

THE LANGUAGE OF SUPREME COURT BRIEFS: A LARGE-SCALE QUANTITATIVE INVESTIGATION

Brady Coleman* and Quy Phung**

I. BACKGROUND AND INTRODUCTION

In 2003, we initiated a long-term project to investigate empirically the language used in United States Supreme Court briefs.¹ The exploratory stage was open-ended, largely without any particular results initially sought or predicted. We wanted to collect and categorize as much linguistic data from Supreme Court briefs as possible, and analyze such data as thoroughly as we could—and let the results lead to possible topics for publication, rather than vice versa. Indeed, at times we hoped (admittedly with quite a bit of skepticism) that we might find statistically significant differences between the linguistic styles of winning and losing briefs, and be able to offer profitable

* Professor of Human Rights, Faculty of Law, International Islamic University, Islamabad, Pakistan; M.A. (Applied Linguistics, University of Leicester 1998); J.D. (Harlan Fiske Stone Scholar, Columbia University 1989); B.A. (summa cum laude, Amherst College 1985).

** J.D. (South Texas College of Law, 2006); Ph.D. (Molecular Biology, Johns Hopkins 2000); B.A. (University of Pennsylvania, 1995).

1. This project would not have been possible without the generous funding and foresight provided by South Texas College of Law, which saw a natural link between its award-winning and top-ranked advocacy and legal writing programs, and the first-ever large-scale empirical project devoted to a linguistic analysis of written advocacy before the United States Supreme Court. Much of the funding came in the form of compensation for student research assistants to download briefs and to develop software programs designed specifically for this project. The project also won a competitive scholarship award of \$2,500 in 2007 from the Association of Legal Writing Directors which made the final stages of the project possible. We are indebted not only to STCL and the ALWD, but also to Chng Huang Hoon, Connor Graham, Joseph Kimble, Kim Seon Kyoung, Monica Ortale, Philip Page, Jim Paulsen, Lisa Sotir, Dru Stevenson, Cherie Taylor, John Worley, Roland Yap, and Andrew Solomon, and research assistants Jerry Canady, DeAnna Carlson, Zahra Jivani, Shakeb Syed, Mary Vu, Sarah Willms, and Emily Byrum, as well as those STCL faculty present at a 2005 forum to critically discuss this project.

advice for practitioners based on such information. But even without any unrealistic Holy Grail outcomes, we nonetheless were confident such a study would be able to provide useful advice to legal practitioners and educators, as well as possibly interesting outcomes for scholars of legal advocacy or linguistics. Our first publication, in the *American Journal of Trial Advocacy*,² was based on a less complete database, and was narrower in scope, because it focused on the language of only one short component of the brief, the question presented. Still, this earlier article did find interesting relationships between linguistic and other variables (time, party, and the like) in Supreme Court briefs, and concluded with advice for Court advocates.³

The scope of the current article is more extensive. Our database consists of nearly every brief on the merits presented to the Court for the thirty-five years between 1969 and 2004.⁴ We initially downloaded about 9,000 briefs, and then chopped them up for analysis into about a quarter of a million separate brief components such as Table of Contents, Table of Authorities, Summary of Argument, and the like. To clean up and analyze the briefs, eight original PERL software programs were written for this project. We decided to download every brief, rather than a smaller number based on an appropriate statistical sampling, for two reasons. For one, downloading every brief allowed us to sidestep any sampling concerns in the first place. But more importantly, although our database is comprehensive for our purposes, we were curious about how style might vary depending on a large number of legal issues, and of course even with a full set of briefs over thirty-five years, some legal issues appear rarely (or not at all).

2. Brady Coleman, Quy Phung, DeAnna Carlson, Zahra Jivani & Shakeb Syed, *Grammatical and Structural Choices in Issue Framing: A Quantitative Analysis of 'Questions Presented' from a Half Century of Supreme Court Briefs*, 29 *Amer. J. Tr. Advoc.* 327 (2005).

3. *See generally id.*

4. We excluded amicus briefs and reply briefs, and any merit briefs that were not included in WestLaw's otherwise comprehensive database. Our database ends with the 2004 term because the ALLCOURT database also ended at this time when we were conducting the research. That database has of course been updated since our research was completed.

A. Other Empirical Studies of the Language of Legal Advocacy

Our project is certainly not the first to use quantitative methods to investigate the language of written (or oral) legal advocacy. The first published work that applied computational linguistics to analyze the language of judicial briefs focused on the University of Michigan affirmative action litigation, as decided by the Supreme Court in *Gratz v. Bollinger*⁵ and *Grutter v. Bollinger*.⁶ Over a hundred amicus briefs had been filed in these companion cases, so the authors had a healthy corpus of advocacy language for analysis. Using programs that counted the appearance of key words in each brief, they were able to show that quantitative methods alone could successfully predict the policy positions that were being advocated; statistically significant differences were found in the language of amicus briefs supporting the respondent, as opposed to amicus briefs supporting the petitioner.⁷ In other work, scholars have usefully polled large numbers of active judges to ascertain what stylistic factors in appellate briefs are most favored and disfavored by decisionmakers.⁸ Empirical work has also found positive relationships between success and attorney qualifications in oral arguments before the Supreme Court of Canada.⁹

5. 539 U.S. 244 (2003).

6. 539 U.S. 306 (2003).

7. Michael Evans, Wayne McIntosh, Cynthia Cates & Jimmy Lin, *Recounting the Courts? Toward A Text-Centered Computational Approach to Understanding the Dynamics of the Judicial System* (unpublished ms., presented at Annual Meeting, Midwest Political Sci. Assn., Chicago, Apr. 9, 2005) (available at <http://www.umiacs.umd.edu/~digidock/publications/Evans-et-al-MPSA2005.pdf>) (accessed Mar. 3, 2010; copy on file with Journal of Appellate Practice and Process); see also Katherine O’Harra Vigilante, Virginia Hettinger & Christopher Zorn, *Legal Arguments and Supreme Court Decision Making: An Experimental Approach* (unpublished ms., presented at Annual Meeting, Midwest Political Sci. Assn., Chicago, Apr. 21, 2001) (copy on file with authors).

8. See e.g. David Lewis, *What’s the Difference? Comparing the Advocacy Preferences of State and Federal Appellate Judges*, 7 J. App. Prac. & Process 335 (2005); David Lewis, *Common Knowledge about Appellate Briefs: True or False?* 6 J. App. Prac. & Process 331 (2004); Charles A. Bird & Webster Burke Kinnaird, *Objective Analysis of Advocacy Preferences and Prevalent Mythologies in One California Appellate Court*, 4 J. App. Prac. & Process 141 (2002); Robert W. Benson & Joan B. Kessler, *Legalese v. Plain English: An Empirical Study of Persuasion and Credibility in Appellate Brief Writing*, 20 Loy. L.A. L. Rev. 301 (1987).

9. See John Szmer, Susan W. Johnson & Tammy A. Sarver, *Convincing the Court: Two Studies of Advocacy: Does the Lawyer Matter? Influencing the Outcomes on the Supreme Court of Canada*, 41 L. & Socy. Rev. 279 (2007). These studies did not separate

Only recently has the United States Supreme Court allowed oral argument transcripts to be released that identify a Justice by name in recording questions they pose to advocates. Previously, they had been identified generically in the transcripts (as “Justice”). With such information, we are attempting (in a separate study) to determine if the language used in oral argument questions can be quantified to predict later judicial votes. This study was initiated based on the relatively uncontroversial assumption that judicial attitudes towards litigant positions (as revealed through linguistic variables) are often well established before oral arguments are held, and so the use of language by each Justice during oral argument questioning should reveal psychological biases.¹⁰ But even before such transcripts were available, Supreme Court scholars had been using more labor intensive methods to empirically compare the language used in oral arguments to other variables.¹¹ In sum, the current project might be viewed as a natural follow up to existing work: Large scale empirical studies of the language of legal advocacy have just become achievable at low cost; until recently, our project would have been extremely time-consuming and/or very expensive. We were able to benefit from access to comparatively large (and importantly, freely provided to legal academics) databases on Westlaw, combined with computer programs we created to clean up, categorize, and analyze such briefs, to automate the processing of millions of bits of data.¹²

style from substance, so the results may be a function of argument (content) rather than language.

