515, 523)

The Van Wyk case and Judge Bassler’s ruling demonstrate the timely need for a sound and tested methodology for language-based author identification techniques.

Further, this case reveals that the Courts want the scientist-experts to self-regulate in terms of developing and testing techniques and in terms of certifying or qualifying experts. It is the linguists’ responsibility to create the theoretically sound hypotheses, test these hypotheses and perform the empirical evaluation of our own methods. It is the linguists’ responsibility to recognize junk science before it gets to court. These are especially interesting obligations in the case of forensic linguistics, because, as any linguist knows, everyone has something to say about language, and linguistics has many sister-disciplines within academia.

In this article, I present the results of experiments testing the empirical reliability of language-based author identification techniques currently in use in the United States.³ These techniques can be divided into three groups.

In the first group are two techniques – syntactically classified punctuation and syntactic analysis of phrase structure – which withstand the scrutiny of experimental testing and statistical analysis. These two techniques are rooted in linguistic theory, are underpinned by factual ideas about linguistic performance, and demonstrate the utility of linguistics in forensic applications. Further, since these two techniques quantify linguistic patterns, they are amenable to statistical testing and thus allow us

straightforward ways to decide the authorship question and to calculate error rates. Finally, these two linguistics-based techniques correctly differentiate different authors and cluster the questioned document with its actual author with a high level of accuracy.

In the second group are several techniques – sentential complexity, vocabulary richness, readability, content analysis – which quantify linguistic patterns, and are amenable to statistical testing. Error rates are easily calculable. These techniques derive from language-related disciplines, but unfortunately they rest on ideas about linguistic performance which do not hold within the forensic context. For instance, the type-token ratio is a venerable idea with predictable results in the context of Shakespeare studies, but within the confines of a forensic investigation with its typically short documents, the type-token ratio loses its utility. Finally, these techniques fail to differentiate authors and/or cluster documents correctly with a high level of accuracy.

In the third group are ‘forensic stylistics’ techniques – spelling errors, punctuation errors, word form errors, grammatical errors – which are rooted in handwriting identification and prescriptive grammar. These techniques do not quantify linguistic patterns; they are not amenable to statistical testing nor the calculation of error rates. Further, these prescriptive techniques rest upon factually incorrect ideas about individuality in language performance and violate theoretical principles of modern linguistics. Finally, these techniques fail to differentiate between documents authored by different writers and/or fail to cluster documents authored by the same writer with a high level of accuracy.

There is another approach to language-based author identification exemplified by literary critics such as Donald Foster. In an interview with Professor Foster, *The New York Times* (November 19, 1997 Metro Section) published a section of Professor Foster’s analysis of the JonBenet Ramsey ransom note. From the first two words of the note (‘Listen carefully!’), Professor Foster deduced, ‘diction associated with films like “Ransom,” “Dirty Harry,” and “Speed.” ... diction associated with a chief executive officer, day-to-day business concerns or computer equipment, possibly indicated a businessperson as author, and/or someone wishing to implicate John Ramsey.’ Foster’s method apparently relies completely on the literary critic’s subjective word associations. It is impossible to replicate the method, since one person’s subjective word association is naturally different from another’s subjective word association. I, for one, do not associate the phrase ‘listen carefully!’ to action movies or hi-tech business operations. Because this approach is purely speculative and relies completely on the literary critic’s subjective response to the words, it is impossible to test scientifically and, therefore, it will not be considered any further in this article.

Experimental method and task

Ten falsifiable hypotheses about language-based author identification were tested separately on one set of documents written by four subjects matched as well as possible for sociolinguistic features. Each hypothesis was tested to determine if (a) it could accurately differentiate documents written by the four different subjects and (b) if it could correctly assign an unknown document to one of the four subjects. The documents were drawn from the author's Writing Sample Database. A full description of the Writing Sample Database, demographic features of the writers and topics of the writing samples can be found in Appendix 1.

These experiments were intended to simulate as much as possible the actual task of a forensic examination. The analyst was given a set of known documents and a questioned document. These documents were unedited, actual writings; they contain different numbers of words, and they are rather short. These experiments are intended to test the *lower* limits of text length and quantity because forensically significant documents are often short and cannot be amplified; indeed, even known documents are often short in length and limited in quantity. The experiments use data reflecting the lower limits so that the results apply to actual casework. If the tested technique can work on the lower limits, it can certainly work on cases with abundant words in abundant documents. On the other hand, if the tested technique cannot work accurately on the lower limits, the technique's supporters must demonstrate that the technique can work when the limits are raised to include longer documents or more documents and accept that the technique cannot work under the typical, lower restraints of actual cases.

The subjects, or in casework the suspects, are matched as well as possible on sociolinguistically relevant features such as age, sex, race, dialect and educational level. In much casework, the suspects will share at least some if not all of these features. Again, the experiments use data reflecting the most restrictive type of case so that the techniques can be tested as rigorously as possible. Since dialectal features are relatively well-documented and easy to spot, the more difficult case is differentiating among documents from the same dialect. If the tested technique can differentiate the authors of documents which share dialectal features, then it can certainly work on documents which do not share dialectal features. On the other hand, if the tested technique can only work when the documents arise from different dialects, the technique probably only distinguishes dialectal characteristics and not idiolectal characteristics.

The text-types are also controlled in the experiments much as they would be controlled in actual casework. Variations in communicative functions cause easily recognizable differences in text-types such as essays, business letters, and love letters. Again, since these register differences are relatively well-documented and easy to spot, the more difficult case is

differentiating among documents from the same register. If the tested technique can differentiate the authors of documents from the same register, then it can certainly work on documents from different registers. On the other hand, if the tested technique can only differentiate different documents arising from different registers, the technique probably only distinguishes register characteristics and not idiolectal characteristics of the authors.

There is one important way in which the experiments do not simulate actual casework. In actual case work, the forensic linguist may want to use the three possible conclusions which are standard to many forensic sciences: (1) the known and questioned items belong to the same person (identification/inclusion); (2) the known and questioned items belong to different persons (elimination/exclusion); and (3) the analyst is unable to state a conclusion about authorship based on the current set of known items (inconclusive evidence/no conclusion). These experiments, however, allow for only two possible conclusions. In these experiments, the tested hypothesis either (a) accurately shows that documents are written by different authors; (b) accurately shows that the questioned document is written by the one, correct author. These experiments do not allow for the third conclusion that the tested hypothesis is unable to determine the authorship on the basis of the texts, because we are testing each hypothesis against the same set of texts. The purpose of these experiments is to test the hypotheses' discrimination and clustering ability against one set of data, not to test the kind of data needed for each hypothesis. But interestingly, the type of data needed for each hypothesis, especially the length of texts, comes to light when we consider the accuracy of each hypothesis' discrimination and grouping ability.

Test subjects

The tests were conducted on a set of documents authored by four women in their forties, two Euro-American, two Afro-American, with two-to-three years of college education and long-term exposure to the Delmarva dialect of the Eastern United States. These women, identified only by the database identifier, were subjects 001, 009, 016, and 080.

In order to demonstrate the validity or invalidity of certain hypotheses about individuality of language use in writing, it was sometimes useful to include additional subjects. These subjects are always referred to by their database identifier. Pertinent sociolinguistic features of any additional subjects are given at the time each test result is reported; the full sociolinguistic and demographic features of the Writing Sample Database can be found in Appendix 1.

Writing sample data

Documents written by the test subjects were coded by the author's

database identifier and task identification (e.g. 001-01). One unidentified document was labelled SQD2. Document SQD2 was selected from the writing samples of one of the four test subjects and typed by a research intern so that the analyst would have no experience with it prior to the experimental analyses. The actual documents are reproduced in Appendix 2.

Table 1 shows the number of words in the documents from the four women test subjects.

Table 1 Number of words in documents from four women writers

Subject ID	Number of words in text 1	Number of words in text 2	Number of words in text 3	Total number of words in all texts
001	223	121	187	531
009	361	265	372	998
016	344	556		900
080	239	93	103	345
SQD2				341

Experimental task

The experimental objective is to discover if any language-based author identification technique correctly discriminates between documents written by different authors and correctly clusters SQD2 with documents written by the actual author. All of the techniques have been re-created as accurately as possible from their original sources in questioned document examination, literary criticism, textual criticism, communications and linguistics. The analyst applied one language-based author identification technique at a time to all the documents in the set. Based on the technique, the analyst assigned the documents to authorship sets.

If the techniques measure and quantify the data, these measurements are tested statistically to determine the probability that the differing measures for each document set either come from or do not come from the same source.

Interpretation of results

After the application of the techniques, the research intern revealed to the analyst that document SQD2 was actually authored by subject 016. This fact allowed us to interpret the test results in terms of the replicability of the techniques. All of the techniques have been previously used by sources in questioned document examination, literary criticism, textual criticism, communications and linguistics. Therefore, if a technique produced the correct assignment of documents to authors in this experimental task, it is

a replicable technique, but if it produced incorrect assignments, then it is a non-replicable, unreliable technique.

Error rates

In this experimental task, the discrimination accuracy of the technique is calculated as a percentage of the times that the technique correctly differentiates documents from different authors and correctly assigns the questioned document to the correct author. Conversely, the discrimination error rate is the percentage of times that the technique fails to differentiate documents from different authors and incorrectly assigns the questioned document to an incorrect author. In most tests, four women writers are used. Each writer is paired with each other writer and the questioned document to test for similarity or difference: this results in ten pairings. If the technique performs absolutely perfectly, each writer will be differentiated from each other writer (six pairings), non-authors will be differentiated from the questioned document (three pairings) and the questioned document will only be clustered with its actual author (one pairing). So a technique would have a zero error rate if it correctly differentiated nine pairings and correctly clustered one pairing. Since we will only see this occur once, it is worth considering the alternative. If a technique fails to differentiate three of the nine pairings, then it has a discrimination accuracy of 6 out of 9 or 67 per cent, and a discrimination error rate of 3 out of 9 or 33 per cent. If the technique fails to cluster the questioned document with subject 016, it simply fails (with no number attached).

EXPERIMENTAL RESULTS: STATISTICALLY TESTABLE TECHNIQUES WITH GOOD REPLICATION RESULTS

In my empirical testing of language-based author identification techniques, two techniques look very promising. These are syntactic phrase structures and syntactically classified punctuation patterns. Both of these techniques require expertise in modern linguistic analysis. While punctuation patterns may seem to be an obvious kind of textual phenomenon which both the American high-school graduate and the language expert would pay attention to, the way that punctuation patterns are used in the empirical tests that follow requires knowledge of syntactic structures and statistics. So while any juror or judge may notice that one document contains lots of hyphens while another does not, they may not notice that the hyphens in the one document are always syntactically conditioned in ways that are not available in the other document. In other words, even such an obvious feature as punctuation has to be handled in a non-obvious way in order to yield reliable results for author identification. Syntactic structures, on the other hand, are the kind of phenomenon which is not obvious to the American high-school graduate or the language expert who has not been trained in syntactic theory and analysis.

Syntactic analysis of phrase structure

The fundamental idea about language individuality underlying this technique is that syntactic processing is automatized, unconscious behaviour and therefore is difficult either to disguise or imitate (for more discussion see Chaski 1998a).

Hypothesis:	Abstract syntactic structures differentiate and identify authors.
Sources:	Chaski (1997a, 1997b, 1998b).
Methodology:	Parse text using a generalized phrase structure grammar. Count structures and calculate ratios between structures of related type. Test for differences between texts statistically.
Tools:	Knowledge of phrase structure grammars. ALIAS® computer program (not necessary, but helpful). ⁴ Knowledge and use of χ^2 statistic.

Since the analyst (with or without the help of ALIAS®) parses each word of a document, and each phrase of a document, many syntactic features are available for analysis. For brevity's sake I will show only one set of parse results: auxiliary verbal patterns.⁵ These verbal features include modal, progressive, passive, perfective, negative inversion, infinitive, verb-particles, and the elliptical verb; (the causative does not occur in the test data). All of these features modify the main verb and so are subsumed under the grammatical heading auxiliary (Aux). The results are presented in Table 2.

Table 2 Raw frequencies of verbal features in sets 016 and 080

Aux. verbal features	001	009	016	080	qd2
modal	10	18	17	18	9
v-progressive	6	14	13	10	3
v-passive	2	2	8	1	0
v-participle	1	18	16	1	2
modal + v-participle	2	0	1	0	0
v-participle v-passive	0	1	1	0	0
v-negative inverted	9	7	6	0	3
v-infinitive	18	30	27	11	1
v-infinitive v-passive	0	0	1	0	0
v v-particle	41	9	71	44	39
vp[empty]	0	1	1	1	0

The statistical analysis Chaski (2000; 1998; 1997) employed is a standard technique for categorical frequency data, the chi-square test (for description, see Seigel and Castellan 1988; Woods, Fletcher and Hughes 1986; Williams 1968). This statistical technique has been independently reviewed and used in authorship attribution cases by Svartik (1968), Dreher and Young (1969) and Smith (1994). The chi-square technique has also been used by O'Brien and Darnell (1982), two economists who studied the authorship of economics texts.

The chi-square tests the independence of two groups of categorical frequencies for which we cannot assume or do not want to imply a normal distribution. While the chi-square statistic is not the most high-powered statistic, it suits our purposes because we are measuring categorical frequencies for which we are unsure of distributional properties and which we hope to match or not match to other sets of categorical frequencies.

Two points about applying and interpreting the chi-square results are in order. First, in applying the chi-square statistic we have to consider the size of our observed frequencies because these are used to calculate the expected frequencies. The practical rule allows no more than 20 per cent of the expected frequencies generated during the calculation to be less than 5, while none of the expected frequencies can be less than 1 (Seigel and Castellan 1988; Woods, Fletcher and Hughes 1986; Williams 1968). As small as some of the observed frequencies in Table 2 are, all of the test results pass this practical rule.

Second, the interpretation of the probability is a bit more interesting than the standard 'significant difference if below .05' interpretation of statistical results because we are asking two questions. We are asking the result to tell us if the different authors are distinguished and also if the unknown document is clustered with the correct author. It is possible to interpret the probabilities to answer both questions (as is also done in Dreher and Young (1969)). Since the chi-square is testing for difference between groups, the null hypothesis states that there is no difference between the groups. We accept this null hypothesis of sameness if the probability associated with the chi-square result is greater than .05. We reject the null hypothesis, accepting instead the alternative hypothesis of difference, if the probability is less than or equal to .05. So, in standard parlance, there is a statistically significant difference between two groups of authors/documents, if the probability falls equal to or less than .05, while there is no difference if the probability rises above .05.

Now let us apply that to our case results, as shown in Table 3. When two different writers are paired for testing, we want the probability of their sameness to be very small, to be less than .05, since they are actually different. When the questioned document and its non-authors are paired for testing, again we want the probability of their sameness to be very small, less than .05, since the unknown document was not written by these

authors. But when the questioned document is paired with its actual author, we want the probability of their sameness to be large, at least greater than .05. A low p-value demonstrates significant difference, while a high p-value indicates similarity or at least consistency.

