Plot trends over time for narrator analysis

Adrian Barnett

15 September 2022

This document contains trend plots for our analysis of language over time in papers published on *PubMed*. In total there were 20,876,733 abstracts.

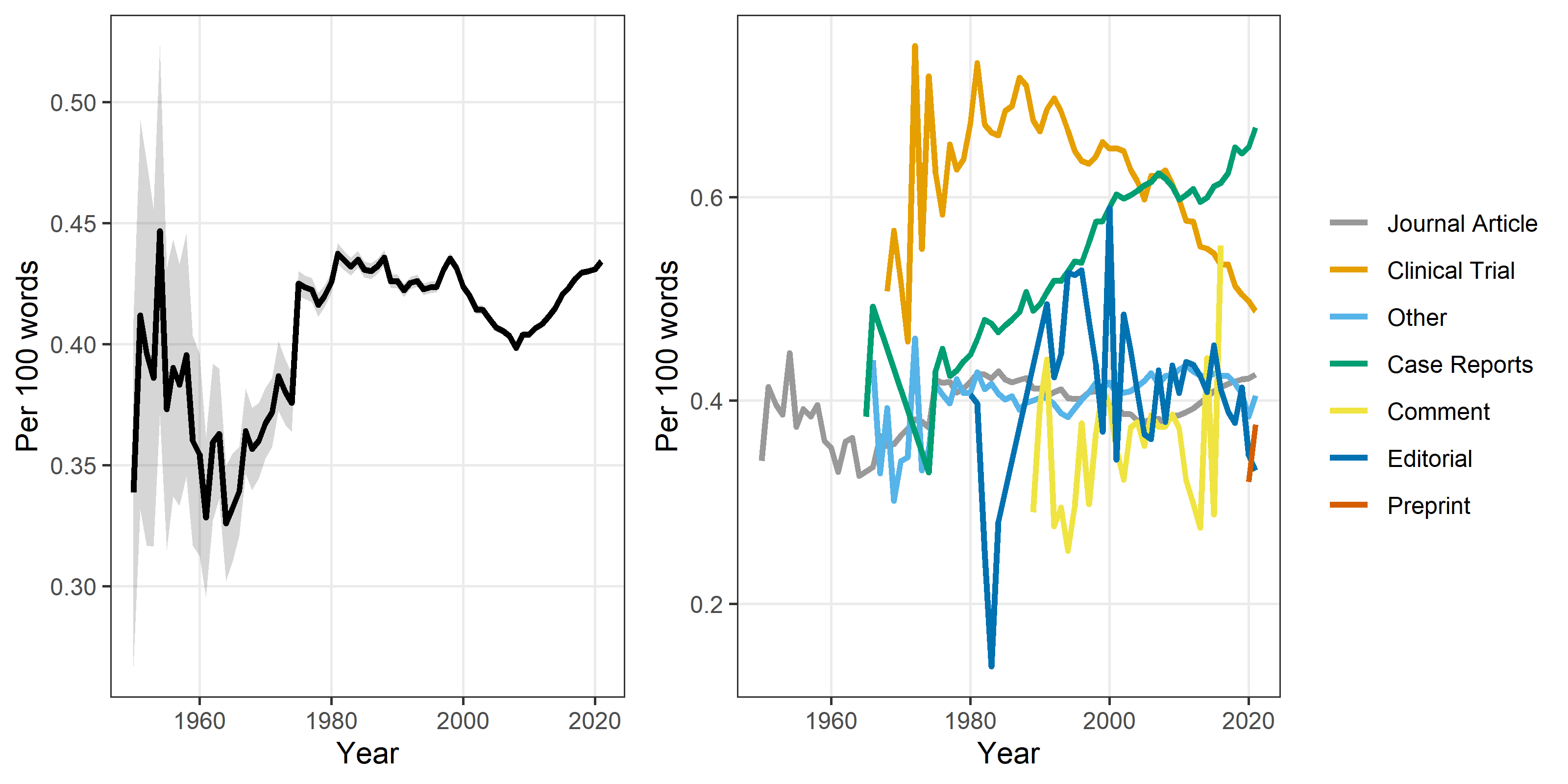
The plots show the mean per year and the 95% confidence interval for the mean (shaded area).

There are two plots per outcome:

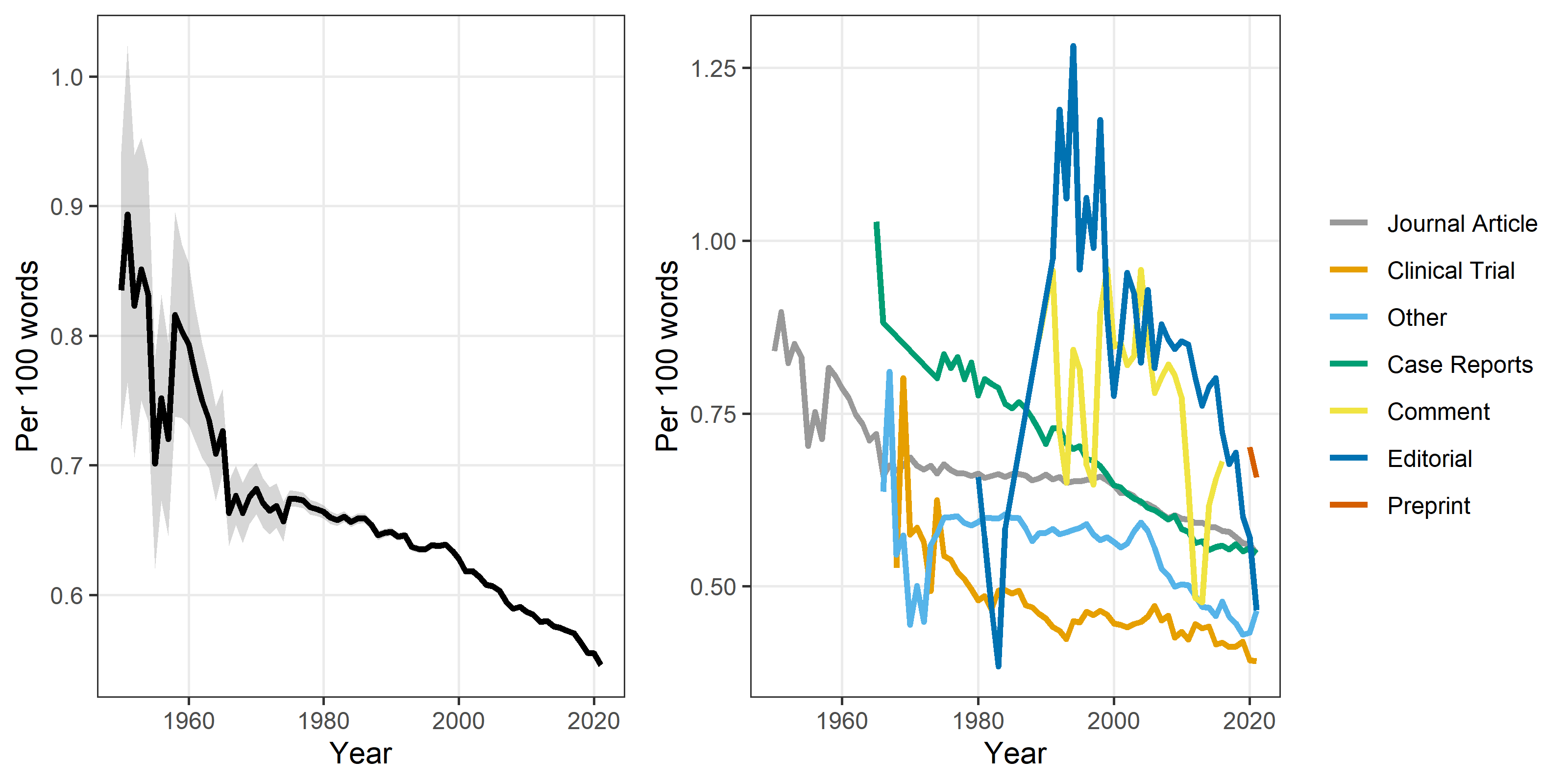
* All articles types combined
* Split by article type

For the results by article type, we excluded years with fewer than 50 abstracts. We did not add confidence intervals to this plot as this made the plot too crowded.