10. Wayne McIntosh, Michael Evans, Brady Coleman & Mary Vu, *Predicting Supreme Court Decisions from a Linguistic Analysis of Justice Questions during Oral Arguments* (forthcoming). These databases have also been mined to determine which Justice elicits the most laughter during oral argument. See Jay D. Wexler, *Laugh Track*, 9 Green Bag 2d 59 (2005).

11. See e.g. Timothy R. Johnson, *Information, Oral Arguments, and Supreme Court Decision Making*, 29 Am. Pol. Research 331 (2001); see also Sarah Levien Shullman, *The Illusion of Devil’s Advocacy: How the Justices of the Supreme Court Foreshadow their Decisions during Oral Argument*, 6 J. App. Prac. & Process 271 (2004); cf. Kathleen A. Bergin, *Sexualized Advocacy: The Ascendant Backlash against Female Lawyers*, 18 Yale J. L. & Feminism 191 (2006) (relying on empirical work to analyze communicative advocacy in extra-linguistic forms).

12. Although not concerned with legal language used persuasively, quantitative studies in the general area of legal language continue to proliferate with the explosive growth of digitized corpora. See e.g. Kevin T. McGuire & Georg Vanberg, *Mapping the Policies of*

B. Databases and Methodology

1. Databases

We began by downloading the briefs and tagging them with unique identification numbers so that we could precisely link sets of variables to each other. Our database included two categories of variables: those generated with our own software, and those imported from The Supreme Court Database, better known as ALLCOURT.¹³ The ALLCOURT database, funded by the National Science Foundation, contains final vote data from the beginning of the Warren Court (in 1953) and is updated periodically to include the current Court's last complete term.

The variables coded by the database are so specific that they allow Supreme Court scholars to study a great array of empirical issues (the relative unanimity of Court decisions under different Chief Justices; the positive or negative votes of a particular Justice when the petitioner is a consumer, a creditor, or a criminal defendant; the relative success of respondents in civil rights cases; and so on). Published work that has made use of the database (predicting, for example, outcomes of decisions based on earlier voting patterns of Justices) generally supports a view of the Supreme Court decisionmaking process as ideological rather than legal.¹⁴ In any case, we recognized that

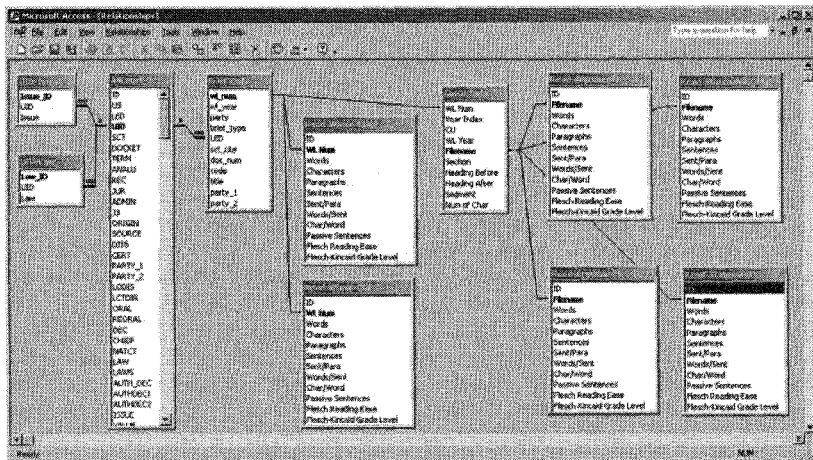
the U.S. Supreme Court: Data, Opinions and Constitutional Law, (unpublished ms., presented at Annual Meeting, Am. Political Sci. Assn., Washington, Sept. 1, 2005) (copy on file with Journal of Appellate Practice and Process) (using recent innovations in content analysis to estimate judicial ideology from an algorithm that counts certain key terms in justices' opinions); Stephen J. Choi & G. Mitu Gulati, *Which Judges Write Their Opinions (and Should We Care?)*, 32 Fla. St. U.L. Rev. 1077 (2005) (attempting to determine whether judges or their law clerks write judicial opinions by measuring stylistic variables); Brady Coleman, *Lord Denning and Justice Cardozo: The Judge as Poet-Philosopher*, 32 Rutgers L.J. 485 (2001) (comparing quantitatively measured judicial style to judicial philosophy); Martin Horwitz & Orlando do Campo, *When and How the Supreme Court Found Democracy—A Computer Study*, 14 Quinnipiac L. Rev. 1 (1994) (counting the number of times the term "democracy" appears in decades of Supreme Court opinions to form hypotheses about judicial ideology and American political history).

13. *The Supreme Court Database*, <http://scdb.wustl.edu> (noting that ALLCOURTS contains "over two hundred pieces of information about each case decided by the Court," including "the identity of the court whose decision the Supreme Court reviewed, the parties to the suit, the legal provisions considered in the case, and the votes of the Justices") (accessed Mar. 8, 2010; title page on file with Journal of Appellate Practice and Process).

14. Theodore W. Ruger, Pauline T. Kim, Andrew D. Martin & Kevin M. Quinn, *The Supreme Court Forecasting Project: Legal and Political Science Approaches to Predicting*

the database could also be very useful for a linguistic—rather than only a political science or legal—investigation, because it would allow us to compare variations in language (including readability) to the variables that the ALLCOURT database had teased out of decades of Supreme Court opinions (with significant effort and expenditure). Since the freely downloadable ALLCOURT database had already coded Supreme Court opinions as involving certain legal issues, with certain outcomes, and certain histories, we could link our database to it and then ask questions about the relationship of readability to these procedural, substantive, outcome, and other relevant variables. Thus, to offer one illustration, because the ALLCOURT database includes a code for the vote ultimately taken in each case, linking to ALLCOURT allowed us to automatically generate data comparing the readability of briefs to the vote count of the decision. Figure 1 illustrates graphically the links between different sets of variables for analysis in our MS-EXCEL relational database management system.

FIGURE 1
Screen Capture: Linked Variables



Supreme Court Decisionmaking, 104 Colum. L. Rev. 1150 (2004); Howard Gillman, *What's Law Got to Do with It? Judicial Behaviorists Test the "Legal Model" of Judicial Decision Making*, 26 L. & Soc. Inquiry 465 (2001); Harold J. Spaeth & Jeffrey A. Segal, *Majority Rule or Minority Will: Adherence to Precedent on the U.S. Supreme Court* (Cambridge U. Press 1999).

2. Methodology

After downloading the briefs, we faced two chief methodological challenges to prepare them for our own (not the ALLCOURT) database. First, we realized that any investigation into linguistic style that was going to automatically work with the millions of words in our database (to derive average sentence length, number of letters per word, and so on) would not be accurate if it attempted to include the many citations to authority that appear in Supreme Court briefs. How would our automated program count abbreviations for statutes, cases, and the like? As complete words? As something else? And how would the punctuation in string citations, for example, be interpreted for purposes of determining where sentences began and ended? Rather than struggle with this problem, particularly because it seemed irrelevant to the question of readability itself, we decided upon an automated process to eliminate the citations. But in addition to eliminating the citations, we wanted to keep a place marker for where the citations had been, so we ultimately developed a program to replace every citation with the term “scite.” Because “scite” has five letters (which is also consistently the average length of words in briefs), including it does not significantly influence our results.

The second major challenge was to develop a program to automatically separate the briefs into different components for analysis. We hypothesized that even if certain linguistic quirks were not apparent in the Argument sections of briefs, they might be revealed in more idiosyncratic patterns found in the Statement of Facts component, to provide one illustration. Or maybe information about the average length of the Summary of Argument section (which is not specified in the Supreme Court’s rules) would prove useful to advocates writing such sections, to offer another illustration.

