

Economical Writing (or, “Think Hemingway”)

Andrés Marroquín and Julio H. Cole
Universidad Francisco Marroquín (Guatemala)

ABSTRACT

This paper is based on an argument proposed by Salant (1969), who complained that on many occasions he found the writing of his colleagues “nearly incomprehensible,” and made suggestions to improve economists’ writing skills. Among other things, he argued that a “simple way of avoiding clumsiness is to prefer the short word to the long one and to avoid the unfamiliar word if a familiar one can be found that is equally correct, specific, and concrete” (p. 556). We call this “the Salant hypothesis,” and use ANOVA to test this hypothesis by comparing the average length of words used by Nobel laureates in their banquet speeches. We find that Literature laureates tend to use shorter words than laureates in other disciplines, and the difference is statistically significant. These results support Salant’s idea that words should be used efficiently. This includes using short words instead of longer ones whenever possible. In short, good writing is also “economical writing.”

Key words: word length, economical writing, Salant hypothesis, ANOVA

Corresponding Author:

Dr. Andrés Marroquín, andresmg@ufm.edu

First draft: Dec 17, 2011

Revised: Jan 23, 2012

1. INTRODUCTION.

The motivation for this paper is an argument proposed by Walter Salant in a paper entitled “Writing and Reading in Economics” (Salant 1969). Salant voiced many criticisms of the economics profession in this paper, but his chief complaint was that on many occasions he found the writing of his colleagues “nearly incomprehensible.” Too often, he said, it is difficult to understand the main message that an economist is trying to communicate: “Crimes of violence are committed daily against the English language and the helpless reader is often frustrated in his efforts to understand the message” (p. 545).

Salant also gave some advice to improve economists’ writing. Good writing, he argued, starts with good logic and a prior outline of what one wants to say:

... the first requirement of clear writing is that the writer must know what he wants to say and say it in logical order. I would have assumed that this elementary fact did not need to be stressed if I had not seen evidence that it does. I have read manuscripts in which important points are in footnotes and unimportant ones are in the text. If your thoughts do not proceed in a logical sequence, no elegance or grace of language will make your writing clear I think it should be taken for granted that before one begins to write, he should have in mind—and most of us need to have on paper—an outline of what he is going to say. That will enable him to see his article or book as a whole and to see if the points he wants to make are relevant to his message and are stated in logical order. If, while writing, the author becomes aware that he has deviated much from his outline, he should make an outline of what he actually has written and see if it, although different, is equally logical (pp. 546–47).

Salant argued that mathematical formulas should be used only when they convey a message that cannot be said in words, or when doing so would require a very large amount of words and sentences:

I now come to the problem which I consider most serious in manuscripts prepared by modern economists. It is generally referred to as the use of mathematics in the exposition, but the word “use” in that description is rather loose. Very often the mathematics is not

really being used; all that is being used is mathematical notation. In other cases the mathematics is being used, but there is the question of whether its use is justified. Both abuses raise serious problems of communication (p. 553).¹

Salant also stressed that using a word that does not fit exactly the meaning of the thoughts or concepts one wants to express is a “misallocation of scarce verbal resources” (p. 547). Economists, he argued, often ignore subtle but important distinctions among words, and as a consequence the message is confusing.² He also criticized what he calls “elephantine language,” i.e., economists’ excessive use of nouns to modify nouns:

I have been reading manuscripts that assaulted the reader with three, four, and even five nouns placed consecutively. All but the last were intended as adjectives or parts of an adjectival phrase. For example, in one manuscript, I found within a few pages of each other all the following: “high risk flood plain lands,” which presumably means plain lands in which the risk of floods is high; “aircraft speed class sequencing,” which uses three nouns to modify a word that might be a noun if it existed but does not really exist; and then, to top it off, “terminal traffic control program category,”

