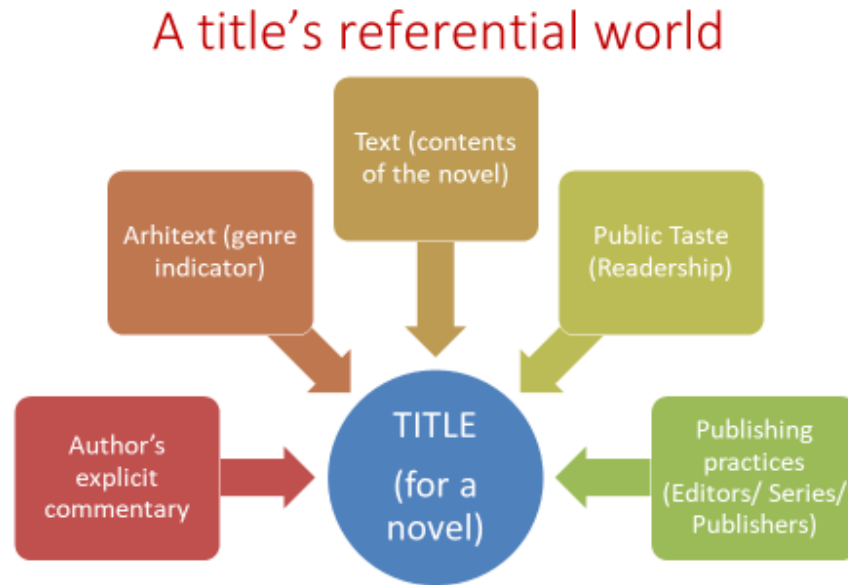


Some thoughts on our “title story”

Adapting Kant’s blunt questions to our “title story” – What can I know? What should I do? What may I hope? – I’d start with a barbarian schema on a title’s possible co-references, which will enable me develop some thoughts on the limits of knowledge and on the infinity of hope:



Of course, I could have gone for a bipolar visualization, inspired by Genettes’s metaphor of “thresholds:” ELTeC titles in between *text-world*, *fiction-reality* and so forth. But do we now anything about either the historical-social context or the textual dimension of our novels? And if we knew (from literary histories or anywhere else¹), would we be in right to treat either of them as homogeneeneous? Truth is that the linguistical heterogeneity of texts as well as their diverse deep contingencies make the dichotomy “text-world” approach rather slippery. My intuition is that a larger range of co-referents could help us understand which one is more prominently expressed in titles. So, I’d ask first how much information from the 5 possible co-referents is actually poured into a title and what means/ strategies/ practices there are to emphasize one or more referential sources. *Note 1: publishing practices* and *arhitext (genre indicator)* can overlap to a certain degree but I’d rather take them separately. Our data has shown so far that some of our categories tend to be more dynamic (persons), and some other tend to be more stable (places).

¹ I’m not sure if we will ever have a properly documented historical point of view for all collections: my point is that people who will annotate will also have to document possible trends/ concepts (for instance, “le roman personnel”) which haven’t probably been categorized as such in their literary tradition. Even if I could say that bildungsroman + other 1st person novels (or very lyrical ones) can be considered “romans personnels”, I’m not sure that the Romanian literary historians have used this phrase as such.

Indeed, we can't trust very dynamic categories because our research has 3 *methodological concerns* that, in spite of all cautions, will make our results very volatile and debatable. In what follows I'd like to propose several solutions to address these risks:

1. Collections will grow, so we can't hope to state anything about ELTeC in general or to make a possible correlation between titling practices and sampling principles; even if ENG, FRA, GERM, and perhaps POR collections are quite complete, the overall results on ELTeC dynamics will be problematic. My idea for addressing this is to focus on categories that tend to be more stable (placeEntities), then try to identify what are the constant features of personEntities and otherEntities. In any case, for persons and other entities we can't hope for more than a descriptive approach of data (social status is indeed the most interesting).
2. We haven't planned to correlate our findings with the title's coreferentials or at least to choose those ones that are more frequently poured in our titles: text, author, arhitext (genre indicator), public taste (readership), publishing practices. Don't worry, some of them will be hard to track in ELTeC titles (publishing practices and readership), others are problematic (we know a lot about titles but we don't know a thing about what they stand for), but there's a lot of hope with *arhitext* (genre indicator) and probably *author's comment*, the latter only if the explicitness + specificity of titles was analyzed. For instance, a proper name (e.g. "Lorenzo Benoni") shows a strong relationship between title and text because we expect that Lorenzo Benoni will be in the text; the same goes with all PersonEntities, but I guess it's not valid for PlaceEntities and OtherEntities too. I'll have to compute on text-title co-referencing if places or other entities are mentioned, but I'll skip it for the moment. Since texts will keep on being "the great unknown" for us — thus we may not hope that the title-text relationship pops out from the analysis of a few ELTeC titles — maybe we should try to get a deeper knowledge of the attributive density of titles and actually find out what is the author already saying/ explaining about his character, places and other entities: Lorenzo Benoni is, first of all, "un italiano," thus a specimen of his race, nation, etc.
3. Titles evolve because of changing editorial practices. Paul Nulty's article shows that Hathi Trust, ECCO, and BL 19th-century corpus illustrate various editorial practices, the most conservative (thus having the longest titles, that is, the original ones) being the BL corpus. We can sort out this issue by making a correction to our list of titles in order to have, at least for this study, a homogeneous editorial practice in titling.

A note on Nulty's article: Moretti claims that titles tend to become shorter as the book-markets get more crowded. A consequence of title-shrinking is that fewer verbs occur, so that *the use* (who's use? here I could trace a hint about "the author's comment") of nouns, adjectives, and articles turns important. However, we must keep in mind that his 7,000 fiction items have been pre-processed to the HathiTrust "trimming" standard, in the sense that the BL subtitles and meta-textual information (e.g. "in three volumes", "a classic tale") was excluded from analysis. Nulty does pretty much the same stating the following: "Is the syntactic pattern of the title related to the types of nouns in the title? The table below shows the frequency with which nouns occur in titles that contain adjectives compared to their frequency in titles without adjectives. The data is from short titles (four words or fewer) in the Hathi metadata. Meta-textual words (such as 'story', 'poem', 'novel', 'tales') have been removed from the data." I checked Nulty's lists of nouns and what do I find? Among the first 50 entries of fewer-than-four-words titles there are a lot of genre indicators: "legends" (pos. 17), "drama" (pos. 19), "history" (pos. 21), "melodies" (pos. 22), "plays" (pos.

28), “sonnets” (pos. 29), “pieces” (pos. 31), “romance” (pos. 32), “dramas” (pos. 36), “musings” (pos. 42), scenes (pos. 44), theatre (pos. 45).

Cutting this story short:

What can we know about ELTeC titles? We can know more about genre indicators (title-genre) and author’s comments (attribution density), and lesser about how titles relate with what they stand for or with the persons they are meant for (title-text; title-the text’s contingency)

What should we do? We can chose to stick to what does not seem to change (more stable categories of analysis) in an equation that already has 3 variables (a. ELTeC growing; b. no analysis of dynamic between title-text; c. title changes over time)

What may we hope? This one is my favorite because I can formulate, at last, the humble PURPOSE of this research: this research aims to know more about categories/ types of titles² + practices of titling and hopes to devise a solution for sampling a representative title population (for title lists longer than ELTeC).

² By “categories of titles” I don’t understand only “short titles” or “long titles” (which as proven my Moretti and Nulty are very relative), but also: a. “person-driven titles”, “place-driven titles”, “other-driven titles”;
b. “specific/ explicit titles”, “non-specific/ non-explicit titles”
c. “self-referential titles”, “referential titles”, “non-referential (??) titles”