Economical Writing (or, "Think Hemingway")

Andrés Marroquín
Professor of Economics
Tel: +(502) 3023-3482
andresmg@ufm.edu

Julio H. Cole
Professor of Economics
Tel: +(502) 2338-7918
jhcole@ufm.edu

Universidad Francisco Marroquín (Guatemala)
Facultad de Ciencias Económicas
Calle Manuel F. Ayau (6 Calle Final), Zona 10
Guatemala, Guatemala 01010

Corresponding Author: Dr. Andrés Marroquín

ABSTRACT

Salant (1969) complained that on many occasions he found the writing of his fellow economists "nearly incomprehensible," and made suggestions to improve economists' writing skills (and, by extension, those of natural and social scientists in general). Among other things, he argued that good writers tend to use shorter words. We call this "the Salant hypothesis," and use standard statistical techniques to test this claim 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."

First draft: Dec 17, 2011
2nd revision: Aug 18, 2012
3rd revision: Nov 20, 2014

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 argued, it is difficult to understand the main message that an economist is trying to communicate.

Salant also gave some advice to improve economists' writing. He stressed, for instance, 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. 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 of 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," which contains five consecutive nouns. I leave it to you to figure out which of these words modifies which (p. 549).

In castigating economists for their bad writing, Salant did not neglect to mention the excessive use of mathematics, noting that "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). Salant argued that mathematical formulas should be used sparingly, and 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. Note that Salant was writing in 1969. Given the trend towards ever-increasing mathematization in the economics literature (Debreu 1991; Coelho and McClure 2005; Sutter and Pjesky 2007), one can only wonder what he would think about the typical economics journal article some 45 years later!

2. Grace and Force in Writing.

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 being

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 thoughts 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 model 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).

TABLE 1. HEMINGWAY VS. THE ECONOMISTS (from Salant 1969, p. 557).					
	PERCENTAGE OF WORDS HAVING				NUMBER OF SYLLABLES PER 100 WORDS
WRITER	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 noted that "... 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), adding that

> … [i]f 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. Literature Review.

The idea that short words make for better writing has long been a major feature of books on how to improve writing style (*e.g*., Strunk and White 1959; Zinsser 1990). Indeed, it is Rule No. 2 in George Orwell's famous list of six rules on how to avoid bad English prose: "Never use a long word where a short one will do" (Orwell 1968 [1946], p. 139).

Of course, it is one thing to recommend using short words as a normative standard, but quite another to provide evidence supporting the stronger claim that good writers really do use this rule in practice. We will call this claim the "Salant hypothesis." To our knowledge there are no previous empirical studies that directly address this issue, though there are quite a few recent studies, in a wide range of disciplines, that deal with the related issue of "readability."

"Readability scores" purportedly measure the ease with which texts can be read and understood. The most commonly used measure, the so-called Flesch index, is based on a formula that combines two simple statistics derived from the text being evaluated: average word length (in syllables per word) and average sentence length (in words per sentence) (Flesch 1948; Klare 1974; Tekfi 1987). The Flesch index and its variants correlate highly with reading comprehension based on reading tests (DuBay 2006).

Hartley, Sotto and Pennebaker (2002) found that papers judged to be "highly influential" in psychology had significantly higher Flesch scores than a matched control sample of articles that were published in the same journal issues as the influential ones. (Since only the overall Flesch scores are reported, it is not possible to determine how much of the higher scores for the more influential authors is due to shorter words, and how much to shorter sentences.) In a similar study, Sawyer, Laran and Xu (2008) found that readability scores for award-winning articles in four top-level marketing journals were significantly higher than for their matched controls.

Hartley, Sotto and Pennebaker (2002) also assessed readability in two samples of highly *cited* articles in psychology and found that, on average, there was no difference in the readability scores of highly cited articles as compared to their matched controls. Thus, it appears that highly *rated* articles are more readable than the average journal article, but highly *cited* articles are not.

This latter result is compatible with the findings of a large literature dealing with the characteristics of highly cited scientific articles. A consistent finding in many of these studies is that readability and/or clarity of expression have little or no impact on citation counts (Sternberg and Gordeeva 1996; Aksnes 2003; Nieminen *et al*. 2006; Stremersch,

Verniers and Verhoef 2007). In fact, some recent studies find that the readability of an article (or of its abstract) is actually *negatively* related to citation counts (Gazni 2011; Didegah and Thelwall 2013; van Wesel, Wyatt and ten Haaf 2014).

Scientific and academic journals in general score very low on readability, and the problem seems to have been getting worse through time: Bottle *et al*. (1983) documented a noticeable decline in the readability of scientific papers over a 100 year period, and noted that by 1980 it appeared to have "bottomed out" at the "very difficult" level. A similar trend was later documented for French scientific literature (Bottle and Tekfi 1988).

