Plot trends over time for narrator analysis. Version stratified by country of affiliation.

Adrian Barnett

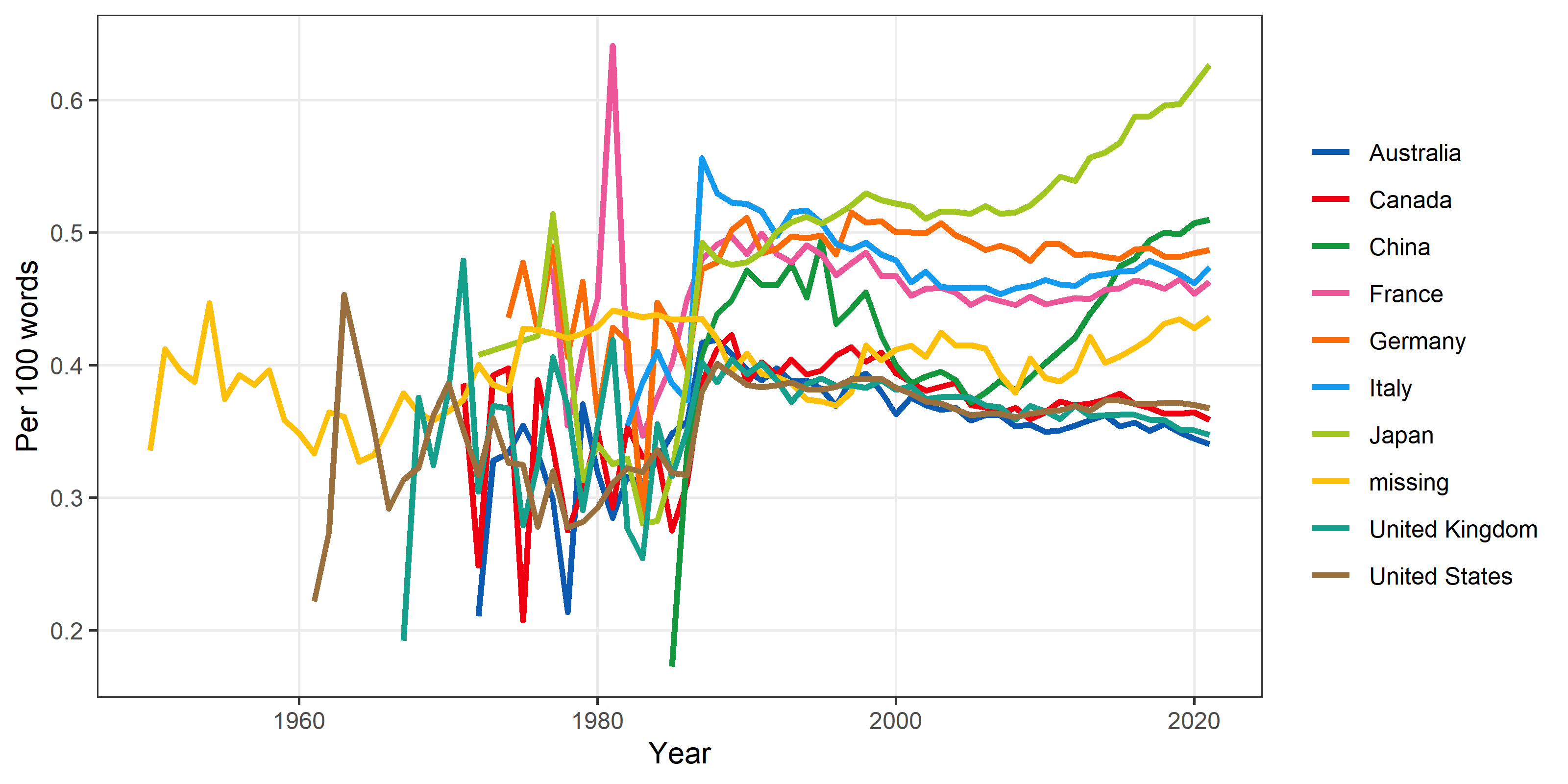
15 September 2022

This document contains trend plots for our analysis of language over time in papers published on *PubMed*. In this analysis we look at trends over time by the top ten countries. In total there were 15,628,561 abstracts after restricting to the top ten countries.

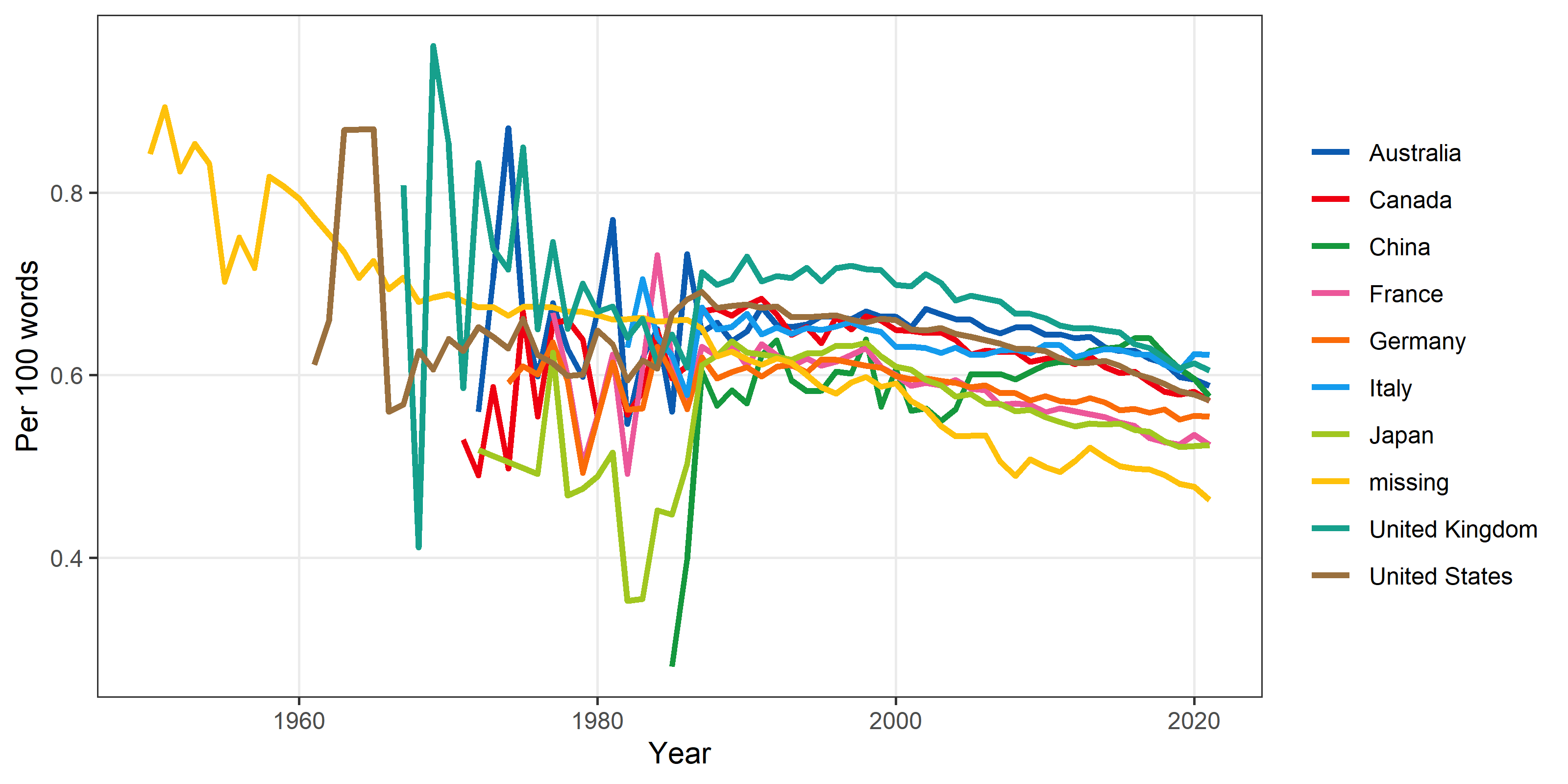
The plots show the mean per year. We do not show the 95% confidence interval for the mean, because the plot would be too busy.

We exclude the yearly results for a country if there were fewer than 50 abstracts that year.