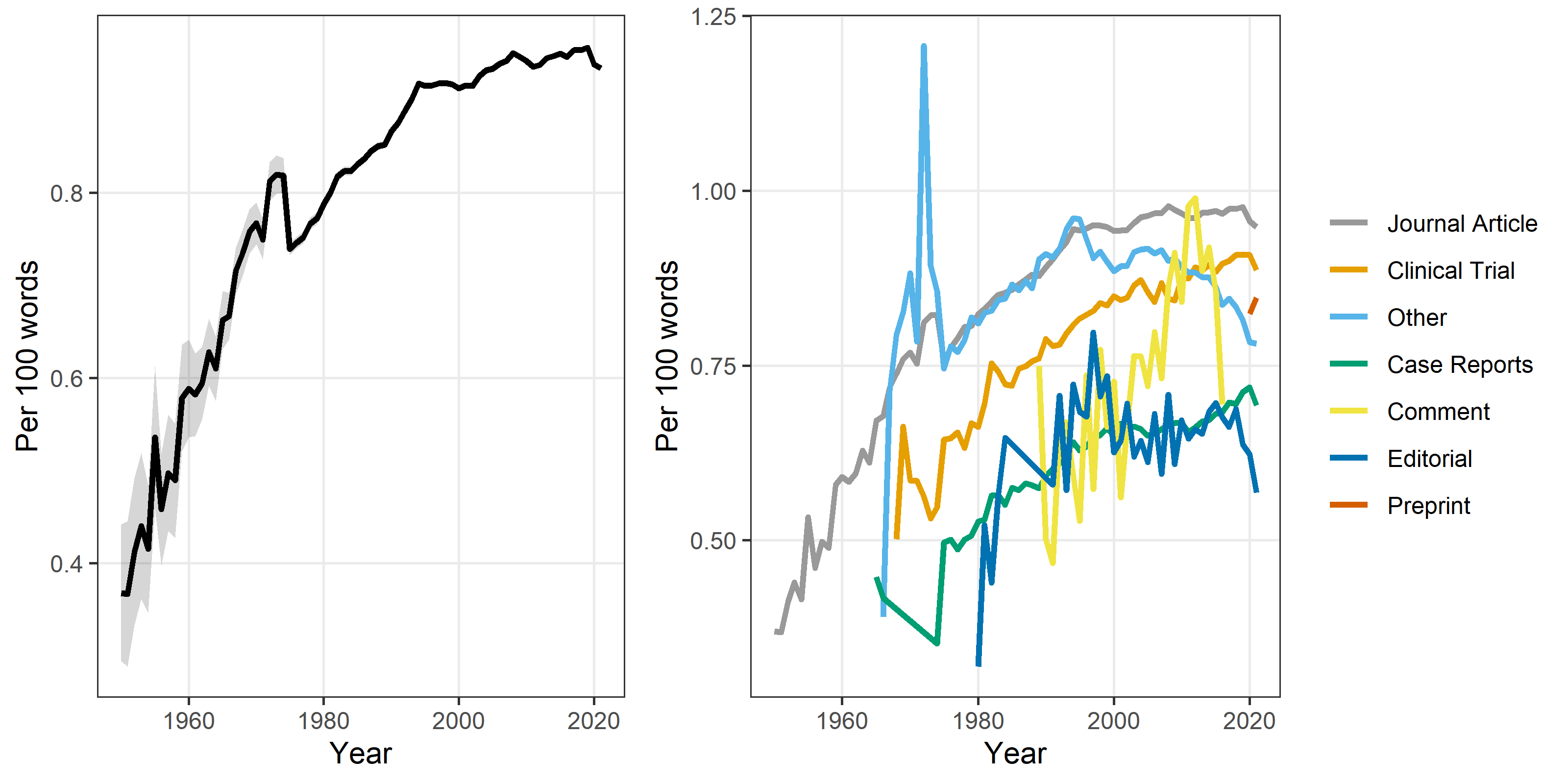
# Signposting



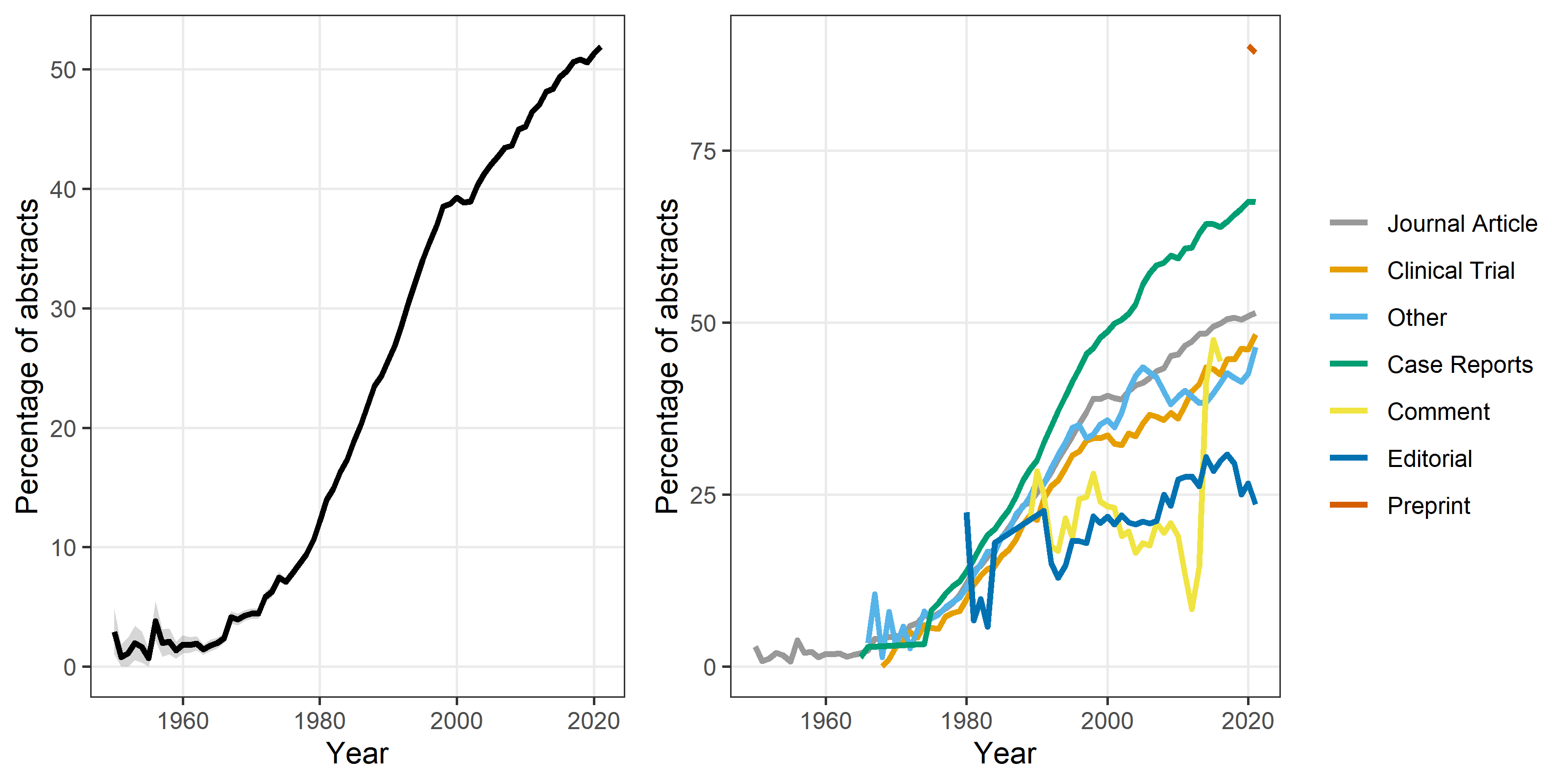
# Hedging



# Noun chunks

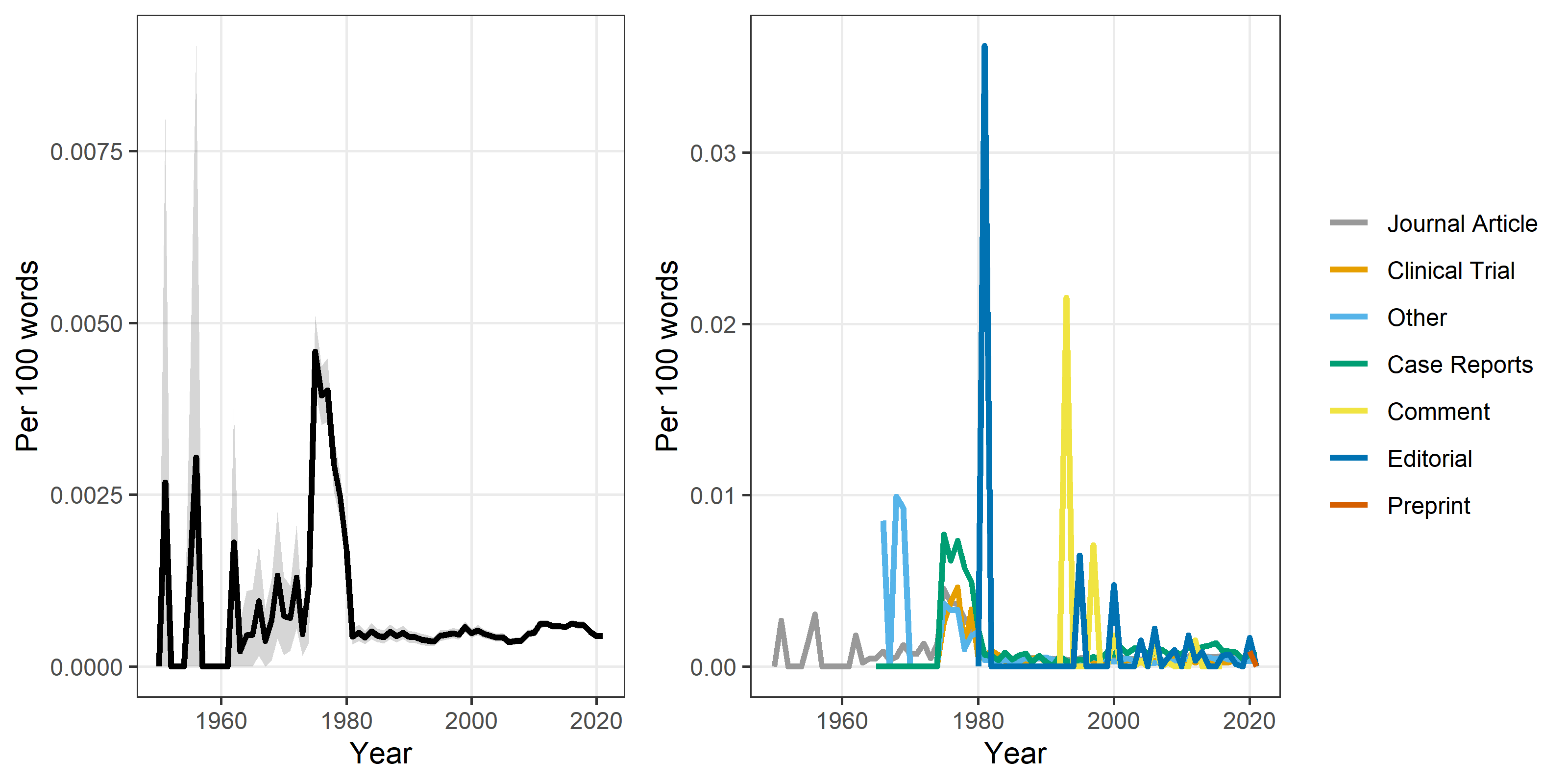


# Any narrator

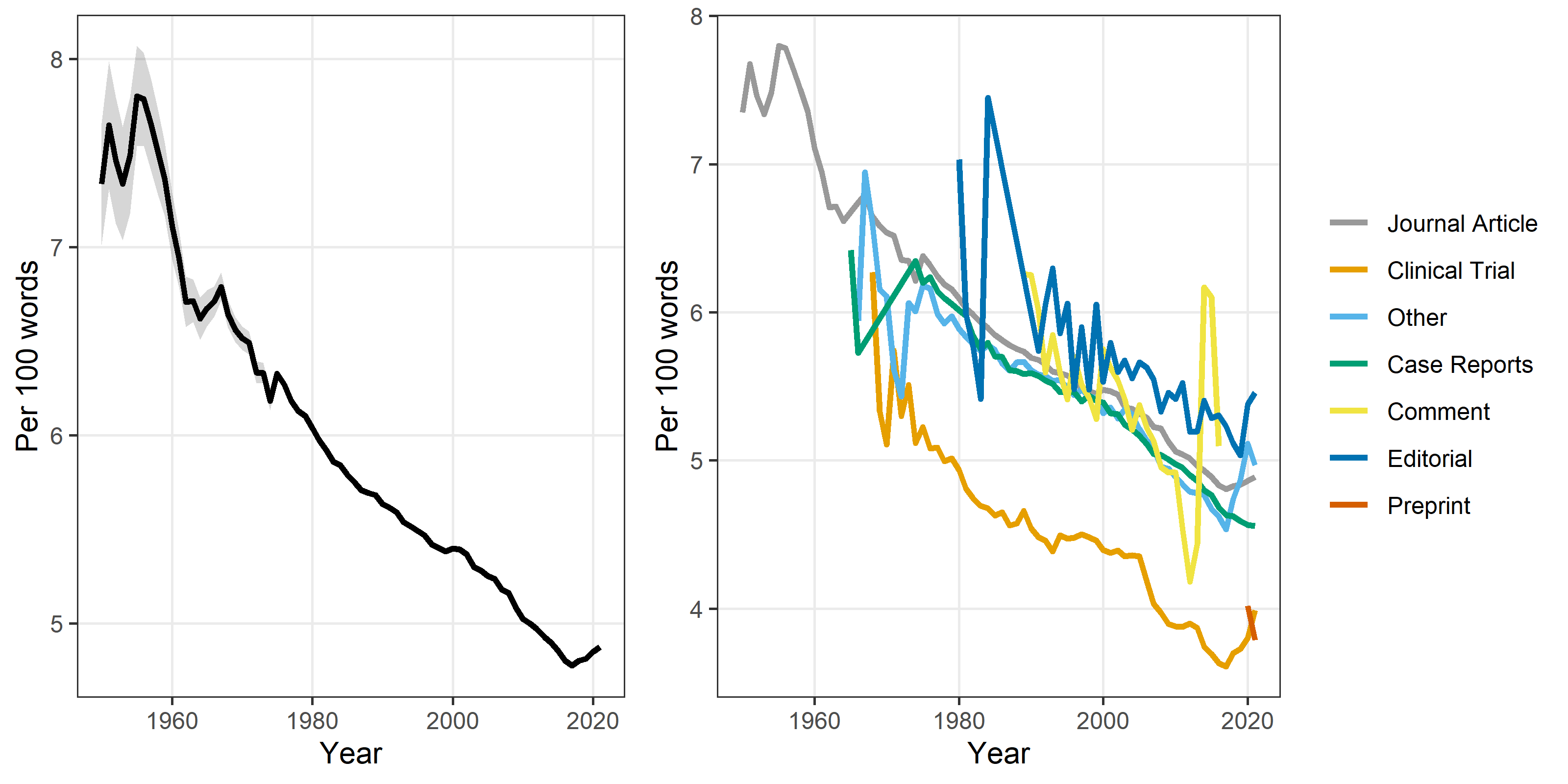


This is a binary yes or no outcome, whereas the other outcomes are counts.

