



## Identifying implicit stereotypical views in natural language through automatic linguistic analyses.

### Applicants

Eligible proposals must have two (and only two) applicants from different disciplines within the Network Institute.

Supervisor Name	Department/Group	Faculty
1. Dr. Camiel Beukeboom	Communication Science	FSW
2. Dr. Antske Fokkens	Computational Lexicology & Terminology Lab (CLTL) (and Web & Media group)	FGW (and FEW)

### Project description

Provide a brief description of the project (max. 300 words) [288 words]

Stereotypical views about social groups (e.g., Muslims, Germans, women, immigrants) play a pervasive role in how we perceive and interact with people. Stereotypes emerge from the way we communicate about categorized people and their behavior both privately and in the media. For instance, the stereotypical associations we bring to mind with people categorized as “Muslims” have changed after terrorist attacks have frequently been linked to Islam (e.g., with terms like ‘Muslim terrorism’). Such generalized negative associations can promote unjustified prejudice and discrimination against individual Muslims. It is therefore valuable to learn about the exact (linguistic) means through which such (negative) stereotypical views become shared within communities. Both applicants investigate these linguistic means in their current work:

Beukeboom [1, 2] proposed a theoretical model that distinguishes three types of specified linguistic biases that subtly reflect stereotypical expectancies. Fokkens [3, 4] provided a formal perspectives model that allows researchers to compare alternative points of view expressed in text by means of linguistic annotations. These models share the vision that the stereotypical perspective displayed in text is a combination of *what* we talk about (content and selection of information) and *how* we talk about it (choices in formulation).

The present project aims to merge these complementary approaches, and thereby develop a methodology for using NLP<sup>1</sup> to automatically identify content and strength of stereotypes about specific groups shared within communities.

---

<sup>1</sup> Natural Language Processing

The academy assistants will assist in taking a first significant step in this direction by specifying how stereotype-biased linguistic cues can be identified in open text, and how this can be achieved automatically using NLP tools. This method can be used to investigate the presence and strength of (implicit) stereotypes in texts, like news media messages and (social media) conversations.

## Project Organization

*Each proposal requests two Academy Assistants from different disciplines. Describe their roles and describe the skills and expertise required from them. (max. 300 words). 290 wdn + 10 in footnote*

Beukeboom's model assumes that stereotypical views are shared in subtle ways in communication, i.e., the information we provide when talking about categorized people and their behavior, and specified ways to express this information (e.g., labels, language abstraction, negations etc.). Following this assumption, the project team will address the following research questions:

- Which stereotype-reflecting linguistic cues can be identified in natural language, and how can these be automatically detected using NLP tools?

In order to address these questions, we need two types of assistants:

**A Communication Science or Social Psychology (CS/SP) student** with expertise on stereotypes in general and the role of media messages and language use in spreading and/or maintaining stereotypical views.

**A (Computational) Linguistics (CL) student** with expertise on (lexical) semantics or sentiment analysis and, preferably, experience with NLP tools for Dutch.

In close collaboration with the applicants, they will take Beukeboom's model (so far mainly based on artificially created linguistic stimuli) and determine how stereotype-reflecting linguistic cues are displayed in natural language and, in particular, which cues can be identified with current NLP tools. An initial method to identify these linguistic cues will be tested using existing constructed examples from experiments on biased language use. Next, the methodology is employed to study the presence of biases -and their linguistic manifestation- in Dutch newspaper articles (e.g., representations of Muslims in texts referring to terrorism; obtained from LexisNexis).<sup>2</sup> The students will focus on manual annotations, where they first design annotation guidelines while jointly working on a development set. In the final stage, they will separately annotate a new data set following their guidelines. This new set will be used to determine Inter-Annotator Agreement, which provides an indication of the reliability and reproducibility of annotation results.

## Collaboration

*Describe how your research improves collaboration and cross-pollination between the disciplines involved (max. 300 words) (wdn 297)*

Stereotypical beliefs have been shown to be shared through largely implicit linguistic biases that reflect and confirm (negative) stereotypes about people, yet this research mainly relies on somewhat artificial experimental studies. This leaves the question whether

---

<sup>2</sup> Use of LexisNexis data has been arranged with Pim Stouten.

these biases function similarly in natural language. As the present interdisciplinary project is a first step towards processing large quantities of natural language, it bears the potential of providing a major leap forward in research on biased language use. The possibility of employing automatic processing tools to detect implicit stereotyping in natural language will not only provide unique knowledge to verify the validity of theoretical models of biased language use, it also opens up major opportunities for research in a variety of applied contexts.

