Toward the Automatic Identification of Neopronouns

Lain Nelson, Claire Cardie

Department of Computer Science, Cornell University, Ithaca, New York, 14850

Introduction:

The field of natural language processing (NLP), explores the intersection of human languages and artificial intelligence (AI). While incredible progress has been made in this field, such as the advent of ChatGPT, human biases are often replicated and even exacerbated in language models. Language models are the result of utilizing machine learning with instances of human language such as newspaper articles. Gender bias is also present in these language models, especially as it relates to individuals who use neopronouns, or "new pronouns" (e.g. ey/em). These pronouns often fail to even be identified as such.

What are neopronouns?

Neopronouns, or "new pronouns" are typically third person, singular pronouns. Neopronouns are most commonly utilized by non-binary or transgender individuals. They may be subject, object, dependent possessive (determiner), independent possessive, or reflexive. Neopronouns typically come as a set of these five types, and neopronoun users may list the entire set (e.g. xe/xem/xyr/xyrs/xemself). Often, users will list only a partial set (e.g. xe/xyr), expecting people to infer and utilize the rest based on grammatical norms. People will sometimes utilize pronouns from different pronoun sets. For instance, someone may list their pronouns as they/xe/sun and use they/them/their/theirs/themself, xe/xem/xyr/xyrs/xemself, and sun/sun/suns/suns/sunself pronouns.

What are some neopronouns?

xe/xem

ze/hir/hirs/hirself

Nounself Pronouns:

sun/suns/sunself

moon/moon's/moonself

Emojiself Pronouns:

🍓/🍓s / 🍓self

💅/💅s / 💅self

Nameself Pronouns:

Bob/Bobs/Bobself

Numberself Pronouns:

0/0s/0self

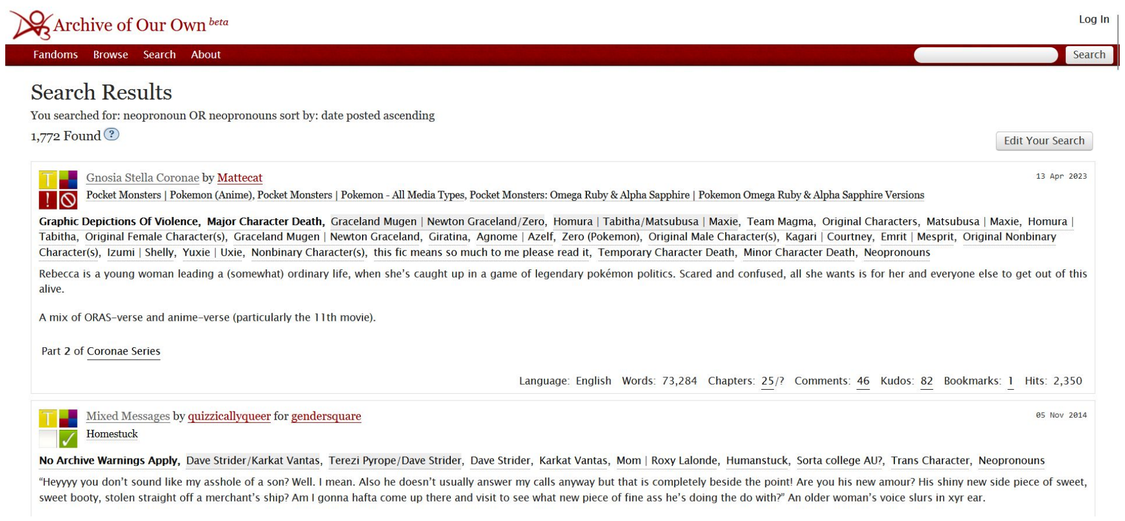
Borrowed Pronouns:

elle (from Spanish)

iel (from French)

Methods:

I realized that because neopronouns have largely been ignored by NLP, natural language models which have been used to determine the part-of-speech (e.g. verb or adjective) of words may not even be able to identify them as pronouns. Looking at part-of-speech tagging is also an effective way to analyze the treatment of neopronouns, whereas with other, more complex, natural language inference tasks it is more difficult to determine how the presence of neopronouns is affecting performance on the task. I found that textual examples of neopronouns were found most easily on Archive of Our Own (AO3), a fanfiction website. Using the boolean search "neopronoun OR neopronouns", I found 1,772 works.



I wrote code to scrape the text of works from these works and put it into Google Drive files, skipping fanfictions with adult content. This scraping is still in progress; I currently have the works from eight of the 89 search pages. In the meantime, I began to run part-of-speech taggers supplied by SpaCy on this text.

Results:

My initial explorations revealed that neopronouns are often misidentified, commonly as either nouns or proper nouns.

```
[('Sol', 'PROPN'), ('stood', 'VERB'), ('suddenly', 'ADV')]
```

```
('sol', 'NOUN'), ('said', 'VERB'), ('sheepishly', 'ADV'), (',', 'PUNCT'),
('shrugging', 'VERB'), ('at', 'ADP'), ('suns', 'NOUN'),
```

These first two examples come from a work utilizing sol/sun/suns/sunself pronouns. These pronouns are incorrectly labeled as a proper noun or noun when they appear.

```
[('Ey', 'PRON'), ('goes', 'VERB'), ('back', 'ADV'), ('to', 'ADP'), ('eir', 'DET'), ('tea', 'NOUN'), ('.', 'PUNCT')]
```

The above comes from a fanfiction using ey/eir pronouns. "Ey" is identified correctly, but "eir" is wrongly labelled as a determiner.

I have yet to gather and perform analysis on the entire corpus of fanfictions, but I plan to produce more concrete statistics on the treatment of neopronouns and explore how part-of-speech tagging of neopronouns can be improved.

References:

Aloof_Introvert. (2016, March 22). Neverending. Archive of Our Own. https://archiveofourown.org/works/6319618?view_full_work=true

Storytelling_Sparrow. (2019, March 30). After School Special. Archive of Our Own. https://archiveofourown.org/works/18290111

Tinkler, E. (2023). Neopronoun Buttons. Medium. The History of Neopronouns. Retrieved September 15, 2023, from https://medium.com/matthews-place/the-history-of-neopronouns-366b1fee48c4.

Anne Lauscher, Archie Crowley, and Dirk Hovy. 2022. Welcome to the Modern World of Pronouns: Identity-Inclusive Natural Language Processing beyond Gender. In Proceedings of the 29th International Conference on Computational Linguistics, pages 1221–1232, Gyeongju, Republic of Korea. International Committee on Computational Linguistics.