

Dumbing Down? Trends in the Complexity of Political Communication*

Kenneth Benoit[†]

Kevin Munger[‡]

Arthur Spirling[§]

March 18, 2017

Abstract

Commentators and some political scholars claim to have observed a “dumbing down” of the level of sophistication of political language, leading to anxiety over the quality of democratic deliberation, knowledge, and policy design and implementation. This work typically focuses on the President’s State of the Union addresses. Using quantitative indicators of textual complexity, we measure trends since 1790 in that and other key political corpora, including rulings of the Supreme Court, the Congressional Record, and Presidential executive orders. To draw comparative lessons, we also study political texts from the United Kingdom, in the form of party broadcasts and manifestos. Not only do we cast shade on the supposed relentless simplification of the State of the Union corpus, we show that this trend is not evident in other forms of elite political communication, including presidential ones. Finally, we argue that a stylistic—rather than an obviously substantive—shift towards shorter sentences is driving much of the over-time variation we see in traditional measures of political sophistication.

*Prepared as a chapter for a book arising from the SSRC-funded “Anxieties of Democracy” workshop, held at Princeton University, 28-29 October 2016. We are grateful to Burt Monroe and other colleagues present for their very helpful comments.

[†]Professor of Quantitative Social Research Methods, London School of Economics. kbenoit@lse.ac.uk.

[‡]PhD Candidate, Department of Politics, New York University. km2713@nyu.edu.

[§]Associate Professor of Politics and Data Science, New York University. arthur.spirling@nyu.edu

Dumbing Down as an Anxiety of Democracy

Anxiety about the quality of public political discourse is nothing new. Beginning at least with Plato’s call for ‘Philosopher Kings’ to protect citizens from their own base impulses, this fear has its modern incarnation in commentators’ concerns about the supposed ‘dumbing down’ (Gatto, 2002) of politics, culture and public life. Claims that elected officials and their output are less ‘intellectual’ (e.g. Lim, 2008) than they used to be generally rely on the idea that contemporary politicians must appeal to a median voter who demands progressively simpler messages despite, or perhaps because of, the increasing complexity of the modern world and the fact that information is no longer hard to acquire, but rather difficult to escape. According to this view, modern citizens are too busy or too distracted to absorb anything other than facile sound-bites that the news media serves up on a 24-hour rolling basis. And the prognosis is that a bad situation will worsen. Thus, for example, the Guardian newspaper laments that the “state of our union is...dumber”,¹ while various outlets describe Donald Trump’s speeches on the campaign trail as having the standard of a fifth grader,² a fourth grader,³ or even a third grader.⁴

In contrast to such pessimism, more sanguine commentators point to several interrelated trends that suggesting we should not be worried. They argue that as the world has grown more complex, so has the public ability to absorb more—and more complex—information. In terms of formal education, the post-war period has seen a massive expansion in college enrollment, from around 45% of school leavers in 1959, to some 70% by 2009 by US Department of Labor figures. This pattern of increasing attendance, if not at identical rates, is not true for historically privileged groups. It is also seen in women and African Americans, who joined the franchise in the United States much later than did white men. Thus any claims that the extension of suffrage to relatively uneducated groups might be causing a decline in discourse quality (see, Spirling, 2016, for an application of this idea to the UK) are extremely dubious. Even if one does not accept that college

¹<http://www.theguardian.com/world/interactive/2013/feb/12/state-of-the-union-reading-level>

²https://www.washingtonpost.com/news/morning-mix/wp/2016/03/18/trumps-grammar-in-speeches-just-below-6t?utm_term=.175e0dddcfcf

³<http://nypost.com/2016/03/17/donald-trump-speaks-at-a-fourth-grade-reading-level/>

⁴<http://www.politico.com/magazine/story/2015/08/donald-trump-talks-like-a-third-grader-121340>

adds much to students' core cognitive capacities, there is evidence at a more fundamental level that citizens can cope with more sophistication. The 'Flynn Effect' describes the relentless year-on-year increase in IQ-scores in multiple countries and societies, including the US (Flynn, 1999). And in the specific domain of civics—presumably one of the more relevant subjects for political understanding—knowledge levels have been at worst flat in recent times (Galston, 2001).

In this chapter, we examine the evidence for these polar opposite positions in the debates on modern political discussion. In particular, we analyze long-term trends in the sophistication of language using similar metrics, comparing political and cultural texts across time and space. Our conclusions should be relatively anxiety-relieving. Put very simply, the SOTU is probably an outlier as elite time series go, and the changes it has been through—which are by no means as stark as they first appear—are mostly a benign (and sometimes mechanical) consequence of other broader trends in language.

We proceed as follows. Because of its centrality to extant accounts of dumbing down, we re-analyze SOTU in some detail, and consider standard time series diagnostics and what they reveal. We also provide a sentence bootstrap to get a more accurate sense of how that corpus is—and is not—changing over time. We then provide a somewhat general critique of the standard measure(s) of textual complexity, and explain why they might be particularly inappropriate for speeches over such a long historical period. Although obvious to us, perhaps our most important innovation below is to compare SOTU to other time series. In particular, we show that when we look at other branches of US government, comparatively to other countries (the UK), or to other cultural touchstones (such as Nobel literature announcements), there is little or none of the purported dumbing down we see for SOTU. That is, we suspect that SOTU addresses, and presidential texts more broadly, are outliers in data terms. Finally, we look under the hood of the FRE measure, and decompose the patterns in our data into the component deriving from changes in sentence lengths versus the component that is a product of varying word lengths. We demonstrate that, while there is little systematic change in terms of word use, we are an era of ever shortening sentences. While this may well make for clearer, punchier rhetoric, we are unconvinced that it alone should provoke

anxiety about the state of our democracy.

A Noticeable Trend: The Simplification of Presidential Addresses

Since 1790, with only a few exceptions, the President has delivered an annual State of the Union address to Congress. This makes it an unusually regular, if not especially frequent, source of data on the changing priorities of the executive. Importantly for our purposes, there is no requirement that it be a speech. Indeed, from Jefferson until around 1934—and even after on some occasions—it was generally a *written* document. Because the form of spoken English usually differs from that of written English, even for pre-written speeches, we might expect that the change in formats is reflected in differences in the structure and complexity of the language, particularly in terms of the length of sentences. This induces an obvious problem of confounding: it is difficult to know whether changes over time are a product of some latent dynamic or simply the consequence of altering the means of the address. Nonetheless, studying the purported sophistication of this series has proved popular, both for academics (e.g. Lim, 2008) and journalists.