¹Salant did admit, of course, that there is a proper place for mathematics in economics writing: “Casting a statement in mathematical terms is justified if the verbal statement of the equation or the inequality would be more complicated or as unfamiliar to the reader as the mathematical statement. In that case, nothing is lost by using mathematical notation and something may be gained Of course there are times when mathematical operations really have to be used to reach a conclusion. In such cases, the decision about using mathematics in the text should depend on the nature of the intended audience. But it is certainly a good rule for all publications that are not intended only for a professional audience to say in words both the conclusions and whatever can be said about the reasoning, and to confine the mathematics to an appendix. This principle can be followed without much strain. The author who follows it will still have the opportunity, if he wants it, to tell the mathematical reader how he reached his conclusions, and he will gain a wider audience Apparently words can express not only propositions of economic theory that are derived mathematically but, according to the late Stefan Valavanis, those of econometrics, too. In the Preface of his textbook *Econometrics* [Valavanis 1959] he expresses this opinion in one sentence: ‘If anything in Econometrics (or in any other field) makes sense one ought to be able to put it into words.’ Copies of this statement suitable for framing should be made freely available to all members of our profession” (pp. 554-56). Note that Salant wrote this in 1969. One can only wonder what he would have thought about the typical economics journal article four decades later!

²For instance, he claimed that economists usually confuse “expect” with “anticipate.” To anticipate something, he says, is not merely to expect it but to do something as a consequence of the expectation (p. 548).

which contains five consecutive nouns. I leave it to you to figure out which of these words modifies which (p. 549).

In castigating economists' bad writing, Salant did not neglect to mention the appropriate use of commas, adequate use of pronouns, and the importance of precision in writing. To illustrate precision he offered this example:

You must have heard about the steelworker who was teaching an apprentice how to rivet. "I'll hold the rivet," he said, "and when I nod my head, you hit it with the sledge hammer." The apprentice did. His instructor left a widow and three children (p. 552).

2. GRACE AND FORCE IN WRITING (THE SALANT HYPOTHESIS).

In the final section of his paper, Salant argued that economic writing, like all writing, should have grace and force. We should pay attention, for instance, to how the words sound to make the reading of the manuscript more enjoyable:

At a conference in Bellagio, someone asked Fritz Machlup why he had used one word in a draft rather than another, longer one that the questioner thought more appropriate. Machlup said, "Because it is more euphonious." On asked if he really paid much attention to that criterion when he wrote, he said, "Absolutely. I ask myself, 'Does it sing?'" (p. 556).

One of Salant's final ideas on this subject is the one that we are mainly concerned with in this paper: "*Another simple way of avoiding clumsiness is to prefer the short word to the long one and to avoid the unfamiliar word if a familiar one can be found that is equally correct, specific, and concrete*" (p. 556, italics added). Salant took the prose of Ernest Hemingway as a standard of good writing:

That the length of words affects the cleanness and force of writing is not news. The powerful effect of short words hits one most forcibly in Ernest Hemingway's prose. His sentences strike like bullets. All are clear-cut and forceful. It is remarkable how many are built entirely of words of one syllable. Indeed, I found nearly a whole page of his writing that contained hardly a word of more than one syllable (p. 557).

Salant then compared a sample of Hemingway’s prose with selections of writing drawn from three major economists (including “two of the better writers among economists”):

To get a reasonable sample of each man’s writing I have counted the words used in a few paragraphs of each manuscript or published book, one or two from portions that do not deal with technical matters, and one or two from portions that do. I do not claim that the sample was chosen scientifically. The results, for whatever they are worth, are summarized in Table 1 (p. 557).

WRITER	PERCENTAGE OF WORDS HAVING				NUMBER OF SYLLABLES PER 100 WORDS
	One Syllable	One or Two Syllables	Three Syllables	Four or more Syllables	
Economist No. 1	57	76	18	6	175
Economist No. 2	52	81	14	5	173
Economist No. 3	64	81	13	6	163
Ernest Hemingway*	83	96	4	0	122

*From *A Moveable Feast* (Hemingway, 1964).