Table 3 shows the results of running the chi-square statistic on the frequencies in Table 2. The probabilities that the writers 001, 009, 016 and 080 show no difference are all low, less than or equal to 5 per cent. Even the pairing of 001 and 016 meets this test because it could not be rounded up over the .05 bar. From these low p-values, we conclude that the authors are different. The syntactic feature technique differentiates different authors.

The probabilities that the writers 001, 009, 080 show no difference from the questioned document are, again, all very low, less than 5 per cent. From these low p-values, we conclude that these authors are different from the questioned document. The syntactic feature technique differentiates the questioned document from its non-authors.

But the probability that 016 and the QD are not significantly different – are similar – is very high, greater than 5 per cent. From this high p-value, we conclude that 016 and the QD are consistent with each other. The syntactic features technique accurately clusters the questioned document with the correct author.

Table 3 Statistical analysis of verbal features in four women writers

	01/009	01/080	01/016	009/016	009/080	016/080
χ^2	43.92	14.076	10.922	40.215	47.691	14.384
p	.0001	.0071	.053	.0001	.0001	.0133
df	5	4	5	6	4	5
	01/QD	009/QD	080/QD	016/QD		
χ^2	13.866	64.657	8.746	5.618		
p	.0077	.0001	.0329	.2295		
df	4	5	3	4		

Table 3 shows that verbal features correctly differentiate all the writers from each other, correctly differentiate the questioned document from the non-authors, and correctly cluster the questioned document with its actual author. The syntactic phrase structure analysis technique is based in linguistic science, empirically tested on a forensically similar task and successful in both differentiating and clustering documents.

Comments on error rate, sample amount

In this experimental test, syntactic analysis of phrase structure shows a 'zero' error rate on very small, forensically similar documents from writers who are dialectally and demographically similar. While this is a very promising and even exciting result, it is by no means enough evidence to proclaim that the technique actually has a 'zero' error rate. Forensic techniques which claim to have zero error rates, such as current American handwriting identification, are able to make such claims because they have never been empirically and objectively tested; such claims should make any scientist shudder in disbelief. So instead of interpreting the reported result in such an incredible, naïve way, it is prudent to recognize that many more studies which meet the robust standards of short documents by writers with similar dialectal and demographic features need to be completed; with continued funding from the US Department of Justice's National Institute of Justice, such research is currently being conducted.

A better way to interpret this result is to state that, in the context of this experiment, the syntactic feature technique shows a 100 per cent correct matching rate. Since the context of this experiment includes controlling for dialect and document size, this is a rigorous result.

Punctuation frequency and syntactic function

I will present two hypotheses about punctuation marks which focus on frequency and syntactic function. The fundamental idea about language individuality in these hypotheses is that punctuation is a graphic (non-verbal) reflection of syntactic structure.

Hypothesis:	Punctuation mark frequencies discriminate between authors.
Source:	Chaski (1996).
Methodology:	List each punctuation mark. Test statistically.
Tools:	Knowledge of punctuation and syntax; knowledge and use of χ^2 statistic.

Punctuation Analysis 1

One way to examine use and non-use of punctuation marks is to count the frequencies as shown in Table 4. I have added subject 018 to the mix because she also uses hyphens as does subject 016. Subject 018 is a white, thirty-one-year-old female with long-term exposure to the Delmarva dialect who has two years of college education. The results are presented in Table 4.

Table 4 Raw frequencies of punctuation marks in five women writers from Writing Sample Database

Subjects	001	009	080	016	018	QD2	Not usable in χ^2 test
ellipsis	0	0	0	0	0	0	*
period	26	71	24	39	140	14	
comma	8	28	17	67	18	9	
semicolon	1	1	0	1	0	0	*
colon	0	0	0	2	0	1	*
apostrophe	11	7	1	15	15	10	
question mark	0	0	5	2	1	3	*
exclamation point	2	4	0	1	1	0	*
hyphen	0	3	2	50	23	33	
underline	0	0	0	1	0	0	*
quotation (set)	1	3	0	8	1	5	
slash	0	0	0	0	1	0	*
ampersand	0	0	0	0	3	0	*
plus	1	0	0	0	0	0	*
parentheses (set)	0	1	0	2	4	1	*

Given the standard rule about expected frequencies for the chi-square test, many of the categories in Table 4 are actually unusable statistically. These are marked with an asterisk in Table 4. In order to use the chi-square properly, I have run the statistic using only at most five features for different pairings of authors; (the number of features is one more than the degrees of freedom).

Analysis

Assuming the null hypothesis that there is no difference between the punctuation patterns of the five women writers in Table 4, what is the chance that these punctuation patterns come from the same author? What we would expect to see is that there will be, on the one hand, a really small chance (of no difference or sameness) when the documents come from different authors, and, on the other hand, a chance greater than 5 per cent when they come from the same author. Results are shown in Table 5; problematic results are bolded.

Punctuation mark patterns correctly differentiate between nine of the ten pairings, but fail to differentiate between one pair of writers (009/080). Note that the technique distinguishes successfully between writers 016 and 018 even though these two writers superficially look alike because of their hyphens. Further, punctuation mark patterns correctly differentiate the questioned document from the four non-authors, but fail to cluster the questioned document with its actual author (016/QD2). So in terms of discriminating between writers, the technique accurately

Table 5 Statistical analysis of punctuation mark patterns in five women writers from Writing Sample Database

	001/009	001/080	001/016	009/016	009/080	016/080
χ^2	9.84	11.564	40.664	59.932	3.402	22.843
p	.0073	.0031	.0001	.0001	.1825	.0001
df	2	2	4	4	2	3
<hr/>						
	001/018	009/018	016/018	080/018		
χ^2	11.591	19.829	93.955	23.691		
p	.003	.0002	.0001	.0001		
df	2	3	3	2		
<hr/>						
	001/QD	009/QD	080/QD	016/QD	018/QD	
χ^2	33.949	66.988	36.994	17.442	66.988	
p	.0001	.0001	.0001	.0006	.0001	
df	3	3	3	3	3	

discriminates thirteen out of fourteen times or is 92.8 per cent successful, with an error rate of 7.1 per cent for discrimination. But in terms of clustering a document with its actual author, the technique fails. Even so, a technique that holds such promise as 93 per cent discrimination accuracy should be further studied.

Punctuation reflects syntactic structure, thus serving as an alternate means of getting at syntactic structure (see Nunberg 1988, Meyer 1987). Perhaps if we syntactically classify the punctuation marks, we will be able to retain a high level of accuracy in discriminating between authors and also successfully cluster a document with its author.

Hypothesis:	Syntactically classified punctuation discriminates between authors.
Source:	Chaski (1996).
Methodology:	List each punctuation mark. Classify by the mark's syntactic function. Test statistically.
Tools:	Knowledge of punctuation and syntax; knowledge and use of χ^2 statistic.

Syntactically classified punctuation results are shown in Table 6.

Table 6 Frequency data from syntactically classified punctuation analysis of five women writers from Writing Sample Database

Punctuation	s001	s009	s080	s016	s018	SQD2	Too low for χ^2 test
EOS period	24	69	23	37	121	14	
EOS question mark	0	0	5	1	1	3	*
EOS period for question	2	1	1	1	0	0	*
EOS exclamation point	2	4	0	0	1	0	*
EOS hyphen	0	0	0	6	0	4	*
Inter-S comma	1	2	0	12	3	1	
Inter-S colon	0	0	0	0	0	0	*
Inter-S semicolon	0	0	0	0	0	0	*
Inter-S empty	4	1	0	0	1	0	*
Inter-S hyphen	0	1	0	25	5	7	
Inter-C comma	4	9	2	14	3	3	
Inter-C colon	0	0	0	0	0	1	*
Inter-C semicolon	1	1	0	2	0	0	*
Inter-C hyphen	0	1	1	5	8	2	
Inter-C conj comma	0	0	1	3	0	1	*
S-initial phrase comma	3	6	6	7	1	1	
Appositive comma	0	2	2	11	6	0	
Appositive hyphen	0	0	0	8	8	8	
List comma	0	7	6	16	5	2	
List hyphen	0	0	1	2	0	0	*
S-final phrase comma	0	2	0	4	0	1	*
S-final phrase hyphen	0	0	0	4	5	2	*
Parentheses S	0	0	0	1	1	1	*
Parentheses W	0	1	0	1	3	0	*
Quote marks S	1	3	0	6	0	3	*
Quote marks W	0	0	0	2	1	2	*
Apostrophe contraction	10	5	1	11	14	8	
Apostrophe possessive	0	2	0	4	1	2	*
Apostrophe plural	1	0	0	0	0	0	*
+ or & for 'and'	1	0	0	0	3	0	*
Hyphenated W	0	1	0	0	1	0	*
Abbreviation period	0	1	0	0	19	0	
EOP period	0	0	0	1	1	0	*

Note: EOS means end of sentence; S means sentence; C means clause; W means word; EOP means end of phrase.

Again, given the standard rule about expected frequencies for the chi-square test, most categories in Table 6 are actually unusable statistically, due to low frequencies. A standard solution for this problem is to consol-

idate categories, thereby enlarging frequencies (Siegel and Castellan 1988; Woods, Fletcher and Hughes 1986). The finely tuned categories in Table 6 can be logically collapsed into: sentential boundary, clausal boundary, phrasal boundary, appositive or embedding boundary, and word-internal marks, as shown in Table 7. Probabilities are shown in Table 8, with incorrect results bolded.

Table 7 Punctuation functional categories and frequencies of five women writers from Writing Sample Database

Punctuation Boundary Function	001	009	016	018	080	QD2
Sentential	28	74	45	123	29	21
Clausal	10	15	61	20	4	15
Phrasal	3	15	34	12	13	6
Appositive	1	6	29	19	2	14
Word-Internal	12	9	15	38	1	10

Table 8 Statistical analysis of functional categories of punctuation data in five women writers from Writing Sample Database

	01/009	001/080	001/016	009/016	009/080	016/080
χ^2	9.762	2.248	22.224	47.115	4.403	20.577
p	.0207	.1338	.0001	.0001	.1106	.0001
df	3	1	3	4	2	2
	001/018	009/018	016/018	080/018		
χ^2	6.447	12.767	77.964	16.355		
p	.0398	.0125	.0001	.001		
df	2	4	4	3		
	001/QD	009/QD	016/QD	080/QD	018/QD	
χ^2	13.382	23.416	8.62	10.066	21.236	
p	.0096	.0001	.0711	.0065	.0003	
df	4	4	4	2	4	

Syntactically classified punctuation patterns correctly differentiate the questioned document from its non-authors and separate eight pairings of the five writers (for a discrimination accuracy of twelve out of fourteen or 85.7 per cent) but fail to differentiate between two pairings which both include subject 080 (009/080 and 001/080) for an error rate of 14.2 per

cent. Importantly, syntactically classified punctuation patterns correctly cluster the questioned document with its actual author. So, in handling punctuation through syntactic classification (rather than frequency), the discrimination accuracy has decreased from 93 per cent to 86 per cent while the clustering accuracy has changed from wrong to right.

The failure to differentiate between two pairings may not be a matter of the technique, however, as much as of the data. A very important aspect of this experimental result is that it shows us the kind of data that is required to perform this specific technique. When we look at the problematic results in light of the number of words in the documents, as shown in Table 9, (009/080, and 001/080), it is obvious that subject 080 has the smallest amount of text in the test samples.

Table 9 Number of words in documents in tests

Subject ID	Number of words in text 1	Number of words in text 2	Number of words in text 3	Total number of words in all texts
001	223	121	187	531
009	361	265	372	998
016	344	556		900
018	986	810		1796
080	239	93	103	345
SQD2				341

When we pair the 345-word sample of 080 with the 531-word sample of subject 001, there is not enough categorical data to differentiate the two. In fact, in order to maintain the bar of expected frequencies at 20 per cent less than 5, only two categories (sentence boundary and clausal boundary) could be used in the statistical calculation.

But it cannot be merely a matter of how many words are in the documents, because the technique differentiates SQD2, with almost the same number of words as subject 080, from subject 001. Sample size definitely provides a limit for this technique, but the similarity of the syntactically classified punctuation patterns in two authors also matters. Both subjects 009 and 016 have 900-word samples, and the technique is able to differentiate 080 from 016, but not from 009. Subjects 009 and 080 share very low use of appositive punctuation and word-internal punctuation, so that, in order to maintain the bar of expected frequencies at 20 per cent less than 5, only three categories (sentence, clausal and phrasal boundary) were used in the statistical calculation. On the other hand, subject 016 differs from both 009 and 080 in having a relatively high use of appositive punctuation. Only three categories were useable in comparing 016 and

080, while four categories were usable in comparing 016 and 009. In general, gathering more data is always the first way to handle these kinds of problem with both sample size and pairwise similarity, but in forensic situations, gathering more data may be impossible.

It is safe to conclude that, at least in this forensically similar experimental task, quantified punctuation techniques are promising, if not totally successful. The simple punctuation-frequency technique is able to differentiate between different writers most, 92.8 per cent, of the time, but fails to cluster the questioned document with the actual writer. The syntactically classified punctuation technique is able to differentiate between different writers most, 86 per cent, of the time and it also can cluster the documents of one writer.

But, in my opinion, neither measures of punctuation frequency nor measures of syntactically classified punctuation should be used alone in an actual forensic examination, because there is an error-rate (perhaps as high as 14 per cent) associated with these techniques, especially when the sample sizes are less than 250 words.

EXPERIMENTAL RESULTS: STATISTICALLY TESTABLE TECHNIQUES WITH POOR REPLICATION RESULTS

Sentential complexity

Svartvik's (1968) hypothesis that sentential complexity identifies authors exemplifies the kind of precise grounding in linguistics and statistical analysis which is essential for any method in forensic linguistics. Indeed, Svartvik's work serves as the model which all forensic linguists should emulate. This said, the sentential complexity technique showed very high error rates when it was tested on the writing samples used in these replication experiments.

Hypothesis:	Sentential complexity identifies authors. The underlying idea about language individuality in sentential complexity analysis is that some sentence structures are more complex than others and that people will differ in their ability to produce different types of sentential complexity.
Source:	Svartvik (1968).
Methodology:	Classify sentences into sentential categories. Count frequencies of each category. Test statistically.
Tools:	Knowledge of sentential syntactic categories such as simple, compound, complex, and compound-complex or Svartvik's own six clausal categories; knowledge and use of χ^2 statistic.

The results are shown in Table 10.

Table 10 Frequency data of sentence types in four women writers

Subjects	s001	s009	s016	s080	SQD2
sentence fragment	3	0	1	0	2
simple sentence	9	31	5	14	6
compound sentence	3	10	0	1	3
complex sentence	13	29	6	11	7
compound-complex	4	4	15	3	3
Total sentences:	32	74	27	29	21

Analysis

The hypothesis that patterns of sentential complexity differentiates between writers can be tested statistically, and in fact Svartvik (1968) also used the chi-square test. Assuming the null hypothesis that there is no difference between the sentential complexity patterns, what is the chance that these patterns come from the same author? The results are shown in Table 11; incorrect results are bolded. As in earlier statistical tables, we use our standard interpretation of the probabilities as 'low indicating difference, high indicating similarity' and apply this to the actual similarity/difference of authors in the document set. So bolded results mean that the probability suggests the incorrect similarity or difference that is in the actual document set.

