

A dependency-driven workflow for analyzing the agency of men and women in a corpus of historical trial documents

Gideon Kotzé¹, David Zbiral², Robert L.J. Shaw³

^{1, 2, 3}Masaryk University

It has been claimed that in medieval Europe, dissident religious movements might have offered women space for agency less accessible to them in orthodox Christianity. For instance, in records of heresy trials, we find women assuming roles from which they were mostly excluded in the official Church, such as preaching or even performing rituals (Brenon (1998); Benedetti (2004)).

Our growing corpus of inquisition records (ca. 700,000 tokens), mostly Latin-language, ranging from the 1230s to 1520s and from central Italy to England, allows us to study verbal expressions of agency of men and women in individual pieces of trial evidence. This research addresses a gap in the literature, as few studies have attempted to take a data-led, let alone large-scale, approach to the study of the roles of women and men in medieval dissidence (Abels & Harrison (1979); Arnold (1998); McSheffrey (1995)). No study so far has done so using methods of computational text analysis.

We use the Universal Dependencies framework (de Marneffe et al. (2021)) to identify dependency structures. For any root verb, we identify its direct dependent actants (subject, direct/indirect objects, oblique nominals (*nsubj/obj/iobj/obl*)), and perform a combined matching strategy in order to assign gender to all human actants. Depending on the information available in the full text, we match named entities with pre-existing lists of gender-tagged specific person entities, Latin first names, pronouns (including their inflected forms), and nouns describing persons. We also perform rule-based pronoun reference resolution (including unexpressed subject). This allows us to obtain all human actants and, with some degree of success, their gender. We can thus study the frequencies of different actant gender patterns for each individual word, partly even aggregating them into broader categories thanks to hypernymy in a custom-built lexico-semantic network of concepts and actions we are developing. For example, in the trials of the Guglielmites, we note the use of multiple verbs describing women in communicative actions such as criticizing, retorting, exclaiming, and replying.

The initial coverage of actants, limited to direct dependents of the verb root in four different slots, was low. By matching all strings associated with actants in a sentence, however, we are able to identify further syntactic variations to be included, as well as others that should be ignored (such as those including negative modifiers). Those that are not *nsubj/obj/iobj/obl* and are not direct dependents of the root are analyzed further. Using this method, we were able to discover which clauses to match, how compounds are handled, and to what extent pronouns and pronominal phrases are unexpressed, while using the process to improve our heuristic pronoun reference resolution approach. Table 1 produces a few examples of such discovered rules.

We can thus provide an empirical, bottom-up image of male and female agency in heresy trial records. These results can also be filtered by record genre (such as "deposition" or "sentence"). For those records which appear to allow deponents more opportunity for expression, we can also go further than simply looking at the depiction of male and female involvement in actions related to heresy, and also compare the way in which deponents themselves depicted such actions when testifying at trial. Taken together, all this allows us to provide an unprecedented computational analyses of male and female agency in medieval Latin-language inquisition trial records., It thus contributes not only to a better understanding of heresy and its repression, but also offers a rare non-literary window on the broader depiction of gender roles in medieval society and the interconnections between language and gender in medieval Europe.

Table 1: In the first example, the sentence has the root “*dixit*” (“and at the end the said brother Daniel said to them:”), where “*Daniel*” is identified as a nominal modifier with “*frater*” (“*brother*”) as its head. On the right, dependants of the root and their tags show that “*Daniel*” is governed by the root and will therefore not be discovered using our initial approach. In this case, we need to check for nominal subjects (*nsubj*) as well as any dependent *nmods*. In the next example, “*eam*” (“*her*”; “and that she had not told the truth in her other depositions, because they had instructed her that she should not tell the truth about what had been said.”) is the direct object of “*instruxerant*”. Although we check for *obj*, this is not a direct dependant of the root verb. We need to recognize that verbs can be dependants of other verbs as adverbial clause modifiers, and that the former can have direct objects of their own, which in this case is relevant. In the final row, “*sibi*” is gender-ambiguous, but matching the *nsubj* of the previous sentence, and using the list of names to find out that “*Sybilis*” is female, helped us assign gender to the *iobj* “*sibi*”.

| Root | Match_sex_tag_head | Sentence | Dep_tags_of_root |
|---------|------------------------|---|---|
| dixit | Daniel_m_nmod_frater | et in fine dixit dictus frater Daniel eis : | (et, 'et', 'cc'), (fine, 'finis', 'obl'), (frater, 'frater', 'nsubj'), (eis, 'is', 'obl'), (:, ':', 'punct') |
| dixisse | eam_f_obj_instruxerant | et non dixisse veritatem in aliis dictis suis , quia instruxerant eam quod non diceret veritatem de predictis . | (et, 'et', 'cc'), (non, 'non', 'advmod:neg'), (veritatem, 'ueritas', 'obj'), (instruxerant, 'instruxerant', 'advcl'), (., '.', 'punct') |
| dixit | Sybilis_f_nsubj_dixit | quod dictus Andreas Saramita dixit sibi | ('dixit', 'quod', 'obj'), ('dixit', 'Andreas', 'nsubj'), ('dixit', 'sibi', 'iobj') |

References

- Abels, Richard & Ellen Harrison. 1979. The Participation of Women in Languedocian Catharism. *Mediaeval Studies* 41(1). 215–251. doi:10.1484/j.ms.2.306245.
- Arnold, John H. 1998. The Preaching of the Cathars. In Carolyn Muessig (ed.), *Medieval monastic preaching*, 183–205. Leiden – Boston – Köln: Brill.
- Benedetti, Marina. 2004. La predicazione delle donne valdesi. In Dinora Corsi (ed.), *Donne cristiane e sacerdozio: Dalle origini all'età contemporanea*, 135–158. Roma: Viella.
- Brenon, Anne. 1998. The voice of the good women: An essay on the pastoral and sacerdotal role of women in the cathar church. In Beverly Mayne Kienzle & Pamela J. Walker (eds.), *Women Preachers and Prophets through Two Millenia of Christianity*, 114–133. Berkeley – Los Angeles – London: University of California Press.
- de Marneffe, Marie-Catherine, Christopher D. Manning, Joakim Nivre & Daniel Zeman. 2021. Universal Dependencies. *Computational Linguistics* 47(2). 255–308. doi: 10.1162/coli.a.00402.
- McSheffrey, Shannon. 1995. *Gender and Heresy: Women and Men in Lollard Communities, 1420-1530*. University of Pennsylvania Press. <http://www.jstor.org/stable/j.ctt3fj0fs>.