The most common measurement approach is to use Flesch (1948)’s ‘reading ease’ (FRE) statistic. Based on a linear regression of reading comprehension on average sentence length and average number of syllables per word, the FRE has proved durable well beyond its mid-Twentieth Century origins. The FRE consists of a linear combination of average word length (in syllables) and average sentence length, computed as

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

In the initial application, FRE took values between zero and 100, although technically the upper bound of a text’s reading ease is 121.22, if a text consists exclusively of single-syllable, single-word sentences. Scores between 60-70 are considered “plain English”, and easily read and understood by someone with an 8th- or 9th-grade education. In the 1970s, Kincaid et al. (1975) introduced a linear rescaling of the formula designed to correspond more directly to the US grade school levels

required to comprehend a text.

Whether or not FRE is suitable for a given application cannot be answered definitively from first principles. Still, some concerns are obvious. First, FRE is used in a context far from its original domain when applied to SOTU. It is doubtful whether judging the sophistication of SOTU addresses—designed for Congress and adult citizens who may be familiar with political terms in a way that minors are not—is best done with a measure invented for assigning books to school children. And if it is appropriate, the fact that the measure was last calibrated 70 years ago might give us pause. Here, one danger is that other than via syllable information, the FRE does not explicitly take into account the familiarity of the words used in a text. For a given word, this is something which surely varies over time: thus, citizens in 1940 would be have little trouble understanding *locomotive*, relative to voters in 1990, but by contrast may have struggled significantly with a term like *television* which would be trivial for listeners later in the century. It is true that some other approaches do cater for familiarity explicitly (e.g. Dale and Chall, 1948), but in practice the scores for documents are highly correlated with those generated by FRE. This is also the case for the numerous other approaches that have proved less popular than FRE and its variants (e.g. Gunning, 1952; Spache, 1953; McLaughlin, 1969; Coleman and Liao, 1975).

More fundamental perhaps than the linguistic problems with FRE are the *statistical* problems that typically emerge with its use. Notice that for a given document, FRE provides a point ‘estimate.’ This is technically an out-of-sample prediction of the dependent variable (\hat{y}) from the original regression Flesch ran on school children. These predictions/estimates may then be plotted or analyzed. But for any systematic test of claimed changes over time, we also need to know our uncertainty around the points. More precisely, we require the sampling distribution of the FRE such that, for example, we can place 95% confidence intervals around the values. In this way, we can make statements about the whether the decline between two dates is statistical significant or merely due to sampling variation. In the measures presented below, we obtain bounds using the bootstrapping approach of Benoit, Laver and Mikhaylov (2009), which resamples sentences with replacement within texts. This simulates variability while preserving dependencies among

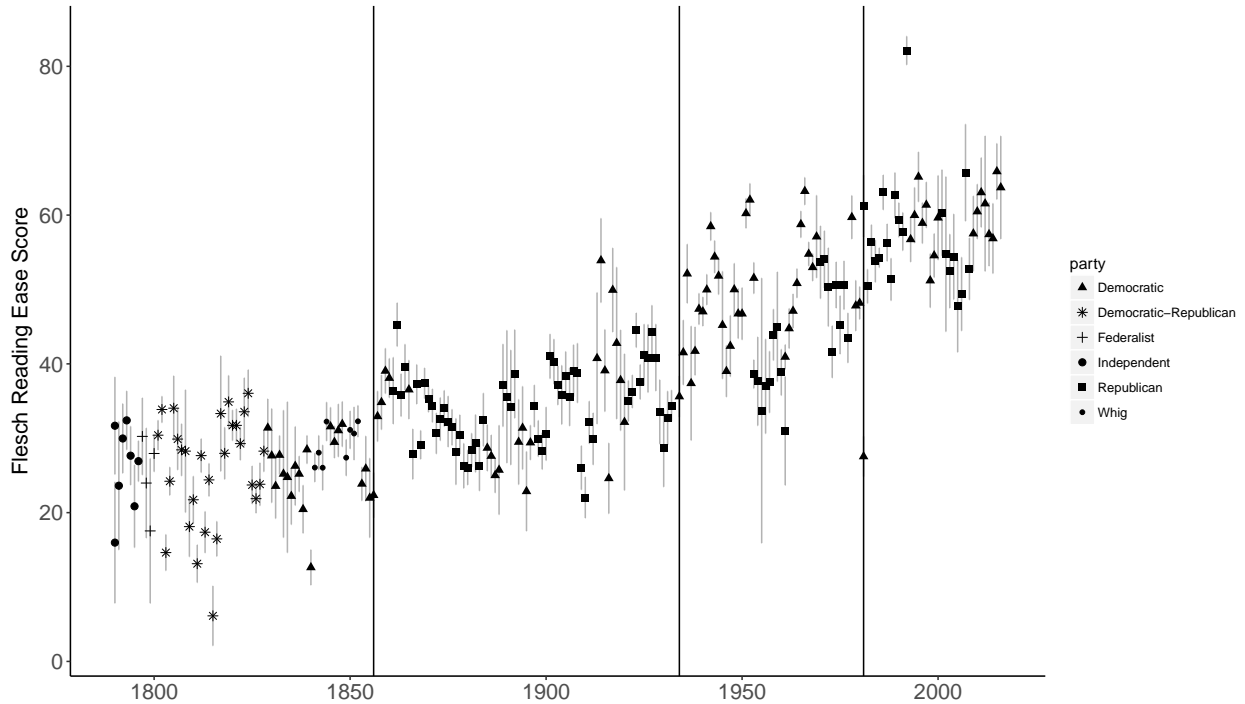


Figure 1: Flesch Reading Ease scores, with 95% bootstrapped confidence intervals, for the State of the Union address. Note that the solid vertical lines are estimated structural breaks.

language within sentence units. For a given text with k sentences, we drew 100 samples (with replacement) of size k of those sentences. From the 100 samples, we computed the 95% confidence interval as the 2.5th and 97.5th quantiles of the replicates. Two factors in a text tend to influence the variability of readability statistics computed in this way. First, the more heterogeneous the sentences of a text in terms of their sentence length and syllabic length, the more variable are the replicates. Second, shorter texts tend to produce larger intervals in the replicates, due to the higher resampling rates of sentences from texts with a smaller pool of sentences from which to draw.

In Figure 1, we plot the FRE point estimates and their associated 95% confidence intervals from this bootstrapping procedure. We also impose structural breaks (in the sense of Bai and Perron, 2003) on the time series. These are marked by solid vertical dark lines running to the height of the plot, while the confidence intervals are the shorter gray lines.