Table 11 Statistical analysis of sentential complexity in four women writers

	001/009	001/080	001/016	009/016	009/080	016/080
χ^2	9.873	5.262	13.765	37.512	2.927	15.682
p	0.0426	0.2615	0.0081	.0001	0.4030	0.0035
df	4	4	4	4	3	4
	001/QD	009/QD	0016/QD	080/QD		
χ^2	.481	9.699	10.922	5.962		
p	0.9754	0.0458	0.0275	0.2020		
df	4	4	4	4		

These probabilities suggest that writer 016 can be clearly differentiated from the three other writers by the sentential complexity method, because the chance of there being no difference between them is so extremely low (8 in 10,000; 1 in 10,000; 3 in 10,000). Writer 009 is differentiated from 001 and 016, but not from 080. Writer 080 can only be differentiated from 016, but cannot be distinguished from 001 and 009. Writer 001 is

differentiated from 009 and 016, but not from 080. Further, the technique differentiates QD2 from only one of the non-authors (009), while it is unable to discriminate the questioned document from the other two non-authors (001 and 080). In terms of discrimination accuracy this technique is correct five out of nine times or 56 per cent ; this is a 44 per cent error rate. Finally, the sentential complexity technique fails to cluster SQD2 with its actual author, subject 016.

Svartvik's measure of sentential complexity separated relative clauses from other types of subordinate clauses and counted compound verb phrases as separate clauses. This categorization of clausal structure may not be completely defensible within current generative grammar, but it highlights the fact that different categorization and grammatical frameworks may lead to different results. Perhaps the replicability of this technique relies specifically on how we categorize sentential complexity, in which case these disappointing experimental results demonstrate that the technique warrants refinement and standardization rather than rejection.

Replication results

The hypothesis that sentential complexity patterns identify authors has failed to be replicated successfully in a forensically similar test; however, this failure to be replicated may be caused by methodological problems in determining how to categorize and count sentential complexity.

Vocabulary measures

Hypothesis:	Vocabulary richness identifies authors.
Sources:	See Holmes (1994) for review and references; Baker (1988) for PACE; Foster (1989).
Methodology:	Count number of total words in text; let N = tokens. Count number of distinct words in text; let V = types. Calculate TTR and PACE for texts of each writer. Compare each writer's TTR and PACE to each other's.
Tools:	Type-Token Ratio and Pace. $TTR = V/N$ $PACE = 1/TTR$

The results are shown in Table 12.

Table 12 Type-token ratio and pace for each writer's texts

Subject ID	Texts	TOKENS*	TYPES*	TTR	PACE
001	3	527	256	0.4858	2.0586
009	3	998	373	0.3737	2.6756
016	2	879	347	0.3948	2.5331
080	3	435	221	0.5080	1.9683
SQD2	1	341	186	0.5455	1.8333

*Note: Due to the small sizes of these texts, all texts written by the author were combined in order to count tokens and types.

Analysis

Based on comparing the similarities of these TTRs, Table 13 shows differentiating and clustering abilities of the TTR technique.

Table 13 Differentiation and cluster results of TTR technique

Author pairings	Differentiated	Clustered
001/009	yes	
001/016	yes	
001/080	no	
009/016	no	
009/080	yes	
016/080	yes	
001/QD	no	
009/QD	yes	
080/QD	no	
016/QD		no

The TTR technique is able to correctly differentiate only five of nine pairings for a discrimination accuracy of only 56 per cent ; this is an error rate of 44 per cent. Finally, the TTR technique is unable to correctly assign 016 as the author of the questioned document.

Not surprisingly, PACE (which is the reciprocal of TTR) leads to the same errors.

Replication results

The hypothesis that vocabulary richness identifies authors has failed to be replicated successfully in a forensically similar test.

Hypothesis: Hapax Legomena identify authors.
 Sources: See Holmes (1994) for review and references; Foster (1989).

Methodology: Count total number of words in text; let N = tokens.
 Count number of words occurring once in text; let $V1$ = types occurring once; that is, $V1$ equals the number of hapax legomena).
 Calculate Ratio of Hapax Legomena to Tokens (HLR) for texts of each writer.
 Compare each writer's HLR to each other's.
 Tools: Hapax Legomena Token Ratio $HLR = V1/N$.

The results are shown in Table 14.

Table 14 Hapax-Legomena-Token Ratio for each writer's texts

Subject ID	Texts	TOKENS*	V1*	HLR
001	3	527	77	0.1461
009	3	998	213	0.2134
016	2	879	214	0.2435
080	3	435	166	0.3816
SQD2	1	341	136	0.3988

Note: Due to the small sizes of these texts, all texts written by the author were combined in order to count tokens and $V1$.

Analysis

Based on comparing the similarities of these HLRs, Table 15 shows differentiating and clustering abilities of the TTR technique.

Table 15 Differentiation and cluster results of HLR technique

Author Pairings	Differentiated	Clustered
001/009	yes	
001/016	yes	
001/080	yes	
009/016	no	
009/080	yes	
016/080	yes	
001/QD	yes	
009/QD	yes	
080/QD	no	
016/QD		no

The HLR technique is able to differentiate seven of nine pairings of different authors, for a discrimination accuracy of 78 per cent, with an error rate of 22 per cent. Further, the HLR technique fails to correctly cluster the questioned document with its actual author, 016.

Replication results

The hypothesis that Hapax Legomena identify authors has failed to be replicated successfully in a forensically similar test.

Readability measures

Readability measures typically rely on sentence length and word length.

Hypothesis:	Readability measures identify authors.
Sources:	See Ellis and Dick (1996) for an example of this hypothesis; see Holmes (1994) for review and references; Foster (1989).
Methodology:	Select readability formula. Apply readability formula manually or by computer (e.g. through word processing programs).
Tools:	Compare grade level, etc. for each text to other texts. Readability formulae, possibly t-test or correlation statistics.

The results are shown in Table 16.

Table 16 Readability formulae results for each writers' text

Subjects(texts)	001(3)	009(3)	016(2)	080(3)	QD2(1)
Passive sentences	9	13	10	5	0
	0	11	41	0	
	9	11		16	
Flesch	74.5	93.1	58.0	80.7	69.9
	56.3	62.5	68.8	73.7	
	71.5	68.5		57.1	
Flesch grade level	7.5	5.6	10.5	6.9	8.4
	11.1	8.7	8.1	7.6*	
	7.8	8.1		10.8	
Flesch-Kincaid	7.9	3.3	13.6	5.4	9.0
	10.4	8.1	12.2	6.8*	
	7.5	7.1		9.5*	
Gunning-Fog	10.6	5.8	16.7	8.6	11.3
	14.3	11.7	15.3	8.3*	
	9.1	9.4		14.2	

* Note: The Microsoft Word version of these readability formulae reports that the asterisked numbers may not be reliable due to insufficient number of words in the texts.

Analysis

Table 16 shows the scores of different readability tests for each document within each author's set. There is variation among the documents in each author's set; for instance, in 001's three documents, the Gunning-Fog scores are 10.6, 14.3 and 9.1, respectively.

Ellis and Dick (1996), in their work on Civil War correspondents, compare the readability scores of different writers by the *t*-test. Using the null hypothesis that there is no difference between the readability scores of different writers, we would expect to see very low probabilities with pairs of different authors and a very high probability with the pairing of the questioned document with its actual author 016. The results of running a paired *t*-test with one-tail probability on the data in Table 16 (except the passive sentence data, since it is not an actual test score but a simple frequency) is shown in Table 17.

Table 17 Paired *t*-test with one-tail probability results for readability measures

	001/009	001/080	001/016	009/016	009/080	016/080
t	-.152	-.041	-.441	-.105	.101	.176
p	.4409	.48	.3364	.0001	.4606	.4325
df	11	11	7	7	11	7
	001/QD	009/QD	080/QD	016/QD		
t	.345	.329	.222	.012		
p	.3765	.382	.4192	.4954		
df	3	3	3	3		

The probabilities show that only one pair of writers (009/016) can be differentiated from each other because their chance of similarity is so low (1 in 10,000). The three other writers cannot be differentiated from each other; that is, the p-values for these pairings are all far above 5 per cent. Further, the questioned document is not differentiated from its non-authors; again, the probabilities for these pairings are all far above 5 per cent. The readability technique is able to differentiate only one of nine pairings of different authors, for a discrimination accuracy of 11 per cent, with an error rate of 89 per cent. But the readability technique does manage to cluster the questioned document with its actual author.

Replication results

The hypothesis that readability measures identify authors has failed to be replicated successfully in a forensically similar test.

Semantic similarities

Hypothesis:	Content Analysis identifies/discriminates between authors.
Source:	Kenneth Litkowski (communication on CORPORA electronic list)
Methodology:	Classify each word in document by semantic category. Analyse statistically the distance between documents.
Tools:	Classification scheme based on semantic categories; linear discriminant functions for statistically computing distance between documents.

This hypothesis first came to my attention in a discussion on the electronic list CORPORA in which Kenneth Litkowski said that he had heard of content analysis as a means of determining authorship. I contacted Mr Litkowski and the experimental data was exchanged. Since Mr Litkowski had previously collaborated with Professor Donald McTavish on Content Analysis, Professor McTavish ran the analysis of the test documents and returned an initial report which was forwarded to me by Mr Litkowski. Portions of this report are quoted below. McTavish's alphabetic identification of the texts is correlated to the Writing Sample Database identifiers in Table 18.

Table 18 Correlating McTavish's identification scheme with Writing Sample Database identifiers

Document ID	Topic	McTavish ID	
001-01	trauma	1	A
001-02	influence	2	B
001-03	goals	3	C
009-01	trauma	4	D
009-02	influence	5	E
009-03	goals	6	F
016-01	trauma	7	G
016-02	influence	8	H
016-03	goals	9	I
080-01	trauma	10	J
080-02	influence	11	K
080-03	goals	12	L
QD2	anger	13	M

Analysis

McTavish's analysis derives both Context-Scores (C-scores) and Emphasis-Scores (E-Scores). McTavish comments on the C-scores that 'four texts (C,F,K,L [001-03, 009-03, 080-02, 080-03]) talk about goals, four talk about terror (A,D,I,G [001-01, 009-01, 016-03, 016-01]), four talk about influential people (B,E,H,J [001-02, 009-02, 016-02, 080-01]) and one (M [QD2]) deals with anger'. The C-Scores clearly identify the semantic similarities in documents from different authors writing on similar topics, but it does not cluster documents on different topics by each author.

The Content Analysis technique relies on the idea that there is semantic similarity in one person's writing even when the person shifts from one topic to another. The following results show that this idea is at least partially true. McTavish notes that there are two distinct 'arcs' in the results. In one arc of texts (B-M-A-E), documents 001-02, QD2, 001-01 and 009-02 are clustered. Two documents from writer 001 are clustered, even though they are ostensibly about different topics. But this first arc also represents a similarity among documents from writers 001 and 009. Further, the first arc clusters QD2 with both 001 and 009. Therefore, whatever semantic similarity enables the Content Analysis to link the documents from writer 001 is not strong enough to distinguish 001's documents from 009 and QD2.

In the other arc (L-G-D-H-J) documents 080-03, 016-01, 009-01, 016-02, and 080-01 are clustered. The second arc represents a similarity between documents from writers 080, 016 and 009. Again, the C-scores can cluster some documents from authors 080 and 016, but the technique is unable to distinguish these different authors from each other and from 009.

The first arc is a conglomeration of 001/009/QD2; the second arc contains 009/016/080. These arcs differentiate the pairings 001/016 and 001/080 but they fail to differentiate the pairings 001/009, 009/016, 009/080, and 016/080. Further, the first arc shows the C-scores fail to differentiate pairings 001/QD, 009/QD although the C-scores do discriminate the questioned document from 080. The C-scores technique is able to differentiate only three of nine pairings of different authors, for a discrimination accuracy of 33 per cent, with an error rate of 67 per cent. Further, the C-scores technique does not cluster the questioned document with its actual author.

The E-Scores appear to cluster texts from all of the writers 'contextually' and two of the writers, 080 and QD2, 'conceptually'. McTavish comments on the E-scores that:

There is some patterning but it doesn't seem to connect well with discriminating authorship ... I can suggest that some texts are more different than [sic] the others (F [009-03], K [080-02], C [001-03], and

perhaps I [016-03], contextually; K [080-02], J [080-01] and M [QD2] conceptually). K seems to be the one that is different in both respects.

Again, none of the documents from different authors are differentiated. Further, the E-scores incorrectly link the questioned document with 080.

Based on these results, the Content Analysis technique has a discrimination accuracy of between 33 per cent and 0 per cent, for an error rate ranging from 67 to 100 per cent. The Content Analysis technique simply fails to cluster the questioned document with its actual author. McTavish himself recognizes that the semantic categorization of texts is not able to discriminate between authors, when he comments that 'there is some patterning but it doesn't seem to connect well with discriminating authorship'.

Replication results

The hypothesis that Content Analysis scores identify authors has failed to be replicated successfully in a forensically similar test. In forensic cases, as in this experiment, the documents share the same text-type (in this case, the narrative essay), but the documents rarely, if ever, share the same topics. It is essential to recognize that the central idea of Content Analysis for author identification purposes must be the idea that the semantic idiolect of a particular author obtains over all sorts of topically different documents. These experimental results have shown that to be true in part, because at least one author's topically different documents were able to be clustered. But these results have also shown that semantic idiolect is, first, not able to cluster all documents of each author; second, not able to differentiate documents of different authors; and third, not able to identify the questioned document with the actual author.

EXPERIMENTAL RESULTS: STATISTICALLY UNTESTABLE TECHNIQUES WITH POOR REPLICATION RESULTS

The techniques evaluated in this section – spelling errors, punctuation errors, word form errors, grammatical errors – derive from handwriting identification and prescriptive grammar. In McMenamin (1993), the techniques are described without any quantification or statistical testing. In evaluating these techniques empirically, I have attempted to quantify the results of applying these techniques as much as possible; I have also attempted to calculate error rates.

Spelling errors

The fundamental idea underlying this technique is that spelling errors are unique to individuals and constant in an individual's performance. Following are three analyses of spelling errors that contradict this idea.

Hypothesis:	Spelling errors identify authors.
Source:	McMenamin 1993.
Methodology:	List each spelling error in texts of each writer. Compare spelling patterns.
Tools:	Spellcheckers or other dictionaries; knowledge of English spelling patterns.

The results are shown in Table 19.