It is often assumed that more prestigious journals tend to be more difficult to read, but Hartley, Trueman and Meadows (1988) examined the relationship between readability and prestige in academic journals in several fields and found little evidence to support this widespread belief: though they found a small negative correlation between prestige rankings and readability, it was not statistically significant, and they concluded that readability levels are not in fact correlated with prestige rankings. Shelley and Schuh (2001) addressed this issue in terms of journal "selectivity" (as measured by the acceptance rate of submitted manuscripts): in a comparison of 17 journals in the field of education studies, the main finding was that average readability in the "top" journals was essentially the same as in less selective journals.

Metoyer-Duran (1993) reported a rather intriguing finding: papers accepted for publication in a library science journal had lower readability scores than papers that were rejected. An even more intriguing result was reported by Hayden (2008): in a sample of papers published in the *British Journal of Surgery*, the mean readability score was significantly lower when English was the first language of the principal author.

4. Testing the Salant Hypothesis.

Though these studies are all very interesting, they relate only indirectly to what we have called the Salant hypothesis: the claim that good writers actually do tend to use shorter words in practice. That highly *cited* authors do not have high readability scores is perhaps not surprising, since no one claims that highly cited authors are necessarily good writers (though the negative impact of readability on citation counts is, frankly, somewhat disturbing). That highly influential and award-winning authors score high on readability is interesting and suggestive, but the fact that they are highly *rated* authors does not necessarily mean that they are good writers, and to simply assume that they are better writers just because they score high on readability would involve circular reasoning for our purposes, since short word usage is part of the *definition* of readability.

One problem, then, with testing the Salant hypothesis is that we need an independent definition of what constitutes a good writer. For purposes of this study we will not attempt to provide a definition of our own. Instead, we will simply defer to the judgment of the Swedish Academy: not every good writer is awarded a Nobel Prize in Literature, but anyone who *is* awarded a Nobel Prize in Literature is certainly a good writer by any reasonable standard.

To test the Salant hypothesis we will compare the Nobel Prize banquet speeches delivered by Literature laureates with those of laureates in other fields. (The Nobel "banquet speeches" are not to be confused with the "Nobel lecture" given by each laureate. In years when several laureates share one of the prizes, each laureate gives a Nobel lecture, but usually only one of them is invited to give the banquet speech.) Most scientific writing will *per force* rely on technical terms and expressions that are, characteristically, based on long words, but here the writers of these speeches were all placed in the same basic context and were addressing the same audience. The Nobel banquet speeches therefore can provide a standard that, we would expect, is less influenced by technical jargon.

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 mapping 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, on the other hand, where spelling is essentially phonetic, the correlation between these two measures would be much higher.) Although they are not exact equivalents, both measures—characters/word and syllables/word—are 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 "Word Counter" software for Macintosh OS X, Version 2.10.1, available for download at http://www.supermagnus.com/mac/Word_Counter/. To assess the accuracy of the word-counting algorithm we took a sample sentence from every tenth speech for each of the disciplines and compared the results with a manual count of syllables and characters. The software's performance was reasonably accurate, with only a few syllable overcounts in some of the more difficult words (usually names of persons and places). The average overcount was about 2.3 %, which we regard as adequate, since it is well known that automated syllable-counting is a very difficult problem, especially in English. (Character-counting, on the other hand, is relatively easy in English, since diacritical marks are rarely used, and most character-counting programs are very accurate.)

Table 2 shows summary statistics for the average word lengths in the Nobel banquet speeches for five categories from 1969 to 2013 (the full dataset is available upon request to the corresponding author). 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). The 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 (1982) there were *three* banquet speeches for the Medicine award. 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. Only banquet speeches in English were included in the analysis.

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, 7.6 % shorter than the average for all non-Literature laureates, and about 6.6 % shorter in terms of characters/word. Though at first glance these differences might not seem very large in practical terms, in fact they can add up to noticeable differences in readability scores (and in terms of the volume of publication, a 6.6 % difference in average characters/word would be equivalent to about 13 pages for a typical 200-page journal issue, *i.e.*, one additional article per issue).