The most evident pattern in Figure 1 is the upward trend in reading ease: from a typical level of 27 in the first period, to about 57 in the latest period. This corresponds to roughly a level requiring

a college degree to understand the test to requiring just having completed junior high school. From the breakpoint analysis, we observe three structural changes: in 1856, in 1934, and a third in 1981. After each break, the mean FRE increases. The first break, contemporaneous with the advent of the Third Party System may, perhaps be a function of the broader-based party coalitions than had previously been the case. The second break corresponds to a shift in means of delivery: FDR spoke his addresses rather than wrote them, and this was increasingly the norm thereafter. The final break occurs just after Carter's last address, which was by far the longest in history, and coincided with a period in which the president had returned (briefly) to delivering the SOTU report in writing.

Judging from the confidence intervals, however, we see that adding estimates of uncertainty about the reading ease statistics throws significant shade on claims about "dumbing down" in recent times. In particular, many of the confidence intervals from the last two periods in the data overlap. Taking this into account, it far from clear that speeches given around 2000 are any dumber, in a statistical sense, than speeches given in the 1940s and 1950s. Equally, it is not clear that the Republican speeches in the early years of the 21st century are statistically significantly different than those made a hundred years earlier.

But even the point estimates may be misleading here. This is because they don't account for the practices of various presidents to deliver written, as opposed to spoken, addresses to Congress. Simply put, we might expect spoken texts to use simpler sentences, and therefore to score more highly on the reading ease index. To examine this claim, in the boxplots in Figure 2, we have distinguished the scores for our four time periods by whether the texts were spoken or written. As expected, spoken addresses tend to be higher scoring on readability. Furthermore, the difference between spoken and written addresses is increasing over time: In the first period (1790–1856) there was no discernible difference in reading ease, regardless of whether the address was written or spoken, a sharp contrast to the era of 1935–1981 when there was about a 10-point difference between written and spoken addresses (a median 39 and 48, respectively).⁵

⁵In the 1935–1981 period, 43 speeches were spoken and only 7 were written. This contrasts to the first and second periods, when the tallies were 56 to 12 and 67 to 10, respectively. In the era since 1982, all addresses have been spoken.

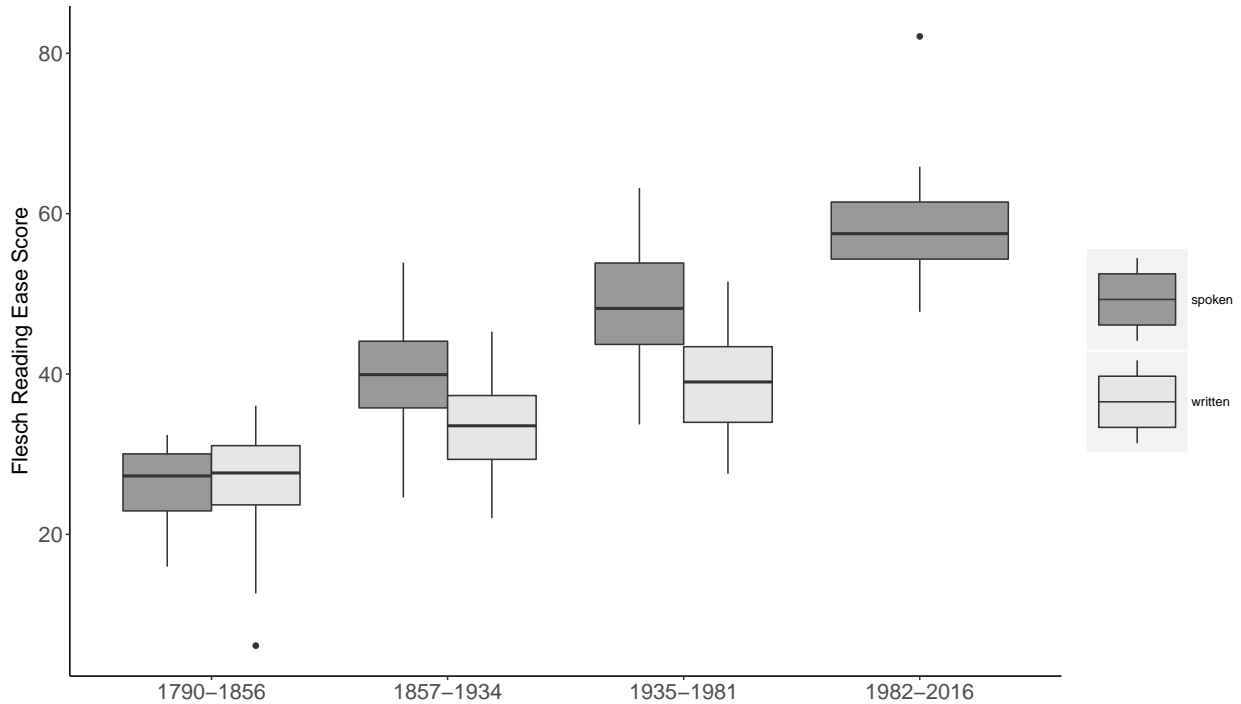


Figure 2: Differences in Reading Ease by Method of Delivery, for State of the Union Addresses.

The difference appears to be that even spoken addresses were designed for audiences with a nearly masochistic ability to endure long, convoluted speeches. While posterity remembers Abraham Lincoln’s two-minute speech on November 19, 1863, far fewer are aware that his immortal Gettysburg’s address followed a two-hour speech by former Secretary of State Edward Everett, a man widely acclaimed for his oratorical ability. In the early years of the Republic, longer texts, containing longer sentences, were far more common. The start of President James Adam’s 1797 (spoken) address, for instance, contained the following sentence, inconceivable today:

Although I can not yet congratulate you on the reestablishment of peace in Europe and the restoration of security to the persons and properties of our citizens from injustice and violence at sea, we have, nevertheless, abundant cause of gratitude to the source of benevolence and influence for interior tranquillity and personal security, for propitious seasons, prosperous agriculture, productive fisheries, and general improvements, and, above all, for a rational spirit of civil and religious liberty and a calm but steady determination to support our sovereignty, as well as our moral and our religious principles, against all open and secret attacks. (James Adams, 1797)

While such passages are highly effective in reducing a text’s reading ease score (it is -57.1 for

Table 1: Summary of the Texts Examined

Corpus	Date Range	Number of Texts	Max Words	Median Words
SOTU	1790–2016	230	31,920	5,716
SCOTUS	1790–2012	6,373	8,731	1,608
Congressional Record	1995–2008	819,978	3,677	36
Executive Orders	1826–2017	4,451	1,263	246
Nobel Prize Announcements	1901–2015	96	2,746	1,156
UK Party Manifestos	1918–2010	71	22,990	5,078
UK Party Election Broadcasts	1964–2001	84	2,202	1,362

that passage, and not even we can understand it!), few commentators would point to the death of such communication styles as a reason to be anxious about the quality of our democratic discourse or of political communication. Put more positively, fewer sentences of this nature are good for the clarity of political life.