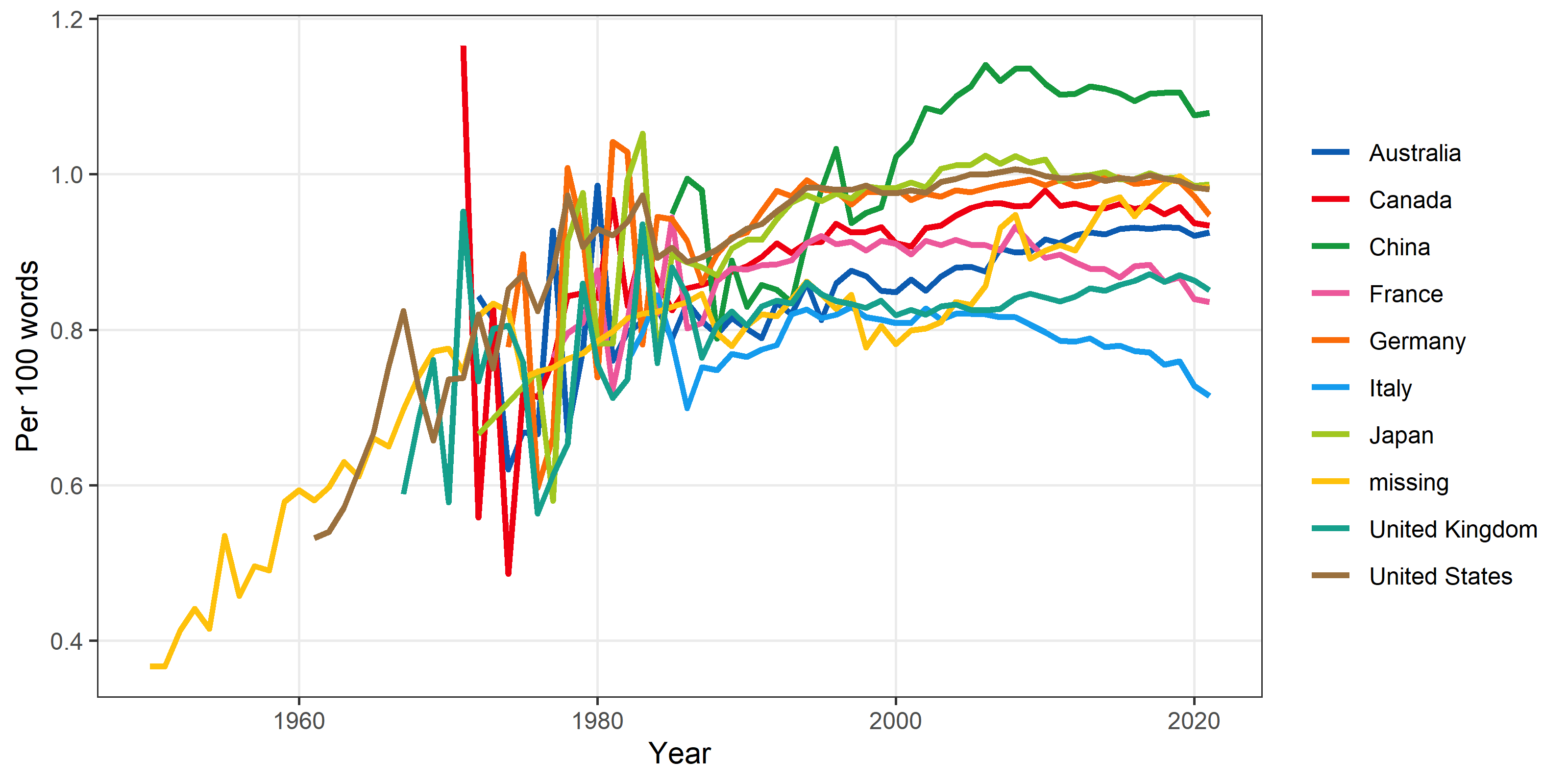
# Signposting



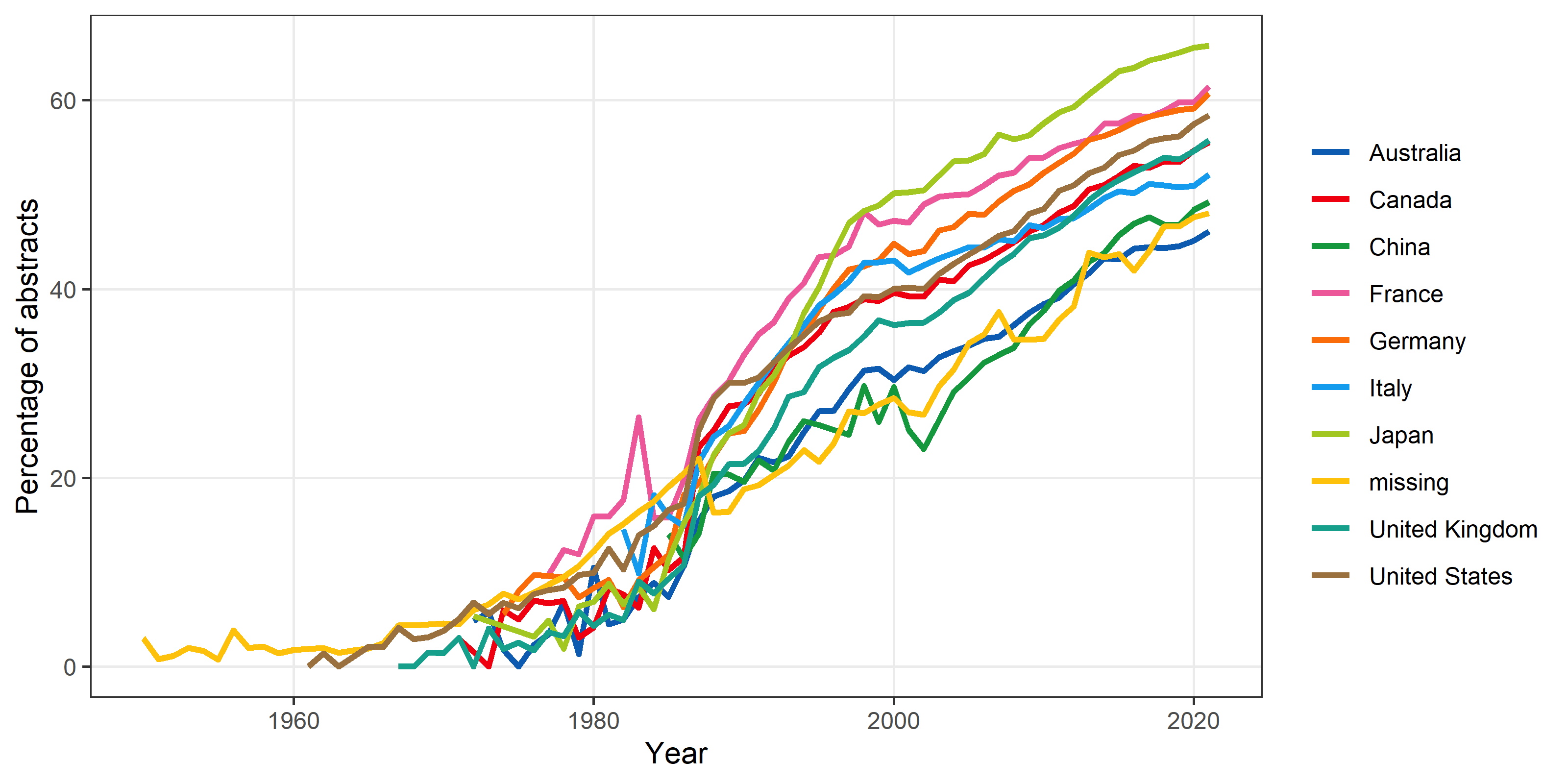
# Hedging



# Noun chunks

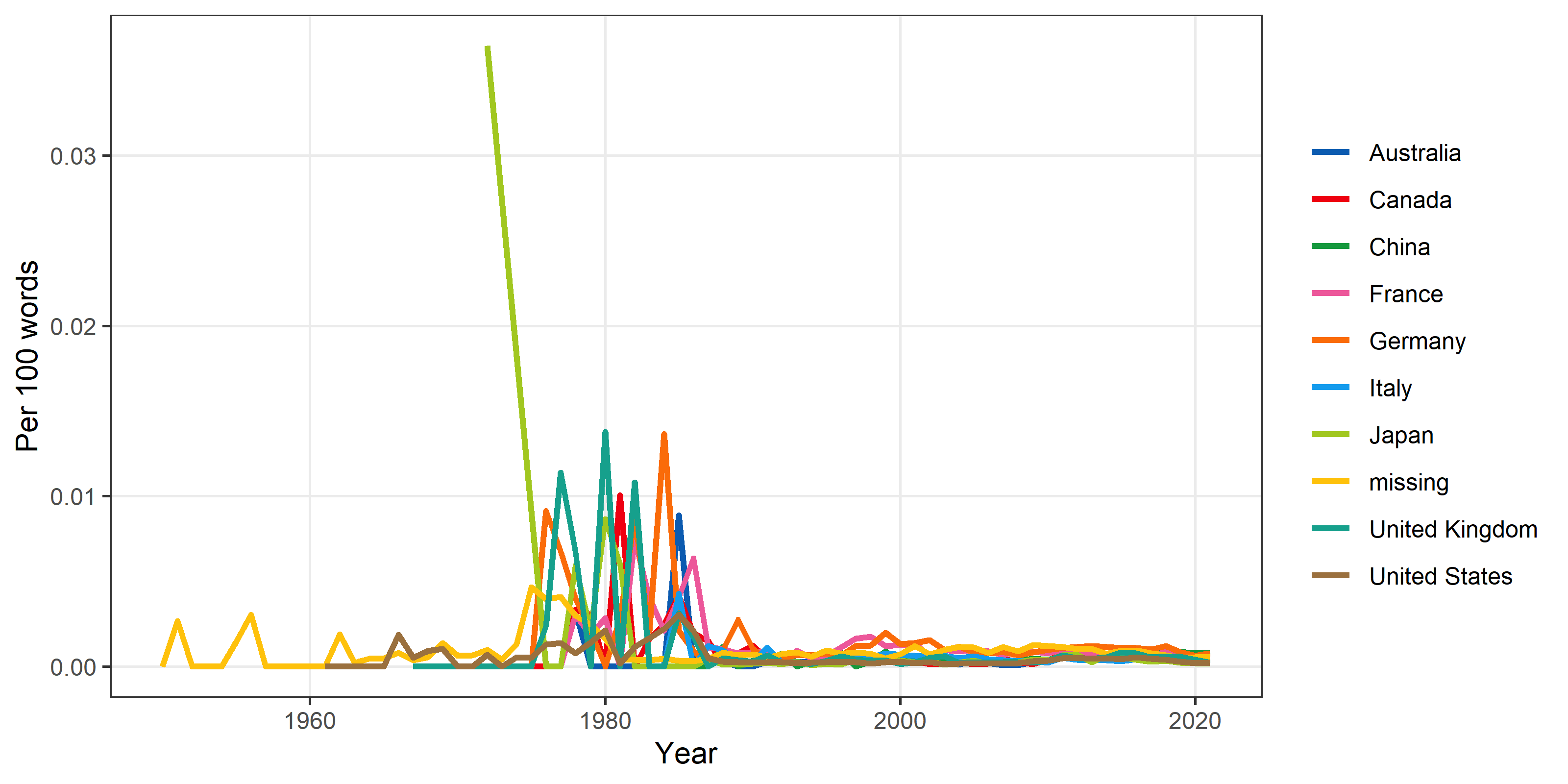