C. Limitations

Many, if not most, modern Supreme Court briefs are the product of more than one authorial style. Several attorneys in a firm or government organization typically work on a brief to be submitted to the high court, and each attorney might draft

different parts of the document, or at least make editorial changes to a brief that was initially drafted by only one attorney. That briefs are a committee product argued against choosing them for comparative stylistic analysis, or so we were concerned, because style might be ironed out with so many contributing authors. By comparison, lower court documents (memoranda in support of motions for summary judgment, for example) would presumably reveal the style of a writer more accurately. Or, rather than choosing advocacy documents, we considered looking at law review articles, or judicial opinions, over time. But while these other variants of legal text have advantages as choices for analysis, they also have disadvantages. The authorship of judicial opinions, for example, is complicated by uncertainties about the role of judicial clerks in the writing process for some percentage of judges.¹⁵ Ultimately, however, there were two decisive reasons for choosing Supreme Court briefs over other possibilities: 1) the low cost and easy availability of Supreme Court briefs covering a fairly long time period (as opposed to lower court documents, for example, at least as of the year we began the project); and 2) our ability to add a wealth of variables to our analysis by linking our briefs to the sophisticated ALLCOURT database.

Moreover, the present article is limited to an investigation of readability scores and related calculations. Inevitably, with so much data, we were interested in a quantitative rather than a qualitative analysis, and readability scores are well-suited to such an analysis because of their algorithmic expression.¹⁶ In addition to readability, however, we were able to generate data on five common measures of plain legal English: (1) stuffy terms that often have plainer counterparts, like “abutting,” “commence,” “multitudinous,” and “necessitous”; (2) compound constructions that usually can be expressed in a plainer way, like “during the period from,” “with regard to,” “because of the fact

15. See Choi & Gulati, *supra* n. 12.

16. Cf. Feng Li, *Annual Report Readability, Current Earnings, and Earnings Persistence* (Ross School of Business Paper No. 1028, 2006) (concluding, from quantitative analysis of readability scores, that corporate executives may make annual reports difficult to read in an attempt to hide adverse information from investors) (available at <http://ssrn.com/abstract=887382> (accessed Mar. 8, 2010; copy on file with Journal of Appellate Practice and Process)).

that”; (3) redundant legal phrases like “false and untrue,” “aid and abet,” and “give, devise and bequeath”; (4) “lawyerisms” like “hereby,” “aforesaid,” and “wherein”; and (5) Latinisms like “id est,” “ad quod damnum,” and “de lege lata”.¹⁷

D. Readability Formulas: Flesch, SMOG, Fog, and Kincaid

For nearly a century, publishers, editors, and writers have developed readability formulas to derive a measure of how difficult it is to read some quantity of text. Of the dozens of such formulas that have been created, many use sentence length and word length as proxies for syntactical and semantic difficulty, while others count syllables or employ a databank of words with a previously determined level of complexity. Naturally, there are monosyllabic words and short sentences that are actually more challenging than their longer counterpoints (and vice versa), but with a large enough sample of written language, such exceptions are ironed out as a general rule. Unsurprisingly, readability scores are heavily used by publishers of academic textbooks, and several of the formulas generate a number representing the grade level required by the reader: A score of twelve would, for example, indicate reading material appropriate for high school seniors, while a score of fourteen would indicate material appropriate for college sophomores.

To take one example of these readability ratings, the formula for the Flesch Reading Ease Score test is expressed as:

$$206.835 - 1.015 \left(\frac{\text{total words}}{\text{total sentences}} \right) - 84.6 \left(\frac{\text{total syllables}}{\text{total words}} \right)$$

Note that in the results of the Flesch test, higher scores indicate material that is easier to read; lower numbers mark harder-to-read passages.¹⁸

17. We hope eventually to write a follow-up article that will present the results of these changes in lexical choice over time and as found in different brief components, and that also will indicate whether there is any relationship between their usage and the legal issue(s) in a case. We have so far been unable to get a count of passive verbs—another common measure of plain English—to work with our database, but we hope to produce such a count before we publish our next article using the same database.

18. Users of Microsoft Word 2007 can check the FRES scores of their documents by clicking the Microsoft Office button; clicking “Word Options”; clicking “Proofing”; and then selecting the check boxes for “Check grammar with spelling” and “Show readability statistics.” See Microsoft Office Online, *Help and How-to, Test Your Document’s*

The “Simple Measure of Gobbledygook” or “SMOG” formula,¹⁹ is expressed below:

$$\sqrt{\text{number of polysyllables} \times \left(\frac{30}{\text{number of sentences}} \right)} + 3$$

When the SMOG algorithm is applied to a sample, it produces a strong correlation to the grade level for which 100-percent comprehension will be achieved.

The Gunning Fog Index (“Fog”) is expressed in the following equation:

$$0.4 * \left(\left(\frac{\text{words}}{\text{sentence}} \right) + 100 \left(\frac{\text{complex words}}{\text{words}} \right) \right)$$

The Fog, like the SMOG, yields results correlated with grade levels. And it can be applied to any piece of writing—not just grade-school textbooks—with comic books scoring at a sixth-grade reading level and *The Atlantic Monthly* at the reading level of a senior in high school.²⁰

The Flesch-Kincaid (“Kincaid” or “Flesch-Kincaid”) grade level formula is written as follows:

$$0.39 \left(\frac{\text{total words}}{\text{total sentences}} \right) + 11.8 \left(\frac{\text{total syllables}}{\text{total words}} \right) - 15.59$$

Again, the result is a number that corresponds with a grade level. For example, a score of 8.2 would indicate that the text is expected to be understandable by a student in eighth grade.²¹

Readability, <http://office.microsoft.com/en-us/help/HP101485061033.aspx> (accessed Apr. 22, 2010; copy on file with Journal of Appellate Practice and Process).

19. See G. Harry McLaughlin, *SMOG Grading—A New Readability Formula*, J. of Reading 639 (May 1969).

20. Robert Gunning, *The Technique of Clear Writing* 40 (rev. ed., McGraw-Hill Book Co. 1968) (including readability chart for general-interest magazines). Readers interested in calculating the Fog Index for their own writing can apply the formula in this way:

One: Jot down the number of words in successive sentences. If the piece is long, you may wish to take several samples of 100 words. . . . Divide the total number of words in the passage by the number of sentences. This gives the average sentence length of the passage.

Two: Count the number of words of three syllables or more per 100 words. Don’t count the words (1) that are proper names, (2) that are combinations of short easy words (like “bookkeeper” . . .), (3) that are verb forms made three syllables by adding –ed or –es (like “created” . . .). This gives you the percentage of hard words in the passage.

Three: To get the Fog Index, total the two factors just counted and multiply by 4.

Id. at 38.

21. Like the FRES, the Flesch-Kincaid can be calculated automatically for documents created in Microsoft Word. See Microsoft Office Online, n. 18, *supra*.

Although readability scores were initially created to improve publishing and editorial decisions about targeting specific markets or audiences, scholars have also employed them for other purposes. A recent large-scale study of corporate reports, for example, found a relationship between the linguistic style used in the reports, and the recent financial performance of the company.²² The study analyzed a sample of more than 55,000 company reporting years using the FOG and Kincaid readability formulas, and theorized that managers, whether consciously or subconsciously, obfuscate poor financial performance with less readable annual reports.²³

Because the different formulas typically generate different scores, we compare three of them, the “Flesch Reading Ease,” the “Gunning Fog,” and the “Flesch-Kincaid,” to analyze the readability of Supreme Court briefs. Moreover, although it might appear an obvious matter, we should justify our assumption that algorithmic measures of readability are related to plain legal English. There are many possibilities about how legal language achieves plainness. Is it the level of brevity, credibility, accuracy, aesthetic appeal, comprehension—even precision²⁴—that makes legal language more or less plain? Does plainness have to do with the length of sentences or the length of words (as both are emphasized in the published readability formulas)? Or is plain legal English more concerned with the amount of jargon, Latinisms, redundant expressions, or even the size of margins on a page? Lacking one established authority that defines the characteristics of plain legal English, we consulted a recent summary of expert views.²⁵ This compilation of characteristics included advice that would not have been detected by our program (use concrete nouns as subjects, punctuate carefully) as well as elements that relate directly to readability scores (keep sentences short, cut unnecessary words).²⁶

22. Li, *supra* n. 16.

23. *Id.* at 3-4.

