

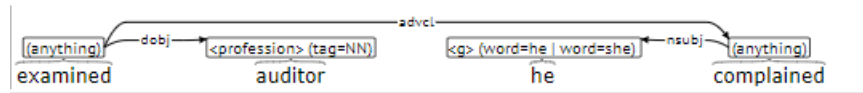
Appendix a – syntactic queries

This document contains the 14 queries we searched SPIKE as mentioned in section 2.1.

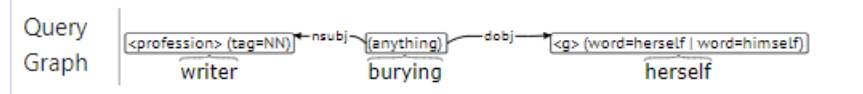
1. the profession:[t]**designer** is known for g:[w=her|his]**her** bright colors and innovative techniques



2. the physician examined the profession:[t]**auditor** for sores because g:[w=he|she]**he** complained



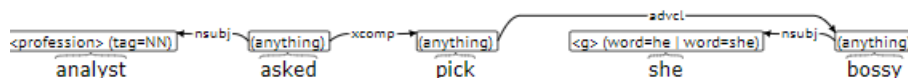
3. the profession:[t]**writer** was burying g:[w=herself|himself]**herself** in writing a new novel



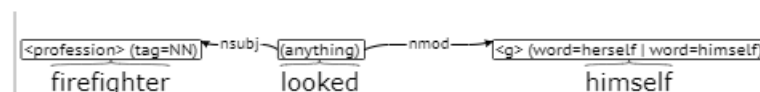
4. The profession:[t]**manager** told the hairdresser that g:[w=he|she]**he** did a good job



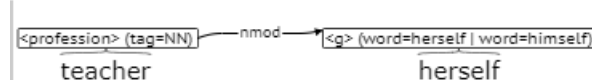
5. The profession:[t]**analyst** asked the cleaner to pick up papers because g:[w=he|she]**she** was bossy



6. the profession:[t]**firefighter** looked at g:[w=herself|himself]**himself** in the mirror



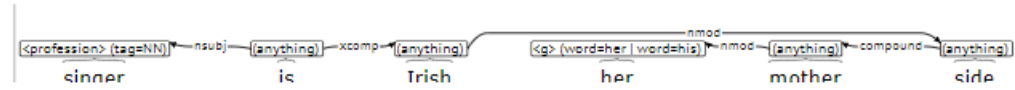
7. the profession:[t]**teacher** g:[w=herself|himself]**herself** wasn't damaged



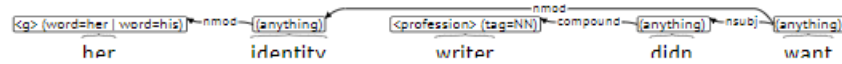
8. The profession:[t]**janitor** cleaned the room and er:[w=he|she]**she** asked for a tip



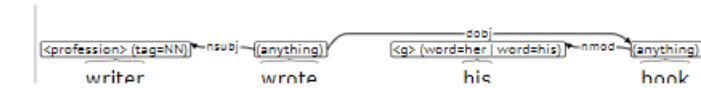
9. The profession:[t]**singer** is Irish from by g:[w=her|his]**her** mother side



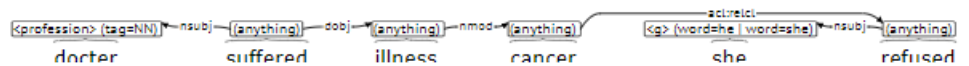
10. Regarding g:[w=her|his]**her** gender identity, the profession:[t]**writer** didn't want to expand in words



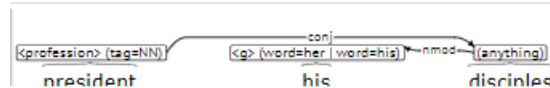
11. the profession:[t]**writer** wrote g:[w=her|his]**his** book when he became sick



12. the profession:[t]**doctor** had suffered a short illness due to pancreatic cancer, for which g:[w=he|she]**she** had refused chemotherapy



13. by the profession:[t]**president** and g:[w=her|his]**his** disciples



14. did the profession:[t]**teacher** want g:[w=her|his]**his** students to talk?



Appendix b - guidelines for human annotation

This document contains the guidelines the human annotators received as mentioned in section 2.3.

Guideline for Sentences' Annotation - Gender Bias Project

First, thank you for helping us understand the data relating to our project.

Please visit this [link](#), where you'll find 3 relevant columns: 'sentence_text', "predicted gender" and "correct?". Please disregard all other columns.

In each sentence, you'll find a marked entity between <ent></ent> tags, for example:
<ent>writer</ent>.

Your job is to decide whether the predicted gender matches your understanding of the entity's gender in the sentence.

If the gender is correct - you need to mark "1" in the "correct?" column, otherwise, mark it as "0".

Examples:

Correct:

1.sentence_text:

In 1774 , the village hired its own <ent>teacher</ent> for the first time , although he only taught in the winter.

predicted gender:

Male

Note that “he” in the sentence refers to the teacher, so we know it’s a male teacher, which matches our predicted gender. This should be marked as “1”.

2.sentence_text:

He managed the band Audio Bullys and also the pop <ent>singer</ent> Lily Allen before she got a record deal .

predicted gender:

Female

Note that “she” in the sentence refers to the singer, so we know it’s a female singer, which matches our predicted gender. This should be marked as “1”.

Incorrect:

1. Sentence_text;

I annoyed the other <ent>editor</ent> because she considered it was his territory .

Predicted gender:

Male

This should be marked as “0”. Note that the editor in the sentence is referred to as “she”. So the gender here should be female.

2. sentence_text:

He needs a <ent>lawyer</ent> for her so he looks up online.

predicted gender:

Male

Note that we don't know the gender of the lawyer in this case. This should be marked as "0".

Ambiguous:

1. sentence_text:

John missed the <ent>Doctor</ent> as he dropped on the ground during the shooting.

predicted_gender:

male

A few of the sentences may read in different ways (Who dropped to the ground? John or the doctor?). Please use your best judgement and mark them as you think is the most likely reading. Here I think it should be marked as "1", as I think that the doctor is the one who dropped to the ground.

Contact

For any further questions feel free to contact me.