# Any narrator

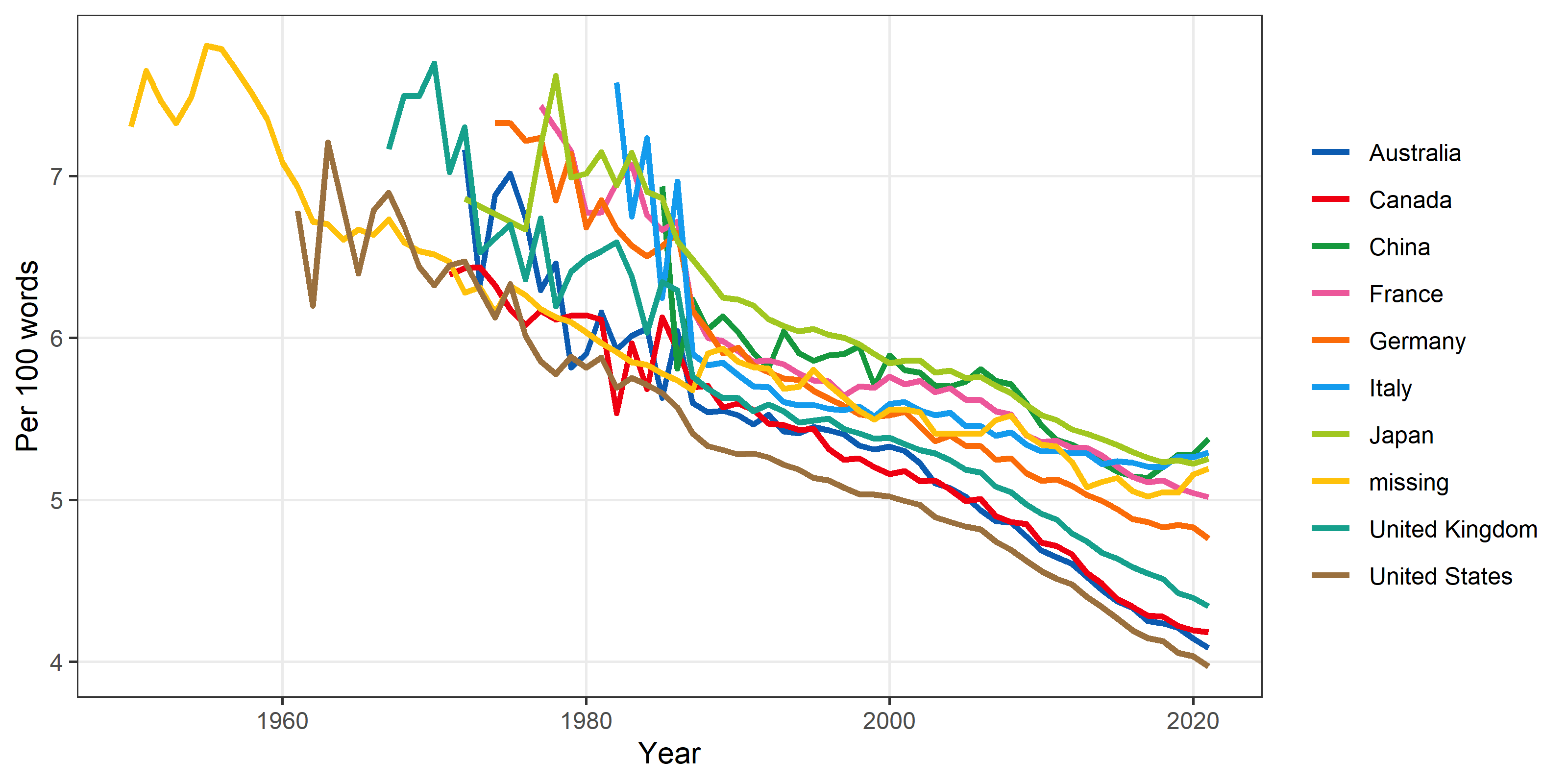


This is a binary yes or no outcome, whereas the other outcomes are counts.