Table 19 Spelling errors in each writer's texts

Subjects	s001	s009	s016	s080	SQD2
Number of texts	3	3	2	3	1
Texts w/errors	2	2	2	0	1
Spelling errors	mos systematicly developement recieve uniqueness	wass	structuring nite arguement		espeically

Spelling analysis 1

Given these lists, 001 and 016 appear to be 'poor spellers' while 080 appears to be a 'good speller' and 009 is probably a 'good speller' who suffered a momentary slip of the pen. 001 texts and 016 texts share one spelling pattern: the [e] before the suffix [ment] in 001's developement and 016's arguement. 001's uniqueness also involves [e] with a suffix but this pattern cannot be related to other patterns outside the 001 set. 001 texts and SQD2 text share a mislinearization of the graphemes [c, i, e] in 001's recieve and SQD2's espeically. 016 texts and SQD2 text show no relation in spelling patterns. Other spelling errors such as 001's systematicly for systematically or mos for months or 016's structuring for structuring and nite for night cannot be related to other patterns in these documents. 009's spelling pattern cannot be related to other documents.

If an analyst relied on spelling error, he might mistakenly conclude that he was dealing with three known writers – the cluster of 001/016, 009 and 080 – rather than four known writers. But even if the four writers were known to be four different writers through external evidence, the analyst would conclude erroneously that the questioned document was authored by 001, rather than the correct conclusion that it was written by subject 016. Only 001 has a spelling error in her writing which is at all similar to the spelling error in QD2, actually authored by 016.

Table 20 shows the differentiation and cluster abilities of the spelling error technique.

Table 20 Differentiation and cluster abilities of the spelling error technique

Author pairings	Differentiated?	Clustered?
001/009	yes	
001/016	no	
001/080	yes	
009/016	yes	
009/080	yes	
016/080	yes	
001/QD	no	
009/QD	yes	
080/QD	yes	
016/QD		no

The spelling errors technique fails to differentiate 001 from 016 and from QD2. Even if we generously and dubiously allow the spelling errors technique the ability to differentiate the other documents on the basis that nothing about their spelling patterns can relate them to each other, the discrimination accuracy for the technique is 78 per cent with an error rate of 22 per cent. The spelling errors technique simply fails to cluster 016 with the questioned document.

Replication results

Based on this empirical test, the forensic stylistic technique of spelling errors has not been replicated.

Spelling errors test 2

Perhaps the spelling error technique requires more writers in the suspect set. In order to allow for this, another spelling errors test was conducted. This time the texts written by the first twenty women in the Writer Sample Database were extracted and each spelling error was listed, as shown in Table 21. The first twenty women range in age from eighteen to forty-nine, so there is less sociolinguistic control in this test.

Table 21 Spelling errors in first twenty women writers from Writing Sample Database

Subject	Spelling variants
001	mos developement systematicly recieve uniqueness
002	terifying licences behide realy registration frount wher
003	somthing wates dispite mostly
004	occured fellas alright
005	tramatic alot differend lattern constitionally beween
006	having togather collasped hospal standrads guidedd opputunities reaily individuaal attudute personaily frightnd potential feild acheive awre crimal venture'es knowing difficult
007	receiving
008	- none -
009	wass
010	- none -
011	occurring aroud prepairing prepared opressed impresionable disfunc- tional beyound habilatation lifes travisty politicaly racialy
012	wasen't, than (for then), every (for ever), reache, twelth, carrers, to (for too), awhile, alot, attitude, mothe, hopefull (for hopefully), thats (for that's),
013	deal (for dealt), pretaining, a (for at), of (for off), payed, truely, your (for you're), throw (for though), some (for so), achived, fill (for feel), than (for then), manged (for managed), biter (for better), a (for and), its (for it's), thats (for that's), motorivates, use (for us), tried (for tired), where (for we're), mine (for mind)
014	its (for it's), off (for of)
015	threw out (for throughout), threw (for through), past (for passed), realy, to (for too) noone, there (for they're), physic (for physique), goal's (for goals), their (for there) your (for you're)
016	structuring, nite, arguement, <i>in documents not used for test:</i> loose (for lose), thru, problem's (for problems)
017	tropper
018	Alot, everyonce, awhile, away (for a way), before (for be for), investorgators, towing (for towering), negitive, thats (for that's), juvieneles, sacrafised, scarafice, know (for now), appropriately, definecey, embarrassed, hormon, juvienel, cauch
019	becasue, grateful, the (for they)
020	through, abanded, thats (for that's), alot, the (for then), its (for it's)

Spelling analysis 2

Writers 002 and 011 share several, very similar spelling error patterns. These are:

1. errors with doubled consonants:

002	terifying	[terrifying]
011	occurring	[occurring]
	opressed	[oppressed]
	impresionable	[impressionable]

2. errors with doubled consonant with suffix [ly]

002	realy	[real + ly > really]
011	politically	[political + ly > politically]
	racially	[racial + ly > racially]

3. errors with vowels preceding nasal consonant

002	behide	[behind]
	frount	[front]
011	aroud	[around]
	beyound	[beyond]

The nasal consonant is dropped in 002's behide for [behind] and 011's aroud for [around]. The vowel preceding the nasal consonant is expanded in 002's frount for [front], and 011's beyound for [beyond].

4. errors with vowel [I] as in 'sit' [sIt]

002	regestration	[registration]
011	disfunctional	[dysfunctional]
	travisty	[travesty]

These spelling patterns are very similar, but they originate from two different authors. If an analyst relied on spelling errors, he would mistakenly conclude that he was were dealing with one known writer – the cluster of 002/011 – rather than two known writers. Likewise, if the common conception of 'poor spelling' is used, writers 002, 006, 011, 013, and 018 would be erroneously thought to be one writer, because these five writers are indeed 'poor spellers'. But these poor spellers are five distinct authors. Similarly, the common conception of 'good spelling' would erroneously lead an analyst to conclude that 008 and 010 are one and the same writer because they are both in fact good spellers, but two good spellers, not one. So the spelling errors technique only distinguishes thirteen of the twenty writers, for a discrimination accuracy of 65 per cent and an error rate of 35 per cent.

The frequency of spelling errors is another issue which should be considered, as Goutsos (1995) pointed out with regard to McMEnamin's (1993)

spelling-based analysis. Even errors that appear to me, subjectively, as rare, such as the behide/aroud pattern, are, objectively, not so odd that they cannot be shared, as shown by writers 002 and 011. Different writers can and do produce the same spelling errors.

Replication results

Again, based on this second empirical test, the forensic stylistic technique of spelling errors has not been replicated.

Spelling analysis 3

A subtle assumption of the spelling error technique is that an incorrect spelling is constant throughout a person's spelling. If it weren't constant, it could certainly not function as 'idiosyncratic' or indicative of authorship. But actually correct and incorrect spellings of the same word do occur in one writer's documents. Consider the data in Table 22, again drawn from the Writing Sample Database.

Table 22 Correct and incorrect spelling of same words in subject 013 texts

Word	Subject 013 errors	Subject 013 correct forms
a [at]:	I went to work <u>a</u> Mountaire	<u>at</u> this present time in my life
a [and]:	bigger <u>a</u> biter things	add to that <u>and</u> say
biter [better]:	bigger a <u>biter</u> things hope for a <u>biter</u> future	much <u>better</u>
motorivates:	daughter <u>motorivates</u> me	I'm determined, <u>motivated</u> , and

These patterns show that the same writer who misspells a word sometimes can correctly spell it other times. Spelling errors are not necessarily constant among several documents written by one author, and even on different lines within one document written by one author, as in the case of subject 013. If the analyst relied on spelling errors, the analyst might erroneously conclude that two documents were not really written by one author or that one document was authored by two writers.

Replication results

Therefore, an underlying premise for the forensic stylistics technique of spelling errors has been shown to be inaccurate, within a forensically similar test.

More on punctuation

In this section, I shall examine two hypotheses about punctuation as an individualizing characteristic of writing. These claims include the use and non-use of particular marks, and use and non-use of sets of punctuation marks. The underlying idea here is that punctuation marks are idiosyncratic.

Hypothesis:	Punctuation mark use and non-use identifies authors.
Sources:	McMenamin (1993).
Methodology:	List punctuation marks used and not-used in documents.
Tools:	The digit 1 signifies use; the digit 0 signifies non-use. Knowledge of American English punctuation marks, word processing or editing search tools.

The results are shown in Table 23.

Table 23 Punctuation use (1) and non-use (0) in five women writers from Writing Sample Database

Punctuation marks	s001	s009	s080	s016	SQD2	s018	Mark clusters
ellipsis	0	0	0	0	0	0	All writers
period	1	1	1	1	1	1	All writers
comma	1	1	1	1	1	1	All writers
semicolon	1	1	0	1	0	0	001/009/016 080/018/QD2
colon	0	0	0	1	1	0	016/QD2 001/009/080/018
apostrophe	1	1	1	1	1	1	All writers
question mark	0	0	1	1	1	1	001/009 080/016/018/QD2
exclamation point	1	1	0	1	0	1	001/009/016/018 080/QD2
hyphen	0	1	1	1	1	1	001 All others
underline	0	0	0	1	0	0	016 All others
quotation (set)	1	1	0	1	1	0	001/009/016/QD2 080/018
slash	0	0	0	0	0	1	018 All others
ampersand	0	0	0	0	0	1	018 All others
plus	1	0	0	0	0	0	001 All others
parentheses (set)	0	1	0	1	1	1	001/080 009/016/018/QD2

Punctuation analysis 1

Table 23 shows the punctuation marks that are used or not used in the documents of five women writers and the questioned document. The final column in Table 23 shows how each punctuation mark's use or non-use by the writers clusters documents. So, for instance, all the documents from all

the writers are clustered by the use and non-use patterns of the period, comma and ellipsis, while the use and non-use patterns of the semi-colon clusters documents from writers 001/009/016 on the one hand and documents from writers 080/018/QD2 on the other. It is apparent that the use or non-use of punctuation marks fails to differentiate most writers from each other most of the time. Table 24 shows the number of punctuation use/non-use features that are shared by the pairings of the five writers; all five writers share at least half of the marks with each other.

Table 24 Differentiation and cluster abilities of punctuation use/non-use technique

Author pairings	Number of 15 punctuation features shared	Differentiated?	016 clustered with QD2?
001/009	12	no	
001/016	9	no	
001/080	8	no	
001/018	7	no	
009/016	12	no	
009/080	10	no	
009/018	10	no	
016/080	9	no	
016/018	9	no	
001/QD	8	no	
009/QD	11	no	
018/QD	10	no	
080/QD	12	no	Yes?
016/QD	12		Yes?

Based on Table 24 we can calculate the discrimination accuracy of the punctuation use/non-use technique as 0 per cent, or 100 per cent error rate, in this experimental task.

When we compare the questioned document patterns to the other patterns, it is apparent that many punctuation marks' use and non-use are shared among these five women writers and the questioned document. If an analyst relied on punctuation use and non-use, they might mistakenly draw two false conclusions from the patterns in Table 23. Most conservatively, the forensic stylistics analyst might conclude that none, including the actual author, wrote the questioned document because none has the same exact pattern of punctuation use and non-use. Alternatively, the forensic stylistics analyst might conclude that either 016 or 080 authored QD2, because they both have the fewest differences – only three – from QD2. Subject 016 and QD2 differ in the semicolon, exclamation point,

and underline. Subject 080 differs from QD2 in the colon, quotation marks and parentheses. Notice that there is nothing in McMenemy's methodology which tells us that one set of three shared features is less important than another set of three shared features. With either of these conclusions, the punctuation use/non-use technique fails to cluster the questioned document exclusively with its actual author, 016.

Replication results

Based on this empirical test, the forensic stylistic technique of punctuation use and non-use has not been replicated.

Punctuation analysis 2

Exact patterns of punctuation use and non-use are shared by different writers.

The results in Table 23 suggest that punctuation use and non-use patterns may be individual, since every writer's pattern is indeed different. The punctuation patterns of forty authors from the Writing Sample Database, the first twenty women and first twenty men, were examined to discover whether exact punctuation patterns – not simply marks – can be shared. Table 25 shows that they can and are.

Table 25 Overlaps of punctuation use (1) and non-use (0) in forty writers from Writing Sample Database

Subjects	Set A	Set B	Set C
	s002/s017	s015/s034s 039/s045/s 051	s008/s033 s037
ellipsis	0	0	0
period	1	1	1
comma	1	1	1
semicolon	0	0	0
colon	0	0	0
apostrophe	1	1	1
question mark	0	0	0
exclamation point	0	0	0
hyphen	1	0	1
underline	0	0	0
quotation (set)	0	0	0
slash	0	0	0
ampersand	0	0	0
plus	1	0	0
parentheses (set)	0	0	0

In Set A, there are two possible misidentifications; in Set B, five; and in Set C, three. In total, ten of the forty authors, or 25 per cent, share their punctuation use and non-use patterns with at least one other person and so would not be differentiated by this technique. Further, Set A differs from Set B by only two features (hyphen and plus), and from Set C by only one feature (plus). Set B differs from Set C by only one feature (hyphen). Let's say that we allow for intra-author variation by allowing for differences up to certain limits, say up to three differences as we said in the earlier analysis. Then all ten of these writers would fall within these limits and be misidentified as one author based on their use and non-use of punctuation.

Replication results

The forensic stylistic technique of punctuation use and non-use has not been replicated, for a second time, in a forensically similar test.

Grammatical errors

The fundamental idea underlying the grammatical errors technique, as far as I can tell from McMenamain (1993), is that prescriptive errors are individual. This idea is essentially flawed: prescriptivism and prescriptive grammar always deal with the dialectal rather than idiolectal level of language use. This is why prescriptive grammarians can list the 'ten most frequent errors' (Berry 1971).

- Hypothesis: Grammatical errors identify authors.
- Source: McMenamain (1993).
- Methodology: List all grammatical errors in text, using school grammar.
Compare errors.
- Tools: Prescriptive grammar books, grammar checker in word processing software.

The results are shown in Table 26.

Table 26 Shared grammatical error types of four women writers

Prescriptive grammar errors	s001	s009	s016	s080	QD2	Authors not differentiated from others	Non-authors not differentiated from QD2
sentence fragment	yes	no	yes	no	yes	001/016 and 009/080	001
run-on sentence	yes	yes	yes	no	yes	001/009/016	001, 009
subject-verb mismatch	yes	no	no	no	no	009/016/080	009, 080
tense shift	yes	no	no	no	no	009/016/080	009, 080
wrong verb form	yes	yes	no	no	no	001/009 and 016/080	080
missing (aux) verb	no	yes	no	no	no	001/016/080	001, 080

The prescriptive grammarian or forensic stylistician would not be able to differentiate 001 from 009, 016 or 080; nor differentiate 009 from 016 or 080; nor differentiate 016 from 080. Further, the prescriptive grammarian or forensic stylistician would not be able to differentiate any of the non-authors from the questioned document. The discrimination accuracy is nil, while the error rate for differentiating authors is 100 per cent. But interestingly, the questioned document would, however, be correctly assigned to 016 on the basis that they share the exact pattern of errors and non-errors.

Replication results

The grammatical errors technique has failed replication of differentiating writers but has succeeded replication of clustering.

Grammatical accident

A variation of the spelling errors technique is the grammatical accident or word form error technique. In this technique the spelling errors are all related to phonetically similar word forms such as their/there/they're. Again, the idea is that these kinds of word form confusion are idiosyncratic, not shared, and constant in a writer's habit.