Salant concludes: “... it is striking that the Hemingway samples have only 122 syllables per 100 words and that only 4 per cent of his words contain three or more syllables” (p. 557).

If we all aspired to be Hemingways, we should have to work at least as hard on our writing as he did. In *A Moveable Feast* Hemingway said it sometimes took him *a whole morning to write a paragraph*. When you read that book you will see why. Although it would be a poor use of resources for economists to spend that much time in polishing, it is obvious that many of us should spend a good deal more time in revising our drafts than we do now. None of us has a right, even at the stage of drafting, to impose on others writing that does not meet the requirement of clarity. He who does so not only irritates the colleagues who must read what he writes, and wastes their time, but also forgoes the larger audience that might otherwise read what he has to say. *He thereby forgoes the influence his work might have* (p. 557, italics added).

3. TESTING THE SALANT HYPOTHESIS.

In this paper we take this finding seriously, and explore the hypothesis that using short words is a characteristic of good English writing. To do so, we compare the Nobel Prize banquet speeches delivered by Literature laureates with those of laureates in other fields (including Economics).³ Banquet speeches give us a standard that, we would expect, is less influenced by technical jargon.⁴ Here, the writers of these speeches were all placed in the same basic context and were addressing the same audience.

The texts for the banquet speeches were obtained from the Nobel Foundation website (<http://www.nobelprize.org/>). In our analysis, in addition to using *syllables* per word to measure word length, we also used *characters* per word. One would expect a positive correlation between these two measures, but the correspondence is not perfect, especially in English, where one can easily provide many examples of one-syllable words that are actually longer, letter-wise, than many two-syllable words. (In languages such as Spanish, for instance, where spelling is essentially phonetic, the correspondence between these two measures would be much closer.) Although they are not exact equivalents, both concepts—characters/word and syllables/word—seem like valid approaches to the more general notion of “word length,” and thus we feel that an exploration based on characters/word is interesting in its own right, and not just as a complement to Salant’s original syllables-based analysis. To measure the average number of syllables and characters per word, we selected the speech texts from the Nobel website, and then we used the “Syllable Counter & Word Count” software available at <http://www.wordcalc.com/>.

The first thing we wanted to investigate was whether Salant was right in choosing Ernest Hemingway as a paragon of good writing (or at least as a model short word user). Table 2 compares Hemingway’s Nobel banquet speech with the

³The Nobel “banquet speeches” are not to be confused with the “Nobel lecture” given by each laureate. In years when the prize is shared by several laureates, each laureate gives a Nobel lecture, but usually only one of them is invited to give a banquet speech.

⁴Most scientific writing (including economics) will *per force* rely on technical terms that are, characteristically, based on long words. Salant recognizes this point in his paper (p. 557).

speeches by the other laureates for 1954. Hemingway's average word length was 1.28 syllables/word and 4.22 characters/word, lower than for any of the other laureates, and almost 11% below the averages for the four non-Literature laureates (1.438 syllables/word and 4.718 characters/word, respectively). Note, in passing, the huge difference between Hemingway's averages and the averages for the Medicine laureate: almost 0.4 additional syllables, and 1.1 additional characters, *per word*!

	Words	Syllables/ word	Characters/ word
Physics (Born)	597	1.40	4.65
Chemistry (Pauling)	374	1.33	4.37
Medicine (Weller)	226	1.66	5.35
Peace (van Heuven Goedhart, UNHCR)	764	1.36	4.50
Literature (Hemingway)	334	1.28	4.22

This comparison is perhaps unfair to the non-Literature laureates, who do not, after all, write for a living. In Table 3 we compare Hemingway with the Literature laureates for two years before, and two years after 1954. The averages for all five laureates were 1.324 syllables/word and 4.418 characters/word, so Hemingway's word length was about 3.3% below average in terms of syllables/word, and about 4.5% below average in terms of characters/word. Salant's choice of Hemingway as a model seems to have been well-founded.