24. See Brady Coleman, *Are Clarity and Precision Compatible Aims in Legal Drafting?* 1998 Singapore J. of Leg. Studies 376 (1998).

25. Wayne Schiess, *What Plain English Really Is*, 9 Scribes J. Leg. Writing 43 (2004).

26. *Id.* at 71-72.

II. THE ANALYSIS: READABILITY AND FILING PARTY

Does the language of legal briefs presented to the Supreme Court, as measured using our quantitative methodology, correlate at all to the issue(s) of law being argued to the Court, or to the parties involved in the litigation? Put differently, did we find evidence to support the presumably intuitive notion that more complex areas of legal subject matter would show different linguistic patterns than less technical legal subjects? Skeptical readers must already be wondering how we could possibly distinguish some areas of Supreme Court jurisprudence—much of it necessarily involving arcane areas of Constitutional law, for example—as especially complex.²⁷ The short response is that we did not so attempt; we approached the problem from the opposite direction, at least initially. That is, we had our database generate a list of legal subjects ranked in order from the most readable to the least readable, using among other linguistic measures such markers as the average length of sentences and the total number of words. And then we considered whether the list suggested any meaningful patterns.²⁸

Overall, as the outcomes below reveal, we found that there was an interesting pattern of relationships between the variations in certain linguistic measures in the briefs, on the one hand, and three sets of variables in the ALLCOURT database, on the other hand: law, legal issue, and party.²⁹ However, we present only

27. *But see* Seth J. Chandler, *The Network Structure of Supreme Court Jurisprudence* (Univ. of Houston Law Center No. 2005-W-01) (arguing that the First Amendment is the most complex area of Supreme Court jurisprudence) (available at http://papers.ssm.com/sol3/papers.cfm?abstract_id=742065 (accessed Mar. 9, 2010; copy on file with Journal of Appellate Practice and Process)); *see also* Scott P. Johnson, *The Influence of Case Complexity on the Opinion Writing of the Rehnquist Court*, 25 Ohio N. U. L. Rev. 45 (1999).

28. These results are presented below in section II.A. And after finding suggestive patterns, we tested several hypotheses, which we present below in section II.B.

29. A subtle distinction between law and legal issue was considered necessary by the creators of the ALLCOURT database. The investigators for the ALLCOURT project ultimately responsible for the coding describe the difference between the two variables as follows:

[The legal issue] variable identifies the issue for each decision. Although criteria for the identification of issues are hard to articulate, the focus here is on the subject matter of the controversy (e.g., sex discrimination, school desegregation, affirmative action) rather than its legal basis (e.g., the equal protection clause).

results based on the party variable in this article, both because of space constraints and because the party variable was related to the law and legal issue variables in generally predictable ways. For example, if briefs written for parties involved in criminal cases tended to be written with fewer complex words, then the same was generally true when the law or legal issue involved criminal matters. Thus, to display results for each of the three variables would have been largely redundant because of the overlap between them.

A. Parties: “Bottom Up” Analysis

The ALLCOURT database codes parties according to dozens upon dozens of possibilities, including everything from the highly specific (e.g., “AMTRAK” and “SENATOR”) to the much more generic and common (e.g., “UNITED STATES,” “DEFENDANT,” and “FEMALE”). We initially ranked the two major brief components (ARGUMENT and STATEMENT OF FACTS) according to different linguistic scores to see which kinds of parties would appear as outliers, by having the highest and lowest scores for each of the linguistic categories in our database. We discovered that the readability scores were similar enough, when investigating their relationship to party, to justify our using only one of them, the Flesch-Kincaid index, as representative of the other two indices (the Flesch and the Fog).³⁰ However, in addition to the Flesch-Kincaid index, we also ranked parties according to their scores on the “average percentage of complex words” linguistic variable. This complex-

[The legal issue] variable identifies issues on the basis of the Court’s own statements as to what the case is about. The objective is to categorize the case from a public policy standpoint, a perspective that the [law variable] commonly disregards.

Unlike the [law variable] where the number of legal provisions at issue has no preordained upper bound, each legal provision should not generally have more than a single issue applied to it. A second issue should apply only when a preference for one rather than the other cannot readily be made. Of the many thousand records in the database, few have a legal basis for decision that applies to a second issue.

Harold Spaeth et.al. *Supreme Court Database Codebook* (Rel. 1 2009) 41, http://scdb.wustl.edu/_brickFiles/2009_01/SCDB_2009_01_codebook.pdf (accessed Apr. 21, 2010; copy on file with Journal of Appellate Practice and Process).

30. We use all three indices for our historical analysis in Section III, below.

words variable had analogous outcomes to the three readability formulas, but the results were different enough to justify presenting it separately.³¹ Complex words are words with three or more syllables, but our program did not count as complex any proper nouns or compound words, nor did it count common suffixes as syllables. Even after the briefs have been sorted into “PARTY” categories, there are still hundreds of records, so we present only the ten outliers on each end of the spectrum.³²

Tables 1 and 2 identify the parties with the lowest percentage of complex words, and the lowest Flesch-Kincaid scores, while Tables 3 and 4 identify parties with the highest percentage of complex words and the highest Flesch-Kincaid scores. Because our data contained separate entries for petitioner and respondent briefs, we include both categories in the tables. Thus, in Table 1, for all petitioner briefs, the party with the lowest percentage of complex words in its fact statement is “Oregon”; for respondent briefs, “Kentucky” has the lowest percentage. So, the table should not be misread into suggesting that the first two entries on the first row involved cases where Oregon and Kentucky were opposing parties in litigation; rather, these parties represent different sets of Supreme Court briefs. And in all Tables, the parties are listed in rank order: The party whose brief contained the lowest percentage of complex words in each category is listed first.

TABLE 1			
Parties with Lowest Percentage of Complex Words in Briefs			
FACTS		ARGUMENT	
Petitioner	Respondent	Petitioner	Respondent
Oregon	Kentucky	Debtor	South Carolina
Arrestee	Arizona	Arrestee	Oregon
Poor Defendant	Louisiana	Oklahoma Commn.	Minority Person
Father	South Carolina	Inventor	U.S. Judge
Juvenile	Wisconsin	Washington Dept.	Wisconsin
Michigan	Florida	Texas School Dist.	Union Member
Texas	Texas	Reporter	Taxpayer
Inventor	Georgia	Minority Job Applicant	Debtor
Out of State	Virginia	Farmer	Georgia
Person Convicted	Mother	Minority Student	Mississippi

31. See pp. 88-90, *infra*.

32. We ignored outcomes that were based on a small sample size of fewer than six briefs.

TABLE 2			
Parties with Lowest Flesch-Kincaid Scores in Briefs			
FACTS		ARGUMENT	
Petitioner	Respondent	Petitioner	Respondent
South Dakota	Arizona	OWCP ³³	Arrestee
Poor Defendant	South Carolina	Voter	South Dakota
Texas City	Wisconsin	Truck	Mine
Missouri	Arrestee	Business	Texas City
Massachusetts Official	Louisiana	Dept. of Justice	Protestor/Picketer
Student	OWCP	Seller	State Commission
Bankrupt	Govt. Contractor	Arizona	Airline
Developer	Child	Illinois Dept.	Out of state
Massachusetts	Dept. of Justice	Minority Student	Nuclear
Spouse	Maryland	Wisconsin	South Carolina

Tables 3 and 4, which show the parties whose briefs were the least readable and contained the most complex words, are strikingly dominated by federal government agencies as parties, or parties typically considered to have more resources (“shipper,” “railroad,” “drug manufacturer”). Tables 1 and 2, on the opposite end of the linguistic spectrum, are heavily overrepresented by parties (1) with presumably fewer resources, (“inventor,” “debtor”), (2) who are members of minority groups (“minority job applicant,” “minority student”), or (3) who are criminal defendants (“arrestee,” “person convicted”). Curiously, moreover, southern states as parties are especially prevalent in Tables 1 and 2. There are scattered exceptions: the Department of Justice appears twice in Table 2 as a party; Table 4 includes both “minority student” and “disabled” as parties. But the basic pattern was strong enough to suggest hypotheses for us to test in a “top down” analysis.