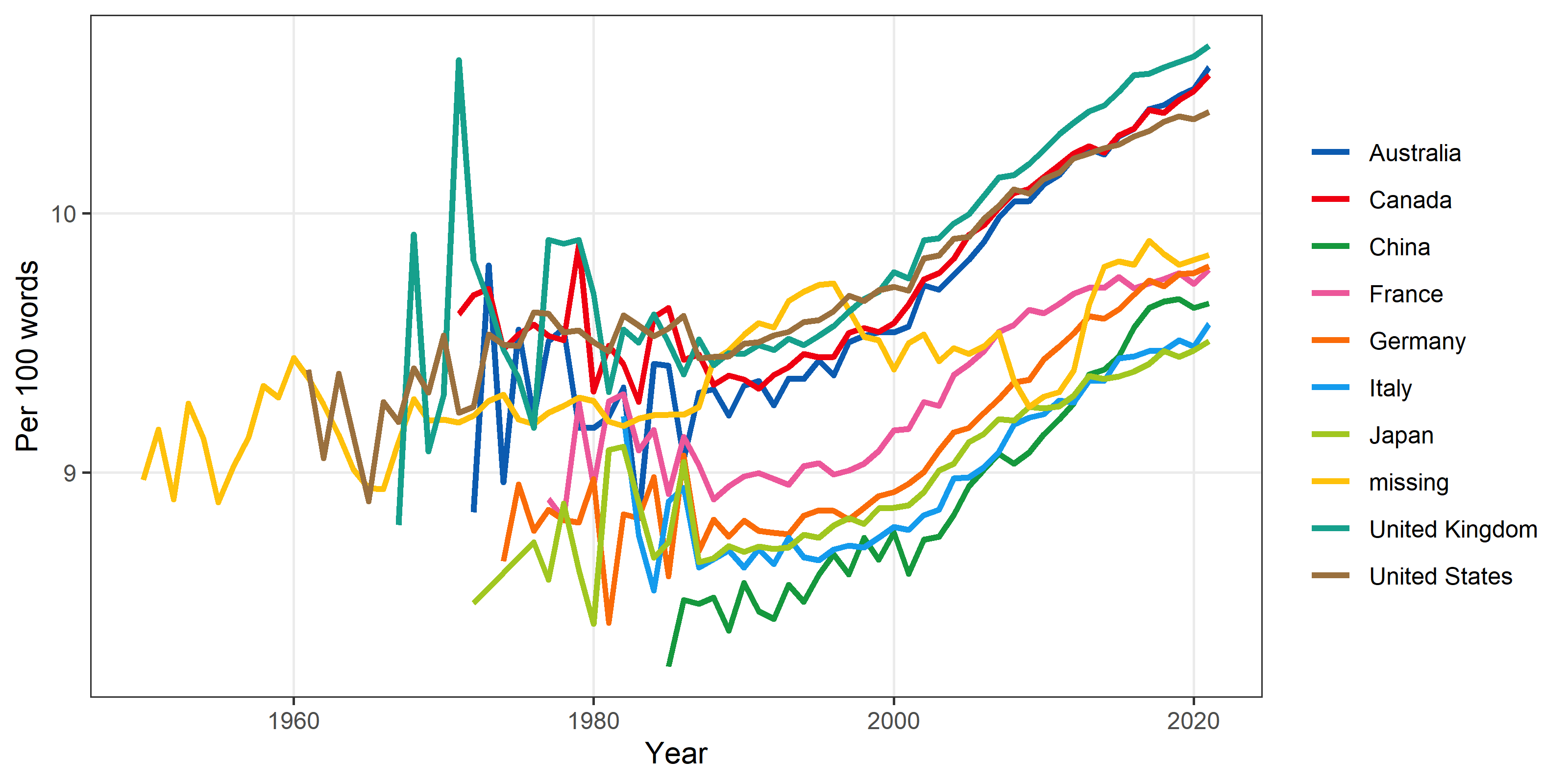
# Misspellings



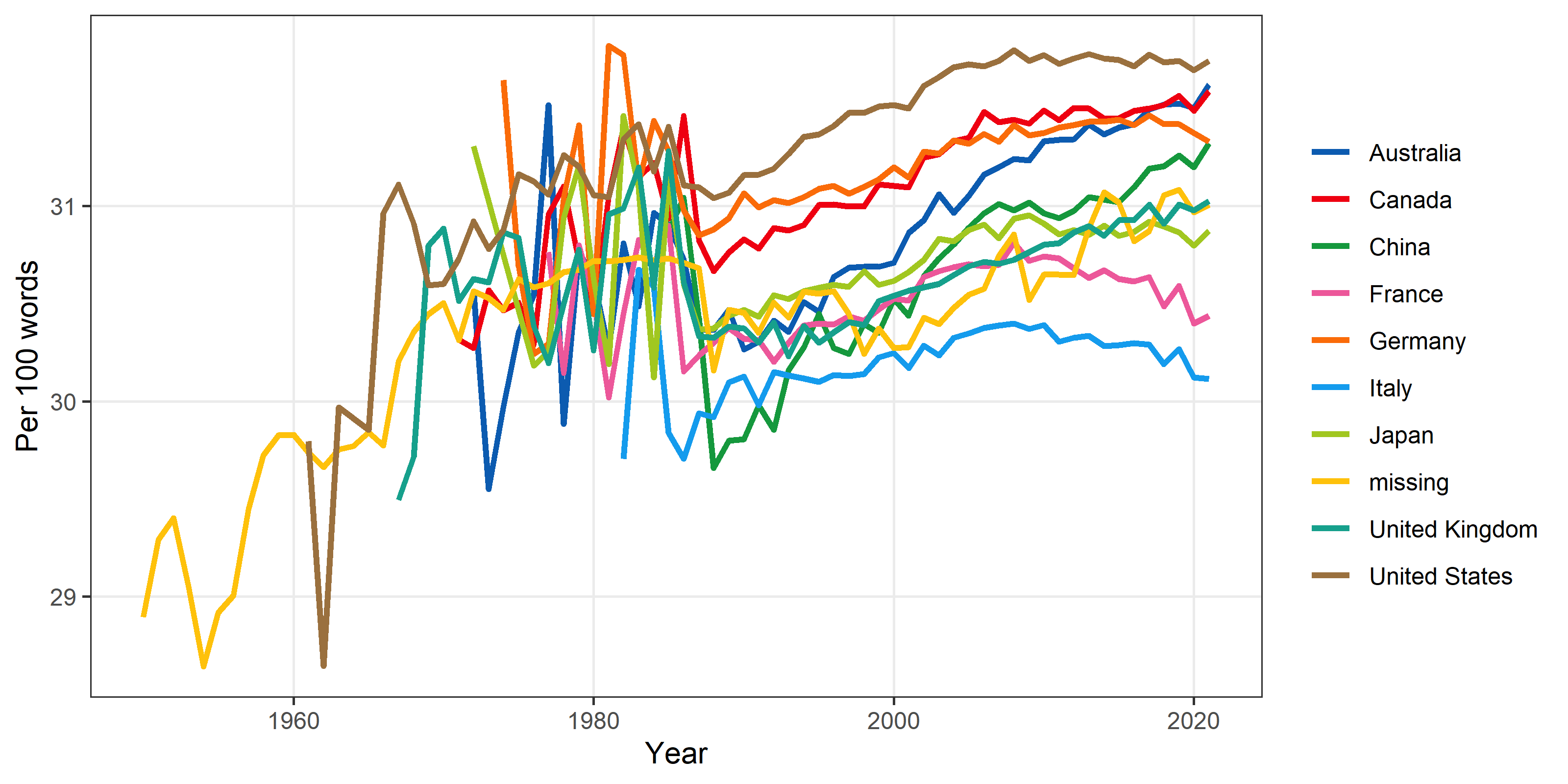
# Negative control



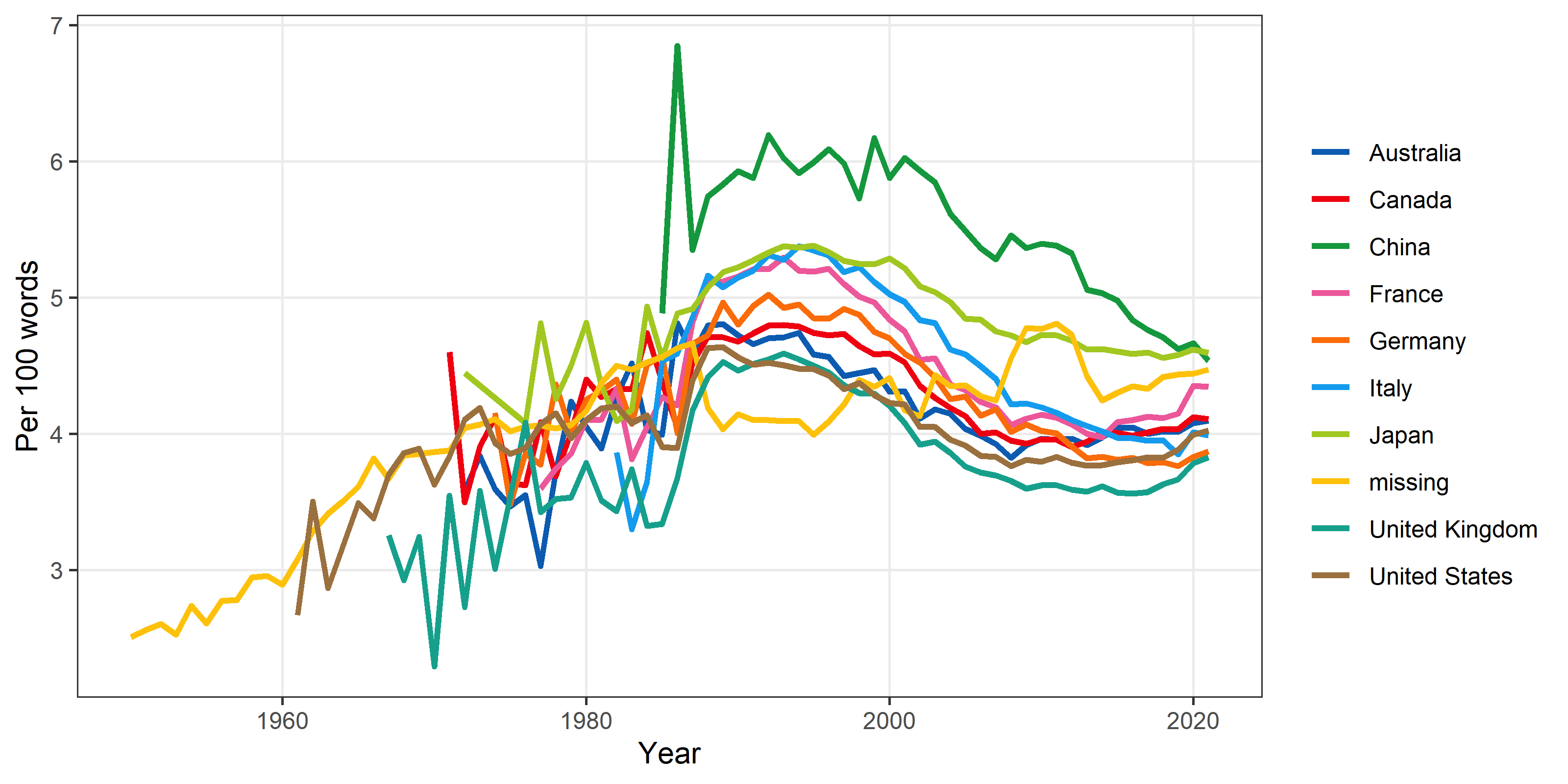