Hypothesis:	Grammatical accident (word form) errors identify authors.
Source:	McMenamin 1993.
Methodology:	List each word usage error in texts of each writer. Compare word usage error patterns.
Tools:	Spellcheckers or grammar checkers, word processing or editing search programs; knowledge of English vocabulary and contraction rules.

None of the four women writers in our experimental task showed any word form confusions. In order to test this hypothesis, the following data are drawn from the first twenty female writers and the first twenty male writers of the Writing Sample Database. The female writers' identifiers range from 001–020; the male identifiers range from 033–058.

The results are shown in Table 27.

Table 27 Form confusions in forty writers from the Writing Sample Database

Form confusion	Shared by writers:	Pattern (frequency)
your/you're	013	your for you're (3)
	015	your for you're (1)
	045	your for you're (2)
here/hear	011	hear for here (2)
their/there/they're	003	their for there (1)
	004	their for there (1)
	006	their for there (1)
	011	their for there (1)
	015	their for there (1)
	015	there for they're (1)
	036	there for their (1)
	036	theres for theirs (1)
	041	there for they (1)
	045	their for there (2)
	046	there for they (1)
	052	their for there (1)
	053	their for they're (1)
it's/its	005	its for it's (1)
	014	its for it's (1)
	035	its for it's (1)
	041	it's for its (1)
	052	its for it's (4)
	054	its for it's (1)
	056	its for it's (2)
than/then	011	then for than (1)
	012	than for then (4)
	013	than for then (1)
	045	then for than (1)
	046	then for than (1)
	054	then for than (1)
thats/that's	012	thats for that's (1)
	013	thats for that's (2)
	018	thats for that's (1)
	020	thats for that's (1)

Word form analysis 1

The data from forty writers demonstrates that these kinds of form confusions are not unique to individuals because they are shared among different writers, both males and females, with a wide age range, and four ethnicities (Afro-American, Euro-American, Hispanic and Native American Indian).

Three writers share the your/you're confusion. One writer from the Writing Sample Database produces hear for here, (and also uses the form here correctly sometimes). Four writers use then for than, and two writers use than for then. Seven writers share the their/there confusion; six writers use its for it's. If an analyst assumed that form confusions are indicative of authorship, he would be unable to differentiate different authors. For example, he might conclude that 011, 045, 046, 054 are all one author because they all confuse then for than. The error rates based on inability to differentiate different authors by shared word form confusions ranges from a low of 7.5 per cent (when only three of forty writers share the word confusion and would thus be indistinguishable from each other) to a high of 27.5 per cent (when eleven of the forty writers share the word form confusion and would thus be indistinguishable from each other).

Twenty of the forty writers, or a full half, make at least one type of form confusion. Of the twenty that make the form confusions listed above, one makes four types, three writers make three types, and another three make two types of confusion (and to make things even more confusing, the forms above are not the only ones that are confusable in English!). Again, if an analyst assumed that the rate of word usage errors is indicative of authorship, he would be unable to differentiate different authors. For example, he might conclude that 046 and 045 are the same author because they both make two types of form confusion.

The technique of confusable word forms is not indicative of authorship: (1) too many people confuse forms, and (2) different writers share the same rates of confusable forms (one type, two types, three types).

Replication results

Therefore, the only conclusion we can draw from Analysis 1 is that the forensic stylistic technique of word usage errors is not replicated as discriminating between authors or clustering documents of authors.

Word form analysis 2

Perhaps Analysis 1 is unfair because what McMenamain means is that the actual errors (rather than the act of the error) are significant indicators of authorship. So maybe it doesn't matter that three of forty writers make two types of form confusion. Maybe what matters is the actual errors.

In this light, consider subjects 015, 045 and 013. All three of these writers share the your/you're mistake. All three of these writers have problems with there/their/they're; 045 and 013 share the same exact error

and 015 has a related one. Subjects 045 and 013 confuse then and than (but in opposite ways). If an analyst assumed that the occurrence of similar word usage errors in different documents is indicative of authorship, he would erroneously conclude subjects 015, 045 and 013 are the same person or, at best, that 045 and 013 are the same person.

Replication results

Therefore, again the only conclusion we can draw from Analysis 2 is that the forensic stylistic technique of word usage errors is not replicated as discriminating between authors.

Word form analysis 3

The data from the Writing Sample Database also allows us to consider the problem of comparability and the subsequent problem of the calculation vs. impression of frequencies. We calculated error rate based on the percentage of writers with word form confusion, so, for instance, three of forty writers. With error rate, we are looking at how many mistakes the analyst might make if she/he is dealing with shared features in a subset of the whole set. So we can say that the analyst might cluster three of the forty writers incorrectly because three writers share a particular feature which makes them indistinguishable.

Error rate is not the same as frequency of the phenomenon within the whole dataset. At first impression, one might be tempted to infer from the data in Table 28 that the your/you're confusion is very rare because only three writers out of forty exhibit the confusion. But it is totally inaccurate to say, for instance, that only three out of forty writers made the your/you're confusion, because 30 writers did not produce your at all, 38 did not produce you're at all, and so we have no idea what their behaviour is. In order to calculate the exact frequencies, we have to take into account that some of these word forms do not occur in some writers' texts at all and we have to exclude those writers from the count. Only two writers used both forms and used the forms your/you're correctly. Only three writers used your for both forms, and none of these three writers used you're. Even if we assume that writers who confuse the your/you're forms do not produce you're consistently, we can only compare two writers to three writers. So three out of five writers producing the your/you're form confuse the forms, implying that we can expect to find the your/you're confusion 60 per cent of the time, and more writers produce the your/you're confusion than do not produce the form confusion. That means that the confusion is actually more prevalent than the correct forms. But generalizing from such a small database of five writers is statistically and linguistically dubious; with such small numbers it would not even be possible to run a chi-square test. The main point is that we cannot generalize from a database of forty writers about the frequency of a lin-

guistic pattern unless we know the behaviour of all forty with regard to this particular linguistic pattern.

Replication results

The hypothesis that form confusion (word usage) errors identify authors may be impossible to implement successfully in a forensically similar test due to the lack of comparable data and the subsequent lack of statistically analyzable datasets, but the data available suggest that these errors are very prevalent.

CONCLUSION

Table 28 shows the error rates of the techniques that have been empirically tested on writers from the Writing Sample Database.

Table 28 Discrimination and cluster error rates of language-based author identification techniques

Language-based author identification technique	Discrimination accuracy	Discrimination error rate	Assign 016 to SQD2?
Phrase Structure Analysis	100%*	0%*	yes
Syntactically Classified Punctuation	93–86%	7–14%	yes
Sentential Complexity	56%	44%	no
Vocabulary Measures – Type/Token	56%	44%	no
Vocabulary Measures – Hapax Legomena	78%	22%	no
Readability Measures	0%	100%*	no
Content Analysis	33–0%	67–100%	no
Spelling Errors 1	78%	22%	no
Spelling Errors 2: twenty writers	65%	35%	–
Punctuation Use & Non-Use	0%	100%	no
Punctuation Use & Non-Use: forty writers	75%	25%	–
Grammatical Errors	0%	100%	yes
Grammatical Accidence Errors	92–72%	8–28%	–

Note: Any 100% accuracy rate must be considered only within the context of those experiments, and subjected to futher testing.

The two techniques which use linguistic analysis are far superior to the other techniques in terms of differentiating between different authors and clustering the questioned document with the actual author. The techniques which draw from other language-related disciplines fail because they are typically not usable on the kinds of short documents that are found in forensic cases or because they simply serve another purpose which cannot be converted to authorship identification.

Finally, techniques which derive from handwriting identification and prescriptive grammar – known as forensic stylistics – fail because they rest on erroneous assumptions about individuality in linguistic performance. Further, these techniques are pernicious to the development of forensic linguistics because they seem so simple to explain to a jury that they bolster the jury's own mistakes. We can be very grateful to Judge Bassler that he recognized the scientific and empirical weakness of forensic stylistics.

There is much work to do, but these empirical evaluations point us in the right direction. To that aim, documents from the Writing Sample Database are reproduced in Appendix 3 to allow further testing and replication studies by other researchers.

APPENDIX 1: THE WRITING SAMPLE DATABASE

In order to test empirically current techniques for language-based author identification, a Writing Sample Database was first assembled, as described below. A set of four writers was extracted from the database in order to control for sociolinguistic factors which we know affect linguistic performance. This subset mimics the kind of data which are actually obtained in real casework, as described below on page 43.

The writing sample database

The database was designed to take into account both general statistical sampling issues and linguistic performance. The decision factors for the writers (or experimental subjects) included the availability of subjects; writing as normal part of the subject's lifestyle; dialect similarity or dialect grouping; generally equivalent educational level; and representation of both genders and several ethnicities. Based on these factors, writing samples were collected from two groups: Criminal Justice majors at a community college; and Business and Nursing majors at a private four-year college. Table 29 shows the sex, age and race distributions of subjects in the current Writing Sample Database.

The decision factors for the writing samples (or experimental tasks) included: genre or text-type parameters; similarity to actual types of questioned documents, e.g. suicide notes, threatening/anonymous letters, etc.; emotional level and home dialect. We know that the social context and communicative goal of a message affect its form. There are differences between the speech and the writing of each individual, differences between language behaviour at home and at work, differences between language in a letter to a friend and an essay. Based on these factors, subjects wrote, at their leisure, on ten topics, some of which are meant to elicit enough emotion to evoke the home dialect, while others are intended to elicit a more formal or workplace dialect. Topics are listed in Table 30.

Table 29 Distributions of subjects by sex, age and race

Sex male 48							
Age	unreported	To 19	20–25	26–30	31–40	41+	Totals by race
Unreported Race	2	0	0	0	0	0	2
Euro-American	3	12	14	3	2	1	35
Afro-American	0	3	0	1	0	1	5
Afro-Hispanic	0	0	0	1	0	0	1
Afro-Native Am	0	0	0	0	0	0	0
Hispanic	0	1	1	0	0	0	2
Native Am	0	3	0	0	0	0	3
Totals by age	5	19	15	5	2	2	48

Sex Female 44							
Age	unreported	To 19	20–25	26–30	31–40	41+	Totals by race
Unreported Race	1	0	0	0	0	0	1
Euro-American	0	6	5	6	5	3	25
Afro-American	0	6	5	1	3	1	16
Afro-Hispanic	0	0	0	0	0	0	0
Afro-Native Am	0	0	1	0	0	0	1
Hispanic	0	0	1	0	0	0	1
Native Am	0	0	0	0	0	0	0
Totals by age	1	12	12	7	8	4	44

Table 30 Writing topics for Writing Sample Database

Task ID	Topic
1.	Describe a traumatic or terrifying event in your life
2.	Describe someone or some people who have influenced you
3.	What are your career goals and why?
4.	What makes you really angry?
5.	A letter of apology to your best friend
6.	A letter to your sweetheart expressing your feelings
7.	A letter to your insurance company
8.	A letter of complaint about a product or service
9.	A threatening letter to someone you know who has hurt you
10.	A threatening letter to a public official (president, governor, senator, councilman or celebrity).

The testing subset

In real casework, the analyst is typically given the unknown, suspect or questioned document(s), and known writing samples from one or more potential suspects. The task is to eliminate some or all of the suspects as the possible author of the questioned document(s) and, if possible, to identify one of the suspects as the possible author of the questioned document(s). In effect, the analyst must distinguish between documents written by different writers and cluster together documents written by the same writer. Both the questioned and known documents are typically short by word count; the longest document I have ever received is a 4,000-word threatening letter. Since the investigators have already developed suspects for independent reasons in the typical case, the task of author identification in casework is circumscribed by the number of known sets, and the sociolinguistic characteristics of the known writers such as age, race, sex, and education.

The parameters of real casework have determined the design of the empirical tests. First, the task in all the empirical tests that follow is the same: to distinguish between different writers and to identify documents by the same writer, some known and one unknown, using one particular technique.

Second, the known writing samples were selected on the basis of demographic characteristics which would make the writers similar enough to qualify as a list of suspects. Also, from a theoretical perspective, we know that certain demographic characteristics affect linguistic performance, so a group of people sharing these sociolinguistically significant characteristics would very likely share dialect features. By selecting our 'list of suspects' so that they share group or dialect features, we can test a language-based identification technique's ability to go to the individual (or idiolectal) rather than group (or dialectal) level of linguistic performance. Based on both investigative practice and sociolinguistic fact, four writers were selected, from the Writing Sample Database, to form the Testing Subset. The subject identification numbers and sociolinguistic characteristics of the four writers are shown in Table 31.

Third, as in actual casework, the writing samples from these four subjects are short. The shortest text contains only 93 words, the longest, 556. Three texts were used from subjects 001, 009 and 080, while only two were used from subject 016, in order to keep the number of words from the subjects relatively comparable. In this way, subjects 001 and 080, and subjects 016 and 009, respectively, produced a comparable number of words. Since most questioned documents are short, the goal is to test techniques on short documents. In fact, it is important to develop techniques which can operate successfully on short documents, as the worst case scenario, even if long documents are available in particular cases. The textual characteristics of the Testing Subset are shown in Table 32.

Table 32 also shows the number of words in the questioned document (QD). The QD text was selected by an intern at the USFDOJ's National Institute of Justice from the documents generated by the four writers, typed into the computer, identified as SQD2. The true identity of SQD2 was not revealed until after the empirical tests were conducted. So, as the analyst, I knew that the document was authored by one of the four writers but not which one.

Table 31 Subjects in the testing subset

Subject ID	Sex	Race	Age	Educational level	Dialect information
001	F	Black	40	College 2	US Delmarva
009	F	Black	47	College 2	US Delmarva
016	F	White	40	College 1	US New England & Delmarva
080	F	White	48	College 3	US Delmarva

Table 32 Text characteristics of subjects in testing subset

Subject ID	Number of texts used	Number of words in text of topic 1	Number of words in text of topic 2	Number of words in text of topic 3	Number of words in text of topic 4	Total number of words in texts
001	3	223	121	187		531
009	3	361	265	372		998
016	2	344	556			900
080	3	239	93	103		345
SQD2	1				341	341

APPENDIX 2: THE DOCUMENTS USED IN THE EXPERIMENTAL TESTS

001-01

Giving birth to my 4th child, 3 mos too early I was in a detox center and premature labor began. First of all, I should state I was in a detox center so I could give birth to a healthy child. I was gripped with unbelievable terror at the thought that my child was coming that early I didn't feel like he would have the opportunity to survive because I was an alcoholic and a crack cocaine user through out the whole pregnancy. The hospital did what they could to save the child but because of his low birth weight and under development he didn't stand a chance. The whole ordeal took 12 hours from the onset of labor until the actual time of death and he died in my arms. I was helpless and totally powerless to do anything to help or ease his suffering. The doctors said that he didn't suffer, but really how do they know!! At that moment in time I believe I would have given my own life to save his. But now as I think who would have taken care of him or my other small children. I'm a single parent of 3 children. I believe my son gave his life so I could live and that's how I go on and stay clean and chemical free.

