

Exploring titles in ELTeC-deu. First thoughts

Carolin Odebrecht

March 26, 2020

1 Context

Corpus design enables empiric studies and sets a formal and (hopefully) theoretical motivated ground. However, it limits the scope of investigation at the same time.

The context of data creation needs to be considered in this study. Among many things, the following is crucial: ELTeC's text selection and proportion criteria are not based on any title characteristics or theoretical frameworks concerning para-texts. Furthermore, we do not cover any statistical representative population of either novels nor title types (for the 19th century).

2 Perspective

The following ideas and thought are grounded in a corpus linguistic perspective which is trained to see, model and calculate patterns. It is further informed by a non-professional reader's perspective (meaning being not trained in literary scholar). Therefore, I will focus on surface information and keep out the historical and literary context in the next sections. It might be a chance to detect hints and features that are not directly linked to existing theoretical frameworks. It will be necessary to link them afterwards of course.

By surface information, I do not refer to means as title length (counting words or characters) or solely counts most frequent parts of speech categories or ngrams. By surface information I refer on syntactic and semantic structures that can be modelled by analysing the linguistic material. It is a focus to investigate how for example a noun is modified, not only if.

3 Focus I Le roman personnel

It seems to be common to use a reference to a person, whether direct or indirect, in titles. It ranges from bare surnames (*Anna*, DEU018; *Sibylle* DEU020), full names (*Jürg Jenatsch* DEU032; *Mathilde Möhring* DEU088), status or professional (*Der Sonnenwirt* DEU005) or social titles (*Graf Petöfy* DEU053) or the combination of both (*Meister Timpe* DEU069).

Next to the references of single persons, social groups (*Weisse Sklaven oder die Leiden des Volkes* DEU0001) can also be found.

When we focus for the moment on references to persons, we might be interested in what we learn about these persons by only reading the title. A linguistic pattern that realises such information is modification. The way of modification can either add characteristics to the person being referred to or add information about possessive relations (genitive attributes or prepositional attributes with e.g. *von*, apposition *Judith die Kluswirtin* DEU047). However, the picture is more complex. If we got complex nominal phrases containing references to persons, it might be interesting to distinguish how this reference is embedded. Consider the following examples:

- I *Die Mappe meines Urgroßvaters* (DEU007)
- II *Der Zauberer von Rom* (DEU024)
- III *Aus dem Leben einer Frau* (DEU012)
- IV *Clemens Brentanos Frühlingskranz* (DEU022)
- V *Lotti, die Uhrmacherin* (DEU058)
- VI *Rachel Gray* (ENG18552)
- VII *Why Paul Ferroll Killed His Wife* (ENG18600)
- VIII *Lady Audley's Secret* (ENG18621)

All examples contain references to persons (definite or indefinite). However, only the noun phrases of example II, V and VI refer to persons. Example VII contains two references to persons whose status or roles are also formulated. The noun phrases of example I, III, IV and VIII refer to objects not persons. A hypothesis might be that the head of the NP represents a kind of thematic /topic focus within the title as many titles contain NPs (which we should classify in general). In some cases it is the person herself or in other case we find objects that are often related to persons (e.g. by possession). Therefore, a syntactic analysis of titles containing references to person could reveal how the author uses these references to set e.g. a focus to the novel. In addition, such an analysis could also consider pronominal references to persons in more details (not only counting e.g. possessive pronouns) in the following way: we could analyse if we can identify anaphoric resolutions either within the title or with the help of the knowledge about the novel (e.g. we know who has a grandpa in I).

This approach might be even more interesting within a multi-lingual perspective. We need to consider linguistic modification strategies across languages and need to map them to semantic frames. We need to ask for example: How is the semantic category *possession* linguistically realised in different languages? Can we assume that we got a kind of either genitive or pronoun that carries this meaning?

Categories might be:

- reference to person: y/n
- count of person: 1,2,3
- for each reference to a person:
 - head of title-NP person: y/n
 - syntactic status of the reference to a person: genitive attribute, head of NP, prepositional attribute, apposition
 - surname: y/n and which
 - last name: y/n and which
 - status: y/n and which
 - profession: y/n and which
 - attribute: y/n and which

4 Focus II Title structure and genre indicator

Related to the first focus is the structure of the title. By structure, I first refer to the linguistic entity which is used to formulate the title: NP, AP, VP, sentence. If two independent phrases are used, we can analyse them as two content elements (e.g title and subtitle or similar in theoretical frameworks).

Focus I can indicate what the title let the reader know about persons within the novel. Focus II looks at if the title is self-referential: explaining what the title covers, e.g. *Fragment eines Romans* (DEU006). Again, similar to focus I , we need to consider what is being said and referred to with the reference to the novel itself. Here, a distinction between actual references such as *Roman/novel* or *Geschichte/history* and indirect references such as *Aus dem Leben einer Frau* need to be motivated and categorised. A reader (and author?) might infer a genre with an actual reference to it.

A quite complex pattern might reflect, how titles refer to their sub-genre. Again, a reader might interpret *Alice's Adventures in Wonderland* (ENG18652) as an adventure novel and *The mysteries of London* (ENG18460) as a mystery novel. Whether both is true, is another question. Again, it is not trivial to categorise such interpretative levels.

We might be also try to interpret title such as *Zum Zeitvertreib* (DEU065) as direct speech to the reader?

Again, a cross language comparison would be more than interesting.