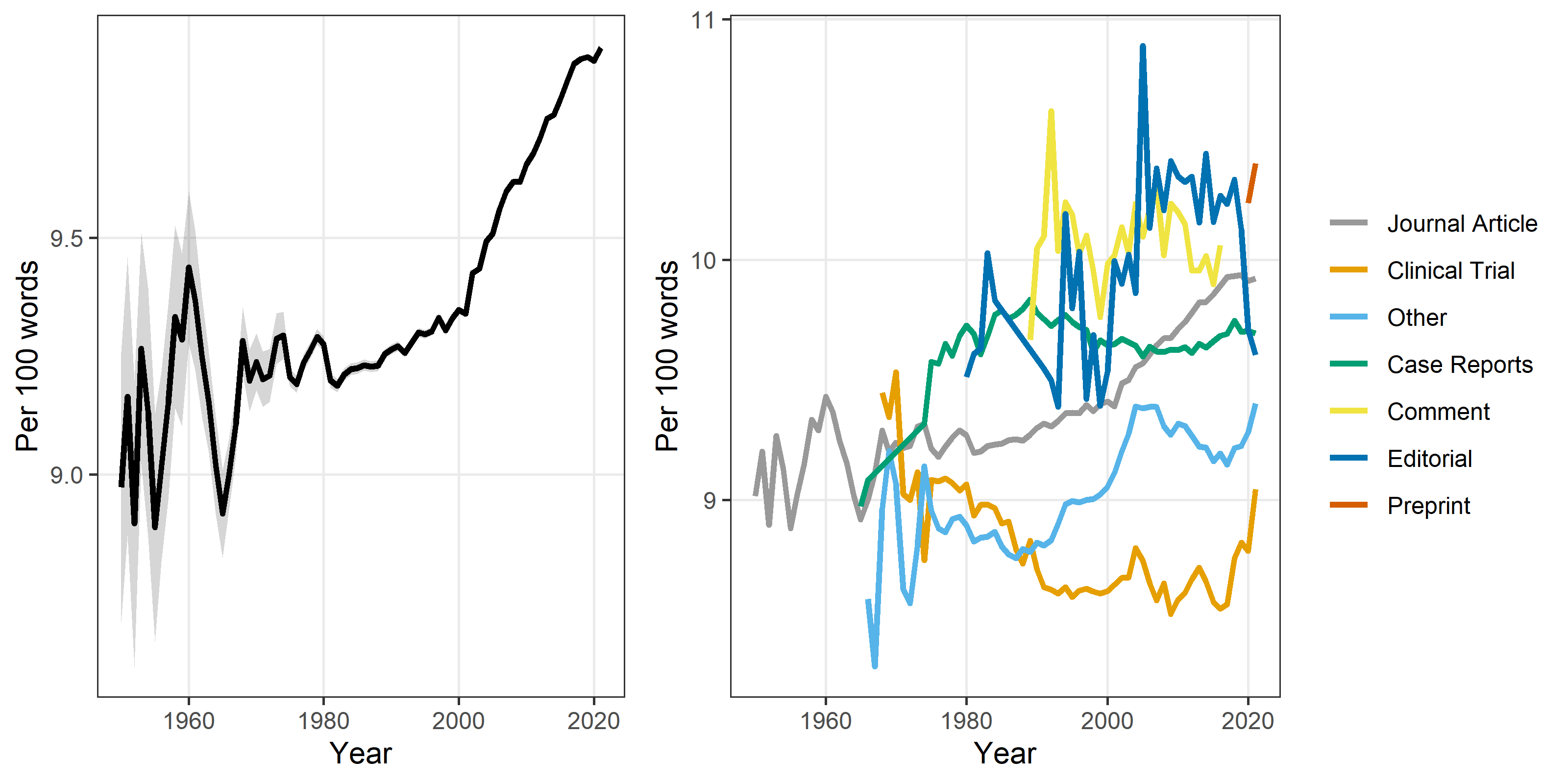
# Misspellings



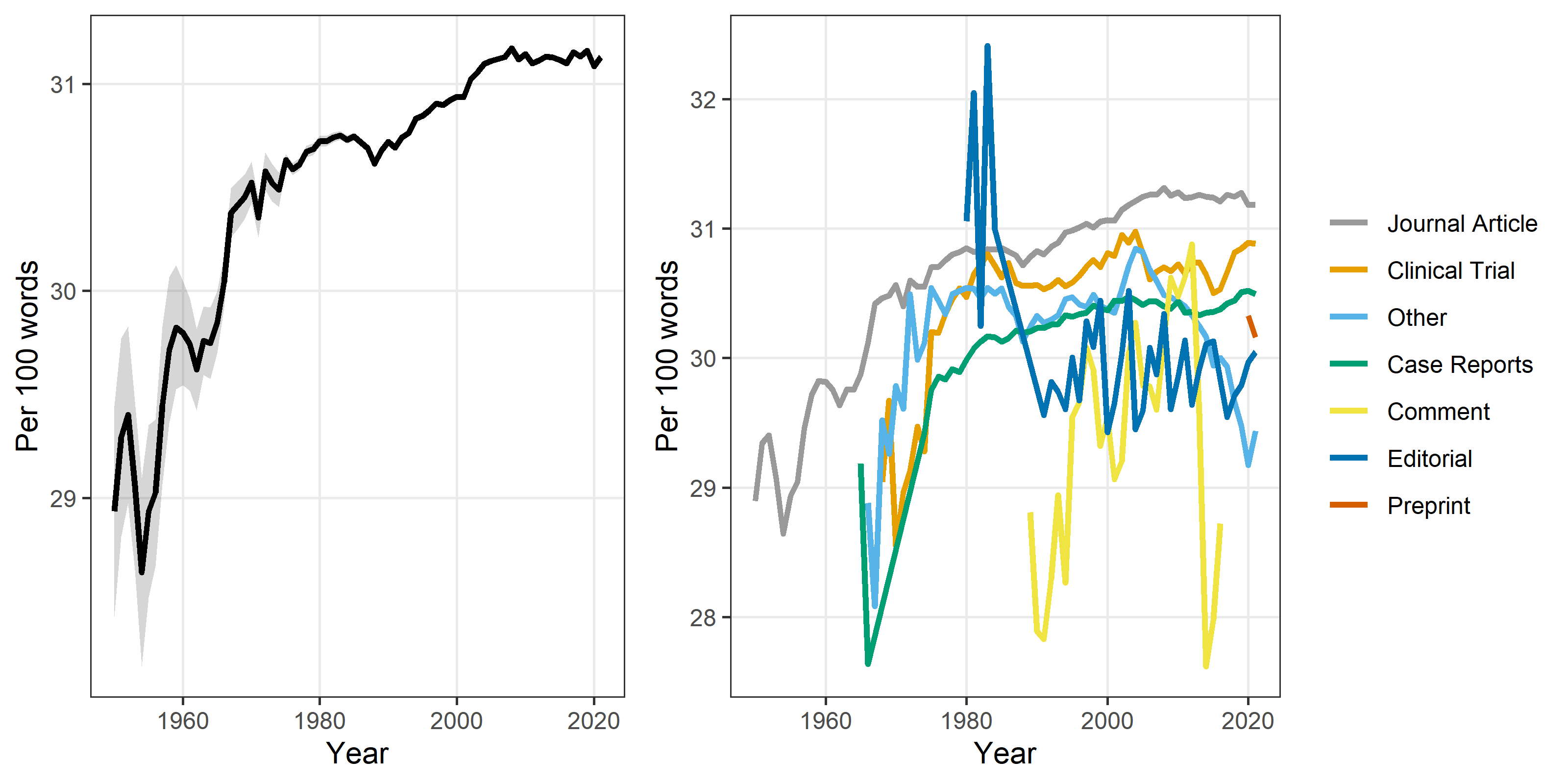
# Negative control



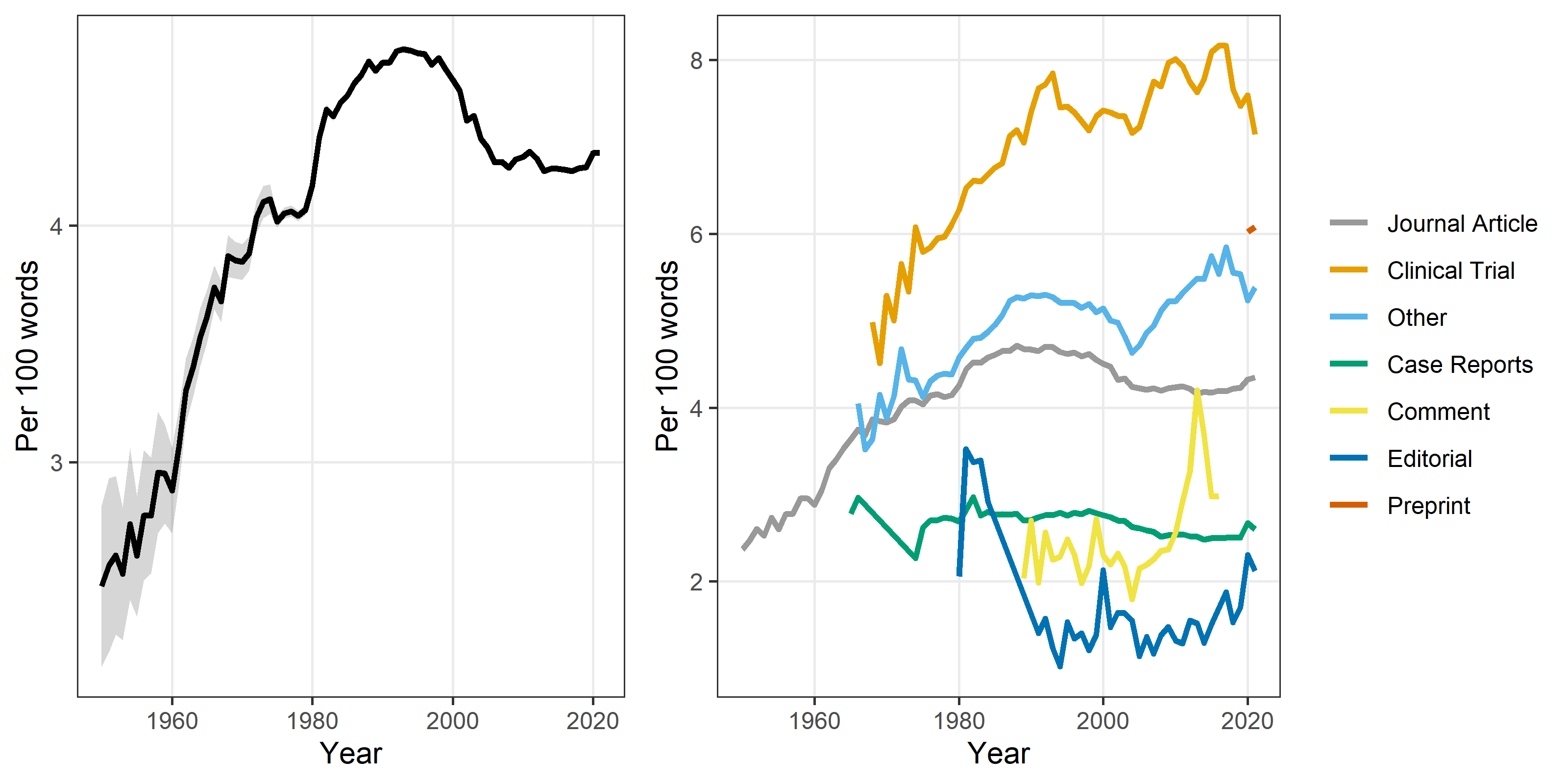
