MedicalExpress

LETTER TO THE EDITOR

A comment on "Writing Good English: Is scientific English a Latin Language in Disguise?"

Carlos Alberto de Bragança Pereira[I]

[I] Universidade de São Paulo, Instituto de Matemática e Estatística e Faculdade de Higiene e Saúde Publica, São Paulo, SP, Brazil

Pereira CAB. A comment "Writing Good English: Is scientific English a Latin Language in Disguise?" MedicalExpress (São Paulo, online). 2018;5:ml18002

Received for Publication on January 26, 2018; **Accepted for publication** on February 1, 2018; **Online on** February 14, 2018

E-mail: cadebp@gmail.com

Dear Editor

I have read an article recently published in "Medical Express", "Writing Good English: Is scientific English a Latin Language in Disguise?"

With respect to the statistical analyses, I have some comments. I would like to point out that ANOVA, the statistical treatment used to compare the incidence of Latin/Greek words in English against Portuguese texts, is not appropriate for the data analyzed. The reason is very simple: the sample space of a normal distribution is the whole real set of numbers extending from $-\infty$ to $+\infty$. In contrast, the sample space of a proportion π extends from zero to one. Another point is that the beta distribution is symmetric only when (π = ½ - beta). The beta distribution, considered here, is one of the adequate distributions to analyze proportions. Asymmetry grows as π moves away from ½.

The author kindly provided me with the raw data allowing me to reanalyze the results within his focus. I came up with the results shown in Figure 3A which may be advantageously compared with Figure 3 of the article. Adding another technical argument and having adopted the beta distribution, I found that the log-odds, $\ln(\frac{\pi}{1-\pi})$, are normally distributed allowing us to apply standard statistical methods. My decision was to use confidence intervals to clearly illustrate the differences between English and Portuguese on the use of Latin/Greek words. I call the attention to the fact that the intervals for Portuguese are more asymmetric than the ones for English which are close to ½. To obtain the intervals for proportions the inverse transformations were taken from log-odds to proportions. These transformations through the log-odds made the distribution of the proportions be described as Logistic-Normal, as painstakingly discussed in: Aitchison J. (2003), The Statistical Analysis of Compositional Data, Blackburn Press.

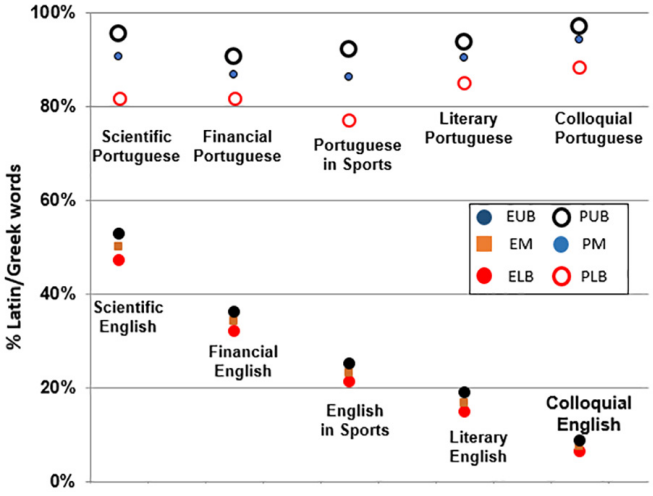

Figure 3A. 99% Confidence intervals for the proportions of Latin/Greek Words. EM, ELB, and EUB: mean, Lower and Upper bounds for English. PM, PLB, and PUB: mean, Lower and Upper bounds for Portuguese. No overlap between any of the English genres. Overlap for all Portuguese genres.

See also article: Writing good English: is scientific English a Latin language in disguise?
Available from: http://www.medicalexpress.net.br/details/315/writing-good-english--is-scientific-english-a-latin-language-in-disguise-

DOI: 10.5935/MedicalExpress.2018.ml.001