Comparing SOTU Trends to Other Texts

We have thus far made the claim that the standard measures of complexity are not ideal, and that the point estimates they yield—at least as regards SOTU addresses—may be misleading. Still, one could make the case that the aggregate nature of SOTU, as a representative sample of elite discourse, is concerning. To assess whether the State of the Union addresses is unusual, or part of a general troubling trend in increasing textual simplicity, we now compare the SOTU trends against several other benchmark corpora, observed over time. These texts include two other political series from the United States and two from a reasonably similar system, the United Kingdom.

The primary US corpora we examined are from Supreme Court (SCOTUS) opinions from 1790 to 2012 and s from 1826 to 2017 (see Table 1).⁶ The SCOTUS opinions not only represent legal language from the highest court in the United States, but also capture changing trends in both language and in subject matter. This is because changes in technology or political complexity

⁶For the SCOTUS corpus, we took a weighted sample of a total of 38,872 SCOTUS decisions. For all years in which there were fewer than 30 SCOTUS decisions recorded, we include all those from that year, but took a random sample of the decisions from years with more than 30. We also cleaned the text so that (e.g.) *Roe v. Wade* and other abbreviations such as *U. S.* would not be counted as multiple sentences, and we eliminated any “sentence” appearing to contain fewer than four words.

may have affected the content of the document in ways that might also have affected the SOTU corpus. The Executive Orders, some 4,451 in total, offer an alternative glimpse into presidential language, albeit in a different context. If the trends we observe in SOTU are absent from the judicial and alternative executive texts, we can perhaps dismiss claims about a general dumbing down of democratic discourse in the American context. We also have a short time series of machine readable text from the *Congressional Record*: in particular, Representatives and Senators speaking in floor debates (with the Speaker of either chamber removed) for the years 1995–2008. This allows us to compare executive texts with those from the legislature.

As an additional benchmark against texts that were in a similar but different political context, we also examined readability trends in two corpora from the United Kingdom: a set of 71 election manifestos from the Labour and Conservative parties, from 1918 to 2010, and a set of 84 party election broadcasts from 1964 to 2001. The latter are transcriptions of officially regulated radio and television broadcasts used by political parties to advertise their positions during election campaigns, in lieu of political advertising.⁷ Both are aimed at the general citizenry and designed to broadcast party positions and to attract voters, although manifestos are written documents and party election broadcasts spoken.

Finally, we compare the SOTU addresses to the announcements of the winner for the Nobel Prize for Literature. These announcements typically explain why the committee is awarding the prize, and thus describe the merits and scope of the relevant author's work. We chose this time series because it is relatively long and regular, and captures a broader, cultural spirit which is hopefully less explicitly political than the other data we have.

Comparing the SOTU trend to the other American political texts, we see in Figure 3 that unlike the SOTU texts, the Supreme Court opinions remained relatively constant in readability since 1790, at a Flesch score of around 40 for most of the period (with a dip around 1900). Even in the twenty-first century, this trend remained unchanged. Looking at the corpus of Executive Orders, moreover,

⁷Party election broadcasts replace political advertisements in the United Kingdom, which are prohibited by law. Parties are allocated broadcast slots of around five minutes each, free of charges, using a formula set by Parliament. We obtained the broadcast transcripts from <http://www.politicsresources.net/area/uk/peb.htm>.

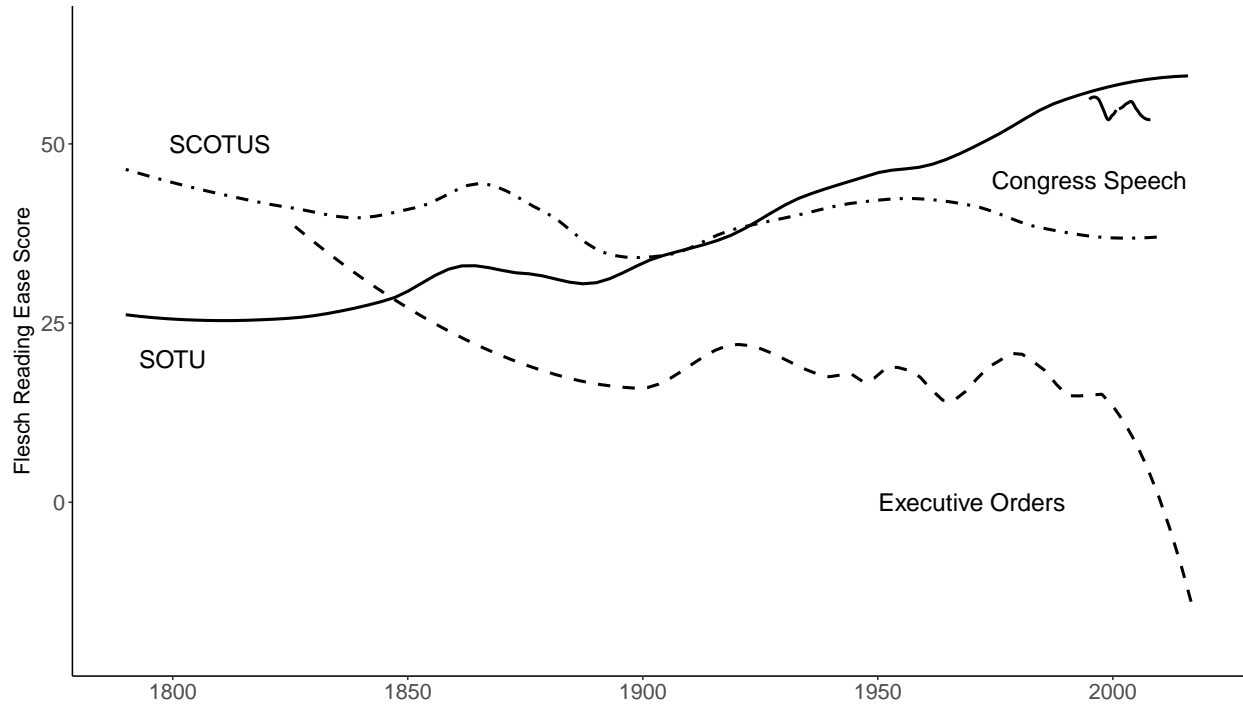


Figure 3: Comparing SOTU reading ease trends to other governmental texts: Supreme Court decisions, Congressional speeches and Executive Orders.