# Verbs



# Nouns

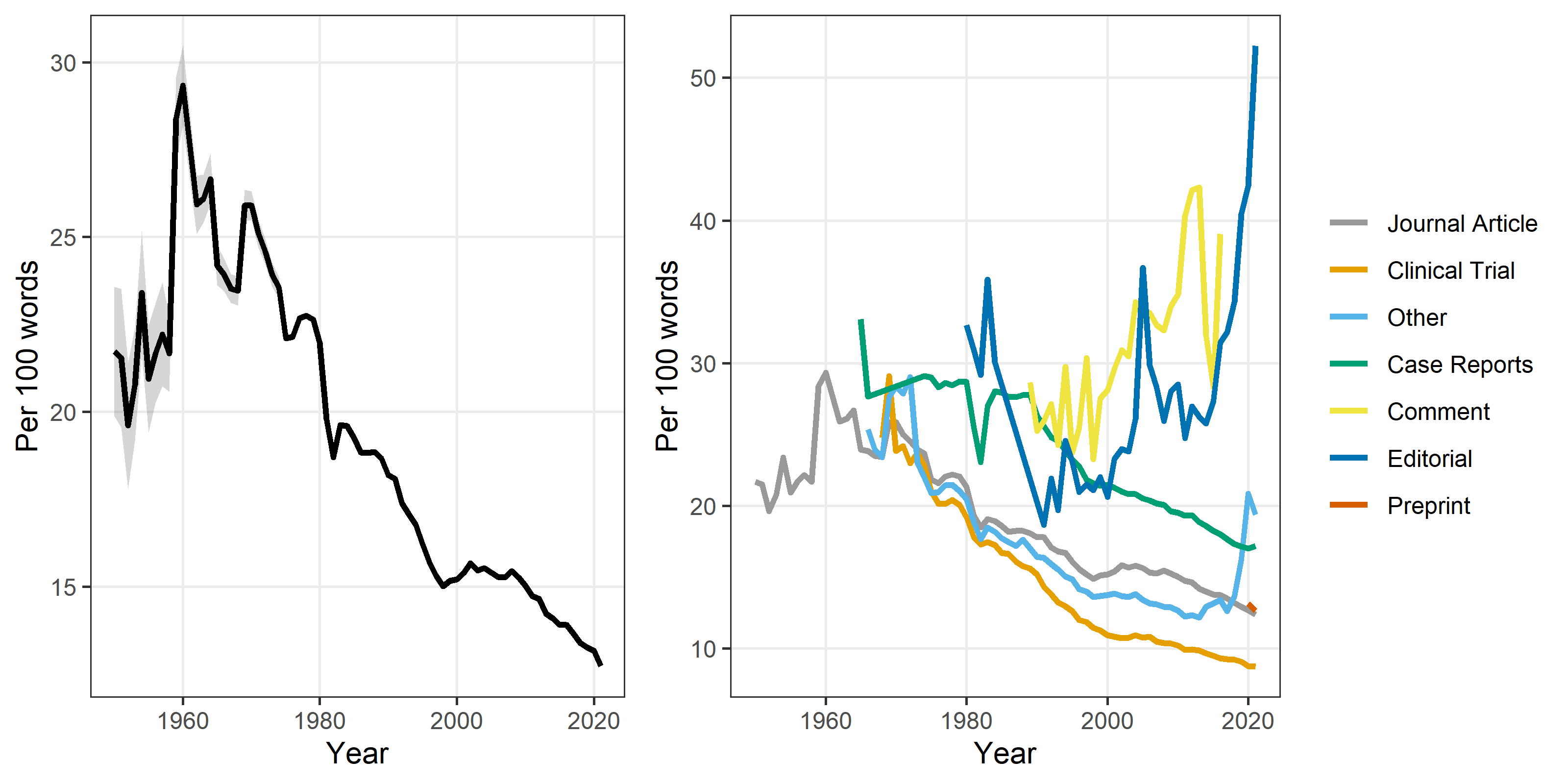


# Numbers

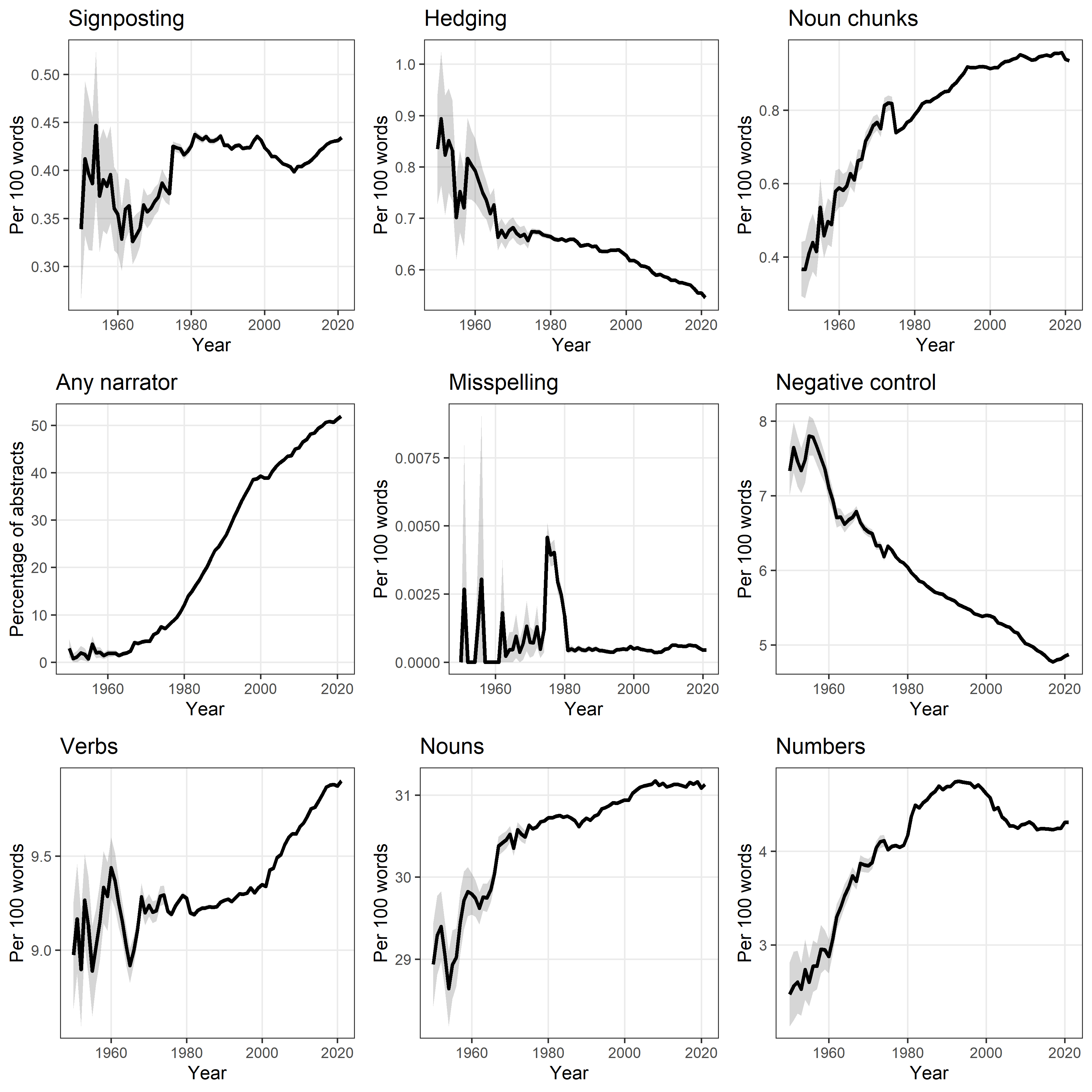


Labelled “numbers”, but includes numerals such as “11:00” or “one”.