001-02

Numerous people and events influence me everyday in different ways. As far as me returning to school, I guess it would have to be wanting a better quality of life for my children and myself. The only way that I knew how to accomplish this is to return to school; and continue my education and show my children how important an education is now, so they don't have to wait until they are adults to get their education. Also the current job market had a high impact on my decision to get a degree, because there are no jobs available that would allow me to support my family effectively. We needed some financial security that a job at McDonald can't provide.

001-03

My Career goals is to achieve a BA in Behavioral Science Although I don't view it that way. I take it systematically one thing at a time and one step at a time. First I will receive an AA in CJ May 96. Then I plan to switch to Wilmington College where I plan to earn my BA who knows may I go further and get a MA also. I hunger for the knowledge in this field because not only do I learn of the human condition and diversity of culture, I also learn of myself and how to handle every day problems. We are all connected by some mannerism either by our uniqueness or likenesses, also there is a thin line between the two. I like knowing the why's and that there is not one answer to certain questions. The more I learn the more I

realize I don't know so it keeps me coming back. I like systematic approaches and the deviations to problems and solutions. This field has broaden my awareness that allow for trial + error. Fairness and 'that's just the way it is.'

009-01

One of the most terrifying events in my life was being held at gunpoint and told to get in the car by two men. All I could feel was dying without Christ in my life. I had a chance to run or get in the car. I was scared. I knew if I dying I would to go to hell and had not made peace with God. I am from a Christian background. So many things ran across my mind. All I could see was this big gun that looked as if it was a cannon. I got in the car, one drove and the other held the gun on me and told me not to look at them. The one guy told me if I looked he would kill me. By the way, the one that was doing all the talking didn't rape me, but made the other guy do it. I believe he was a pervert. I was too scared to cry but wanted the event to end. At that time, I lived in Baltimore and girls were being raped, killed and thrown out on the expressway or beltway. When he told me I should take you to New Jersey, I almost lost it. I remembered my background started praying. They finally let me go. He told me to get out and don't look back. I ran and ran until I reached an apartment with a light. No one would answer the door. I knocked on the door still no one would answer. I don't know how I arrived at my apartment, but I did. I jumped in the shower trying to wash his hands off but kept feeling his touch and remembering what had happened. I tried to tell my husband what happened but he was too high to listen. I didn't call the police because I felt I would be taken through the 3rd degree. I had seen it happen to too many women and nothing done. So I lived with it. I think about it sometimes now, but because Christ is in my life- that is what makes the difference! He has taken the hurt away.

009-02

I have been influenced by many people. A boss I had was very educated, independent, and aggressive. She was very successful and knew what she wanted and how to obtain it. She was a go-getter, not afraid to talk to anyone. When she appeared in a room, no matter what she was wearing, you could see the authority she had. Most women have to wear a suit to have that type of authority. My mother and father both have influenced me because they always succeeded at anything they went after. They taught me never to give up- 'a winner is not a quitter' and a 'quitter is not a winner.' Anything you strive after you can obtain, if you work hard enough. Even though they were unable to receive a proper education, they instilled in me the importance of an education. Honesty and integrity as well as respecting other feelings were also important. There are other

people who influenced me, especially those who have had great obstacles and other factors but still went on in spite of. There was a deaf lady that received a Master's Degree that influenced me because she had been a hearing person before which is much tougher than being born that way. She had developed a disease and lost her hearing but against all odds she received a Master's. According to her, she had no encouragement from outsiders but her family was very supportive. To me, this is most important. Family is an important factor in everyone life! Many more people would be successful if they only had family support.

009-03

My ultimate goal is receive a BS degree in Criminal Justice. With this degree my plans are to work extensively with juveniles and addicts. Since I have started I have mixed emotions about exactly what to do because I have found so many avenues to pursue in this field. I love people and concerned about their well being. Since I was involved in many things in the past but overcame them; I feel can be an asset to many people. Counseling has always been a desire but I had a family and they were more important at the time. People have always felt comfortable talking to me and relating their problems. I feel comfortable talking to anyone. I never been afraid to start a conversation. Therefore, counseling would be ideal. Another career goal is to own a bookstore with coffee shop (Gourmet) and a boutique. I love to shop but I hate to see too many of the same kind. Boutiques are unique, since they usually only have one or two of the same item, so much different from a department store. I would like to return to my first goal, education is priceless. Many times jobs are not obtained due to lack of education. I always have told my siblings, 'don't ever give a person an opportunity not to hire you because of the lack of education or qualification.' My oldest daughter obeyed my advice and completed. My son enlisted in the army, married and then entered workforce. Now he is pursuing his career in criminal justice. My youngest daughter has enlisted in the army after in the workforce for a few years. Maybe she will also take my advice and pursue a career and attend College. More important than all of the above I must be a success in my ministry. I would like to be a success in leading many teens, or anyone hurting, to Christ. After all is done, career, family etc. we all must give an account to Jesus as to what have we done for him and with him when He was offered to them. Our goals are only temporal to get us through this life! Most important, where will you spend eternity. God Bless!

016-01

I guess my most terrifying feeling is not being here for my two sons. My own mother died when I was 30, and I've always thought that I've sheltered and protected my sons as much, or more, than mom did me, I was

the youngest of 4, and if I left this world early, I'm not sure how my boys would function. Both emotionally and physically. Emotionally, we are a very close threesome, relying and depending almost solely on one another, with me being a focal point for problems they find themselves unable to deal with. We talk about everything together, and I always find it amazing when their peers say things like 'my mom doesn't treat me like yours'- I treat my kids as people who need structuring, raising and guidance- not as kids who 'belong' to me. I wonder -if I die- who my boys would hash over the week's happenings with. Who would they turn to for guidance and understanding- my family is of little help because I've raised my sons so differently- the boys father's family is of no help- they're far away and don't even know the two guys. Physically -my boys have been sheltered, once again- from the cruel realities of today's world. At the ages of 16 and 20, they are only now becoming financially responsible, I have raised them to respect a dollar, but they are only now beginning to learn where that dollar has to go before it can go where they want it to go. If I left my children now- they would be alone in that I have kept them mine- I have not involved them in financial matters, I have not forced them to accept and be with family members who do not see our 'way'- my kids would survive -I have taught them that- but it would not be an easy survival- I worry for them- jobs are scarce, cost of living rises more each day- Being a parent is a very real fear.

016-02

I think my mother influenced me more than anyone. As a child, we were taught a lot of values but in ways that most kids couldn't pick up on. Like -we were seldom told 'no'- we were told things like 'if you choose to do such and such, these are the results, you make your decision. As teenagers we were given the choice to hang out where we wanted, with whom we wanted but we were told things like 'if your grandmother sees you there, would she be proud and say 'hi'?' Or, 'you are who you're seen with'- We were also seldom threatened, she did just as she said she was going to do- we knew that if she said she was going to pour cold water on us next time we didn't get up out of bed on time- that is exactly what she would do- no second chance. Two stories that stick out in my mind are: she got tired of my sister and I arguing over who's turn it was to do the dishes; she said if we couldn't decide, she would solve the problem and decide for us- as kids, we never seem to learn, so the next nite, the same old s__t, and the next thing we knew- mom had opened up the window next to the table and thrown all of the dinner dishes out the window onto the lawn- she turned to us and said 'now neither one of you have to do dishes- there are none left to wash- your only problem now is to explain this to your father when he gets home' (he was a truck driver.) The other thing I remember well is: I seldom 'thought' to hang up my coat when I got home from

school, it was always laying on a chair, or on the couch, -anywhere but where it should have been- She kept telling me to take care of it -finally she told me if I didn't, she was throwing it out in the snow. Well, one morning in January, I asked her where my coat was, and, you guessed it- in the Snowbank outside the kitchen door- left there from the nite before- I was born and raised in Houlton, Maine- in January, in Maine, it's pretty damn cold- Mom taught us to stand up for our beliefs, try to walk away from an arguement, and to treat others as you want to be treated. The other two people who have influenced my life are my 2 sons- I have raised them by myself and it has been interesting, heartbreaking, thankless, and one hell of an experience. But I wouldn't trade that experience for a ship full of hundred dollar bills. They have taught me to laugh from the inside, to look at the world from the ground up, and to never loose sight of who I am and who I'll be. Having those 2 has taught me to respect my own feelings, to show them (my feelings) in a way I can be comfortable with later- and to hold onto my goals- never loose sight of the future- the past is what made us what we are today- and mom was right 'Someday I'll thank her for what she did'.

080-01

The scariest thing in my life was when the doctor told me I had to have a hysterectomy because my pap smear revealed positive cancer cells. My fear and the unknowing were awful. Would I have to have chemotherapy or radiation? Would I lose my hair. Would I die, and if so, how much would I suffer? I guess he noticed the fear in my eyes and tried to assure me that the cells were probably localized, but I was not buying this. He tried to assure me and calm my fears by stating that by removing my uterus, the cancer cells would not spread. The two weeks waiting for the surgery were hell. How would my children be if I died? Who would be there for them? I loved them so much and wanted to see them grow into adults. Most of the time I was scared- couldn't concentrate and cried when I was alone. At other times, I felt guilty for being so selfish. I would scold myself and tell myself that I had no control over this and it was out of my hands and I should just accept whatever happened. But the fear of the unknown is stronger than rational thought, and would rear its ugly head. Years later, I guess my scariest moment was unfounded- but who knows for sure? The scariest thing in my life so far has been the question of immortality.

080-02

My third grade teacher influenced me greatly. She was very intelligent, warm, and funny. She encouraged me and in so doing instilled confidence in me which up to that point was lacking. Because of her, I became a better student and proud of my accomplishments. Because of her quiet and praising manner, I loved going to school and tried harder to please her so

she would bestow her warmth and praise on me. Through her guidance, I excelled that year, and became more aware of what I could achieve if I applied myself.

080-03

My career goal is to land a position where I could become free of working two jobs as I have in the past. I would like this to be a management position as I enjoy this. In addition, I am fond of travel, so this would be an asset as I am willing to relocate. Office management or human resource management are areas of interest to me. My goal is obtaining either of these positions with a corporation providing employee benefits. Primarily, however, I am interested in a Monday to Friday job that would provide an adequate salary so I could enjoy weekends.

SQD2

A lot of things anger me but nothing makes me really angry. I've pondered this question for a couple of hours and can't come up with one single factor. I can describe lots of small, irritating examples - but no one large 'thing'. Injustice makes me angry- treating all people the same in any system- people are all different- all circumstances are different- no one person is exactly like another- stereo typing people- that makes me angry- commercials on TV that ask for money to feed starving kids over seas makes me angry (Sally Struthers looks like she could give up a meal or two)- has anyone really looked in their own neighborhood lately? What about those kids down the street? Maybe they're hungry, too. People who are capable of working but dont - or won't- make me angry- kids who say 'I can't' make me angry- people who live in perfect worlds- created by money- make me angry. Disease -espeically cancer- makes me angry. Cancer stole my mother at 52, and she never harmed a single living thing- and bore such pain, never complained- her death made me very angry- Families who don't appreciate one another make me angry. Wives who take advantage of their mate- and vice versa- make me angry. Our country's system of child support paying makes me angry- one person suffers, one person gains- and the kid gets nothing -is often the case. Or like my children- no support at all- and no help from welfare- because I lived with my parents, or because I 'make too much money:- is that after taxes? No, that's before Uncle Sam takes his share- Incompetence in the work place makes me angry. If you can;t do the job- let someone who can do it, do it- Blacks who use 'prejudice' like the term 'thank-you' make me angry. Whites who can't envision a black president make me angry. people who don't vote make me angry- Seaford's school system makes me angry. Kids who go to college and goof off, make me angry.

APPENDIX 3: DOCUMENTS FROM THE WRITING SAMPLE DATABASE FOR FURTHER TESTING

016-04

To the makers of Always Sanitary napkins. Why do you people keep changing your product? Where do you get these women who tell you what 's best and do any of them have a physical labor job working with a high ratio of men? And just where are you suppose to put the used napkin while you 're putting the new one on before you can roll the used one up in that sticky plastic? Do your test women have extra hands? Just when I thought I had finally found a pad that didn't feel like a 10 pound brick in my underwear, you jokers changed the design - for bigger and better service - O.K. - but what's with the 'Wings' - are these things suppose to fly - or serve a purpose? Not every women in America hemorrhages on a monthly basis - some of us actually have a mild flow for a couple of days, and then a spot or two - and then we're done - we d n't all bleed half to death and need a full size diaper complete with wings and wonder weave. And while I'm at it - what 's with the yard long pad? Who really needs all that length? If you really want to please women - talk to some who work side by side with men - ask them if there's a solution to how to get these things to the bathroom unnoticed - and why do the darn things smell so bad? I know some odor is bound to be present - but, geez, - some of these things knock over a Lysol can Before you wear them! And, good God, the price. Logically speaking - you know we're gonna buy 'em - so why the hell do they cost more than a 6 pack of beer? Can it really cost that much to make them? Thank God you haven't decided to recycle them, tho' - I will say that much. But that dry weave stuff - it was a good idea - but ... the length and holding thing - Don't you think you got more than a little carried away? Maybe you should stick with packaging, disguise that sucker - or something - I work closely, in a factory, with a lot of men - I don't carry a purse - where am I suppose to put that thing to get it to the bathroom? God knows, the ones from the bathroom despinser are as rough as sand paper, as thick as Black's Law Dictionary, and as comfortable as the old fashioned belt thing my mother used to wear. Come to think of it - maybe you have made some improvements, after all.

016-05

I am writing to find out why you have denied the enclosed claim for the emergency room. On Thursday, September 10, I awoke with a headache and what I assumed was an earache, caused by chronic sinus problems. I work at DuPont in Seaford, Swing shift, and I had just worked 5 nites of 12 midnite - 8 a.m., with 2 more to go. I awoke @ around 5 p.m. - not feeling well, so I took 2 tylenol, assumed I had not got my sleep out, and went back to bed. I slept, very uneasily, for another hour or so, and still

feeling bad, decided to lay down on the sofa in the living room for awhile. I got back up at around 6:30 p.m- layed on the sofa- and dozed for a while- probably another half hour - or so- by 8 o'clock, the whole left side of my face was tingly, and by 8:15- I couldn't hold my head up, I was so weak and shaking. My vision was blurred, my left side was going numb, and my fiance loaded -literally carried- me to the emergency room. He signed me in- I don't remember much about it- and the nurses took me back to the examining office- I was given a shot for pain- examined and sent home to see my dr. in the a.m- they told me at the emergency room I had an acute ear infection- all nite long, every 4 hours- I awoke in horrible pain- and begged my fiance to give me more pain killers. In the a.m., he rushed me to my allergy-sinus specialist and I was diagnosed as having a severe migraine headache- at this point- if they would have shot me, I would have been happy. My ENT doctor gave me a double shot of a pain killer, and some medicine intravenously to open up constricted vessells that cause migraines. The 2 shots worked, but I -quite literally- slept for 3 days- which is why I didn't call your office within 48 hours- and my fiance knew nothing of any of this calling business- I pay for this insurance- I was very ill- and I believe this claim should be paid, in full, with or without a phone call to your office within 48 hours. I will pursue this issue- I believe my rights have been violated- please attend to this matter immediately.