the reading ease of these presidential texts not only did not show an increase similar to the SOTU texts, but actually declined in the 19th century, remaining constant throughout the 20th century, and then declined sharply after 2000. So, while annual presidential addresses to Congress—and through live broadcast, to the public—became markedly more comprehensible, presidential text in the form of Executive Orders became markedly more dense. Thus, there seems to be a clear divergence in text that communicates goals and broad plans to a general audience (SOTU) and texts that specifies legally binding instructions from the President to the officers and agencies of the federal government (Executive Orders). As Executive Orders have increased in frequency and scope since the founding of the republic, they have clearly also grown in textual complexity. Executive political communication to the masses may have gotten markedly simpler, but Executive Orders are clearly exhibiting a tendency quite opposite to the “dumbing down” thesis. To the extent we can draw firm conclusions from the Congressional corpus, we see two things. First, the mean readability of this speech is high: certainly well above that of SCOTUS and the executive orders,

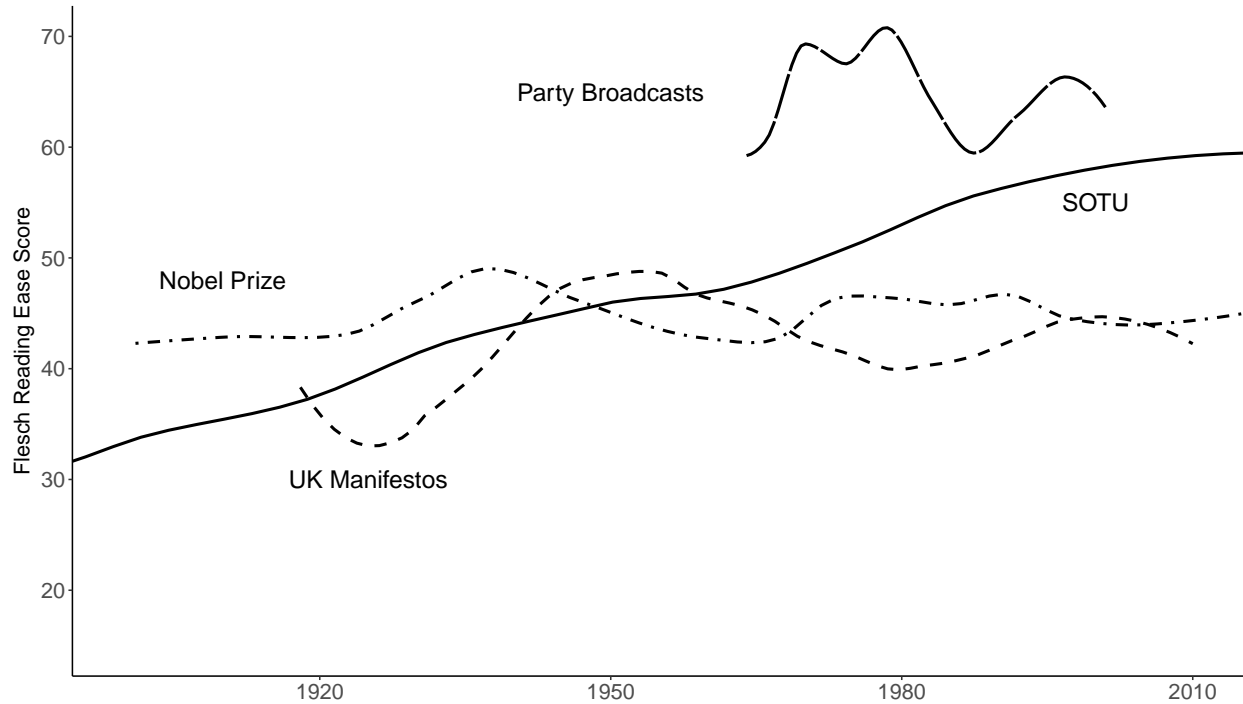


Figure 4: Comparing SOTU reading ease trends to other text series: UK Conservative and Labour Party Manifestos, UK Party Political Broadcasts, and Nobel Literature Prize Presentation Speeches.

and on a par with SOTU addresses around 1975. This perhaps makes sense given the texts here are spoken rather than written. But, perhaps more interestingly any trend in simplicity from Congress is downwards in the direction of *more complexity* not less. Recall that we have a great deal of data from Congress cross-sectionally (although not much over time). When we perform a *t*-test of means for the first and last year of the *Record* data, we easily reject the null of equal difficulty over time ($p < 0.01$).

Comparing the US speeches to the UK context and the international texts from Nobel prize announcements, we also see relative stability in these patterns compared to the trend to greater reading ease from SOTU. Figure 4 plots the time trends for UK manifestos, the party election broadcasts, and the Nobel Prize announcements. While Nobel Prize announcements gained in reading ease slightly from about 1930 to 1950, their scores remained relatively constant across that period, at around 50, similar to the relatively stable UK election manifestos. The stability in the reading ease of the UK manifestos is in many ways a remarkable contrast to the SOTU trend,

because of significant changes in the character of the electorate and of political campaigning during this period. Finally, party broadcasts, while at a much higher reading ease level than the other texts generally, also showed relative stability within the 60-75 point range, the level at which texts are considered easily comprehended with only 8 or 9 years of formal education.

Overall, the conclusions from our benchmarking are clear: Compared to judicial trends, Congress, and to the other executive text, the trend to greater simplicity of the SOTU texts appears to be a somewhat isolated phenomenon. Compared to political texts aimed at mass publics, such as the UK manifestos and party election broadcasts, the singular nature of the SOTU trend is upheld. Even the texts honoring some of the brightest minds in the world, as indicated through the Nobel Prize texts, showed no real changes in complexity. Something unique is happening in the SOTU texts. To investigate what this might be, in the next section we decompose the drivers of the reading ease score. This allows us to assess whether the observed pattern appears to be coming from a change in sentence length or rather, because Presidential addresses are using syllabically simpler words. If the latter, we might consider this be a form of “dumbing down,” but if the former, then it probably reflects primarily stylistic differences in an evolving preference for shorter sentences.

Digging deeper into the components of reading ease

The index of reading ease which we have used to measure the sophistication of political communication, the Flesch index, consists of a linear combination of sentence length and word lengths, as explained above. For each additional increase in the mean word length of a text’s average sentence, the index decreases by 1.015, and it decreases by 84.6 for each one unit increase in a text’s average word’s number of syllables. While these weights are specific to the Flesch index, the components, or some form of them, are common to nearly every index of readability.

