

Children and Adults as Amateur and Professional Authors: A Comparative Quantitative Analysis

Yulia Petukhova

Ruprecht-Karls-Universität Heidelberg

While some books written by amateurs and teenagers are published as hard copies, other literary works remain on the web forever. Do the latter suffer from the lack of managerial effort of their authors or are there formal features which make them somehow unfit for the book trade? It is assumed that learning creative writing helps eliminate typically amateur features and acquire the techniques of professional writers. In this project, I would like to challenge creative writing theory and see if those tools are what in fact distinguishes an amateur from a professional.

My doctoral research is aimed at finding formal characteristics of amateur adults' and children's writing which differentiate them from that of professional authors. This poster will show the results of the comparative analysis of one of the literary genres, poetry, authored by three groups of people: children and teenagers (age 5 to 18), adult amateur poets, and professional poets who are generally thought of as paragons of contemporary poetry.

The method of the analysis involves the use of Python for parsing the web-sites with the poems in question and create three corpora for lemmatization and POS-tagging, which prepares the texts for statistical analysis by the relevant features, derived mainly from creative writing textbooks. These features include, for example, frequency of adverb usage or frequency of the use of 'adjective plus noun' pattern, which are regarded as simplistic when it comes to poetry in creative writing classes. The question is whether professional authors are, in fact, characterized by a lower percentage of these parameters.

An essential method used in this project is topic modeling. Defining the semantic structures hidden in the texts provides the opportunity to see what concerns different groups of people divided by age and level of proficiency. It allows to see the trends that exist among younger writers who are much more uniform in the choice of topics than professional poets. While professional authors use indirect ways of approaching a topic, amateurs are more straightforward in conveying a message.

A simple look at the distribution of different parts of speech in the three corpora shows an interesting result. While professional authors are more likely to incorporate more nouns than adverbs, pronouns or verbs, amateur authors tend to use more pronouns and verbs. A close reading of the texts shows that professional authors do, in fact, use imagery based on nouns while amateur poets often concentrate on the expression of their feelings, using pronouns and verbs to describe interpersonal relationships.

Overall, NLTK marks the three corpora as neutral. Yet the results of the sentiment analysis show that the corpus of children's poems has the least extremely negative words and the most positive words while the amateur corpus demonstrates the opposite trend. It is indeed supported by a close reading of some of the poems from the two corpora.

One of the major challenges of this project is the quantitative analysis of figurative use of language. Currently, there are no efficient computational methods of approaching figures of speech. The use of figurative language is an important feature that differs in children's and

adults' writing. While its statistical analysis would be invaluable for this project, computer recognition of metaphors is an area that requires further research. In this project, it is partially tackled by examination of semantic trees. Children tend to stay within the same topic throughout the whole poem while adult professional writers incorporate unexpected semantic groups which appear to be juxtaposed with the branches of the main topics.

This research is a work in progress which is supposed to elucidate the underlying features that make a professional writer and those features that can strike the eye as childish or amateur. It is also aimed at challenging the current state of the creative writing theory.

Bibliography:

Archer, Jodie and Jockers, Matthew L. *The Bestseller Code: Anatomy of the Blockbuster Novel*. St. Martins Press, 2016

Groff, Patrick J. "Figures of Speech in Poems by Children" *The Elementary School Journal*. Vol.63, The Univ. of Chicago Press, 1964, pp. 136-40

Harper, Graeme, ed. *Teaching Creative Writing*. Continuum, 2006

Jockers, Matthew L. *Text Analysis with R for Students of Literature*. Springer, 2014

Mills, Paul. *The Routledge Creative Writing Coursebook*. Routledge, 2006