TABLE 2. NOBEL PRIZE BANQUET SPEECHES, 1969-2013.						
(a) Syllables per word						
	LIT	ECON	PHYSICS	CHEM	MED	COMBINED NON-LIT
Mean	**1.493**	1.613	1.598	1.633	1.622	1.616
Median	**1.480**	1.610	1.600	1.645	1.630	1.620
Std. Dev.	**0.099**	0.117	0.108	0.095	0.091	0.103
(b) Characters per word						
	LIT	ECON	PHYSICS	CHEM	MED	COMBINED NON-LIT
Mean	**4.473**	4.752	4.750	4.830	4.827	4.790
Median	**4.480**	4.770	4.760	4.795	4.830	4.790
Std. Dev.	**0.244**	0.298	0.281	0.242	0.248	0.269
Observations	29	45	47	44	49	185

The differences reported in Table 2 are statistically significant. Analysis of variance (ANOVA) tests for both syllables/word and characters/word reject the hypothesis of equal means (Table 3), which implies that at least one of the groups is significantly different from the others. On the other hand, ANOVA tests for only the non-Literature prizes (Table 4) do not reject the hypothesis of equal means: average word length is essentially the same for the 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.

(The average word length in syllables/word for the Economics laureates in Table 2 is somewhat lower than the values reported by Salant for his three unnamed economists in Table 1. However, Salant's word samples were drawn from books and articles by economists writing in a professional capacity, so we would expect longer words, on average, due to a greater preponderance of technical terms. In a study comparing word samples drawn from 15 major economics textbooks [Gallagher and Thompson 1981] the

average word lengths ranged from 1.558 to 1.826 syllables/word, with an overall average of 1.677, so it would appear that Salant's "unscientific" sample was actually quite representative of professional writing in economics.)

Table 3. ANALYSIS OF VARIANCE (all groups).		
Syllables per word: *F* = 9.719 (*p* = 0.0000)		
Characters per word: *F* = 9.968 (*p* = 0.0000)		
df	5% critical value	
4, 209	2.415	Null hypothesis is rejected

Table 4. ANALYSIS OF VARIANCE (non-Literature prizes).		
Syllables per word: *F* = 0.971 (*p* = 0.4076)		
Characters per word: *F* = 1.282 (p = 0.2821)		
df	5% critical value	
3, 181	2.655	Null hypothesis is not rejected

(a) Syllables per word				
F	Degrees of freedom	5% critical value	p-value	
0.971	3, 181	2.655	0.4076	Null hypothesis is not rejected
(b) Characters per word				
F	Degrees of freedom	5% critical value	p-value	
1.282	3, 181	2.655	0.2821	Null hypothesis is not rejected

One possible distorting factor is that, even though the speeches we have studied were all written in English, not all of the laureates are native English speakers. To control for this "mother tongue factor," we repeated the analysis, but excluding all laureates who were born in countries where English is not the official language. ("Native English speakers" were thus defined as laureates born in the United States, the United Kingdom, Ireland, Canada, Australia, New Zealand, India, Pakistan, Nigeria, South Africa and the English-speaking Caribbean islands.)

Table 5 reports the summary statistics and ANOVA results for the restricted sample of native English-speaking laureates. The results are essentially the same, and the only difference is that the significance levels for the ANOVA tests are lower (which is to be expected given the smaller sample sizes). The average word lengths are actually quite

similar for both Tables 2 and 5, which suggests that our basic conclusion is unaffected by mother-tongue effects: good writers tend to use shorter words when writing in English, whether they are native English speakers or not.

TABLE 5. NOBEL PRIZE BANQUET SPEECHES, 1969-2011 (native English speakers).						
(a) Syllables per word						
	LIT	ECON	PHYSICS	CHEM	MED	COMBINED NON-LIT
Mean	**1.516**	1.613	1.597	1.641	1.620	1.618
Std. Dev.	**0.136**	0.112	0.112	0.108	0.086	0.104
ANOVA	Degrees of freedom	5% critical value	p-value	Null hypothesis is rejected		
$F = 2.685$	4, 126	2.444	0.0344			
(b) Characters per word						
	LIT	ECON	PHYSICS	CHEM	MED	COMBINED NON-LIT
Mean	**4.557**	4.753	4.751	4.853	4.829	4.796
Std. Dev.	**0.333**	0.291	0.296	0.278	0.232	0.274
ANOVA	Degrees of freedom	5% critical value	p-value	Null hypothesis is rejected		
$F = 2.483$	4, 126	2.444	0.0470			
Observations	10	35	25	28	33	121

5. Conclusion.

Salant (1969) argued that the use of short words is a characteristic of good English writing. We found support for this hypothesis by comparing the banquet speeches of Nobel laureates. To be sure, not every nuance of elegance and clarity in writing can be captured by simple word and sentence numerics, and word length is only one dimension of what makes for "good writing." But it seems to be a necessary dimension.