Investigating biased language use in natural language through NLP is only possible by combining expertise from involved domains: Communication scientists provide expertise about the hypothesized means, antecedents and consequences of various types of biased language use in context. (Computational) Linguists provide expertise on how language expresses such biases (e.g., role of word meanings, context, world-knowledge), and expertise on using NLP tools.

The collaboration opens many new research possibilities for both disciplines. This research is an initial step in what is intended to become a longterm collaboration.

- Gains for communication science: access to linguistic expertise directly relevant for biased language use, and connection to tools that enable studying natural language on scales that are impossible using manual annotations alone.
- Gains for computational linguistics: insight into what part of linguistic annotations are relevant for other researchers, new possibilities of evaluating their analyses and tools.
- Gains for both: insights into a 'digital social science': i.e., methodological insights by combining different levels of analyses, and insights into the potential of using automatic linguistic annotations to study social science hypotheses.

## **Deliverables**

*Enumerate intended project results: papers, research proposals or otherwise. (max 200 words) (wdn 174)*

While working on the question which linguistic means reflect stereotypical views in natural texts, the applicants and student assistants will develop a methodology to (manually) label these linguistic cues in text. This methodology will be specified in annotation guidelines (Dutch: *code boek*) that allows for reproducible labeling efforts of data. These guidelines will both reflect and inform the theoretical model on biased language use, as well as specify the annotations to be provided by NLP tools.

During the project we will work towards a publication in a scientific journal. Depending on the outcome and insights this might focus on the development of the methodology, or on a study using the developed methodology. The student assistants will be (at least) co-author of this publication.

We also plan to work towards new project proposals. In case of interest, the students can use the work as a basis for a PhD proposal. Beukeboom and Fokkens will also look at other possibilities for proposals on joint work, such as the NLeSc center proposals that were offered this year.<sup>3</sup>

---

<sup>3</sup> <https://www.esciencecenter.nl/project-calls>

## Planning

*Provide a breakdown of the project into phases with tentative timing (max 150 words) (wdn 147)*

The team will meet bi-weekly. The planning is based on previous experience with academic assistants. Work on research proposals starts in month 5.

**Month 1-2:** Background research in individual fields and interactive meetings to get to know each other's fields. Immediate focus on textual examples to develop methodology (i.e., annotation guidelines).

**Month 3-4:** Investigating linguistic cues and output of NLP tools on constructed examples (mainly the CL student). The CS/SP student starts to explore biases in newspaper text.

**Month 5-6:** First version of annotation guidelines based on constructed examples (CL student works on linguistic annotation, CS/SP student on guidelines for an overall presence of a bias).

**Month 7:** Joint work on annotation guidelines using development data

**Month 8:** Textual annotations on new data and inter-annotator agreement.

**Month 9-10:** work out and write up the results for a publication, set up PhD proposal (if applicable).

***Please respect the word count limits: proposals that exceed the stated limits will not be eligible.***

## References

[1] Beukeboom, C. J. (2014). Mechanisms of linguistic bias: How words reflect and maintain stereotypic expectancies (Chapt.). In J. Laszlo, J. Forgas, & O. Vincze (Eds.), *Social Cognition and Communication* (pp. 313-330). New York, NY: Psychology Press.

[2] Beukeboom, C.J. & Burgers, C. How stereotypes become shared knowledge: Biased language use in communication about categorized individuals (manuscript in preparation).

[3] Fokkens, A.S., Rospocher, M., Vossen, P.T.J.M., et al. A Generic framework for representing Perspectives. (manuscript in preparation).

[4] Zuhaitz Beloki, German Rigau, Aitor Soroa, Antske Fokkens, Kees Verstoep, Piek Vossen, Marco Rospocher, Francesco Corcoglioniti, Roldano Cattoni, Stefan Verhoeven, Mathijs Kattenberg. System Design. NewsReader Deliverable 2.2. (Section 6 provides an outlook to Fokkens et al. in prep)