TABLE 3			
Parties with Highest Percentage of Complex Words in Briefs			
FACTS		ARGUMENT	
Petitioner	Respondent	Petitioner	Respondent
FEC	FCC	FEC	FLRA
Television	Public Utility	FPC	FEC
Dept of State	State Commission	AGRI	FERC
HEW	Insurance Company	INS	Public Utility
Dept. of Trans.	Reporter	DOS	FCC
INS	FERC	FCC	GSA

33. Office of Worker’s Compensation Programs.

TABLE 3 (continued)			
Parties with Highest Percentage of Complex Words in Briefs			
FACTS		ARGUMENT	
Petitioner	Respondent	Petitioner	Respondent
FERC	FLRA	LBR	Hospital
LABR	Hospital	Attorney General	Reporter
Kentucky Official	Iowa Commission	Dept. of Trans.	HHS
Bookseller/Printer	Medical Claimant	HEW	COMN

TABLE 4			
Parties with Highest Flesch-Kincaid Scores in Briefs			
FACTS		ARGUMENT	
Petitioner	Respondent	Petitioner	Respondent
Shipper	Minority student	FPC	Bookseller/Printer
Father	Offeror	Los Angeles Official	Oklahoma
Seller	Parent	Michigan	Gay
Railroad	State Commission	Heir	Disabled
Fishing entity	FCC	California County	Hospital
New Jersey	Theater	Television	FERC
Union Member	New York	California Commission	Public Utility
Illinois City	Insurer	Drug Manufacturer	HEW
California County	Public Utility	Network	Govt. Employee
Builder	School	Public Utility	Wife

B. Parties: Top Down Analysis

We tested five different hypotheses, derived from our initial “bottom up” analysis, and approaching the “parties” variable from the opposite direction to confirm apparent patterns. Again, in the “bottom up” analysis, above, we queried our database to rank parties in ascending order based on the numerical value of linguistic variables (such as our three readability indices). For the “top down” analysis, on the other hand, we identified parties that fell within a certain area of interest, and identified which of those parties fell below or above the median for all parties for a particular linguistic variable. For example, half of all parties will fall below the median score on each of our readability indices, necessarily, and half will fall above the median. Our test for significance was relatively straightforward, if difficult to satisfy: If all of the parties in a certain category fell consistently above the median (or below, as the case may be), then we considered the outcome meaningful, if we had a large enough sample size of briefs for such party.

Hypothesis 1: Parties involved in the criminal justice system are more likely to have briefs written in a less complex style.

Our list of parties involved in the criminal justice system, taken from the codes in the ALLCOURT database, includes: **AC** (person accused, indicted, or suspected of a crime), **ARRESTEE** (arrested person, or pretrial detainee), **CC** (person convicted of crime), **CRIM INS** (person allegedly criminally insane or mentally incompetent to stand trial), **POOR D** (indigent defendant), **PRISONER** (prisoner, inmate of penal institution) and **PROBATIO**, (probationer, or parolee). We put in boldface all parties with briefs with significantly less complex language. We found that every party involved in a criminal matter, except probationers and parolees, had briefs written in a significantly less complex style. The average scores for probationers and parolees were sometimes above the median, and sometimes below it, depending on which linguistic variable we were looking at, and whether they were respondents or petitioners. In sum, our hypothesis was proven.

*Hypothesis 2: Parties with presumably fewer financial resources are more likely to have briefs written in a less complex style.*³⁴

Our list of parties with presumably fewer financial resources, taken from the codes in the ALLCOURT database, includes: *ALIEN* (alien, person subject to a denaturalization proceeding, or one whose citizenship is revoked), *DISABLED* (disabled person or disability benefit claimant), *HANDICAP* (handicapped individual, or organization of devoted to), *ICMP* (involuntary committed mental patient), *INDIAN*, *POOR* (indigent, needy, welfare recipient), *RETARDED* (retarded person, or mental incompetent), and *UNEMPLOY* (unemployed person or unemployment compensation applicant or claimant). We put in boldface all parties with briefs with significantly less complex language. We put in italics all parties

34. We were not able to conduct an independent statistical analysis to determine whether these parties had fewer financial resources as a matter of fact. In other words, we selected these parties based solely on anecdotal information. We also considered using a separate group of parties “with presumably greater financial resources,” but this turned out to be too slippery a category.

with briefs with significantly more complex language. The parties listed above that are neither in boldface nor italics did not show a significant placement above or below the median for our linguistic variables. In sum, our hypothesis was not proven.

Hypothesis 3: The federal government as a party is more likely to have briefs written in a more complex style.

Our list of parties as federal governmental bodies (or persons so representing), taken from the codes in the ALLCOURT database, includes: **AF** (Secretary or administrative unit or personnel of the U.S. Air Force), **AG** (attorney general of the United States, or his office), **AGRI** (Department or Secretary of Agriculture), **CIA**, **CPSC** (Consumer Products Safety Commission), **DOD** (Department or Secretary of Defense), **DOS** (Department or Secretary of State), **DOT** (Department or Secretary of Transportation), **EEOC**, **EPA**, **FBI**, **FCC**, **FEC** (Federal Election Commission), **FERC** (Federal Energy Regulatory Commission), **HEW**, **HHS** (Department or Secretary of Health and Human Services), **INS**, **IRS**, **LBR** (Department or Secretary of Labor), and **NLRB** (National Labor Relations Board, or regional office or officer). This list includes in boldface all parties with briefs using significantly more complex language. The only party in this group filing briefs written in a less complex style was the NLRB. In sum, our hypothesis was overwhelmingly supported.

Hypothesis 4: Southern states are more likely to have briefs written in a less complex style.

The states shown in boldface below filed briefs written in a significantly less complex style. The states in italics had briefs written in a significantly more complex style. Not every state is listed because some states did not have a high enough numerical representation in our database (i.e., they had five or fewer briefs), or they had enough briefs to represent a strong sample, but had mixed results (i.e., with some scores above the median and some below). **ALABAMA**, **ARKANSAS**, **ARIZONA**, **CALIFORNIA**, *CONNECTICUT*, **GEORGIA**, **ILLINOIS**, *INDIANA*, **MASSACHUSETTS**, **MARYLAND**, **MAINE**, *MINNESOTA*, **MISSOURI**, **MONTANA**, *NEW JERSEY*, *NEW*

YORK, OHIO, OKLAHOMA, OREGON, PENNSYLVANIA, SOUTH DAKOTA, TENNESSEE, WASHINGTON, and WISCONSIN. Fifteen states are in boldface, and nine states are in italics. The boldface states (with less complex language) include states that represent a range of geographical areas in the United States. Nonetheless, of those nine states with more complex language, all are “northern” states in a compass-direction sense of the word. We may in a later article lay out the exact linguistic averages on a shaded map of the United States to get a more accurate understanding of the relationship between geography and linguistic complexity, but that exercise is beyond the scope of this article. In sum, our hypothesis was neither proven nor disproven, but our results suggest room for further analysis.

Hypothesis 5: The gender or race of a party was related to the complexity of the language in their briefs.

Our list of parties categorized based on gender or race, taken from the codes in the ALLCOURT database, includes: GOFEE (female governmental employee or job applicant), GOMEI (minority governmental employee or job applicant), FATHER, **FEE** (female employee or job applicant), MEE (racial or ethnic minority employee or job applicant), **MFEE** (minority female employee or job applicant), MOTHER, **RAMI** (racial or ethnic minority), RAMIPROT (person or organization protesting racial or ethnic segregation or discrimination), RAMISTU (racial or ethnic minority student or applicant for admission to an educational organization), and WIFE. We put in boldface all parties with briefs with significantly less complex language. Parties without boldface did not have enough briefs to represent a legitimate sample size (or had no briefs at all). In sum, our hypothesis was neither proven nor unproven, because of sampling deficiencies.