Figure 5 plots the long-term trends in the three American text series, for the mean sentence length in words. Here we see a steady change in the measure for SOTU that mirrors the trend in the Flesch index that contains the mean sentence length as a component: from a typical sentence

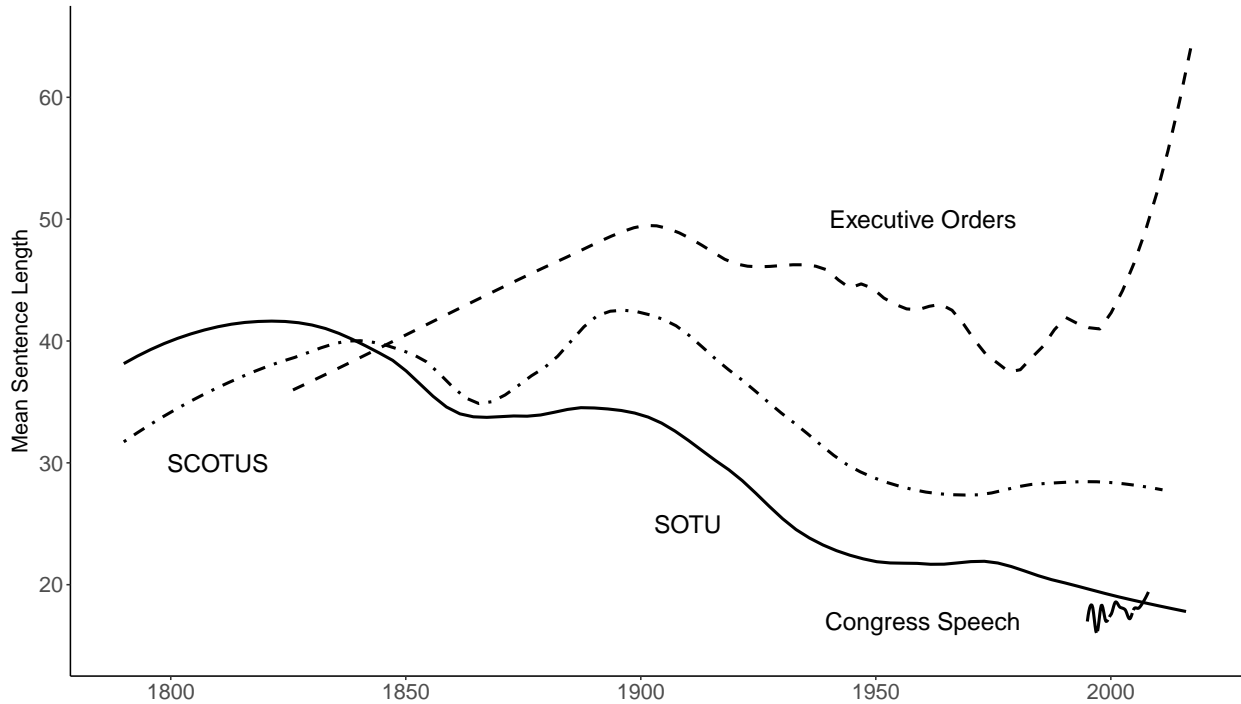


Figure 5: Comparing mean sentence length trends to other governmental texts: Supreme Court decisions and Executive Orders.

of 40 words around 1800, to a typical sentence of around 20 words or fewer by 2000. As measured by the Flesch index, a 20-word length change corresponds to an increase in the Flesch score of 20.3, which would explain two-thirds of the increase observed between the median texts in the first and last periods (Figure 1). If we compare the SOTU address with the longest sentences, James Madison in 1815 with an average sentence length of nearly 53.5 words, to an address with some of the shortest sentences, Donald Trump’s speech in 2017 at around 17.5 words per sentence, this alone accounts for a difference in the Flesch score of over 36.5.

Examining the other corpora measured in Figure 5, we see generally similar trends to their Flesch scores. A key difference is a noticeable decline in mean sentence length for the SCOTUS texts between 1900 and 1950, of about 15 words on average. The sentence length of the executive orders increases steadily between 1825 and 1900, then holds relatively steady until increasing sharply following 2000. The latter shift in particular explains the sharp drop in the Flesch scores of the executive order texts from about 2000. As with the SOTU addresses, most of the changes

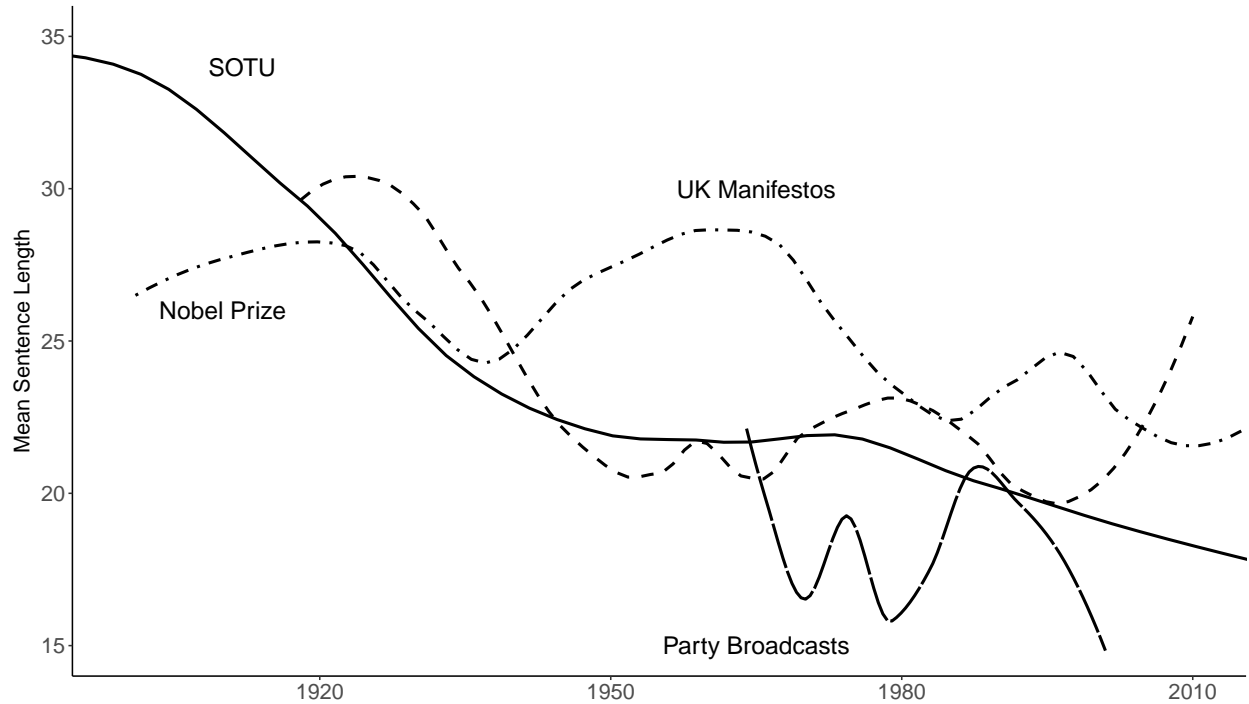


Figure 6: Comparing mean word length trends to other text series: UK Conservative and Labour Party Manifestos, UK Party Political Broadcasts, and Nobel Literature Prize Presentation Speeches

appear to be driven by changes in sentence length. For the Congressional speeches we are, as usual, hampered by a very short time series. Nonetheless, the trend looks approximately constant, with a very small recent increase in sentence length.