016-06

I am really sorry that things have ended up this way. I know that even tho you and Chris have your differences, you still love him- the saying of blood is thicker than water will always be true. As for he and Deena, well... I can say this- Deena is the one who holds a grudge- Chris is willing to let dead dogs lie- he'll never say he's sorry- because he meant every word he said- so he's not sorry- I am- because I didn't expect that when your daughter moved home from Florida, that I would lose my best friend. I have no guilt fellings- personally- over what happened, I was not even there and have heard only Chris' side- and part of your mom's- but I do know- if you and Chris are able to get along- regardless- why is it Deena- who also- wasn't there- is allowed to carry on so rediculously? She's living in your house, eating your food, saving no money- what makes her situation any different than Chris'? And why did she drag me into it? Because I'm his girlfriend? Or is she just being an ass? I feel bad that we re no longer friends. I miss our laughs, our shopping trips, the 4 of us- you, les, me + Chris- or even you, Deena, me + Chris- playing cards. I hate not going to visit you because Deena is rude to me- and to Chris- I hate not being able to go places with you, because Deena is rude- Do you see a pattern in this? I know that you think Chris is an ass- but ... are your daughter's actions any better? As for the way you treat Chris- he assures me it's alrite- not that he likes or accepts it- just for me to let it slide. You

know that, after three years, I'm going to take his side- right? I may not like his point of view- but I'll still back him and support him. You used to accept that- but that was pre- Deena home from Florida, I guess. I'm sorry that she has been allowed to put such a wedge in your family. And she has, you know- and I'll tell you this- I blame her for a hell of a lot more than I blame Chris- her conscience- if she has one- should bother her when she goes to bed at nite. She ain't so hot either, you know. No one is perfect- just forgiven.

016-07

I am very happy to be a part of your life. I know that it is hard for you to let anyone get close to you, I'm glad you chose me to be some one special. We have our faults and disagreements, our relationship is not perfect, and neither are we- but I think we can overcome our differences and become a couple- with patience, understanding, caring, and time. I don't expect much from you- I like to be held, hugged, and talked to- I like to know that I'm special to you- and I like to be told that you care. You should never assume that I know- because just when you assume I know- is when I 'll get insecure and feel like you don't care. And please, don't take our relationship for granted- or me- nothing is just handed to you- you have to work for everything you get- and work hard. Attend to my needs, and I'll do my best to attend to yours. Give me my space, and I'll do the same for you. Never give me an order- but always request what you feel I'm capable of. Never 'leave it for me' because I won't 'leave it for you'- I'm not that way- if the trash needs taken out- I do it- I don't 'leave it for anyone.' Make up the bed if you're the last one out of it and it bothers you that it's unmade- because I have far too much on my mind to worry about whether the bed is made or not- I'm just going to get back into it, anyway. I will always take time for you- no matter what else is pressing- or what time it is- please do the same for me- Remember that I have feelings- and I'll remember that you do, too. Never judge my emotions- lots of times I can't control them- and I encourage you to show your emotions- I won't laugh, or criticize- I promise. And it's O.K. to cry- it makes me feel better- it might help you, too. You can be strong and cry- one has nothing to do with the other. Accept my faults- and I'll accept yours- love me for who I am- not what you want me to be. Remember- our backgrounds have made us different- our ancestors have taught us different- be prepared for those differences- Once again, that's what makes us who we are. Don't be afraid to say 'I 'm sorry'- and I'll be the same way- but I won't apologize if I'm right and I know it- I'm very stubborn. I'm very spoiled- but I'm happy- and I want you to be, too.

018-01

I guess the most terrifying thing for me is the passing away of someone.

When I was growing up in a very small community in upstate- N.Y. there was not a lot do, I guess for adults there was probably even less to do than for children. My father turned to alcohol and was very abusive. My Mother, brother & myself would hide under the bed not to be found. We would have to climb out the window when he would go to the celler to get a gun. I can remember him wanting to bring his friends home from the VFW to go to bed with me and I was only 9 yrs. old. My Mother would never let him and I remember those terrible fights. Needless to say my father and I never got along very well but my Mother who is still my best friend to this day held us together. I ended up doing very poor in school and was absent a lot. I mentioned my brother who is 5 yrs. older than me. My father would get drunk and pull him out of bed and try to beat him up because my brother was very athletic and extremely smart. Again my 5 foot Mom would get between my brother and Father who was 6 foot 2 in. Many years later my parents divorced much to my happiness and my brothers dismay. My Mother met a nice Man that treats her the way she always deserved- My brother saw it for 7 very long years of truly breaking the family up. I say 7 years because he wouldn't talk to anyone in the family. He had cut off all ties. My father, besides being a heavy drinker, was also a heavy smoker. He became very ill and was in the hospital alot. I knew what my father had done in the past but he was the best father he could have been. I had always felt bad for him basically growing up with no parents at all. I would visit him, read to him, hold his hand and tell him about his grandchildren that the nurses let me sneak in to see him every once in awhile. My brother lived in Maryland and could never find time to see him, although they had become very close. I remember calling him and telling him how bad he looked but he still wouldn't come to see him. One night I visited kind of late. He looked bad and you know how they say you can smell death- well you could. I held his hand and said I love you and he said he loved me too. That was the first time in 21 yrs. he had said that to me. The next day he died. I handle bad things by sleeping at the drop of a hat -I slept a lot that Nov. and I had dreams of my father and I floating in blue skies and I was always trying to tell him, it was okay, he was dead he needed to go on. He would smile and say he knew. Those kind of dreams are different for me because ever since A child I always had true night terrors -and still do. I still wake up crying, screaming. But those about my father were okay -he seemed happy. Eight years or so later my Gramma who was 99 yrs. old passed away. She was ill for a long time. When I was growing up I would spend weekends there (at her house) to be away from my father. She would sing and make us treats. She could also be kind of mean always yelling over nothing. But she was raising a second family -her son's while he worked. When she passed away I had already been living here for almost three years. I had been back a couple times to see her. I had been playing cards with my husband when I got the call. I said thank you

and continued on playing cards. I was pretty much okay about. She had been suffering and had no idea what was going on. I did become depressed though and I couldn't really put my finger on it. In my dreams these terrible black holes represented my Grandmother. Something just felt so wrong. Going to NY I was feeling somewhat better but the closer we got to NY the more I started to cry not because of grief but because of dread. My Grandma to me seemed upset, unhappy about something. I never felt so relieved getting back to Delaware. But the night terrors seemed worse-almost like my Grandma had opened away for my father to get back. Once I woke up crying and went to go to the bathroom and thought I saw him standing in the hall way waiting for me. Recently I had a dream about my husbands Mother dying. I felt so much pain for her. Maybe I'm afraid to be alone, I don't know. My sons teacher was just killed recently in a car accident and he seems to be dealing with it okay. I am there for him at anytime if he needs to talk. All I know is these dreams about people passing or the night terrors- that I am so awake in my dreams I will get up feeling like I never slept at all and even say why are you waking me up I didnt go to bed. I tend to handle things inwardly. Maybe they come out through my dreams- I have too many people counting on me to not before them. I was sexually assaulted by an adult neighbor and I still think I see him every once in awhile even though I know he is in N.Y. Thank you for letting me get this all out!

018-02

That is a tough one. I have always admired Uncles. I will tell you about a couple of them and why they influenced me. My Uncle Bob, in N.Y. was Police of Chief and Chief of investorgators. I can remember seeing him walking the streets or as I would later find out, he was walking his beat. He was always a happy man for the job he did. He was a towing man who reminded you of a Grandfather figure. We had a local drunk who would just wander. I can remember my Uncle Bob always buying him some food and something to drink. He never expected anything in return. The N.Y. Police (actually) Saratoga Police Dept. did not pay very well at all so he couldn't do much more than that for him except give him friendship and to see that no harm came his way while he was on -and- off duty. I never really remember him talking about his job too much, He never really put much on the negitive things of his job but focused on the positive side and that's what he mostly talked about. He once told me if you really want to help make a difference- work with juvienes or become a cop that works with the community. He has since passed away but my favorite image of him is him walking his beat, stopping & talking to someone and patting them on the back. He also had his card with number on it in case someone needed him. My Other Uncle is my Uncle Jack. He is my mothers brother. We used to spend a lot of time at his house on the weekends because my

father drank. He would take us in to be safe -mentally & physically. Not to mention he was living with his mother because his wife left him and their three small daughters. When his oldest daughter was 17 she became pregnant and he helped her raise the baby. He sacrafised (sp?) as much as I have ever seen anyone scarafice. At one point he had money to buy a small piece of land and to build his own house but because of his mother didn't want to- because it was too far out in the country- he didn't and ended up buying a trailer in a trailer Park. All through this he remained a happy and up beat person. He was someone you could draw strength from. He would say- unfortunately there is always someone else worse off than you. The Other people Who Influence me know, are my Children. I have three very different Children! Sometimes I wonder if they met on the play ground if they would even become friends. Dustin my oldest is 12 1/2 yrs. old. He is probably more like me than the other two. He is very headstrong in his belief of being fair and treating each other appropriately. I'm a little more rigid though. He says I'm a left brain person- someone who is usually serious most of the time. He is much more relaxed and enjoys having fun and telling jokes. I admire that. He just seem very focused and can knows when to be silly and when not to be. He doesn't know it, but he shows me how to be a little more relaxed and maybe be a little more right brain as he puts it. My other son is Nicholas and he is 11 yrs. old. When he was 3 yrs. old- he completely stopped growing and gaining weight. That was the size he was going to be for life until one Dr. Decided to test him for a growth Hormone definecey- which I passed down to him. He went through a couple hospital stays and at the age of 5 yrs. old started growth shots. He spent a lot of time with Kids Who had cancer and would be in the playroom and would be sick- The mother of one was kind of embarrassed and said she was bring her daughter back to the room and it was Nick who said please don't do that and he went over and played with her. the way he treats others that have problems influences others not to be cruel. My youngest child is a girl. She is 9 yrs. old and her name is Jessica. Jessica also takes growth hormon shots and wears a back brace 12 hr. a day. She is 4 feet tall but only weighs about 40 lbs. She is beautiful but other kids in her class tend to pick on her but she stands her ground and tells them that it is their problem. They all influence me to see things in more than just one way and I also get terrific positive strength from them.

020-01

On the night before Christmas Eve, at 9:00 PM, the phone rang. I was living with my Mom and we were getting ready for bed- my mom was in the bathtub. We were talking about everything we needed to do to get ready for Christmas and what last minute shopping we had to do. On the phone was a nurse, saying that my sister had just been brought in to emergency room by Ambulance. She had been in a car accident and the nurse

would have the doctor call as soon as there was more information. The hospital was 1 1/2-2 hours away. Waiting for that phone call back from the doctor was horrible. All kinds of questions ran through our heads. Why was my sister there? How did the accident happen? How bad was she hurt? Would she have to stay in the hospital? Who, if any, family members should we call? My mom and I had no money. We had just gotten paid but didn't get off work in time to make it to the bank. Would we have to stay overnight? It seemed like it had been hours and doctor still hadn't called. We packed an overnight bag just in case. We called our jobs and made arrangements so we wouldn't have to be there the next day. When the doctor finally called- he said that she was going to be okay but she had some head trauma, and that she was in X-Ray and he wouldn't know any more until he got the X-rays back. He gave us directions to the hospital. Luckily mom had filled her car up with gas on her way home. On the way there, I couldn't drive fast enough. I kept thinking -what if something happened to my sister before we got there? What were we in for when we got there. I was very protective of my mom. I saw the heartache she went through when my grandmother died and I didn't want to see her go through that again- but I knew I couldn't keep her from it. The ride took forever. It was dark and cold and abandoned. No one else was on the road- and by the time we got there, it was Christmas Eve. We were in a big city that we didn't know. I didn't look like the good part of town. We found the hospital but couldn't find a parking space. We ended up parking 3 blocks away in a parking garage. Mom and I ran to a hospital door where we saw an ambulance sitting. We both stopped outside the door to get ourselves together- To try to put our fears underneath and to prepare for the worst of what we would find. We walked in the door and immediately ran into a desk. The nurse there told us where my sister was. We starting walking down the hallway, scared, frightened. Just before we opened the door to the room, we heard this voice -faint, scared, quivering, calling 'MOM!' My sister was in the waiting room outside -waiting for us. She couldn't stand being in the room by herself- beatup -scratched, stitched, and just as scared as we were.

020-02

My mother has been the most influential person in my life. I am the oldest child of 3, with a younger brother and sister. We grew up relatively poor and with an alcoholic father. I don't remember too much of my childhood until I was about 9 or 10 years old and that's when I remember my mom becoming a very strong figure in my life. We were living in a very rural area, our nearest neighbor was a quarter mile down the road. Dad started drinking a lot and spending his time away from home. Mom stayed home until my little sister started kindergarten. She found a job making \$50.00 a week, and a job I considered to be well beneath her. She saved money and

bought a car -an old dark green metallic Impala. When she she started working, I had to help out more around the house and with the kids. Mom had to work late one night and Grandmom would come over and be with us kids until Mom got home. Mom found another job and she got a job working for the state as a prison guard. She finally had a decent paying job with benefits. She managed to save enough money to hire a lawyer and file for divorce. The divorce was a tough change, but a good one. Now it was just the 4 of us and we had to really come together. We went through alot of rough times- but Mom was always there. She was there when my brother had asthma attacks and had to be taken to the emergency room. She was there on the first day of school; when we got home she bake home-made chocolate chip cookies. She was there to help us with our homework. She was at the school when we had problems there. I fell in gym class and tore the cartilage in my knee. She came to the school to pick me up, in full uniform. She was there at my brothers boy scout meetings. The only mom amongst a sea of fathers; while my father was 5 blocks away getting drunk. She was there for our band and chorus concerts. She was there during a blizzard, with no electric, no phone, no water; and the state wanting to send a helicopter to take her to work. She was always there. She never yelled, never hit us, never showed her anger, never showed her weak moments, never cried, and never gave up. My father has always said I'm just like my mother- I hope I am. She is my hero and my best friend.

020-03

I would like to someday work with children and victims. I grew up in a home environment that was poverty- level and alcohol -addicted On behalf of my father. Somehow, someway, my mom, my brother, my sister, and myself have managed to come out of those risk factors better people, good people. I want to learn how we managed to do that but more importantly, I would like to help other children and other families going through Crisis. Maybe to share some of the things I have been through. Maybe to let them know that they are not alone. that are people have been and are going through the same things. I would like to be able to show the strength in me that my mother showed of herself. I would like to be able to make a difference in a life- especially a child, the way my mom made a difference for me. I know that if I get a job doing this, its going to be hard to detach from the situation and the people. That is going to be something I'm going to have to work very hard at. Maybe I would like to teach a hard-core bone-head cop that he's making a very big impression of peoples lives and that his actions or inactions are going to be very important to someones life- if not alot of people. I wish that all officers would have some sympathy and empathy and treat people in crisis with respect. But then again, people in Crisis need to realize that an officer can only do so much within

his job. And that officers are people too and they have Crisis themselves. There are alot of things going on in the world today- poverty, addiction, homelessness, abuse, racism, prejudice, and on and on. I would like to help out some small people to deal with all that. To help them work through the maze of all stuff, including what may be going on in their life. To help some small person realize that they are a person too and need to be remembered that they are a person.