Unique linguistic patterns for certain categories of parties can be seen as the result of differences in two more fundamental variables: (1) brief author and (2) brief subject matter (both legal and factual). For criminal matters, both of these variables are likely to influence linguistic outcomes. The authors of briefs in criminal cases may be more likely to lack legal (or any advanced) education, because such authors are sometimes the

parties themselves (e.g., when a prisoner writes a brief from prison). Moreover, when parties involved in criminal matters do rely on counsel to write their briefs, they may require court-appointed counsel, or lawyers from firms that focus on criminal cases, either of whom may bring linguistic patterns to their writing that are not apparent for other brief-writers. The subject matter of criminal cases, especially as shown in the statement of facts, may be more likely to show less linguistic complexity, because of the subject matter of criminal fact patterns.

Where the party is a person, agency, or department of the federal government, the Office of the Solicitor General writes the brief in any litigation before the Supreme Court.³⁵ As our results clearly show, such briefs are written in a significantly more complex style. These results mirror the outcomes found when the “Question Presented” (QP) component of the brief is analyzed for linguistic patterns: The Office of the Solicitor General has a significantly different QP style than all other brief writers.³⁶

In addition to the variables of party, law, and legal issue, we looked for relationships between further variables in the ALLCOURT database, and the linguistic variables we derived from the Supreme Court briefs. More specifically, we compared the readability scores (as well as other foundational data we counted, e.g., the total number of words, sentences, complex words, etc.) of briefs written: (1) for petitioners as opposed to respondents;³⁷ (2) for the party that ultimately won (before the Supreme Court), as opposed to the party that lost; (3) in cases that were decided by a five-four split (as a possible measure of how controversial the litigation had been) as opposed to a unanimous vote; and (4) according to a range of other non-substantive elements available to us in the ALLCOURT database, (the Justice who ultimately authored the opinion, for example). We found no significant relationship between our

35. The Office of the Solicitor General is tasked to conduct all litigation on behalf of the United States in the Supreme Court, and to supervise the handling of litigation in the federal appellate courts. 28 CFR 0.20.

36. Coleman et al., *supra* n. 2, at 332, 335, 340-43, 346-47, 350.

37. Usually it was easy to separate petitioners from respondents, although some cross appeals (in which both sides are petitioners as well as respondents) and direct, non-certiorari, appeals to the Court (where the parties are labeled “appellant” and “appellee”) complicated matters somewhat.

linguistic variables and any of these other factors; this did not surprise us, but it was relatively straightforward to conduct a comprehensive set of queries after the more laborious task of database creation was complete. In sum, the only meaningful variation in linguistic variables related to party, law, and issue—as well as time, which we discuss next.

III. THE ANALYSIS: READABILITY AND TIME

In the following series of graphs, we portray changes in our readability formulas over time, for each of the two longest brief components: the Argument and the Statement of Facts.³⁸ Necessarily, we tried to imagine as many confounding variables as possible that would lead to only the appearance of a broad based increase in the use of plain English (as manifested in readability scores) over time. For example, we considered whether our results were influenced by changes in Supreme Court rules, or were skewed by historical changes in the language of Supreme Court briefs as represented in the percentage of 1) footnotes, 2) citations, 3) cases involving certain areas of law (over others), and so on. We were not able to completely eliminate the possibility of confounding variables, but we are certain that our methodology (using different linguistic measures, and the like), minimized any effect.³⁹

Figures 2A, 2B, and 2C graph the average readability scores of the Argument component of all Supreme Court briefs written on the merits from 1970 to 2004.

38. Both the Summary of Argument and the Conclusion showed no significant changes over time in readability. Other components (for example, the Table of Contents and the Table of Authorities) do not lend themselves to a readability analysis, for obvious reasons.

39. As the previous section found a relationship between certain parties and certain legal issues on the one hand and readability on the other, it is natural to wonder if any apparent changes over time in readability are actually functions of shifts in the percentage of certain parties and issues over time. A representative sampling look at our specific data, however, did not support this theory. Recall that for the general historical shift in readability, the number of briefs we are using is much larger than the number of briefs (per issue or per party) in the previous section. So the overall historical pattern should smother all but the most pronounced shift in parties and legal issues over time, and we did not find such an obvious shift.

FIGURE 2A
Average of Fog for Argument—PERL

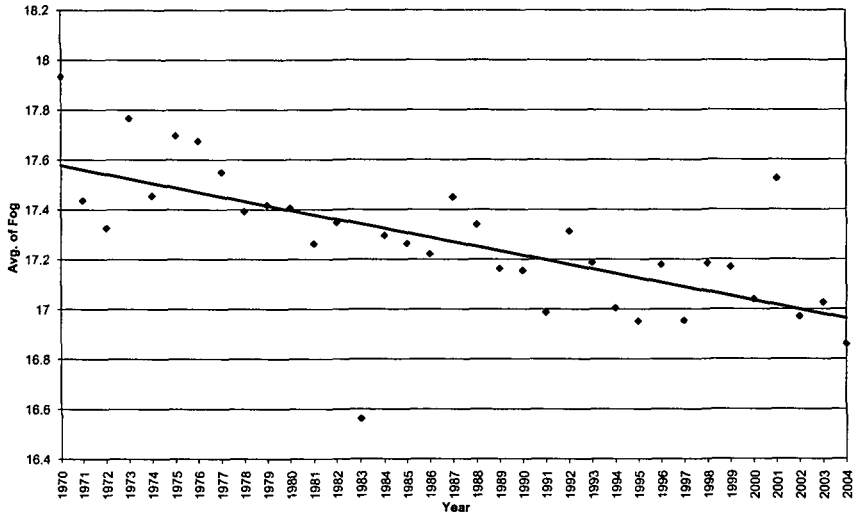


FIGURE 2B
Average of Flesch-Kincaid for Argument—PERL

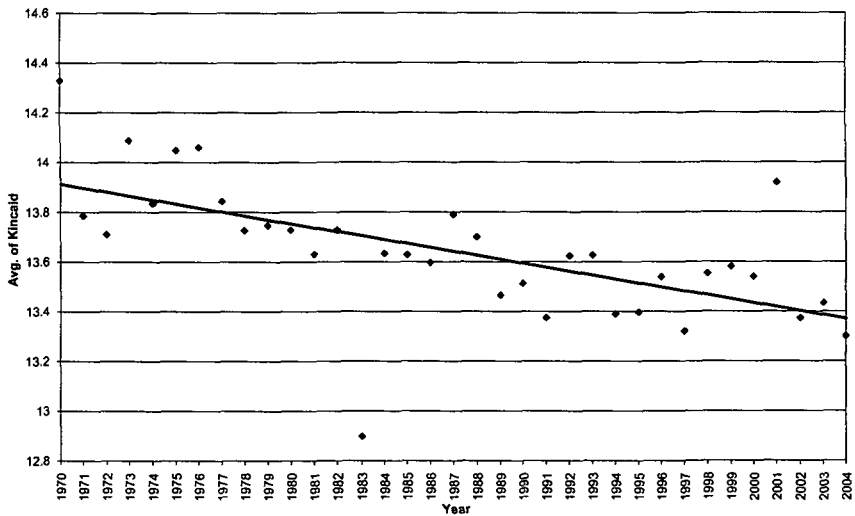
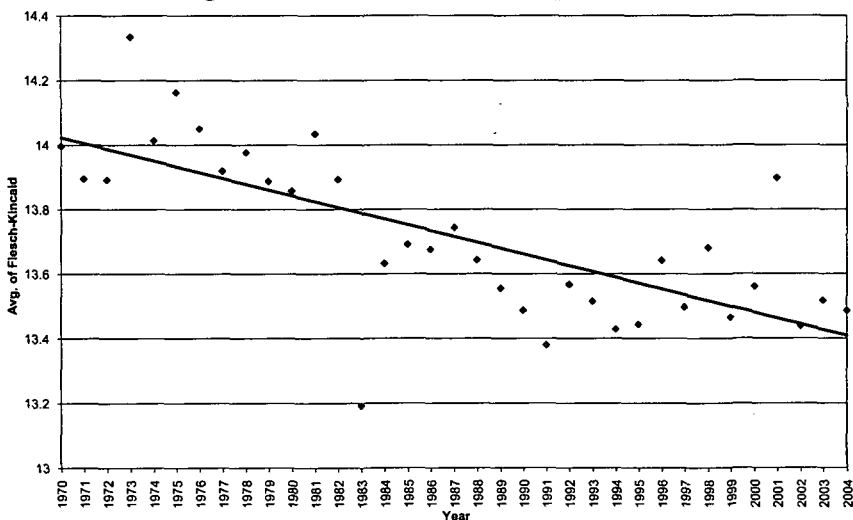