# Number of words per sentence



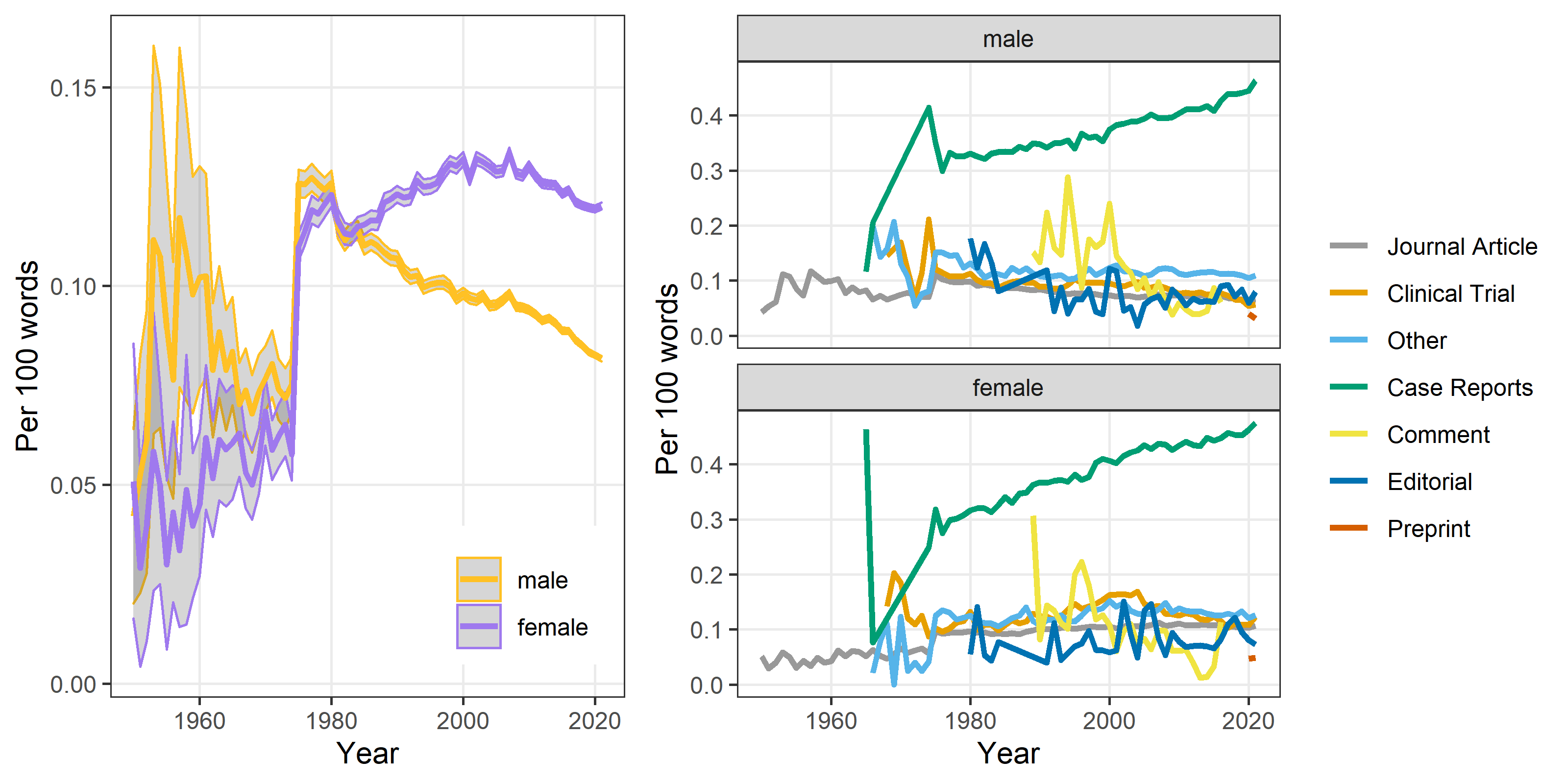
# Summary plot



# Male and Female

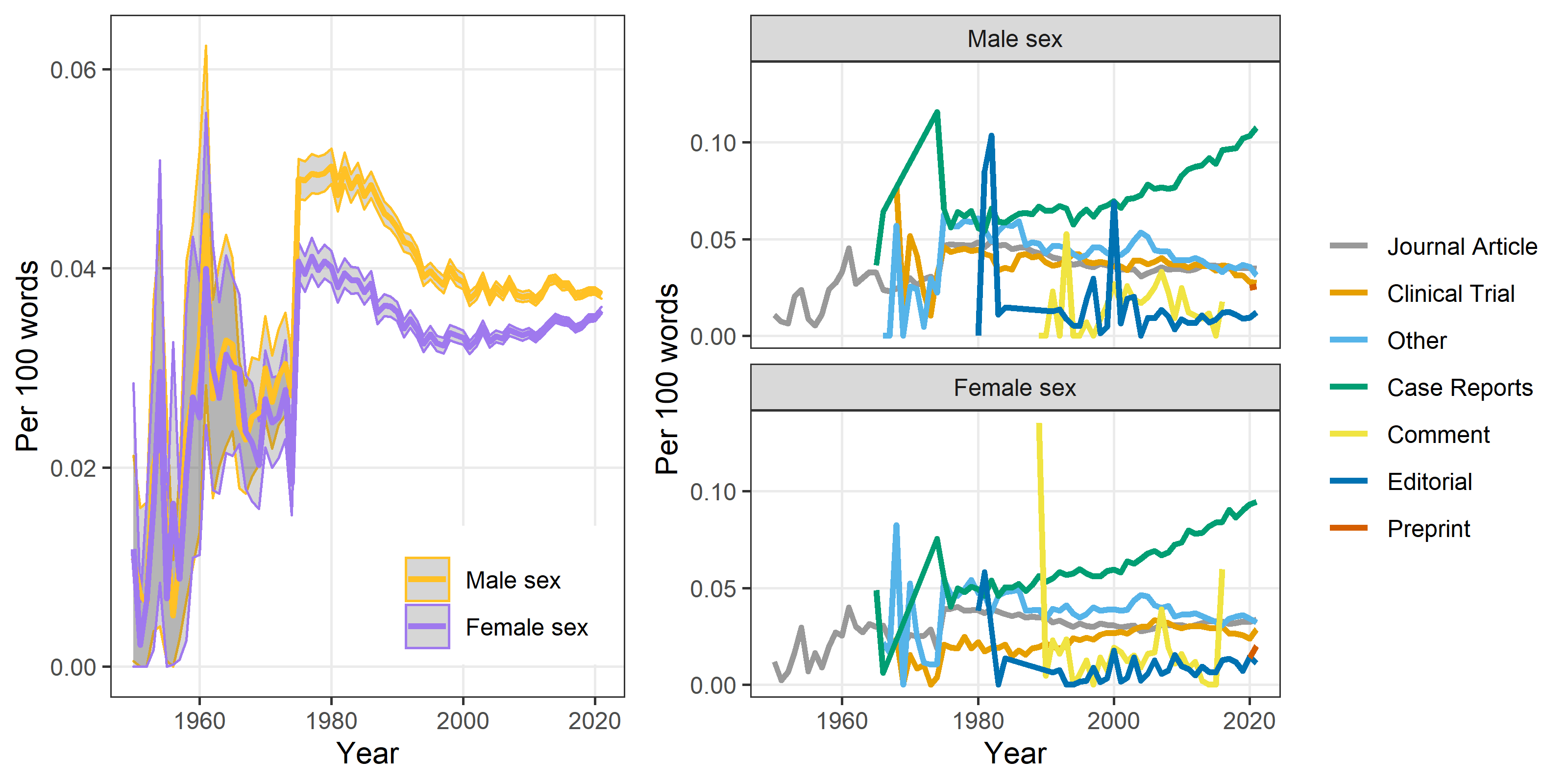
Here we look at the results for male and female words.

### All male and female words



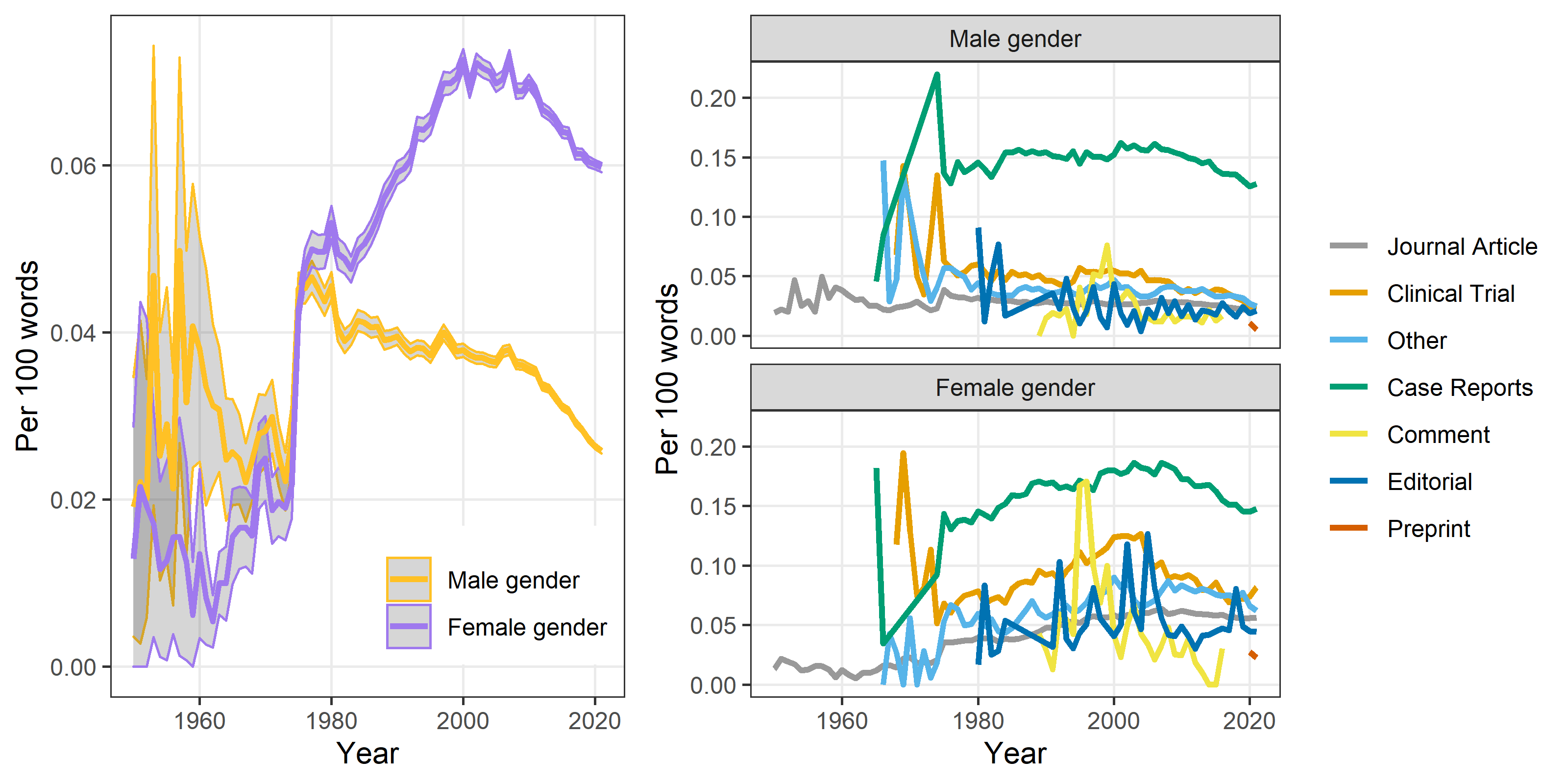
### Sex: Male and female

Includes male, males, female and females.