Getting published is a good thing, but one also wants to be read and (hopefully) understood, and for the reader's sake we should all make an effort to "minimize the volume and maximize the meaning," as Okulicz-Kozaryn (2013) puts it. To this end, words should be used efficiently, and this includes using short words rather than longer ones whenever possible. "Economical writing" does indeed seem to be an important aspect of good writing style.

REFERENCES

Aksnes, D. W. (2003). Characteristics of highly cited papers. *Research Evaluation*, 12(3), 159-170.

Bottle, R. T., Rennie, J. S., Russ, S., & Sardar, Z. (1983). Changes in the communication of chemical information I: Some effects of growth. *Journal of Information Science*, 6(4), 103-108.

Bottle, R. T., & Tekfi, C. (1988). Readability of French scientific texts. *Journal of Documentation*, 44(4), 339-345.

Coelho, P. R. P., & McClure, J. E. (2005). Theory and application: Does complexity crowd out evidence? *Southern Economic Journal*, 71(3), 556-565.

Debreu, G. (1991). The mathematization of economic theory. *American Economic Review*, 81(1), 1-7.

Didegah, F., & Thelwall, M. (2013). Which factors help authors produce the highest impact research? Collaboration, journal and document properties. *Journal of Informetrics*, 7(4), 861-873.

DuBay, W. H. (2006). *Smart language: Readers, readability, and the grading of text*. Costa Mesa, CA: Impact Information.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32(3), 221-233.

Gallagher, D. J., & Thompson, G. R. (1981). A readability analysis of selected introductory economics textbooks. *Journal of Economic Education*, 12(2), 60-63.

Gazni, A. (2011). Are the abstracts of high impact articles more readable? Investigating the evidence from top research institutions in the world. *Journal of Information Science*, 37(3), 273-281.

Hartley, J., Trueman, M., & Meadows, A. J. (1988). Readability and prestige in scientific journals. *Journal of Information Science*, 14(2), 69-75.

Hartley, J., Sotto. E., & Pennebaker, J. (2002) Style and substance in psychology: Are influential articles more readable than less influential ones? *Social Studies of Science*, 32(2), 321-334.

Hayden, J. D. (2008). Readability in the *British Journal of Surgery*. *British Journal of Surgery*, 95(1), 119-24.

Hemingway, E. (1964). *A moveable feast*. New York: Charles Scribner's Sons.

Klare, G. R. (1974). Assessing readability. *Reading Research Quarterly*, 10(1), 62-101.

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

Metoyer-Duran, C. (1993). The readability of published, accepted, and rejected papers appearing in *College & Research Libraries*. *College and Research Libraries*, 54(6), 517-526.

Nieminen, P., Carpenter, J., Rucker, G., & Schumacher, M. (2006). The relationship between quality of research and citation frequency. *BMC Medical Research Methodology*, 6(42), 1-8.

Okulicz-Kozaryn, A. (2013). Cluttered writing: Adjectives and adverbs in academia. *Scientometrics*, 96(3), 679-681.

Orwell, G. (1968 [1946]). Politics and the English language. In I. Angus & S. Orwell (Eds.), *The collected essays, journalism and letters of George Orwell* (vol. 4, pp. 127-40). New York: Harcourt, Brace & World.

Salant, W. S. (1969). Writing and reading in economics. *Journal of Political Economy*, 77(4), 545-58.

Sawyer, A. G., Laran, J., & Xu, J. (2008). The readability of marketing journals: Are award-winning articles better written? *Journal of Marketing*, 72(1), 108-117.

Shelley, M. C., & Schuh, J. H. (2001). Are the best higher education journals really the best? A meta-analysis of writing quality and readability. *Journal of Scholarly Publishing*, 33(1), 11-22.

Sternberg, R. J., & Gordeeva, T. (1996). The anatomy of impact: What makes an article influential? *Psychological Science*, 7(2), 69-75.

Stremersch, S., Verniers, I., & Verhoef, P. C. (2007). The quest for citations: Drivers of article impact. Journal of Marketing, 71(3), 171-193.

Strunk, W., & White, E. B. (1959). *The elements of style*. New York: Macmillan.

Sutter, D., & Pjesky, R. (2007). Where would Adam Smith publish today? The near absence of math-free research in top journals. *Econ Journal Watch*, 4(2), 230-240.

Tekfi, C. (1987). Readability formulas: An overview. *Journal of Documentation*, 43(3), 257-269.

van Wesel, M., Wyatt, S., & ten Haaf, J. (2014). What a difference a colon makes: How superficial factors influence subsequent citation. *Scientometrics*, 93(3), 1601-1615.

Zinsser, W. K. (1990). *On writing well: An informal guide to writing nonfiction*, 4th ed. New York: HarperCollins.