Figure 2A shows a decrease in the average Gunning Fog Index from about 17.6 to about seventeen over our twenty-five-year time period. The average Flesch-Kincaid scores over the same time period fell from approximately 13.9 to approximately

13.4. The scores in figures 2A and 2B were derived from the PERL software scripts that we created. In addition, to cross-check our outcomes, we also used readability programs created by Microsoft. The Microsoft formulas were not able to generate any scores for the Fog Index; however they could produce outcomes for both the Flesch-Kincaid and the Flesch Readability Indexes. Figure 2C graphs the change in the Flesch-Kincaid score using the Microsoft, rather than our PERL, scripts. We do not include graphs of the Flesch Reading Ease scores, because for the Argument component of the briefs, neither our own PERL programs nor Microsoft's programs found significant changes over time in these scores. Comparing figures 2B and 2C, we can see that the scores for the Flesch-Kincaid Index are slightly different using PERL as opposed to Microsoft. This is expected, as the two programs did not use the exact same scripts to measure variables. Each program took a different approach to the challenging problem of quantifying sophisticated linguistic variables (automatically distinguishing periods that appear in abbreviations from periods that end sentences, for example).

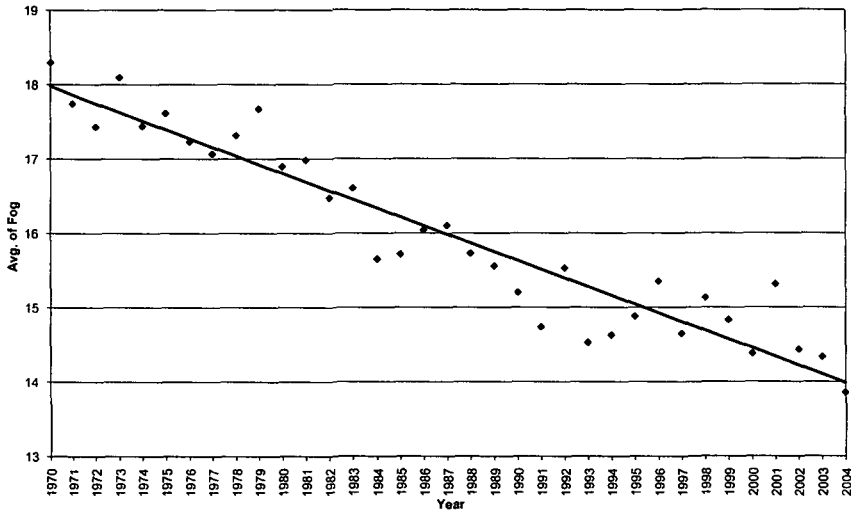
FIGURE 2C
Average of Flesch-Kincaid for Argument—Microsoft



Figures 3A through 3D show changes in the average readability scores of the Statement of Facts component from 1970 to 2004.

Figure 3A reveals that the average Fog Index of the Statement of Facts decreased from about eighteen to about fourteen from 1974 to 2000. Comparing this decrease to that of the Argument section, we see that the average Fog score decreased by significantly more in the Statement of Facts component ($18 - 14 = 4$) than in the Argument component ($17.6 - 17 = .6$).

FIGURE 3A
Average of Fog for Statement of Facts—PERL



Similarly, the average Flesch score for the Statement of Facts has in the same time period gone up from approximately thirty-four to approximately forty-two, as shown in Figure 3B on the next page. Recall that in contrast to the other two indexes we used—the Fog and the Flesch-Kincaid—the values reported by the Flesch score should increase as readability improves. As our Flesch results follow the increasing-value pattern, they show patterns consistent with those revealed by the Fog.

FIGURE 3B
Average of Flesch for Statement of Facts—PERL

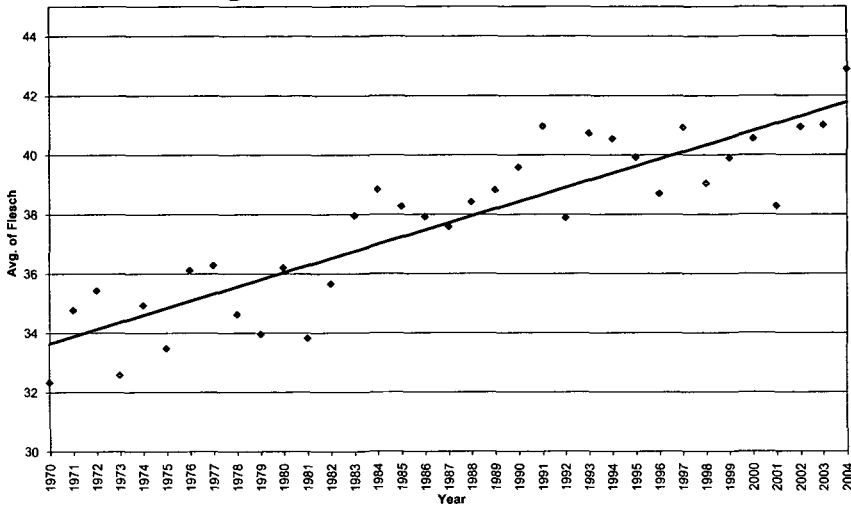


Figure 3C shows that the average Flesch-Kincaid of the Statement of Facts decreased between 1970 and 2004 from about 14.25 to about 10.5. As with the Fog results, the average Flesch-Kincaid decreased more in the Statement of Facts ($14.25 - 10.5 = 3.75$) than in the Argument ($13.9 - 13.4 = .5$).

FIGURE 3C
Average of Flesch-Kincaid for Statement of Facts—PERL

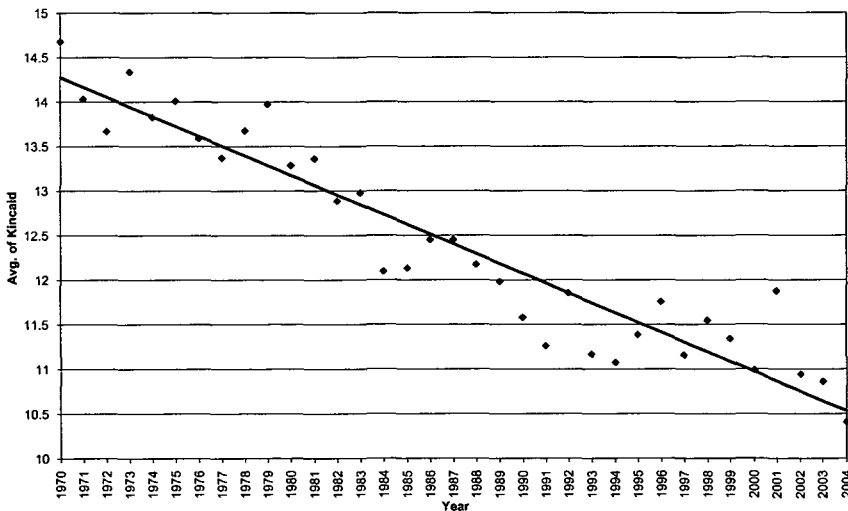


Figure 3D shows that the increase in the average Flesch score using the Microsoft formula is about 6.5 (32.5 to 39). This increase is comparable to the change in the average Flesch score revealed in Figure 3B, using the PERL programs, of about eight (34 to 42). Again, because of the subtle differences in the way the two scripts measure variables, some difference is to be expected, but both formulas show a strong, but steady, change in average readability over the quarter century we studied.