### Gender: Man and woman

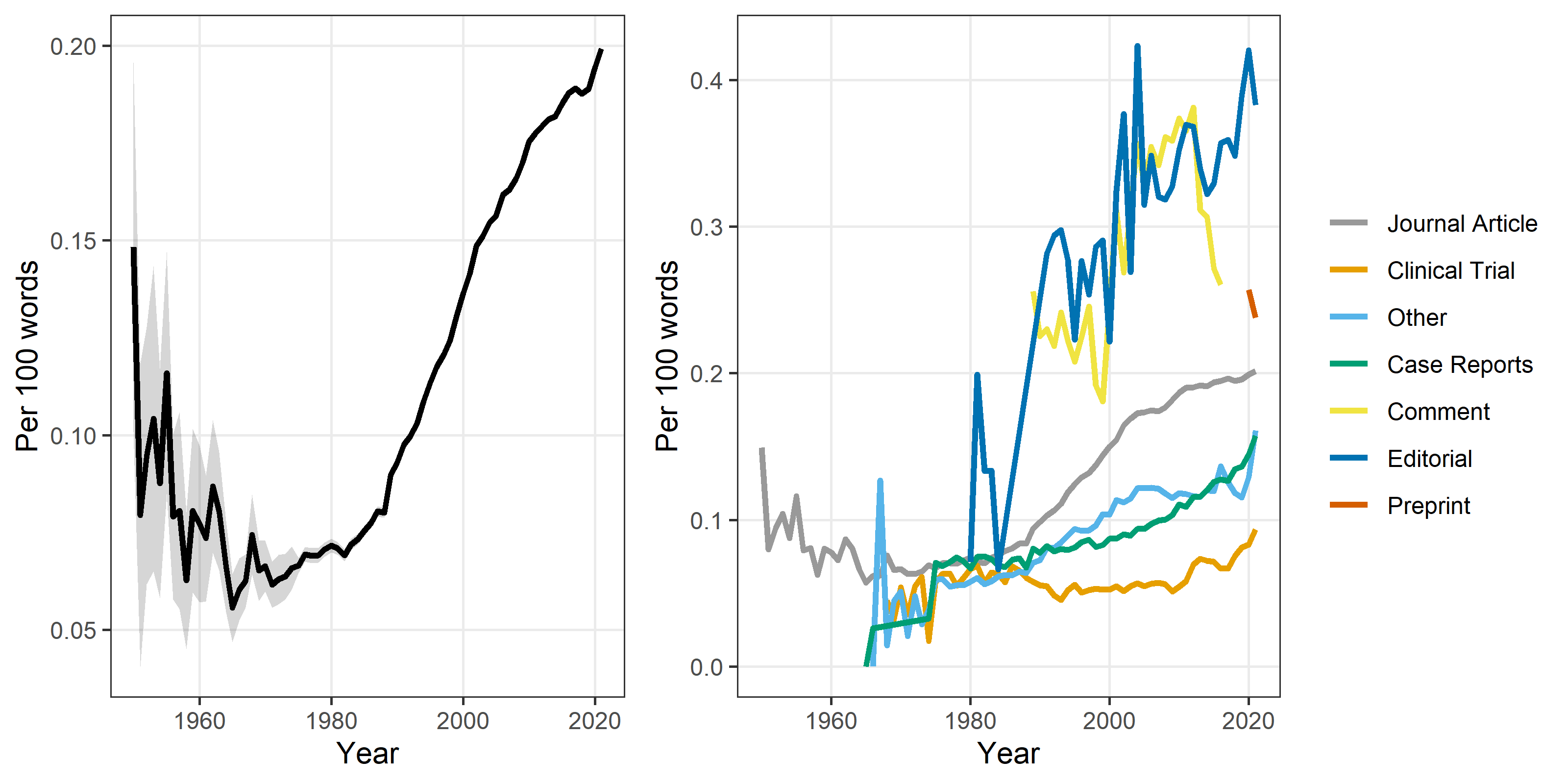
Uses man, men, boy, women, woman, girl and all plurals.



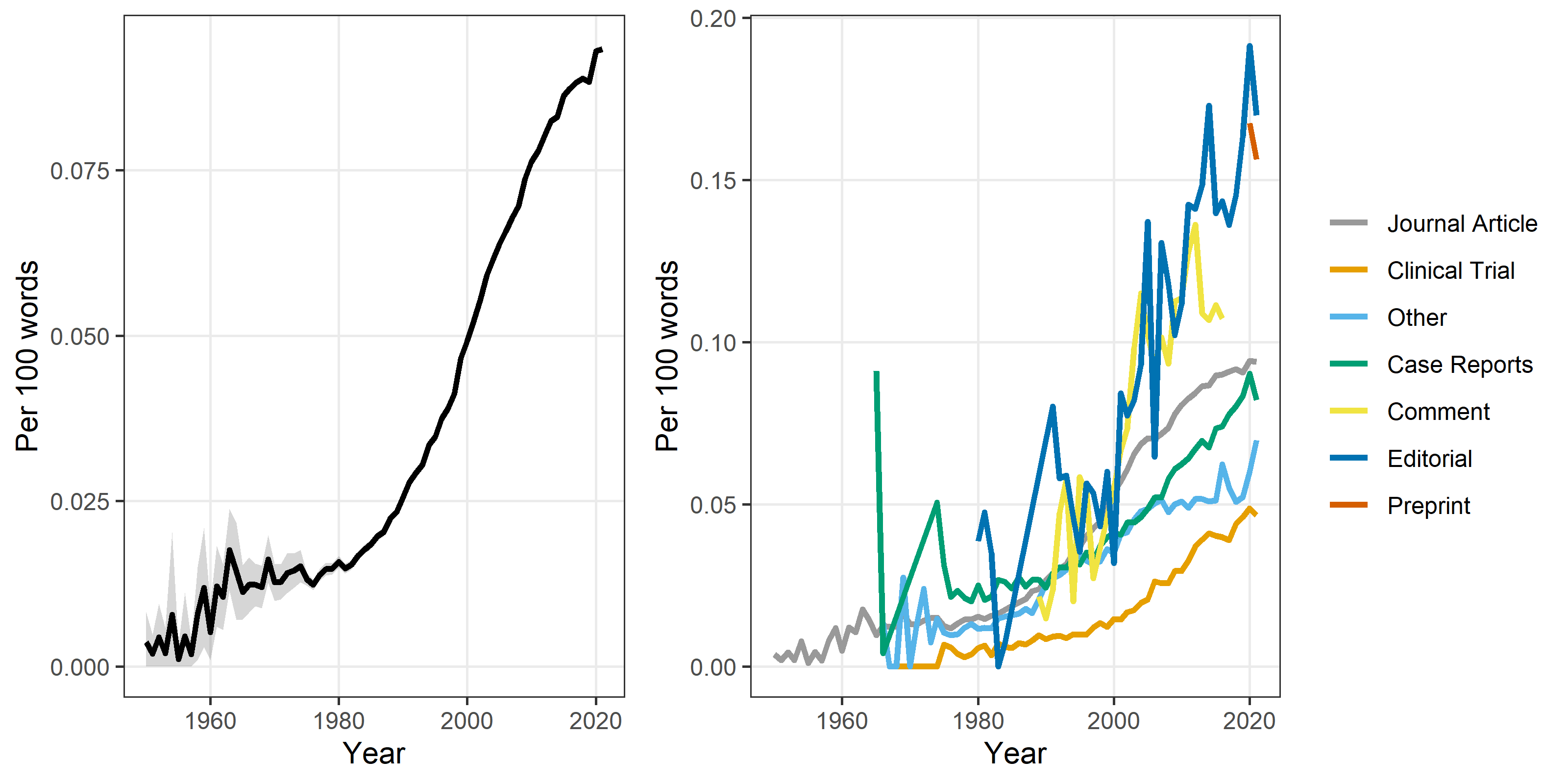
# Hype

Here we examine our definitions of hype.