024-01

The most terrifying situation I was in was a drug raid. I went to go visit a friend I had with w/ the past summer. I did not know his house was being watched for drugs. We were all sitting in the living room. When all of sudden the police came through the door and had a shot guns to our heads yelling everyone on the floor. At first I thought it was just a good practical joke. Then I realized that it was not a joke this was for real. All I could think of is how my mother always telling me that you can always be in the wrong place at the wrong time. We were all lying on the floor and the police officers started handcuffing us. That was one of the most depressing times of my life. All I was worried about was my father finding out. The sad part was all I was doing was visiting a friend and now I'm a suspect in a drug raid. Some people who I had come w/ and I were ask if we were going to help the police w/out calling our lawers, because we had nothing on us and were not in arms reach of anything. Riding in back of the police car on the way to the station all I could think about was what if my name is printed in the newspaper? Will I get kicked out of school? Will other parents not want me hanging around their children? I know I was doing nothing wrong but in a lot of people's eyes it doesn't matter I was there and that is all I know. They would say what were you doing there anyway. I 'm sorry but not all my friends live in high class neighborhoods. I don't pick my friends on where they live. I know a lot of people who are not law abiding citizens. But that doesnt' mean they are not good people. Some people have different value beliefs from me, but that doesn't mean I 'm going to start beleaving their's. We got to the station and the worst thing they handcuffed us to a bench. I felt I had lost all my freedom for nothing. I had done nothing wrong. Then one of the first thoughts that came to my mind was if I was Black would I have been sent to SCI and not even been asked to coroperate at the town level. Well everything worked out o.k. they had nothing on us and we were let go. Comes to find out the local police messed up the whole raid b/c the warrent was not signed and on the wrong date.

024-02

Someone who influenced me is my mom. I grew up watching my mother work all the time. She work a lot so my sister and I could have eveything.

There was a lot of times we had no money but she still seem to get what we needed. I grew up my whole life with people telling me what a wonderful mother I had. I knew it was the true but I still was mad b/c we missed a lot of time together b/c she was always working or had something to do. My mom is a case manager who also used to coach special olyempics. We would try to sit down to eat dinner and the phone would ring all during dinner. I used to get really upset and ask her 'Don't they know you have a family also.' I think it would hurt her and I didn't mean to hurt her. I guess I just wanted some of her atteneetion. When I was young, I didn't realized how much she helped other people and how they depend on her so much. Till I got older and she was nonimated. for a nation wide event for an outstanding employee. Because money was low my mom had to get a second job as a foster parent. The foster child was profoundly retarded, he couldn't even hold him self up. Now it is ten years later and no amount of school could teach me what my mother has. With living with a handycapped people for majority of my life I learned to be thankful that I am not disabled and to help people who are. My mother taught me not to judge people by their race or apperence b/c you never know who they are. She told me a story one time. She was walking out of some building in Washington, D.C. and there was a elderly man dress very poorly and kind of stumbling. She said that he looked like a bum. He had a lot of books in his hands so she asked him if he needed any help. Comes to find out he was a big time scientist for the government. My mom alway would pick-up this oldder man who would always hichhicked in Lewes, and I would alway ask her 'Why do you alway pick him up.' and she tell me. 'There are alot of people who are less fercanite than you and If you can help someone you do b/c one day someone will help you.' My mom instilled alot of beliefs in me that I think children of today need. They were not religious, they were just things that often humans need to do. To this day when someone needs help, I try to help them. When I help someone, I get this glowing inside of me and It makes me feel great about myself b/c I'm giving back to the world. The most important thing she taught me is to treat people an the way you want to be treated.

024-03

My career goal is to one day have my masters in Social Work. A lot of people say why there is no money in that field. I know there is not alot of money, but I don't want to do it for the money. I like to help people and wouldn't it be the best feeling in the world to know you help to create + structure, clean-up peoples lives. To have children grow-up in homes that are friendly and can be understood in. I know it is not easy and it take a lot of emotional strain on a person. I should know I've lived with it my whole life. My mom is a case manager. She was always bring home clients. She works for DMR. So a lot of the clients were serverly handicapped, but my

mom still made them do the same stuff as any non-handicapped person would do. Sometimes people would say she was too hard on her clients but no one would push them to know their true limitations. Right now I'm going to Del tech to get my associate degree in human services. Then I'm going to probably Del State to get my BSW. Then I'll probably get a job to get some experience and work my way through my masters program.

027-01

I would have to say that the most terrifying experience of my life was the day I had to make the admission that I was desperately in need of psychiatric treatment. I had been a severe clinical depressive for approximately 10 years before I could admit to others, and most importantly myself that I had a problem I could not correct on my own. There is such a stigma attached to psychiatric treatment in our society. All I could think about were the ramifications through out the rest of my life if I sought help from a psychiatrist. I am already feeling the negative affects of my decision. I have been informed that if you have ever been under psychiatric treatment, it automatically precludes you from employment as a police officer or a probation/parole officer in this state. Basically, my nightmare is being realized now, because I cannot fulfill a lifelong goal of mine. In my mind I know that the problem I had was purely a medical one, however, government institutions view the problem as me being mentally deficient. I had a very traumatic childhood- my father was terminally ill the majority of my life, and my mother's way of dealing with that was alcoholism and abusiveness. Instead of resorting to a life of crime and using my past as an excuse for committing crimes, I decided to go for help. I will be penalized for that decision for the rest of my life, and I find the whole experience degrading and completely unfair...

027-02

My father was probably the single most influential person in my life. My father was a very hard working, dependable, honest, caring and giving man. He spent his entire life doing without so he could provide for others, including: family, friends, neighbors, acquaintances and strangers. My father lived his entire life by the adage -'what you put out into the lives of others comes back into your own'. He instilled those values in me prior to his death in 1990. My father was a very strong willed, proud man. He was very highly respected by all who knew him, and I sincerely hope that I can grow to become half as good of person as he was. When I was in high school I was useless. I did nothing. More importantly, I never accomplished anything to make my father proud, and I have a great deal of unresolved feelings about that. However, 6 months before my father passed away, he gave me the money to take two correspondence classes to acquire my high school diploma. All he asked of me was that I complete

those classes to the best of my ability. I did. However, as time went by after my father's death, I came to realize some very important things. All my father ever wanted was for me to utilize my potential. To be my own personal best and to be proud of myself. I have made the conscientious decision to come back to college and be the best I can be. To be proud of myself for quite possibly the first time in my life. And if I couldn't do it while my father was here, maybe I still can make him proud even though he has gone. He is what motivates me to face each new day, and probably will continue to do so for the rest of my life.

027-03

I intend to get into forensic psychology. Here at Del Tech I am pursuing an associates degree in criminal justice. After graduation, I intend to pursue a Bachelor's degree in Human Behaviour and a Masters degree in Psychology. Eventually I hope to obtain a Doctorate in Forensic Psychology. I have always been intrigued by the mind of a murderer. I would like to attempt to understand what it is that drives some of us to take another human life, yet prevents others of us from taking that one extra step to murder. I have done a considerable amount of research and reading on this topic, and I find that the more educated I become in this particular area, the more interested I am. This is not a career choice that you can walk right into. Therefore, I intend to start my career in Corrections counseling locally. I have deep rooted personal beliefs about our nation's corrections system. I do not feel as though there is nearly enough emphasis put on rehabilitation in our country, and there is NO emphasis at all in the State of Delaware. Rehabilitation in the State of Delaware is nonexistent. I hope to help inmates in the Delaware Corrections System in a variety of ways. I would like to help them deal with their emotions and impulse control; i.e. anger management, etc. I would like to familiarize inmates with psychiatric treatment. I would also like to familiarize them with the myriad of community resources that are available to them that the majority don't even know exist. My biggest priority is helping them prepare for the transition back into society. I feel as though we (the general public) should put forth a sincere effort to help inmates with reintegration into society. If normal, average people would start giving convicted felons jobs and opportunities and maybe even just a little faith, the recidivism rate would decrease dramatically.

ACKNOWLEDGEMENT

This research has been supported by the Office of Science and Technology, National Institute of Justice, United States Department of Justice through grants 95-IJ-CX-0012, 95-IJ-CX-0012 (S1), and 98-LB-VX-0065. Opinions expressed in this document do not necessarily reflect the official position of the United States Department of Justice. The author wishes to

thank Dr David G. Boyd, Dr Lisa Forman, Dr Richard Rau, practicing attorneys Linda M. Schuett and Glenn Martin, Professors Margaret Berger, Myrna Raeder, Andre Moenssens, Michael J. Saks, Roger Shuy, Gerald McMenam, and especially Lawrence M. Solan, for their support, comments and criticisms. Errors of fact or interpretation remain my own.

NOTES

- 1 In oral argument, the prosecution withdrew its proffer of Fitzgerald as an expert in forensic stylistics and instead proffered him as a text analysis expert; this change in title did not, however, change Fitzgerald's method based on forensic stylistics as represented by McMenam (1993).
- 2 Recent rules and rulings on the admissibility of scientific and technical evidence through the testimony of expert witnesses include Federal Rule of Evidence 702 (FRE 702); Federal Rule of Evidence 403 (FRE 403), and the empirical reliability standard set by the US Supreme Court's opinion in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 US 593 (1993) and extended in *Kumho Tire Co. v. Carmichael*, 526 US 137 (1999). FRE 702 requires that the expert must (a) qualify as an expert by knowledge, skill, experience, training or education; (b) testify to scientific, technical or other specialized knowledge; and (c) must assist the trier of fact. FRE 403 excludes testimony that may introduce the danger of unfair prejudice, confuse the issues or mislead the jury. The *Daubert* criteria include '(1) whether a method consists of a testable hypothesis; (2) whether the method has been subject to peer review; (3) the known or potential rate of error; (4) the existence and maintenance of standards controlling the technique's operation; (5) whether the method is generally accepted; (6) the relationship of the technique to methods that have been established to be reliable; (7) the qualifications of the expert witness based on the methodology; (8) the non-judicial uses of the method (*Daubert*, 509 US 595). Berger (2000) admirably explains the legal complexity of these trends in admissibility; Solan (1999) indicates how these rulings may apply to linguistic experts.
- 3 This research has been supported by grants #95-IJ-CX-0012, #97-LB-VX-0011 to the author and grant #98-LB-VX-0065 to the Institute for Linguistic Evidence, Inc. from the United States Department of Justice, National Institute of Justice, Office of Science and Technology. The opinions of the author do not necessarily reflect the official opinion of the United States Department of Justice.
- 4 ALIAS® (Automated Linguistic Identification of Authorship System) is software I developed during my tenure as a Visiting Research Fellow at the United States Department of Justice, National Institute of Justice, Office of Science and Technology. ALIAS® functions as both a database and a parser; it requires human interaction at various decision points throughout the parsing because accuracy is far more important than speed in forensic work. The software currently runs on both the Macintosh and Windows 95/98 platforms.
- 5 In a complete analysis of known and questioned documents, each document would be analysed for the syntactic patterns of all major parts of speech.

REFERENCES

- Baker, J. C. (1988) 'Pace: A test of authorship based on the rate at which new words enter an author's text', *Literary and Linguistic Computing*, 3(1): 36–9.
- Berger, M. A. (2000) 'The Supreme Court's trilogy on the admissibility of expert testimony', *Reference Manual on Scientific Evidence* (2nd edn), Washington, DC: Federal Judicial Center.
- Berry, T. E. (1971) *The Most Common Mistakes in English Usage*, New York: MacGraw-Hill.
- Chaski, C. E. (1996) 'Linguistic methods of determining authorship', National Institute of Justice Research Seminar. 49th American Academy of Forensic Sciences Meeting, Nashville, TN.
- Chaski, C. E. (1997a) 'Who wrote it? Steps toward a science of authorship', *National Institute of Justice Journal*, Washington, DC: US Department of Justice.
- Chaski, C. E. (1997b) 'An electronic parsing system for document authentication', International Association of Forensic Linguists Biannual Meeting, Durham, NC.
- Chaski, C. E. (1998a) 'Electronic parsing for idiolectal features in suspect documents', Linguistic Society of America Annual Meeting, New York.
- Chaski, C. E. (1998b) 'An automated language-based authorship system for document authentication', Questioned Documents Section, 50th American Academy of Forensic Sciences Annual Meeting, San Francisco, CA.
- Chaski, C. E. (2000) 'Linguistic Authentication and Reliability' National Conference on Science and Law Proceedings, Washington, DC: US Department of Justice, Office of Justice Programs, National Criminal Justice Reference Service NCJ 179630. (www.ncjrs.org)
- Dreher, J. J. and Young, E. (1969) 'Chinese author identification by segment distribution' in Lubomir Dolezel and Richard W. Bailey (eds) *Statistics and Style: Number Six in the Series Mathematical Linguistics and Automatic Language Processing*, New York: Elsevier.
- Ellis, B. G. and Dick, S. J. (1996) 'Who was "Shadow"? The computer knows: applying grammar-program statistics in content analyses to solve mysteries about authorship', *Journalism and Mass Communication Quarterly*, 73(4): 947–62.
- Foster, D. W. (1989) *Elegy by W.S.: A Study in Attribution*, Newark: University of Delaware Press.
- Goutos, D. (1995) Review Article: 'Forensic stylistics', *Journal of Forensic Linguistics*, 2(1): 99–113.
- Holmes, D. I. (1994) 'Authorship attribution', *Computers and the Humanities*, 28: 87–106.
- McMenamin, G. (1993) *Forensic Stylistics*, Amsterdam: Elsevier.

- Meyer, C. F. (1987) *A Linguistic Study of American Punctuation*, New York: Peter Lang.
- Nunberg, G. (1988) *The Linguistics of Punctuation*, Stanford: CSLI.
- O'Brien, D. P. and Darnell, A. C. (1982) *Authorship Puzzles in the History of Economics: A Statistical Approach*, London: Macmillan.
- Siegel, S. and Castellan, N. J. (1988) *Nonparametric Statistics for the Behavioral Sciences* (2nd edn), New York: McGraw-Hill.
- Smith, W. (1994) 'Computers, statistics and disputed authorship' in John Gibbons (ed) *Language and the Law*, New York: Longman.
- Solan, L. M. (1999) 'Can the legal system use experts on meaning?', *Tennessee Law Review*, 66: 4.
- Svarvrik, J. (1968) *The Evans Statements: A Case for Forensic Linguistics*, Stockholm: Almqvist & Wiksell.
- Williams, F. (1968) *Reasoning with Statistics: Simplified Examples in Communications Research*, New York: Holt, Rinehart and Winston.
- Woods, A., Fletcher, P., and Hughes, A. (1986) *Statistics in Language Studies*, New York: Cambridge University Press.