# Verbs



# Nouns

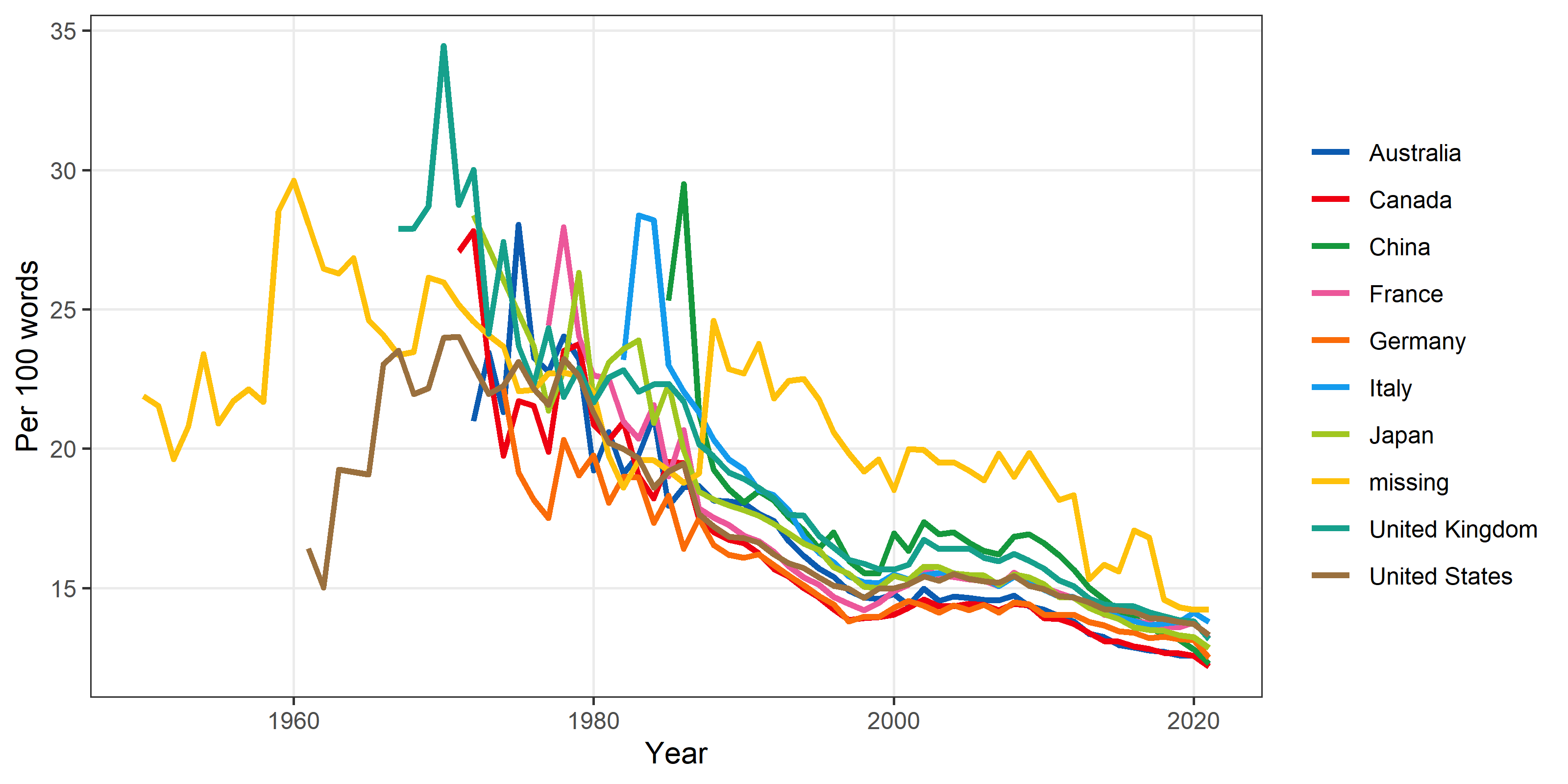


# Numbers

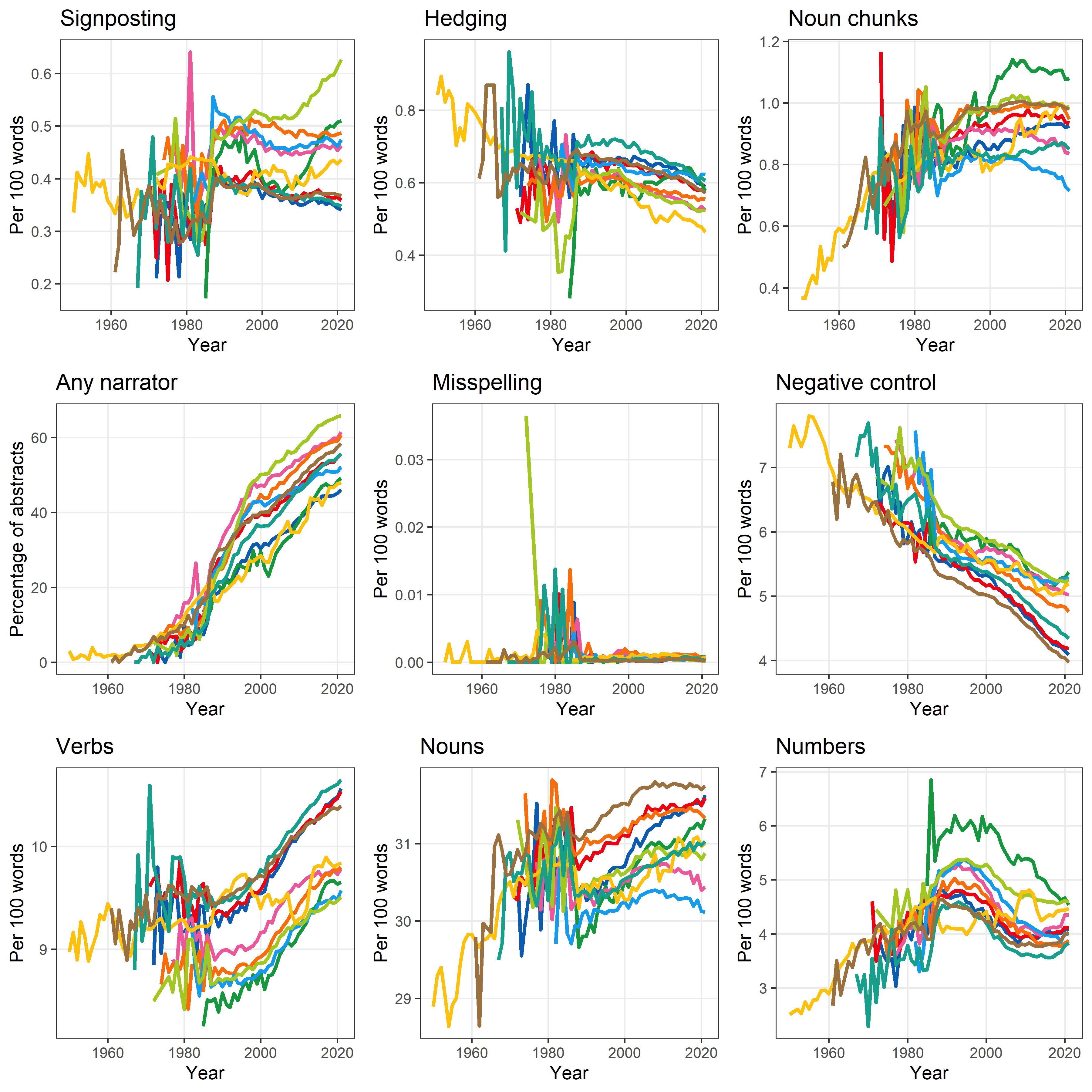


Labelled “numbers”, but includes numerals such as “11:00” or “one”.