To compare trends in sentence length to our non-US benchmark corpora, Figure 6 compares the SOTU texts to the sentence length trends for the UK political texts and the Nobel Prize announcements. While they fluctuate over time, only the Nobel texts from 1920 to 1950 exhibit a strong downward trend, explaining the matching rise in reading ease during this period for the Nobel texts from Figure 4.

Turning to the other component of the reading ease measure, we examine trends in word lengths in Figures 7 and 8. The first shows the trends for our American political texts, and the second for the comparative benchmark corpora. Here the trend for SOTU is relatively flat, until about 1960 when the trend drops to a new plateau around 2000. By contrast, the other three series show steady increases or stasis in mean sentence length. The executive orders have generally longer words



Figure 7: Mean number of syllables per word for four governmental text series: Supreme Court decisions, Congressional Record, Executive Orders and for the State of the Union address.

from about 1920, and SCOTUS uses more syllables (on average) from about 1900. Congressional speeches don't move a great deal in this space: the begin and end in approximately the same place, syllable-wise. Overall, this is a striking contrast: if anything, word lengths are getting longer in US political text, except for the SOTU addresses which have shown a tendency toward shorter words, especially since 1960. Benchmarked against the comparative texts, we see the decline in SOTU word lengths (albeit for a shorter time series) in Figure 8, while the typical word lengths of the other texts has remained relatively constant.

While SOTU texts have declined in both word and sentence lengths, the contribution of gradually shortening sentences is far stronger than the move to progressively shorter words. To assess this, we can compare the coefficients in the Flesch formula that provide weights for each component, against the typical variation in each component among the SOTU speeches. By measuring a "standard deviation" in mean sentence length and mean word length across the texts, we can compare what a change in this standardized unit of variation would yield in scoring terms. For

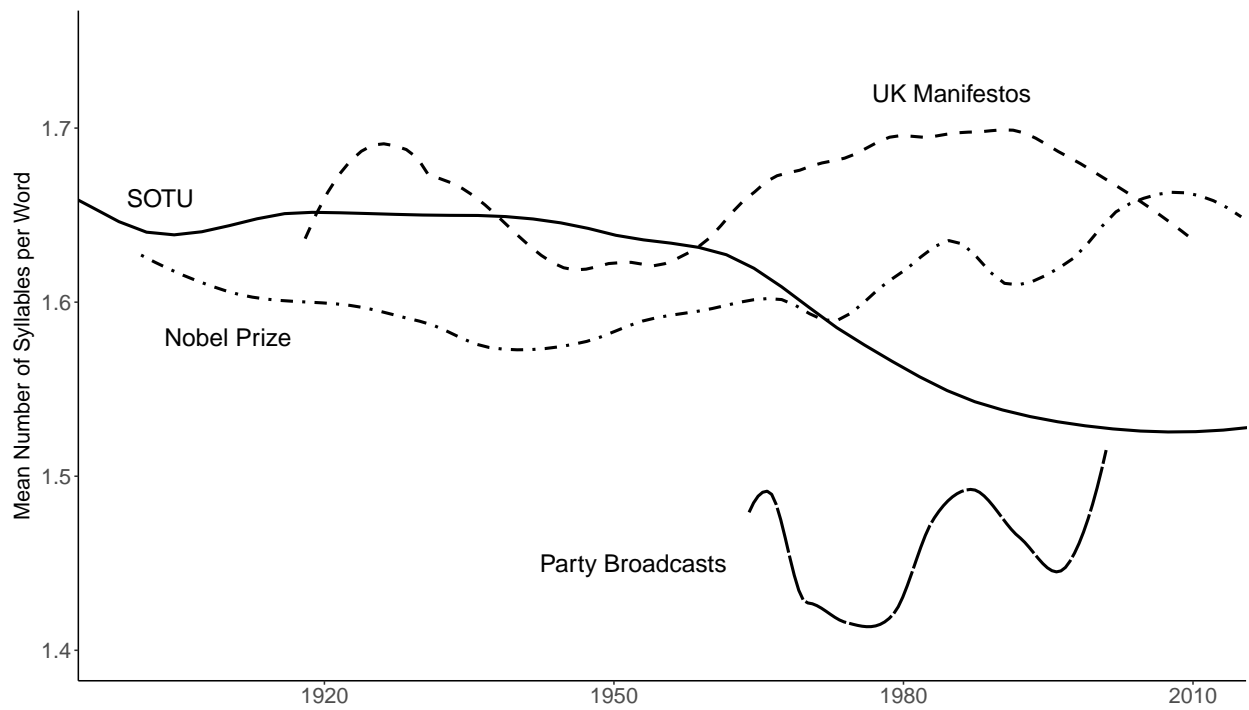


Figure 8: Mean number of syllables per word for four text series: Nobel Literature Prize Presentation Speeches, UK Conservative and Labour Party Manifestos, transcripts of UK Party Political Broadcasts, and for the State of the Union address.

sentence length, the standard deviation is 8.9; for word length, this figure is a much smaller 0.7. This means that this typical change for sentence length results in just over 9 (8.9×1.015) in the Flesch index, while a typical change in word length changes the Flesch index by just under 6 (0.07×84.6). Sentence length changes are weighted 1.5 times more than word length changes for these texts, in terms of the coefficients of the Flesch formula. Given the steady and marked changes in sentence length over the evolution of the US presidential addresses, these changes explain most of the widely noted trend to greater reading ease in this corpus.

Discussion

For some contemporary commentators, ‘anti-intellectualism’ is ever present, as if cultural baked into the historical heritage of public life in the United States (Hofstadter, 1963). For others, today’s “cult of ignorance” represents something new and more terrifying (Asimov, 1980). Thus, while 1968 presidential hopeful George Wallace railed against “pointy-head college professors who can’t even park a bicycle straight”, (current) Secretary of Education Betsy DeVos pushed this rhetoric further by encouraging college students to “fight against the education establishment” whose elites “...tell you what to think.”⁸ It would seem that systematic evidence for evaluating such baleful trends would be easy to come by. And, to the extent that the SOTU addresses appear to show increasing reading ease, scholars and journalists who draw anxious conclusions from that corpus might well have a point. Having assessed the SOTU data in some detail, however, we are skeptical of the need for panic—at least from this time series alone.