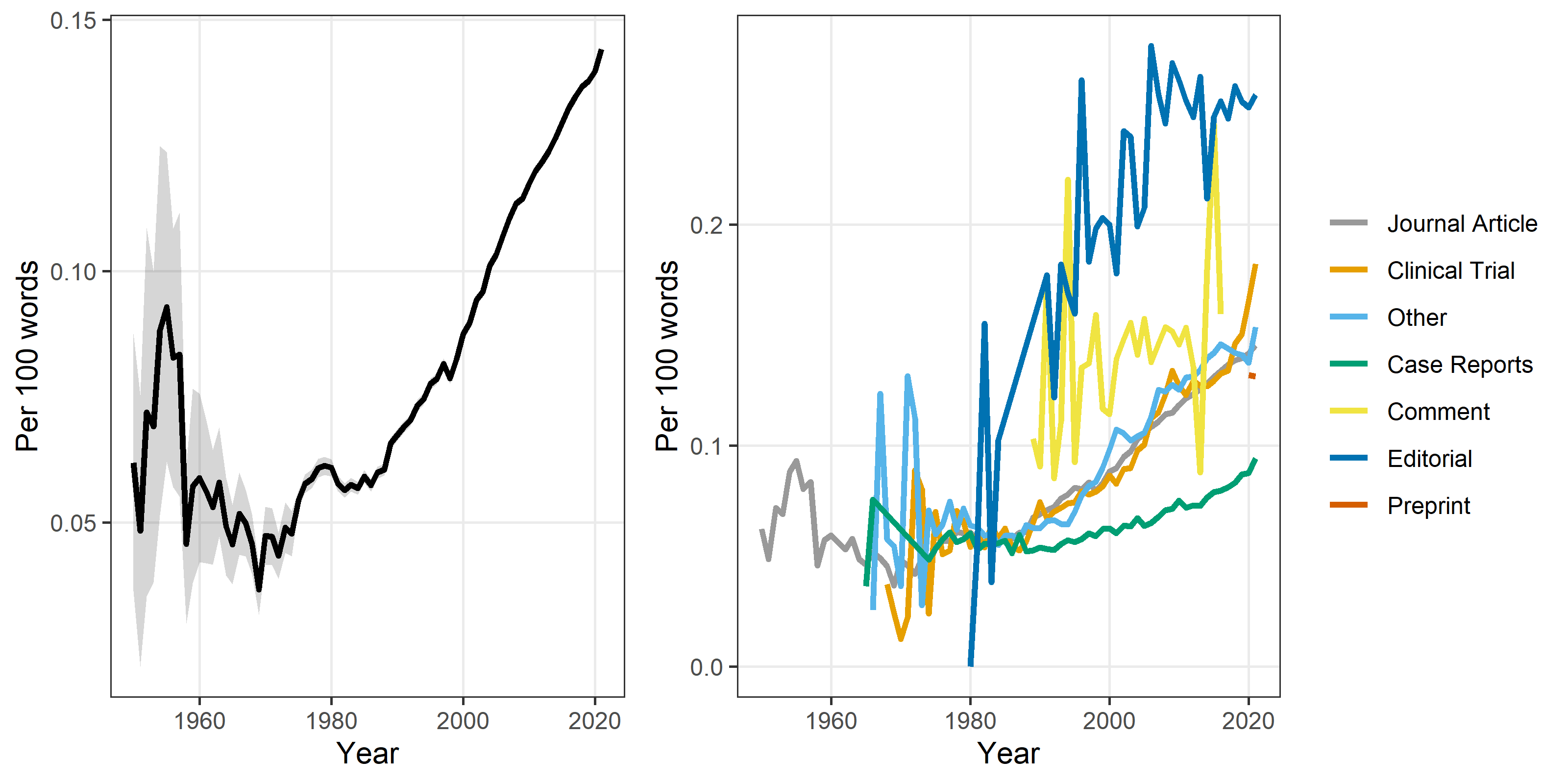
## Importance



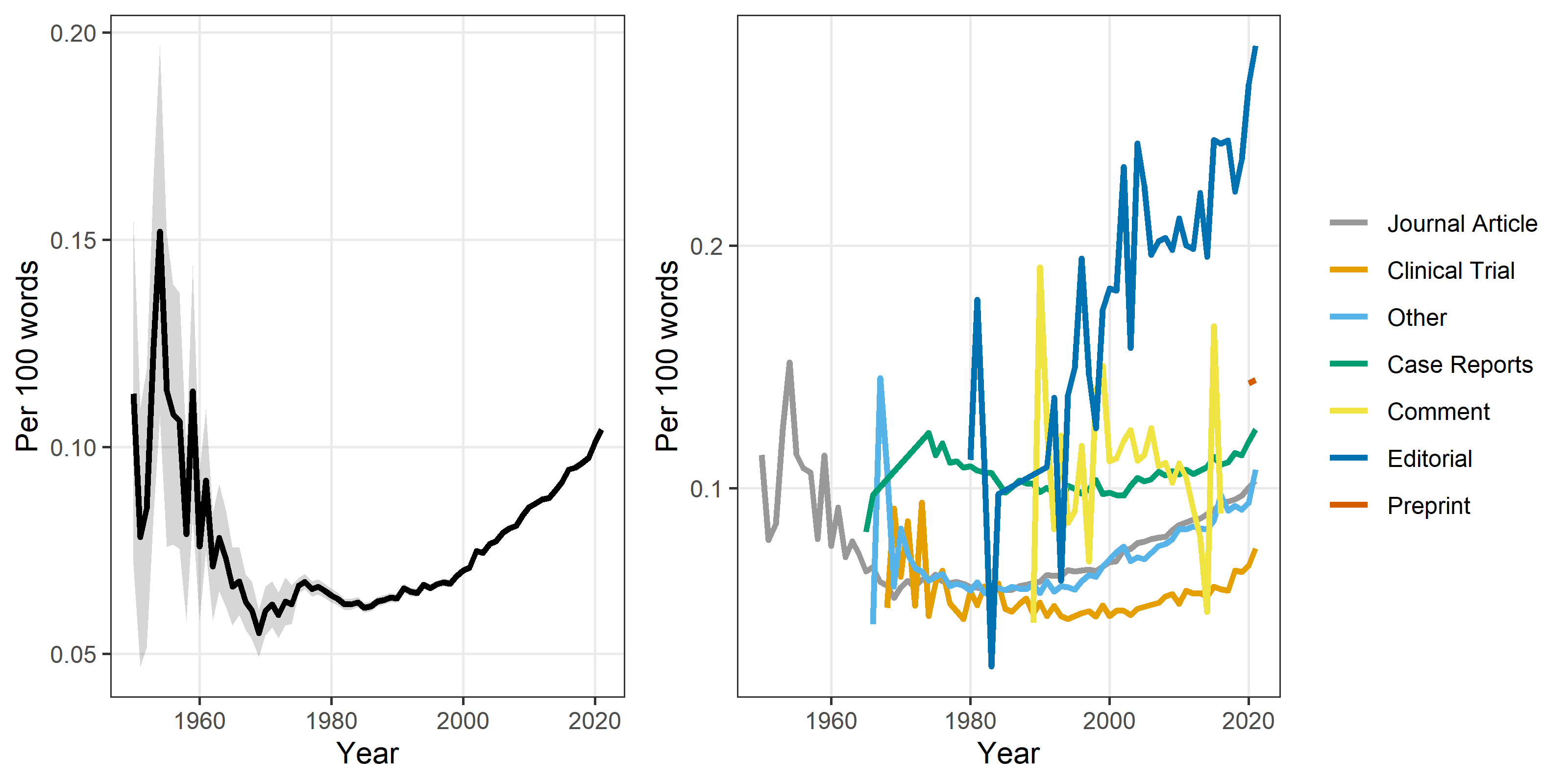
## Novelty



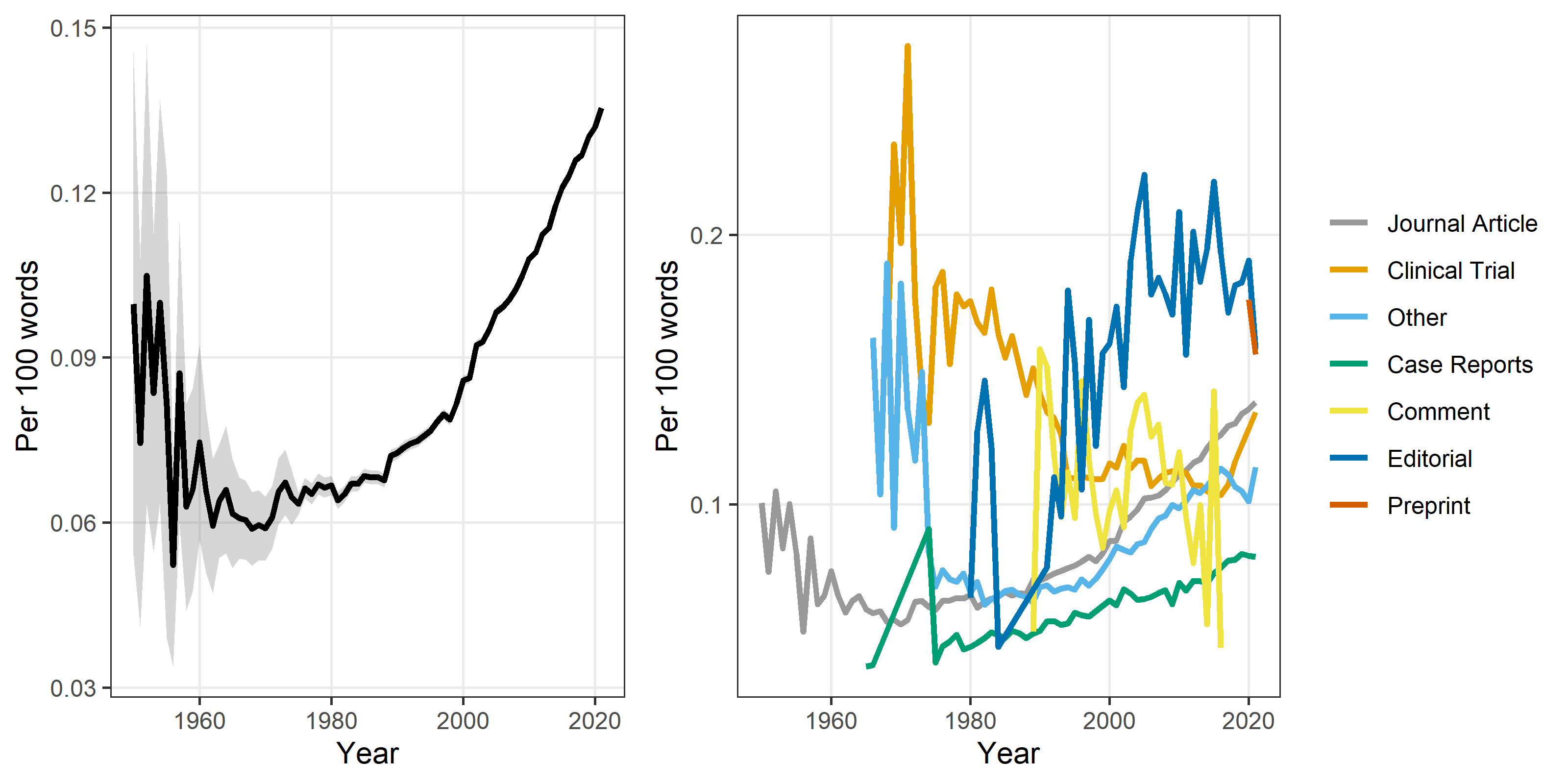
## Rigor



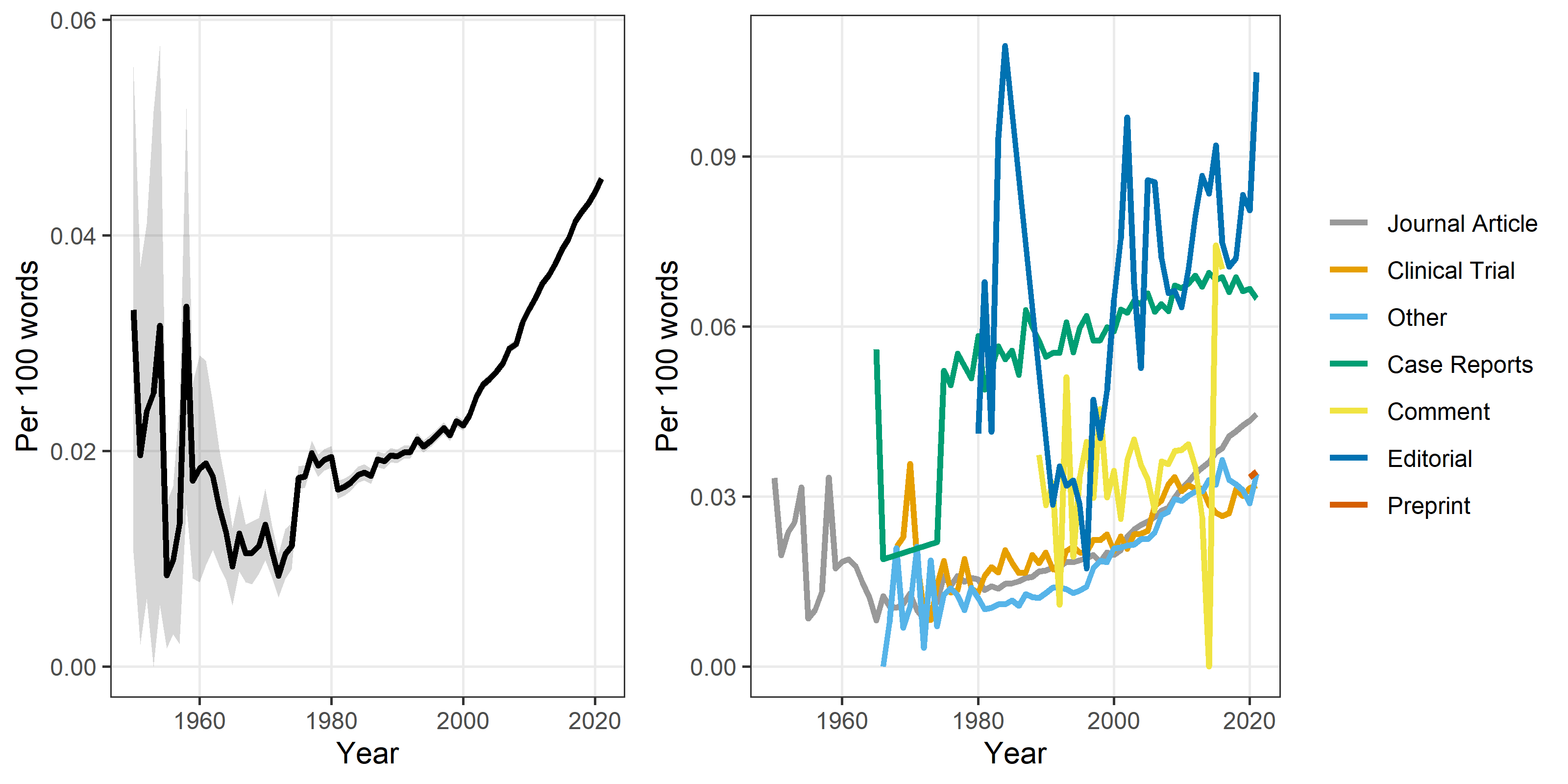