# Number of words per sentence



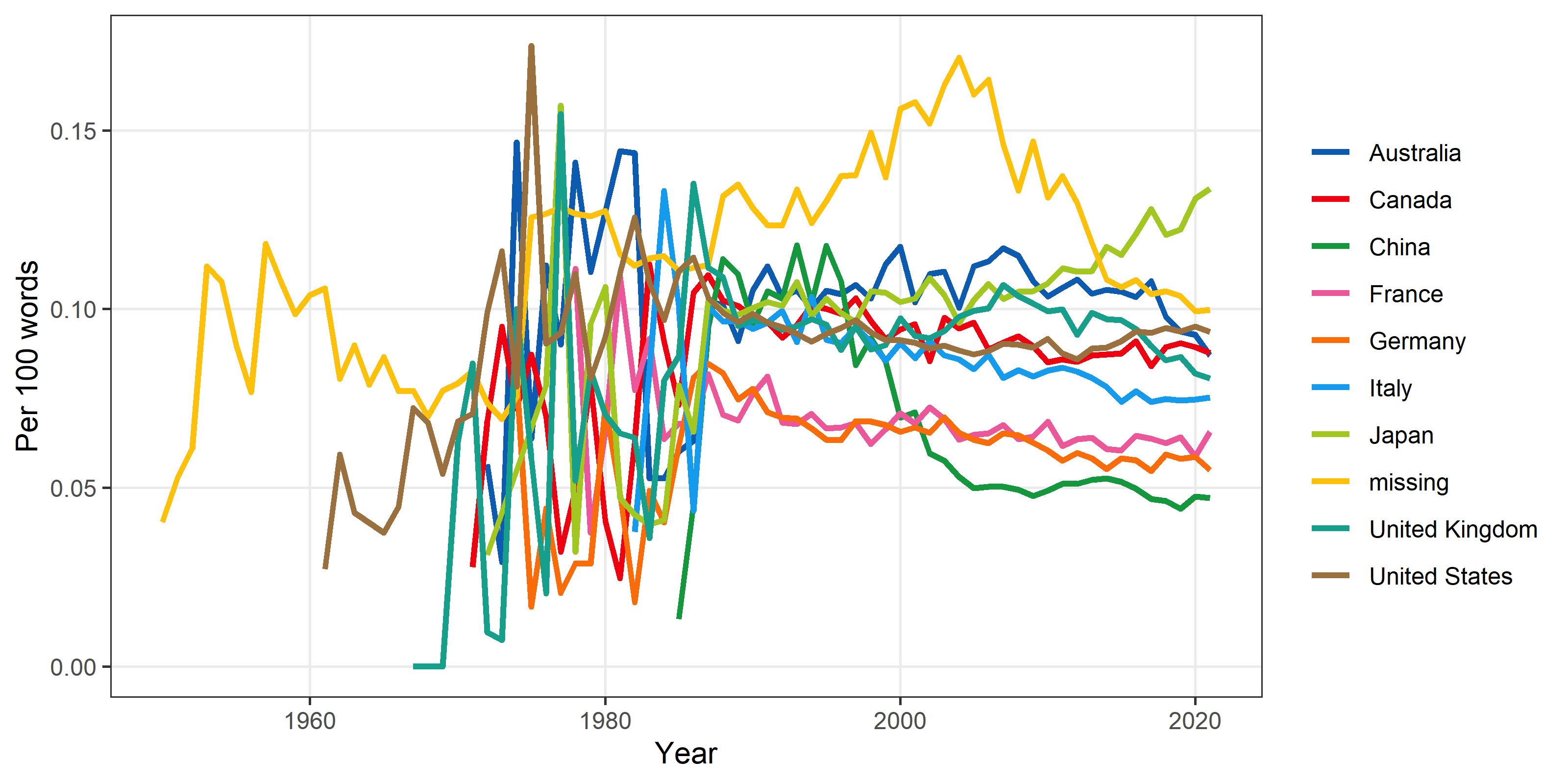
# Summary plot



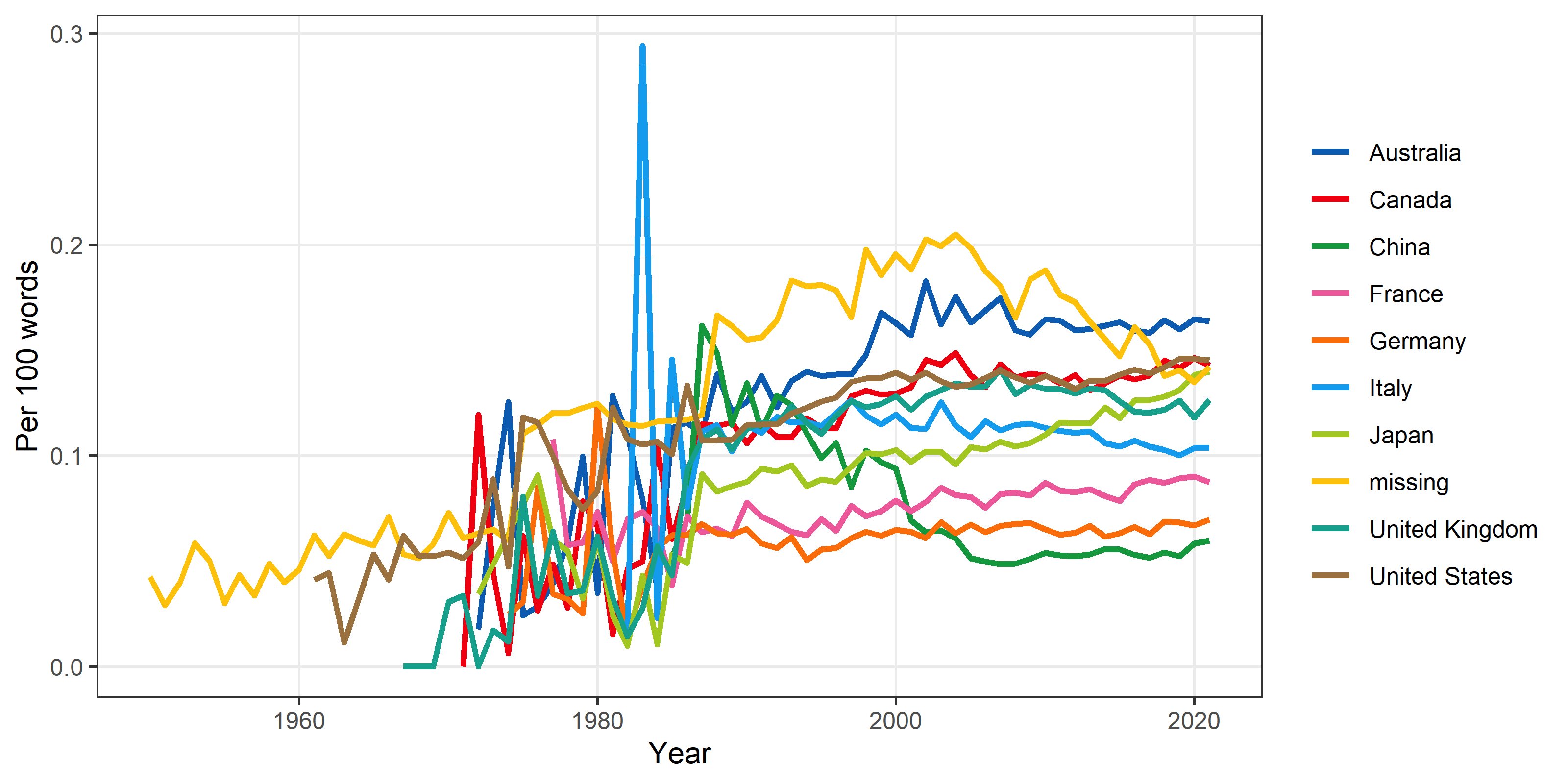
# Male and Female

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