	Words	Syllables/ word	Characters/ word
Mauriac (1952)	1770	1.34	4.49
Churchill (1953)	513	1.39	4.66
Hemingway (1954)	334	1.28	4.22
Laxness (1955)	930	1.29	4.34
Camus (1957)	1668	1.32	4.38

Note: In 1956 Juan Ramón Jiménez could not attend the ceremony. He sent a brief thank you note. It should also be noted that three of the five speeches reported in this table are translations (Mauriac, Laxness and Camus).

The next step in our analysis is to see how economists and natural scientists compare in this regard. Table 4 shows summary statistics for the average word lengths in the banquet speeches of Nobel laureates for five

categories from 1969 to 2011.⁵ The sample period starts in 1969 because that was the first year in which the Economics prize was awarded. The Peace prize was omitted from the analysis because after 1992 the laureates for that category no longer gave banquet speeches (or at least these were no longer published on the Nobel website). To ensure comparability, only banquet speeches in English were included in the analysis.⁶

TABLE 4. NOBEL PRIZE BANQUET SPEECHES, 1969-2011 (summary statistics).					
(a) Syllables per word					
	ECON	PHYSICS	CHEM	MED	LIT
Mean	1.456	1.440	1.470	1.472	1.344
Median	1.46	1.44	1.47	1.47	1.35
Maximum	1.63	1.61	1.68	1.63	1.68
Minimum	1.24	1.21	1.29	1.31	1.19
Std. Dev.	0.106	0.104	0.085	0.088	0.103
(b) Characters per word					
	ECON	PHYSICS	CHEM	MED	LIT
Mean	4.737	4.754	4.845	4.816	4.470
Median	4.75	4.77	4.80	4.82	4.49
Maximum	5.17	5.36	5.54	5.25	5.39
Minimum	3.97	4.05	4.39	4.20	4.09
Std. Dev.	0.297	0.288	0.248	0.250	0.254
Observations	42	45	42	47	27

Note that the lowest average word lengths are for the Literature prize. In terms of syllables/word the Literature laureates' word lengths were, on average, almost 8% shorter than the weighted average for non-Literature laureates, and about 6.6% shorter in terms of characters/word. It is also interesting to note that, with one exception (Derek Walcott, 1992), all of the average word lengths for

⁵The full dataset is shown in the Appendix.

⁶The difference in sample size for each group is due to several factors. In some years there were no observations for some prize categories, either because the laureate did not attend the award ceremony, or because the banquet speech was not in English, or simply because the laureate did not give a banquet speech. On the other hand, in some years there were two banquet speeches in the same category, and in one year there were *three* banquet speeches for the same category (Medicine, 1982). After 1986, the protocol for the award ceremony seems to reflect the rule that when an award is shared by two or three laureates, only one of them is chosen to deliver the banquet speech for that prize category.

Literature laureates were less than 4.8 characters/word, and over half of them were less than 4.5 characters/word. For comparison, the *median* word length for the non-Literature laureates was 4.785 characters/word, and only about 16% had average word lengths below 4.5 characters/word.⁷

These differences are statistically significant. ANOVA tests for both syllables/word and characters/word reject the hypothesis of equal means (Table 5), which implies that at least one of the groups is statistically different from the others.⁸ On the other hand, ANOVA tests for only the non-Literature prizes (Table 6) do not reject the hypothesis of equal means: average word length is essentially the same for Economics, Physics, Chemistry and Medicine laureates. Therefore, it is only the Literature laureates that can be regarded as a separate group when it comes to word length, which is consistent with the Salant hypothesis: the very best writers do indeed tend to use shorter words. Conversely, one can also state this same conclusion by saying that economists and natural scientists tend to use longer words, on average, even when writing in a non-technical vein and for a general audience.

4. CONCLUSION.

Salant (1969) argued that the use of short words is an indication of good writing. We found support for this hypothesis by comparing the banquet speeches of Nobel laureates.