In particular, our reexamination of the upward trend in the SOTU readability patterns has revealed several significant qualifications. First, if one is serious about uncertainty around the measures, the ‘standard’ picture changes: while means tell one fairly clear story, appropriately accounting for sampling variation tells a more confusing tale.

⁸“U.S. Secretary of Education Betsy DeVos Prepared Remarks at the 2017 Conservative Political Action Conference”, <https://www.ed.gov/news/speeches/us-secretary-education-betsy-devos%E2%80%99-prepared-remarks-2017-conservative-political-action-conference>.

Second, and importantly, the SOTU trend towards much greater reading ease is not general in political communication. Indeed, it appears to be unique even among presidential texts, for those we examined. Thus, presidential executive orders, after remaining fairly constant at a college-plus reading ease level, became considerably more complex around 2000 due to the shift toward longer sentences. The language complexity of the Supreme Court has remained almost unchanged since 1800. Other texts, such as the party platforms of the main UK parties, and the Nobel Prize acceptance speeches, are also unchanged in terms of their average simplicity. It is true that UK party broadcasts are less complex than SOTU addresses, but this is not a new phenomenon, and some modern ones are approximately on a par with those from the 1960s.

Third, examining the components of the index, sentence length and word length, we find that most of the changes are being driven by sentence length, although from about 1960 we acknowledge that there is also a trend toward shorter words in the SOTU addresses. In the first decade of SOTU addresses, the typical sentence was twice as long as the typical sentence from those in the 2000s. Part of this trend comes from the shift from written to spoken addresses, but was evident even in those that were written. We suspect that shorter sentences—as noted above, common to most other series—are likely to reflect changing communication styles rather than being *per se* evidence of “dumbing down” of content. In this sense, our efforts here should ameliorate concerns about the quality of political communication, at least from this particular source of democratic anxiety.

Statistical measures notwithstanding, there are substantive reasons to doubt that simpler presidential addresses are a cause for concern for the health of democratic deliberation. In social science, when scholars have applied readability measures to link linguistic sophistication to outcomes, greater clarity is generally associated with beneficial outcomes. Put crudely, if one can understand a speech or document, one may find it easier to adjust behavior and hold the author to account. Jansen (2011), for instance, studies the reading level of communications by four central banks, equating lower reading levels of bank communication with greater clarity, which they link to positive effects on the volatility of returns of financial markets. Likewise, Owens and Wedeking

(2011) and Spriggs (1996) examine the complexity of Supreme Court decisions, pointing to the importance of clarity in court opinions, although their approaches do not rely on measures of reading level. In the context of the British parliament, Spirling (2016) applied readability measures to document the democratizing effects of franchise reform on elite speeches. Finally, as a meta-analysis to defend against charges of elitism and jargon (e.g. Diamond, 2002; Kristof, 2014), Cann, Goelzhauser and Johnson (2014) show that while the reading ease of articles in the top political science journals has declined since 1910, the typical political science article requires less reading ability than the typical article in *Time Magazine* or *Reader's Digest*. Sometimes a complex idea is best communicated using simple, short sentences.

References

- Asimov, Isaac. 1980. "A Cult of Ignorance." *Time Magazine* January(21):19.
- Bai, Jushan and Pierre Perron. 2003. "Computation and Analysis of Multiple Structural Change Models." *Journal of Applied Econometrics* 18:1–22.
- Benoit, Kenneth, Michael Laver and Slava Mikhaylov. 2009. "Treating Words as Data with Error: Uncertainty in Text Statements of Policy Positions." *American Journal of Political Science* 53(2):495–513.
- Cann, Damon, Greg Goetzhauser and Kaylee Johnson. 2014. "Analyzing Text Complexity in Political Science Research." *PS: Political Science & Politics* 47:663–666.
- Coleman, M and T Liau. 1975. "A computer readability formula designed for machine scoring." *Journal of Applied Psychology* 60(2):283–284.
- Dale, Edgar and Jeanne Chall. 1948. "A Formula for Predicting Readability." *Educational Research Bulletin* 27(1):11–20.
- Diamond, Larry. 2002. "What Political Science Owes the World." *PS: Political Science & Politics Online Forum* pp. 113–27.
- Flesch, Rudolph. 1948. "A new readability yardstick." *Journal of Applied Psychology* 32(3):221–233.
- Flynn, James R. 1999. "Searching for Justice: The Discovery of IQ Gains over Time." *American Psychologist* 54(1):5–20.
- Galston, William. 2001. "Political Knowledge, Political Engagement, and Civic Education." *Annual Review of Political Science* 4:217–234.
- Gatto, John Taylor. 2002. *Dumbing us down: The hidden curriculum of compulsory schooling*. Vancouver: New Society Publishers.
- Gunning, Robert. 1952. *The Technique of Clear Writing*. New York: McGraw-Hill.
- Hofstadter, Richard. 1963. *Anti-Intellectualism in American Life*. New York: Knopf Books.
- Jansen, David-Jan. 2011. "Does the Clarity of Central Bank Communication Affect Volatility in Financial Markets? Evidence from Humphrey-Hawkins Testimonies." *Contemporary Economic Policy* 29(4).
- Kincaid, J Peter, Robert Fishburne, Richard Rogers and Brad Chissom. 1975. *Derivation of New Readability Formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease formula) for Navy Enlisted Personnel*. Vol. Research Branch Report 8-75 Naval Air Station Memphis: Chief of Naval Technical Training.
- Kristof, Nicholas. 2014. "Professors, We Need You!"

- Lim, Elvin. 2008. *The Anti-Intellectual Presidency*. New York: Oxford University Press.
- McLaughlin, Harry. 1969. "SMOG Grading: a New Readability Formula." *Journal of Reading* 12(8):639–646.
- Owens, Ryan and Justin Wedeking. 2011. "Justices and Legal Clarity: Analyzing the Complexity of Supreme Court Opinions." *Law & Society Review* 45(4):1027–1061.
- Spache, George. 1953. "A new readability formula for primary-grade reading materials." *The Elementary School Journal* 53(7):410–413.
- Spirling, Arthur. 2016. "Democratization of Linguistic Complexity: The Effect of Franchise Extension on Parliamentary Discourse, 1832–1915." *Journal of Politics* 78(1):120–136.
- Spriggs, James F. II. 1996. "The Supreme Court and Federal Administrative Agencies: A Resource-Based Theory and Analysis of Judicial Impact." *American Journal of Political Science* 40:1122–1151.