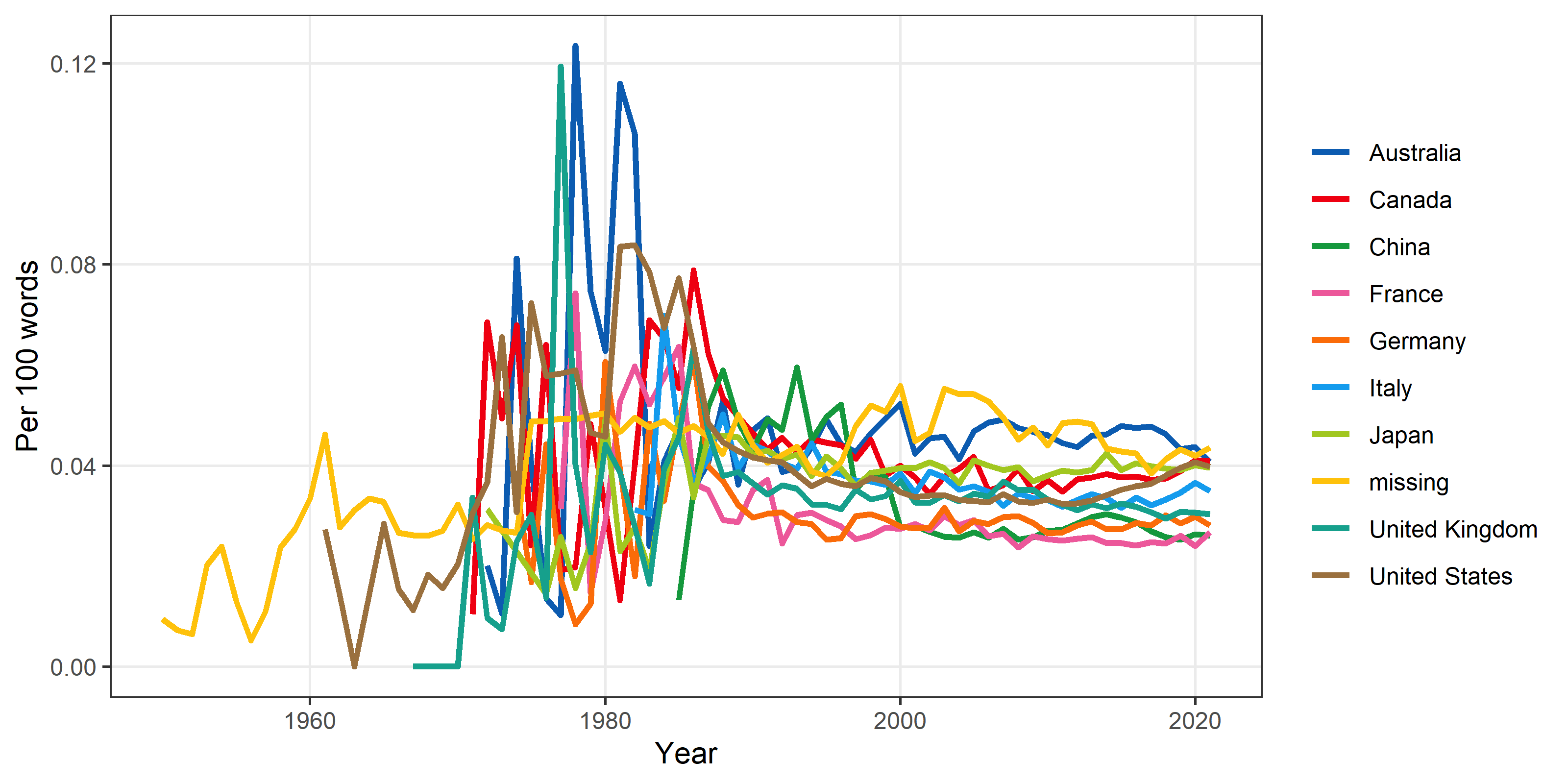
### All male words



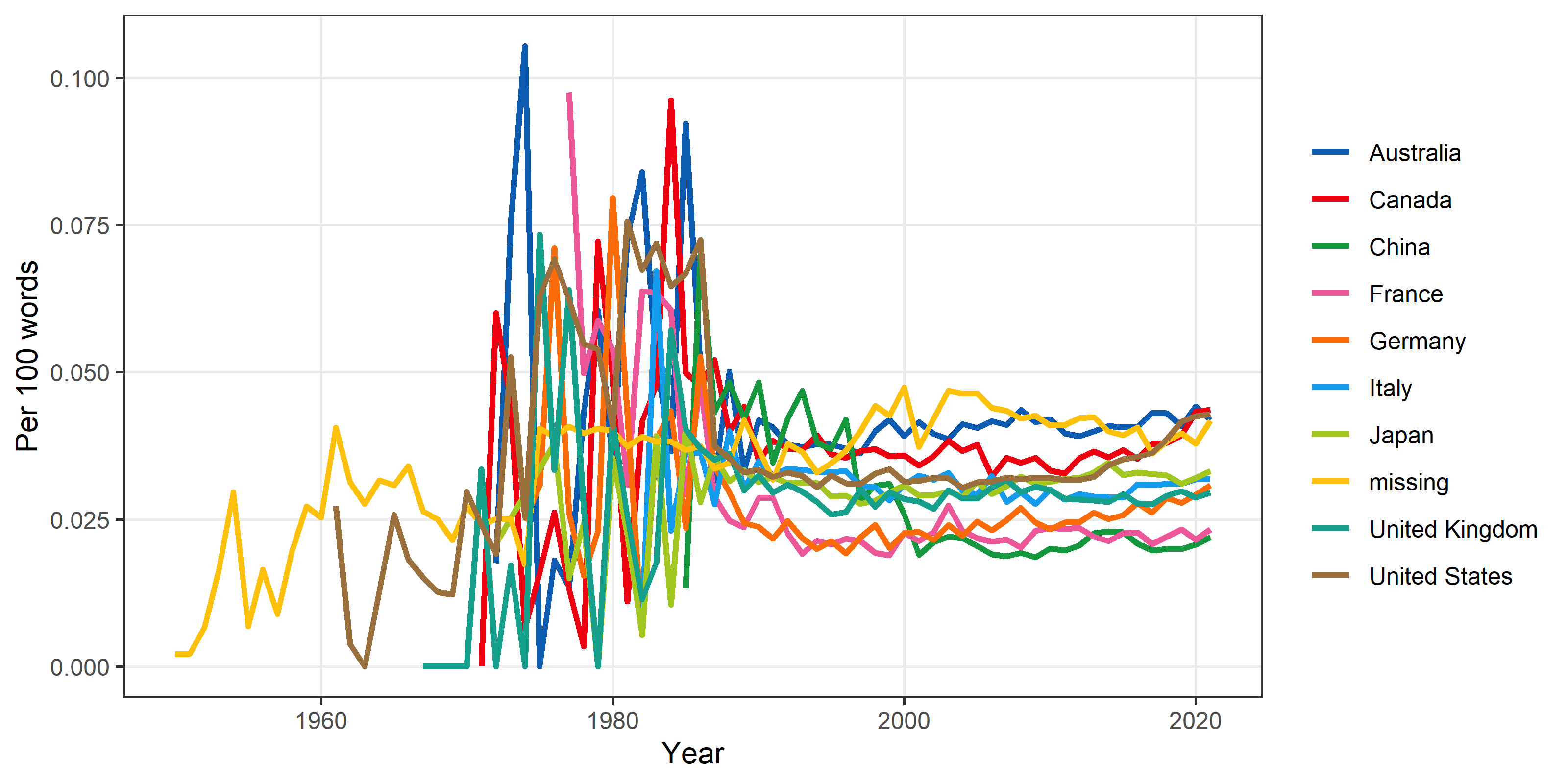
### All female words



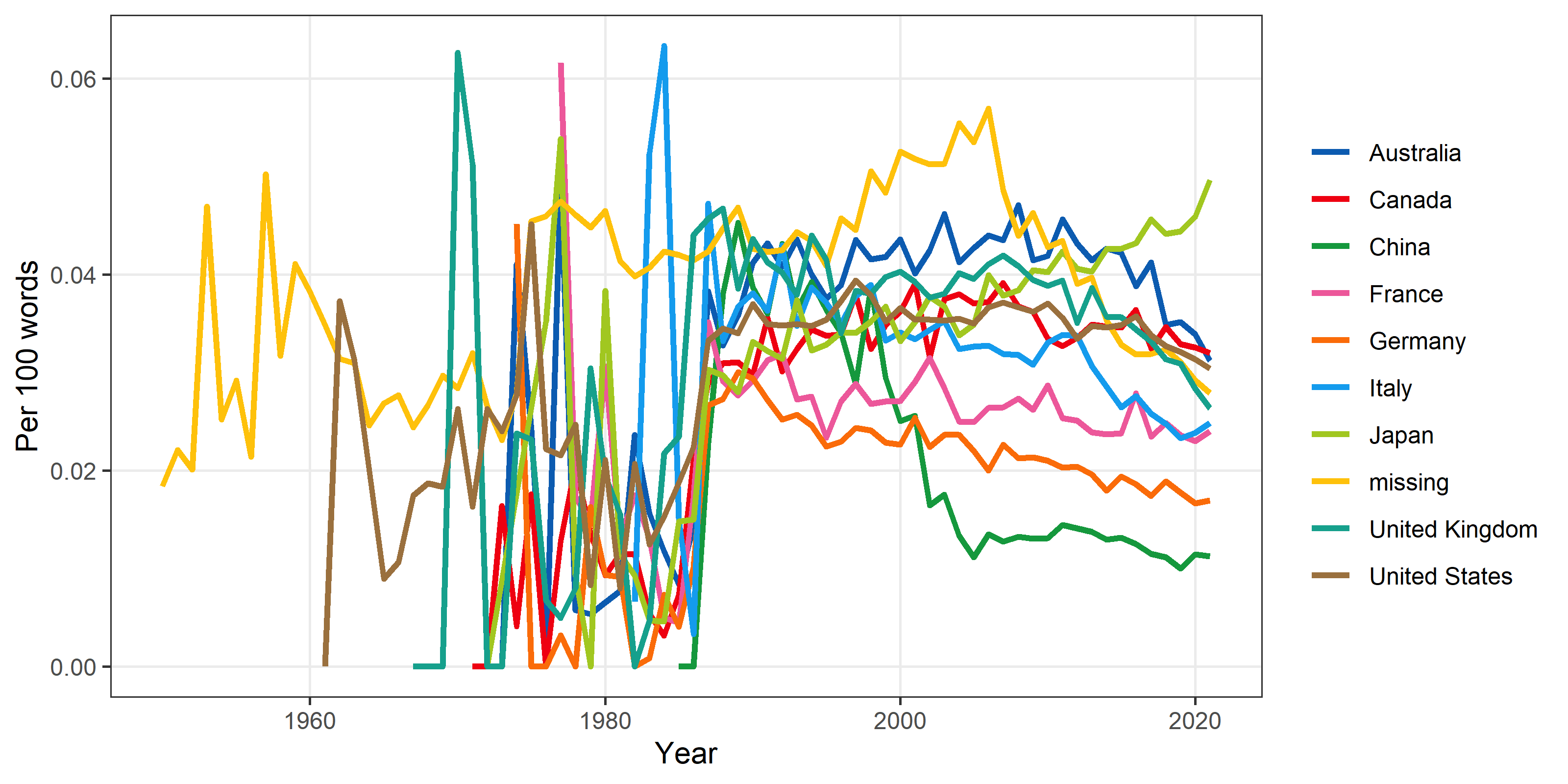
### Sex: Male



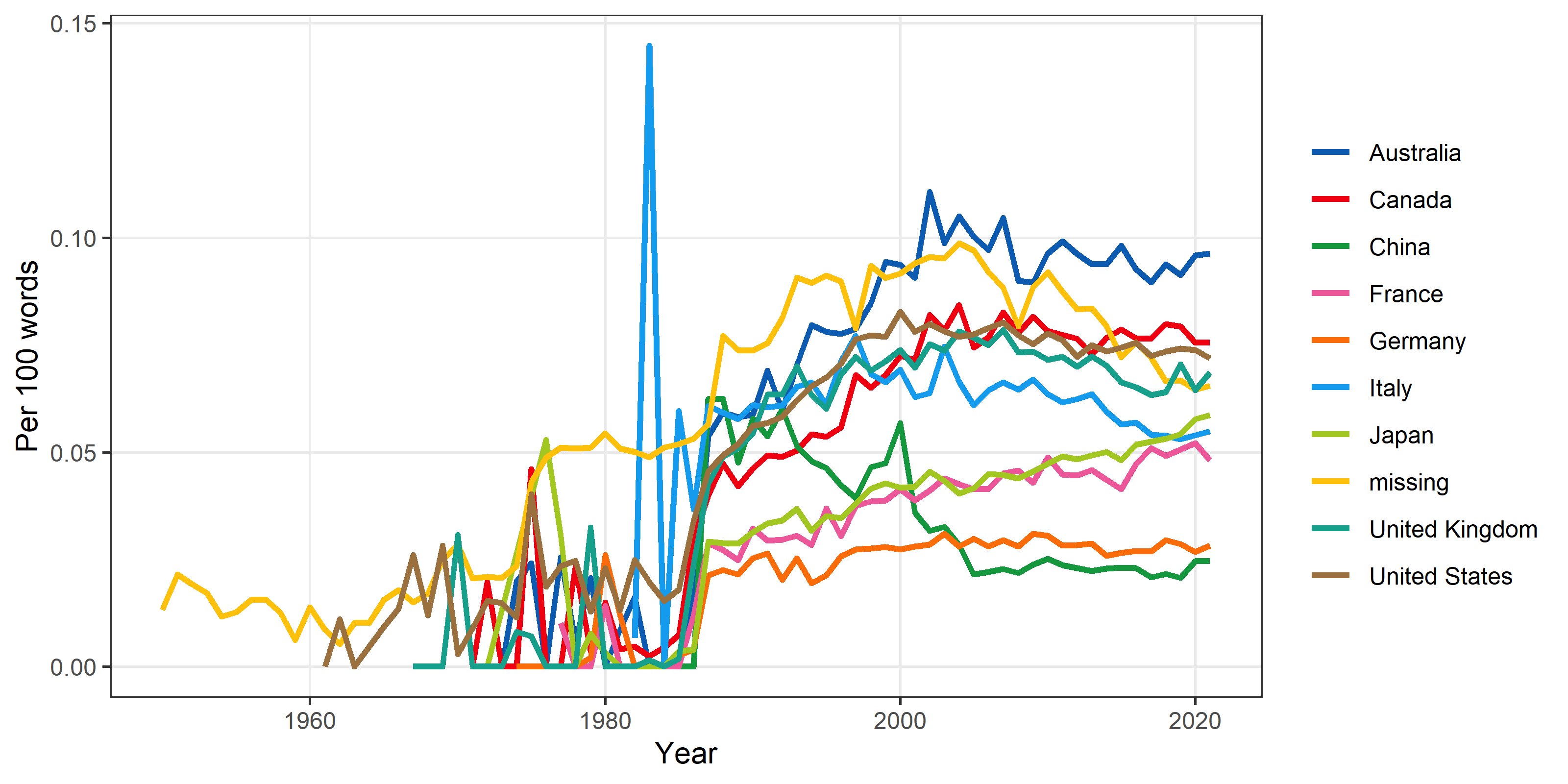