To be sure, word length is only one dimension of what makes for “good writing.” But it seems that it is a necessary dimension. Words should be used efficiently. This includes using short words rather than longer ones, whenever

⁷The “4.8 characters per word” standard is a hard act to follow. The average word length for the 2861 words in this paper (excluding tables and references) is 4.72 characters/word, which at first glance seems pretty good, although it is largely due to our extensive quotations from Salant (1969). Excluding the Salant quotations, our average word length is actually 4.86 characters/word. Salant himself, on the other hand, does quite well: for the 1143 words we quoted from him in this paper, the average word length is an impressive 4.52 characters/word.

⁸A standard ANOVA test assumes equality of within-group variances. As shown in Table 5, Bartlett tests for equality of variances do not reject the hypothesis of equal variances for the data in Table 4.

possible. “Economical” writing might indeed be the key for “good” writing.⁹ We leave for debate the different implications of our paper.

Table 5. ANOVA (all groups).						
(a) Syllables per word						
Source of Variation	SS	DF	MS	F	p-value	F-Crit (5%)
Between Groups	0.339939	4	0.084985	9.031225	0.0000	2.417
Within Groups	1.863202	198	0.009410			
Total	2.203141	202				
Bartlett test (p-value)	0.4870					
(b) Characters per word						
Source of Variation	SS	DF	MS	F	p-value	F-Crit (5%)
Between Groups	2.697770	4	0.674442	9.303732	0.0000	2.417
Within Groups	14.35334	198	0.072492			
Total	17.05111	202				
Bartlett test (p-value)	0.6559					

Table 6. ANOVA (non-Literature prizes).						
(a) Syllables per word						
Source of Variation	SS	DF	MS	F	p-value	F-Crit (5%)
Between Groups	0.030119	3	0.010040	1.086370	0.3564	2.657
Within Groups	1.589535	172	0.009241			
Total	1.619654	175				
Bartlett test (p-value)	0.3549					
(b) Characters per word						
Source of Variation	SS	DF	MS	F	p-value	F-Crit (5%)
Between Groups	0.333378	3	0.111126	1.507571	0.2143	2.657
Within Groups	12.67844	172	0.073712			
Total	13.01182	175				
Bartlett test (p-value)	0.5236					

⁹We have borrowed the expression “economical writing” from McCloskey (2000), though we use it here in a somewhat narrower sense than she does.

REFERENCES

Hemingway, Ernest. 1964. *A Moveable Feast*. New York: Scribner's.

McCloskey, D. N. 2000. *Economical Writing*, 2nd ed. Long Groove, IL: Waveland Press.

Salant, Walter S. 1969. "Writing and Reading in Economics," *Journal of Political Economy*, 77 (4): 545-58.

Valavanis, Stefan. 1959. *Econometrics: An Introduction to Maximum Likelihood Methods*. New York: McGraw-Hill.

APPENDIX – WORD LENGTHS IN NOBEL PRIZE BANQUET SPEECHES, 1969-2011.

(A more detailed dataset is available upon request to: andresmg@ufm.edu)