FIGURE 3D⁴⁰
Average of Flesch for Statement of Facts—Microsoft

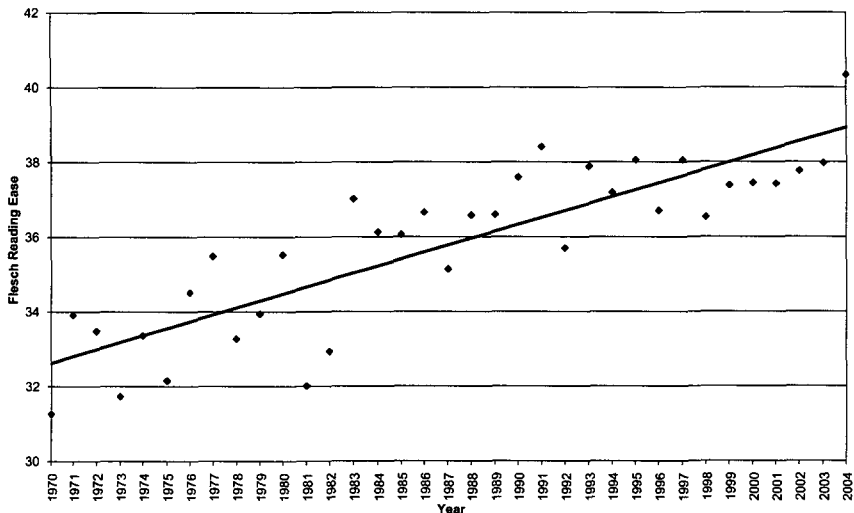
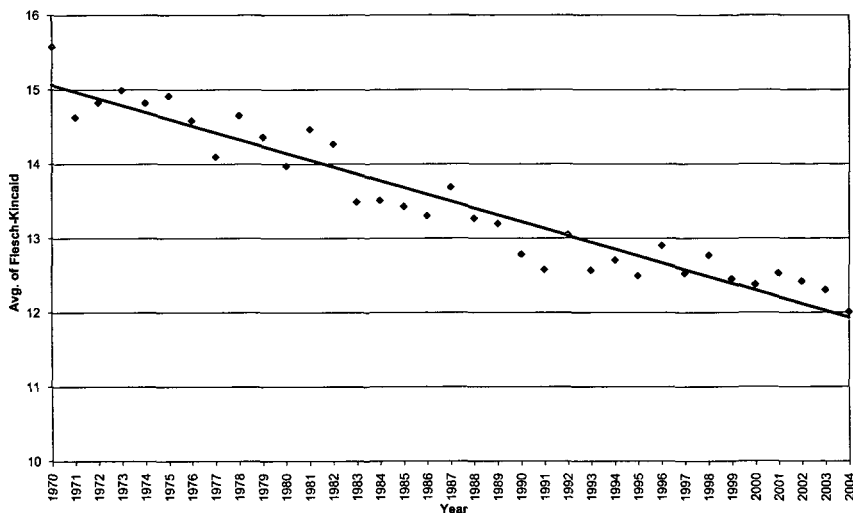


Figure 3E, which appears on the following page, demonstrates that the average Flesch-Kincaid measured by the Microsoft formula decreased by approximately three ($15 - 12 = 3$) over twenty-five years. Compare this with the decrease of 3.75 in the average Flesch-Kincaid using the PERL program, which is shown in Figure 3C.

40. Figures 3D and 3E are taken from Microsoft's readability algorithms, as opposed to our PERL programs (which generated Figures 3A through 3C).

FIGURE 3E
Average of Flesch-Kincaid for Statement of Facts—
Microsoft



As we anticipated, changes in average readability scores were stronger in the Statement of Facts section of Supreme Court briefs than in the Argument section. The Statement of Facts almost always offers legal writers more stylistic and structural flexibility than the Argument section. The factual narrative, typically chronological, yet unencumbered by the constraints of rule-based legal argument and the need for citation to legal authority, should normally reveal more stylistic freedom, and therefore more long-term variation, as a consequence of external forces such as a greater emphasis on writing plainly. Again, though, both the Argument and the Statement of Facts components do reveal parallel changes in readability over the quarter century of our data, even if the latter component provides more striking evidence of this historical shift.

To maintain full accreditation, American law schools and legal educators must devote resources to improving the writing skills of future attorneys. The American Bar Association's curricular demands for law schools requires "writing in a legal context, including at least one rigorous writing experience in the first year and at least one additional rigorous writing experience

after the first year.”⁴¹ According to surveys of hiring attorneys, law schools should be doing even more to produce quality legal writing in their graduates; concerns about writing typically rank at the top of the list in such surveys.⁴² Teaching law students to write well is a multifaceted demand (including much more than instruction in plain English, of course), but a look at the curricula of contemporary legal writing programs, and the texts used in such programs, reveals that “plain legal English” forms an important component of this pedagogy at many law schools.⁴³ But this has certainly not always been the case; the focus on plain English has a long history, but only in recent years has it gained significant strength, and it has paralleled the professionalization (and relative standardization) of the legal writing curriculum itself over the past couple of decades.⁴⁴

If substantial resources are being spent on teaching law students (and requiring practitioners) to write in a plainer way, a natural question asks if the money is being spent successfully. Over the past few decades, has the educational—and regulatory—redirection towards plain legal English made any difference in the way lawyers actually write? One method, as yet untried, to answer that question would be through the use of surveys or experiments testing practitioners who had graduated

41. Sec. of Legal Educ. & Admissions to the Bar, ABA, *Standards and Rules of Procedure for Approval of Law Schools* 21 (ABA 2008). A previous standard was less specific, requiring schools to offer “(a) instruction in substantive law, values and skills, including legal writing; (b) at least one rigorous writing experience; (c) adequate opportunities for instruction in professional skills; and (d) live-client clinical opportunities.” Kenneth D. Chestek, *MacCrate (In)Action: The Case for Enhancing the Upper-Level Writing Requirement in Law Schools*, 78 U. Colo. L. Rev. 115, 122 (2007) (internal quotation marks omitted).

42. See e.g. Bryant G. Garth & Joanne Martin, *Law Schools and the Construction of Competence*, 43 J. Leg. Educ. 469, 473, 477 (1993).

43. James M. Boland, *Legal Writing Programs and Professionalism: Legal Writing Professors Can Join the Academic Club* 18 St. Thomas L. Rev. 711, 715 (2006). Our—admittedly anecdotal—evidence from years of talking to colleagues around the country about curriculum issues is that the vast majority of legal writing professors either require a specific plain English text such as Richard Wydick’s *Plain English for Lawyers* (5th ed. Carolina Academic Press 2005) or use a textbook with some treatment of plain legal English.

44. See Judith D. Fischer, *Why George Orwell’s Ideas about Language Still Matter for Lawyers*, 68 Mont. L. Rev. 129, 133-35 (2007); see also Brady Coleman, *Clarity and Precision in Legislative Drafting: Case Studies from Singapore Law* (unpublished master’s thesis, University of Leicester 1998) (on file with author Coleman).

from law school at different periods of time. We took a different approach, choosing to rely on available data to resolve the issue. To recap: our study chose a corpus of nearly 9,000 Supreme Court briefs, representing a period of three and a half decades in which the focus on plain legal English became increasingly widespread, and used the average readability scores of such briefs as a proxy for plainness in writing.⁴⁵ If our most important assumptions are accepted—that readability offers reliable evidence of plainness, and that Supreme Court briefs provide an acceptable representation of legal writing—then the following conclusion is warranted: A gradual historical trend towards plainer legal writing is revealed over recent decades.



45. Freshly minted law school graduates are less likely to write briefs to the Supreme Court than are more senior attorneys, of course, although junior attorneys employed at the Solicitor General's Office are presumably involved in the brief-writing process at a much earlier stage in their careers, to offer just one possible counter-example. In any case, although there is surely a delay between an education emphasizing plain legal English and the evidence of that education in legal documents, the delay is apparently embedded in the overall trend we found.

