

Writing Songs into Literary History with Digital Text Mining

Lisanne Vroomen MA

Ruusbroec Institute, University of Antwerp

Lisanne.Vroomen@uantwerpen.be

The manuscript Berlin SBB-PK mgo 185 has been written around 1500 and contains 91 devout vernacular songs, which were used by the Sisters of the Common Life in the Dutch city Zwolle. These sisters were part of a larger religious movement, known as the Modern Devotion. In order to position the songs in the textual culture of the Devout, I will compare the content of Berlin 185 with the content of devout prose originating from the same female religious environment. The devout prose that I will use as a reference corpus consists of biographies of the sisters and informal sermons.

Previous research into the textual culture of the Modern Devout has not paid much attention to spiritual songs, a genre that has flourished mainly in female circles (Joldersma 2001, Joldersma 2008 and Van der Poel 2011). Hence, study of song manuscripts can yield new insights into vernacular spiritual literature and the textual culture of Modern Devout. With my research I want to show the unique position of songs within this textual culture.

For the comparison between songs and prose, I use both close-reading and digital methods of texts mining, such as frequency lists, keyword lists and skip grams (WordSmith, Scott 2012; Stylo, Eder e.a. 2014). Several other researchers have shown how these digital methods can shed more light on a specific text, or a specific texts (Arche 2009, Baker 2004, Jockers 2013, Mahlberg 2007 and Stubbs 2005). By combining these methods with close-reading I will be able to get more insight into the texts at a macro level, without losing focus of the micro level.

My hypothesis is that the songs, although they are used in the same environment as the prose, focus on different subjects and have a different function and angle. For example, the songs use the words *Mary*, *Jesus* and *heaven* much more often than the prose, while they give less attention to the concepts of virtues and sins, which are very important in the prose texts. Furthermore, the frequency of certain syntactical categories, such as modal verbs or pronouns, can also inform us about the texts and the way the focus of the genres differences. *Want* is the most frequent word in the songs, and is also a keyword compared to both the biographies and the sermons. For the sermons, however, *shall* is the most important modal verb. This is an indication that the songs are more expressive ('I want') than the sermons, which are didactic in nature ('you shall'). The high frequency of the pronoun *I* in the songs is an indication of the personal character of the songs, and can be contrasted with the use of *we* in the sermons, which indicates a focus on the community as a whole. This is shown in figure 1. In my presentation I will discuss these examples further and combine them with a close-reading.

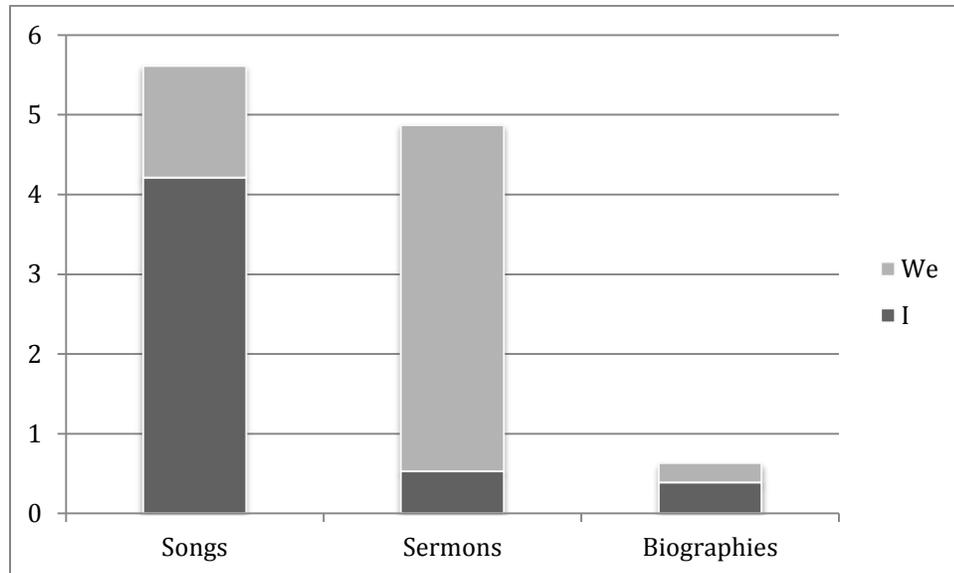


Figure 1: Relative frequency of first person personal pronouns

References

- Arche, D. (2009) *What's in a Word-list? Investigating word frequency and keyword extraction*, Burlington: Ashgate .
- Baker, P (2004) 'Querying Keywords: Questions of Difference, Frequency, and Sense in Keywords Analysis', *Journal of English Linguistics*, vol. 32, pp. 346-359.
- Jockers, M. L. (2013) 'Theme', *Macroanalysis: Digital Methods and Literary History*, Urbana/Chicago etc: University of Illinois Press.
- Joldersma, H. (2001) 'Writing Late-Medieval Women and Song into Literary History', *Tijdschrift voor Nederlandse Taal- en Letterkunde*, vol. 17, pp. 5-26.
- Joldersma, H. (2008) "'Alternative Spiritual Exercises for Weaker Minds'? Vernacular Religious Song in the Lives of Women of the *Devotio Moderna*", *Church History and Religious Culture*, vol. 88, pp. 371-393.
- Mahlberg, M. (2007) 'Clusters, Key Clusters and Local Textual Functions in Dickens', *Corpora*, vol. 2, pp. 1-31.
- Poel, D. van der. (2011) 'Late-Medieval Devout Song: Repertoire, Manuscripts, Function', *Zeitschrift für deutsche Philologie, Dialog mit den Nachbarn. Mittelniederländische Literatur zwischen dem 12. und 15. Jahrhundert*, pp. 67-79.
- Stubbs, M. (2005) 'Conrad in the Computer: Examples of Quantitative Stylistic Methods', *Language and Literature*, vol. 14, pp. 5-24.
- Scott, M., (2012) *WordSmith Tools version 6*, Liverpool: Lexical Analysis Software.
- Eder, M., Rybicki, J. and Kestemont, M. (2014) 'Stylo: a Package for Stylometric Analyses.'