## Scale



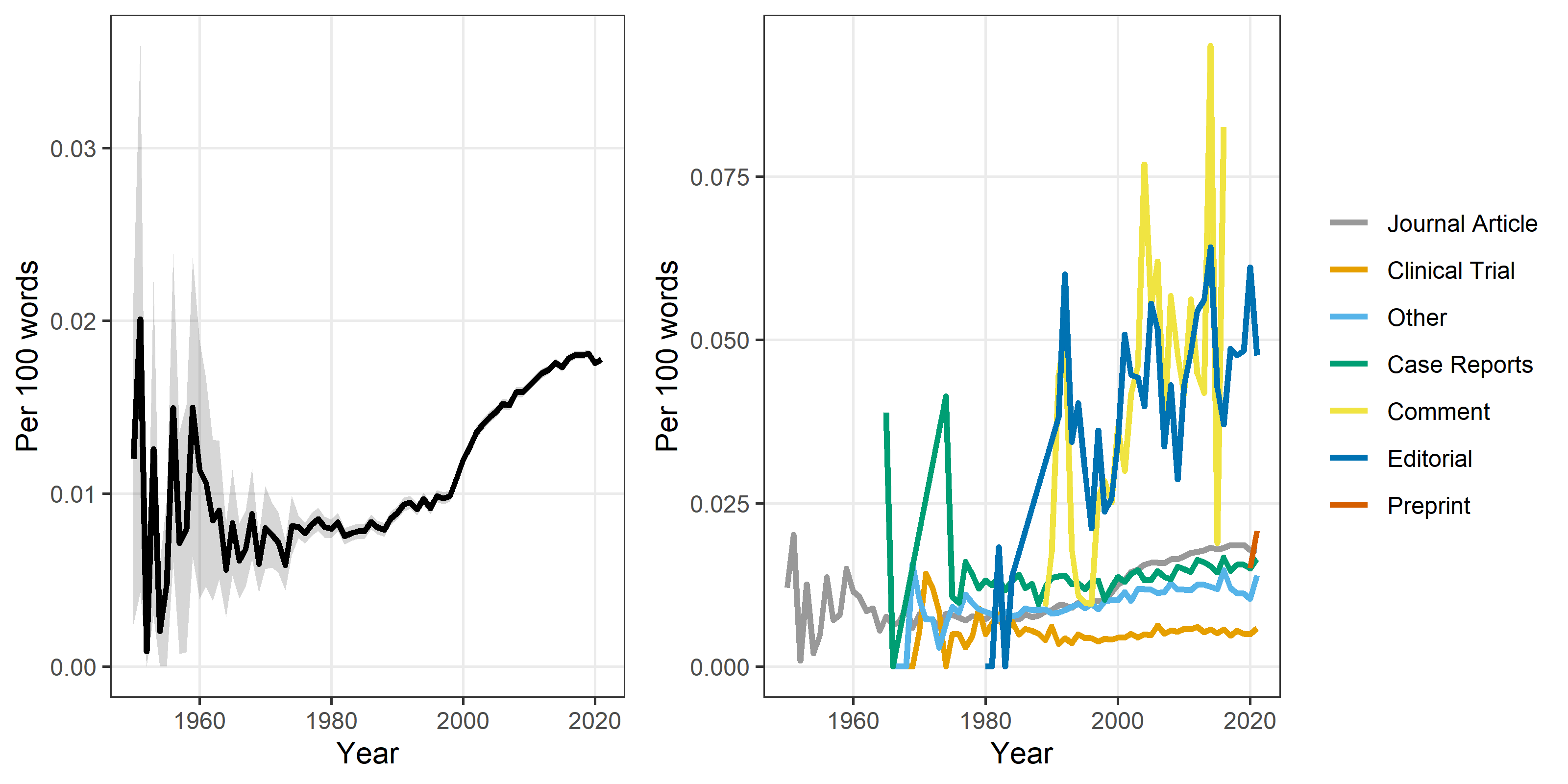
## Utility



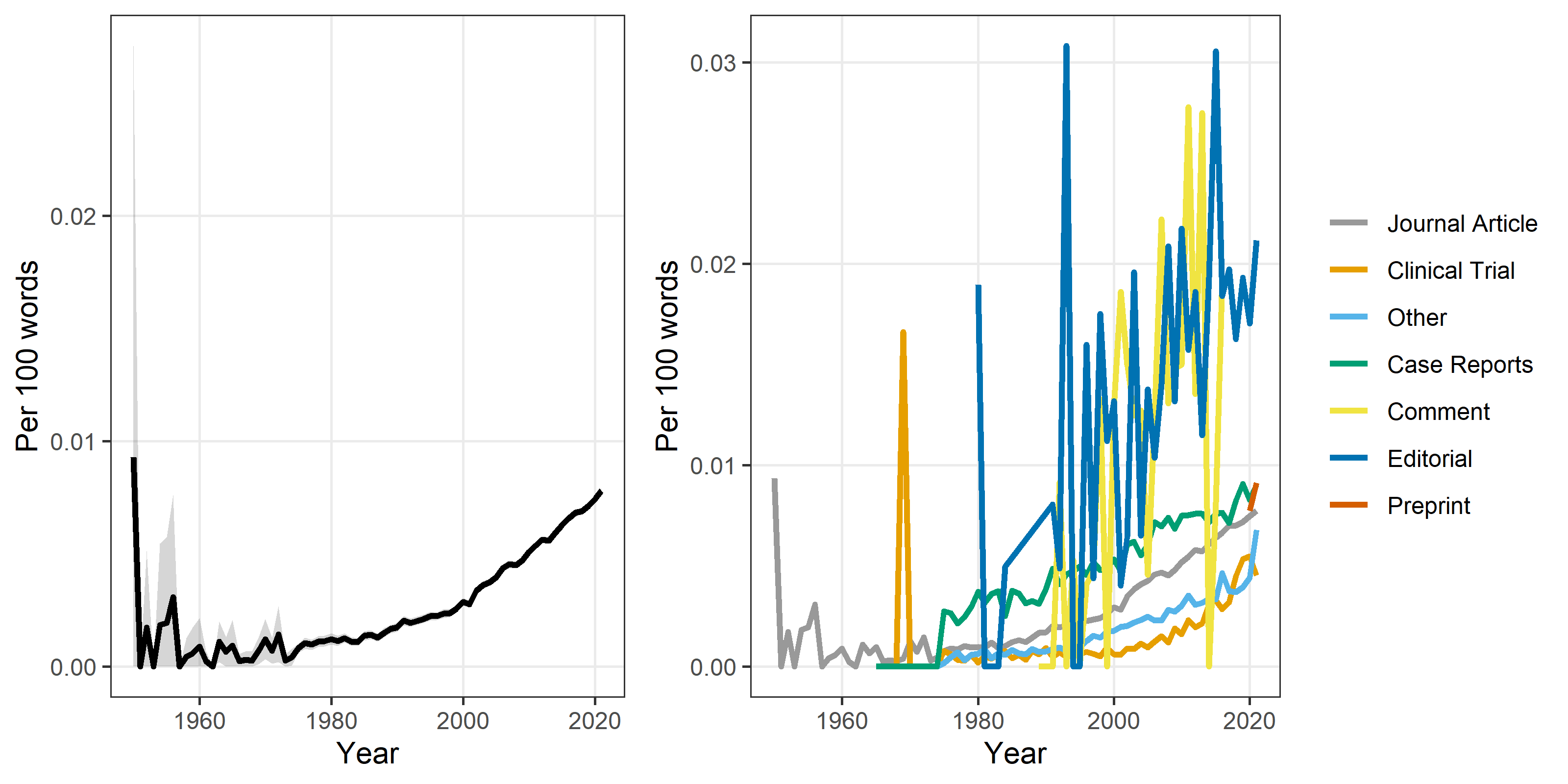
## Quality



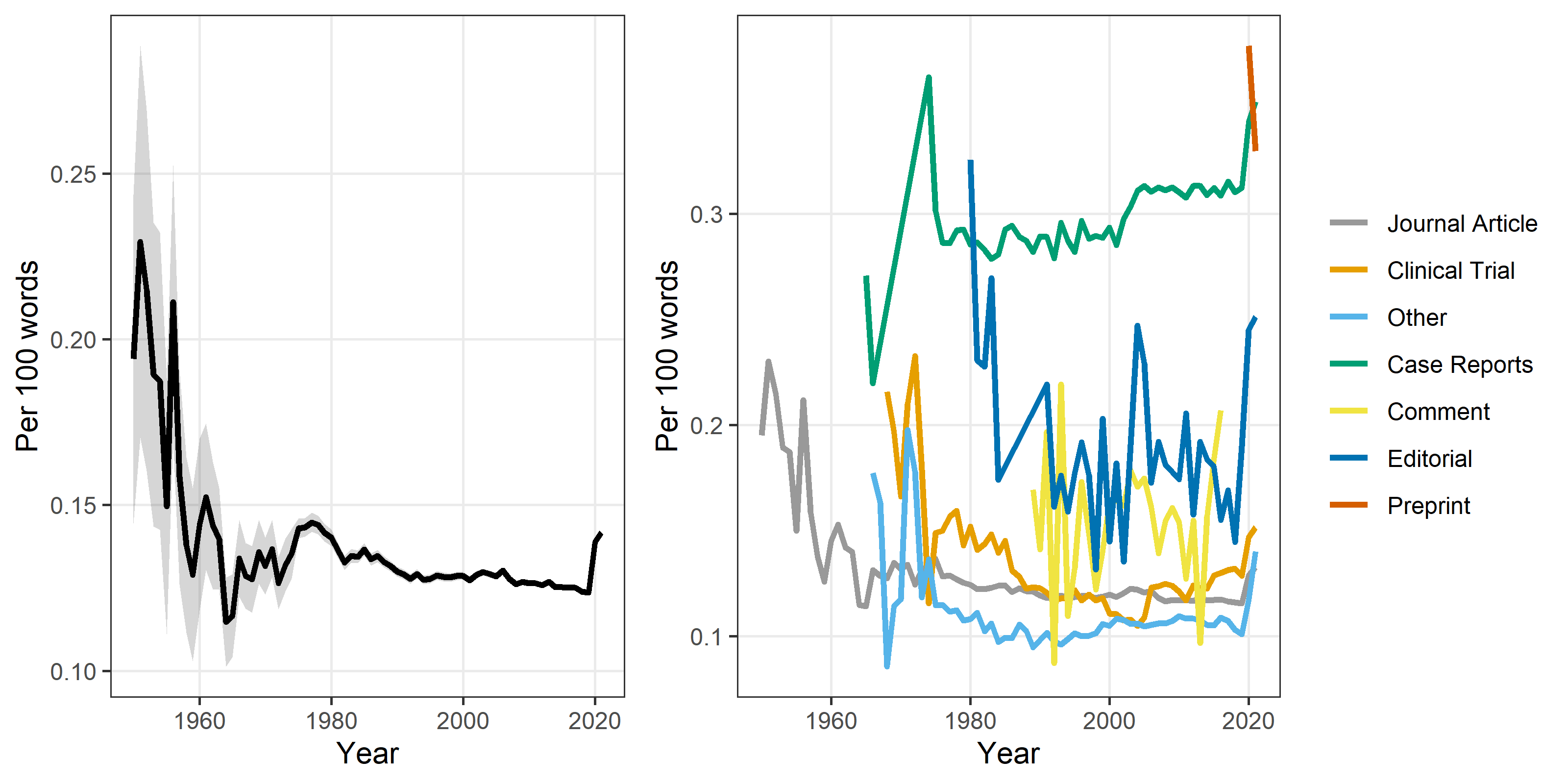
## Attitude



## Problem



## Additional hype words



## Summary plot for hype

