Approaching information structure in historical text corpora: Methodological challenges

Martin Elsig[1] & Esther Rinke[2]

([1]University of Hamburg, Collaborative Research Centre on Multilingualism, [2]Goethe University Frankfurt)

In our contribution, we would like to discuss the methodological challenges we faced in the analysis of information structure in historical texts. With examples from our own research, we illustrate how we dealt with these problems.

The first challenge consists in adopting and operationalizing appropriate notions of topic, comment, focus, and background. A further issue is the potential interdependence of the topic-comment and the focus-background segmentation, especially when it comes to different kinds of foci, e.g. wide vs. narrow focus. Since our main concern is the influence of information structure on word order variation, we considered as most adequate the definition of topic in terms of aboutness (Reinhart 1981) and of focus in terms of that constituent which corresponds to the wh-part of an implicit question (cf. Krifka 2007).

When analyzing authentic literary corpus data, we are confronted with the methodological problem how to identify these functions. Different tests assessing the topic- or focus-hood of individual constituents will be discussed. Even though reliable in many cases, these tests also present a number of critical flaws, as we will show. In addition to this, the analyst has to decide which constituents to consider at all when applying these tests. How to relate information structural notions to adverbs, prepositional phrases, or subordinate clauses?

We will demonstrate that in historical text corpora, the application of these tests always leaves the analyst with a certain residue of sentences in which two or more constituents qualify for one particular information-structural function. In these cases, their relative placement within the sentence is the sole criterion to attribute topic or focus to them. This, however, is problematic in studies which aim at investigating the influence of information structure on word order variation.

We will discuss whether it is possible to compensate for this dilemma by taking into account additional criteria which are known to interact with information structure such as givenness, pronominality, definiteness, and specificity (cf. the cumulative approach by Petrova & Solf 2009). We conclude that in spite of an approximation, a definite and unequivocal solution to these problems still remains to be found.

References

Krifka, Manfred. 2007. "Basic notions of information structure". *The Notions of Information Structure* ed. by Caroline Féry, Gisbert Fanselow and Manfred Krifka (Interdisciplinary Studies on Information Structure, 13-55. Potsdam: Universitätsverlag Potsdam.
Petrova, Svetlana & Michael Solf. 2009. "On the methods of the information-structural analysis of historical texts: a case study on Old High German". *Information Structure and Language Change. New Approaches to Word Order Variation in Germanic* ed. by Roland Hinterhölzl and Svetlana Petrova, 121-160. Berlin, New York: Mouton de Gruyter.

Reinhart, Tanya. 1981. "Pragmatics and linguistics: an analysis of sentence topics". *Philosophica* 27. 53-94.