Year	Syllables per word					Characters per word				
	Econ	Physics	Chem	Med	Lit	Econ	Physics	Chem	Med	Lit
1969	1.38	1.49		1.42		4.40	4.77		4.75	
1970	1.35		1.38	1.36	1.31	4.51		4.49	4.43	4.49
1970				1.55					4.83	
1971	1.53	1.39	1.39	1.44	1.19	5.01	4.67	4.65	4.61	4.14
1972	1.32	1.58	1.52	1.41		4.36	5.13	5.11	4.71	
1973	1.34	1.47	1.49		1.23	4.59	4.92	4.78		4.12
1974	1.47	1.41	1.52	1.34	1.35	4.75	4.65	4.89	4.47	4.49
1974				1.48					4.82	
1975	1.58	1.45	1.29	1.53	1.35	5.07	4.93	4.44	4.84	4.38
1975		1.56					4.88			
1976	1.37	1.58	1.55	1.59	1.40	4.52	5.12	4.97	5.04	4.51
1976				1.49					4.87	
1977	1.61	1.44	1.52	1.32		5.08	4.78	4.84	4.30	
1977				1.44					4.78	
1978	1.48	1.37	1.48	1.63	1.31	4.75	4.57	4.80	5.18	4.41
1978		1.58					5.09			
1979	1.37	1.34		1.38		4.60	4.52		4.52	
1980	1.51	1.57	1.46		1.52	4.94	5.08	4.78		4.79
1980			1.29					4.39		
1981	1.56	1.53	1.45	1.33		5.02	5.04	4.55	4.47	
1981			1.38					4.65		
1982	1.54	1.55	1.40	1.55		4.98	5.26	4.59	5.14	
1982				1.46					4.80	
1982				1.59					5.02	
1983	1.61	1.45	1.41	1.58	1.38	5.17	4.58	4.61	4.97	4.45
1983		1.27					4.21			
1984	1.37	1.44	1.34	1.48		4.54	4.74	4.45	4.81	
1985	1.29	1.38	1.45	1.52		4.29	4.65	4.80	5.10	
1985				1.44					4.69	
1986	1.57	1.54	1.43	1.50	1.40	5.11	5.05	4.63	4.90	4.65
1987	1.35	1.52		1.58	1.31	4.42	4.93		5.01	4.21
1988		1.56	1.42	1.53			4.98	4.88	5.08	
1989	1.39	1.61	1.52	1.37	1.37	4.53	5.36	4.91	4.73	4.53
1990	1.35	1.30	1.54	1.58	1.40	4.30	4.37	5.12	5.19	4.62
1991	1.51	1.21	1.40	1.40	1.38	4.86	4.05	4.74	4.72	4.57
1992	1.60	1.40	1.47	1.38	1.68	5.03	4.72	4.74	4.64	5.39
1993	1.60	1.50	1.45	1.47	1.35	5.02	4.85	4.80	4.68	4.48

(cont.)

Year	Syllables per word					Characters per word				
	Econ	Physics	Chem	Med	Lit	Econ	Physics	Chem	Med	Lit
1994	1.57	1.60	1.58	1.47	1.37	5.15	5.21	5.12	4.96	4.47
1995	1.44	1.36	1.63	1.46	1.38	4.55	4.61	5.26	4.75	4.55
1996	1.30	1.42	1.58	1.35		4.29	4.75	5.13	4.45	
1997	1.44	1.42	1.41	1.52	1.25	4.91	4.72	4.78	5.06	4.20
1998	1.43	1.34	1.51	1.60	1.39	4.73	4.45	4.90	5.06	4.59
1999	1.24	1.24	1.40	1.57		3.97	4.31	4.72	5.13	
2000	1.63	1.44	1.68	1.55		5.14	4.83	5.54	5.08	
2001	1.31	1.49	1.48	1.43	1.24	4.42	4.99	4.87	4.78	4.35
2002	1.51	1.33	1.53	1.31	1.42	5.02	4.53	4.93	4.20	4.58
2003	1.54	1.37	1.55	1.47	1.19	4.95	4.43	5.26	4.90	4.29
2004	1.32	1.43	1.55	1.46		4.48	4.81	5.03	4.86	
2005	1.41	1.32	1.52	1.33		4.64	4.58	5.00	4.41	
2006	1.45	1.25	1.45	1.52	1.22	4.58	4.29	4.73	5.05	4.09
2007	1.47	1.37	1.38	1.39		4.79	4.46	4.80	4.69	
2008	1.44	1.49	1.60	1.55		4.73	4.83	5.28	4.97	
2009	1.54	1.52	1.45	1.46	1.34	4.96	4.77	4.78	4.67	4.57
2010	1.50	1.52	1.42		1.26	4.75	4.86	4.71		4.28
2011	1.56	1.39	1.48	1.60	1.31	5.05	4.59	5.02	5.25	4.50