### Sex: Female



### Gender: Male



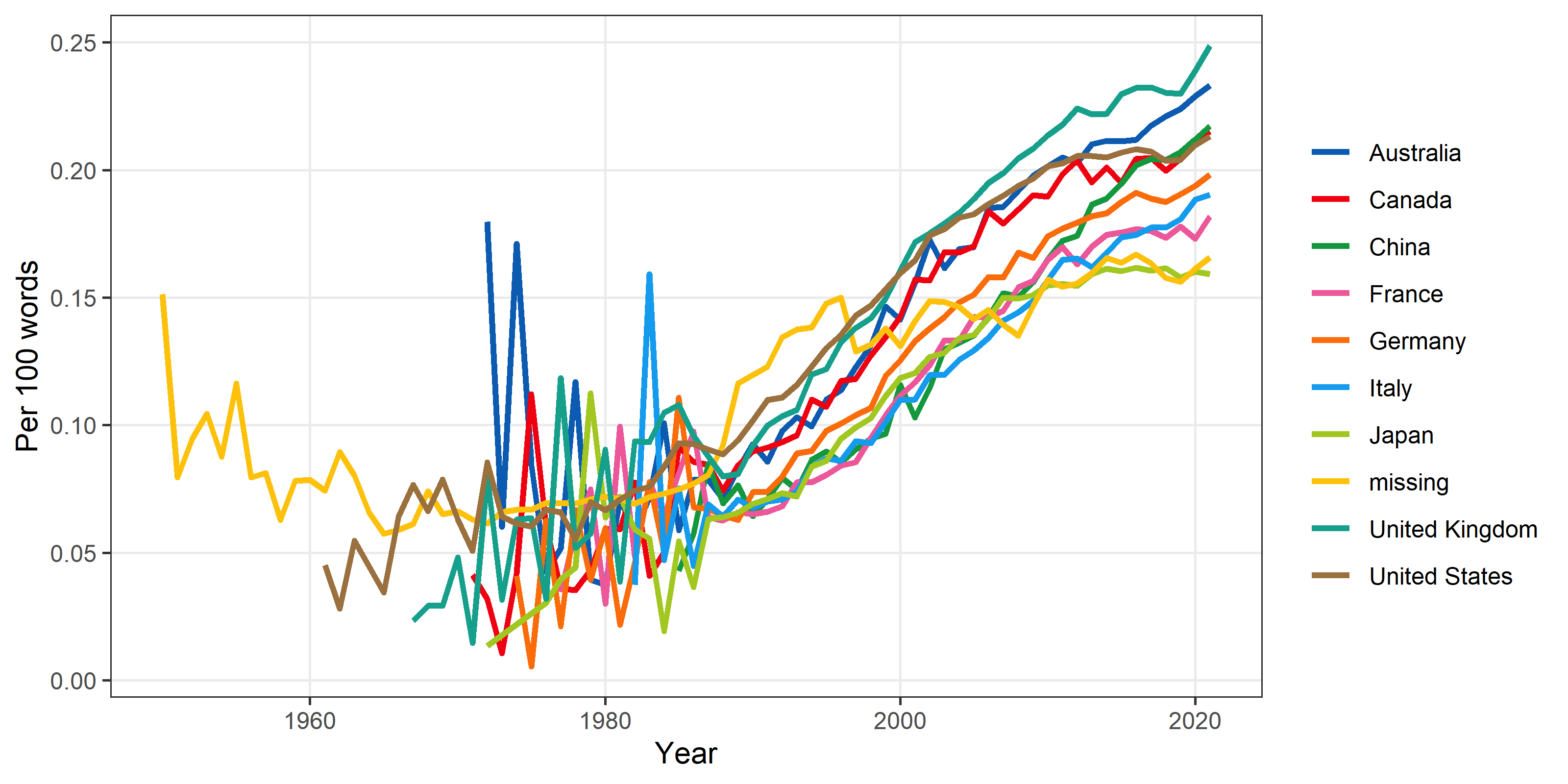
### Gender: Female



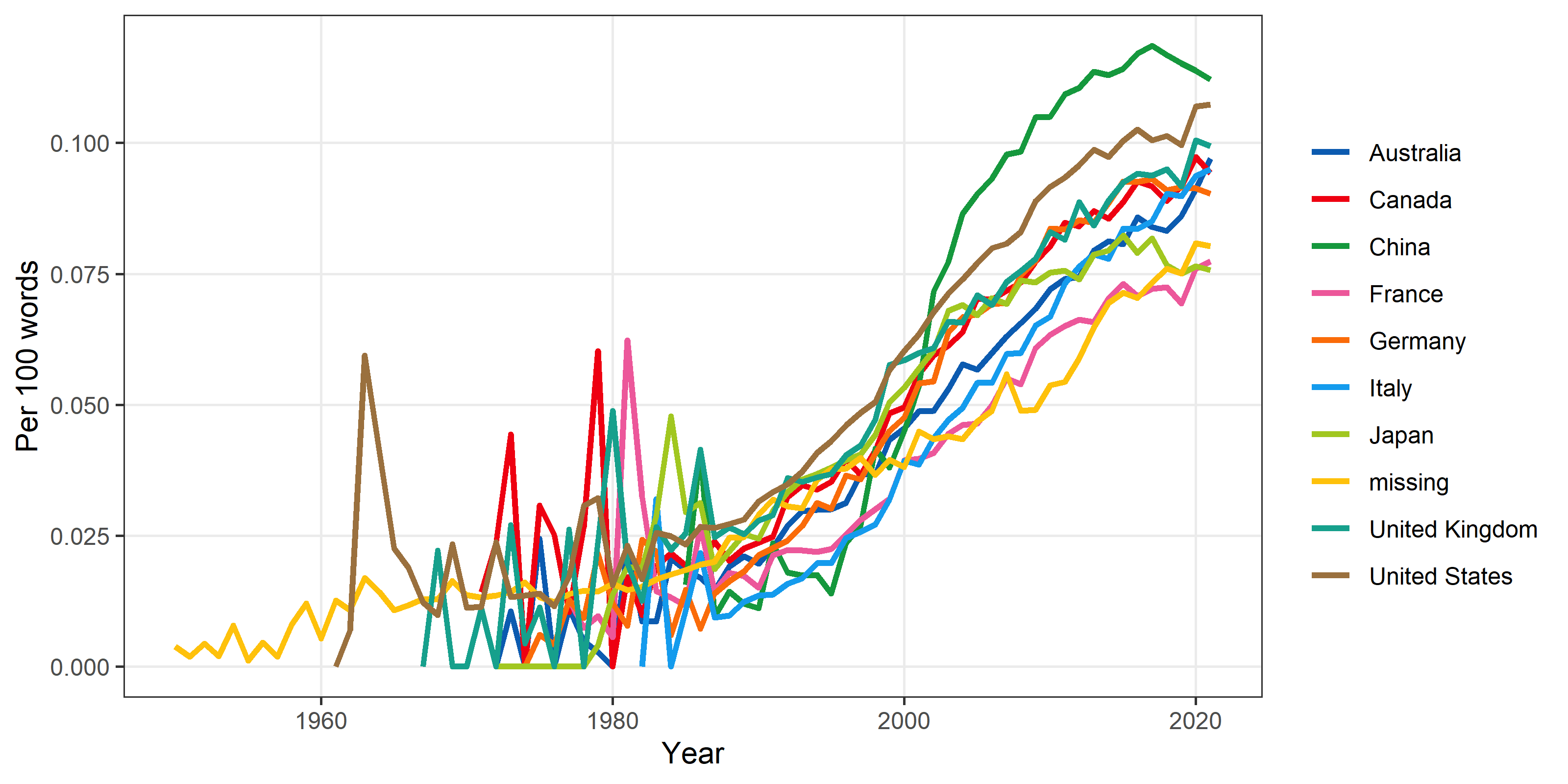
# Hype

Here we examine our definitions of hype.

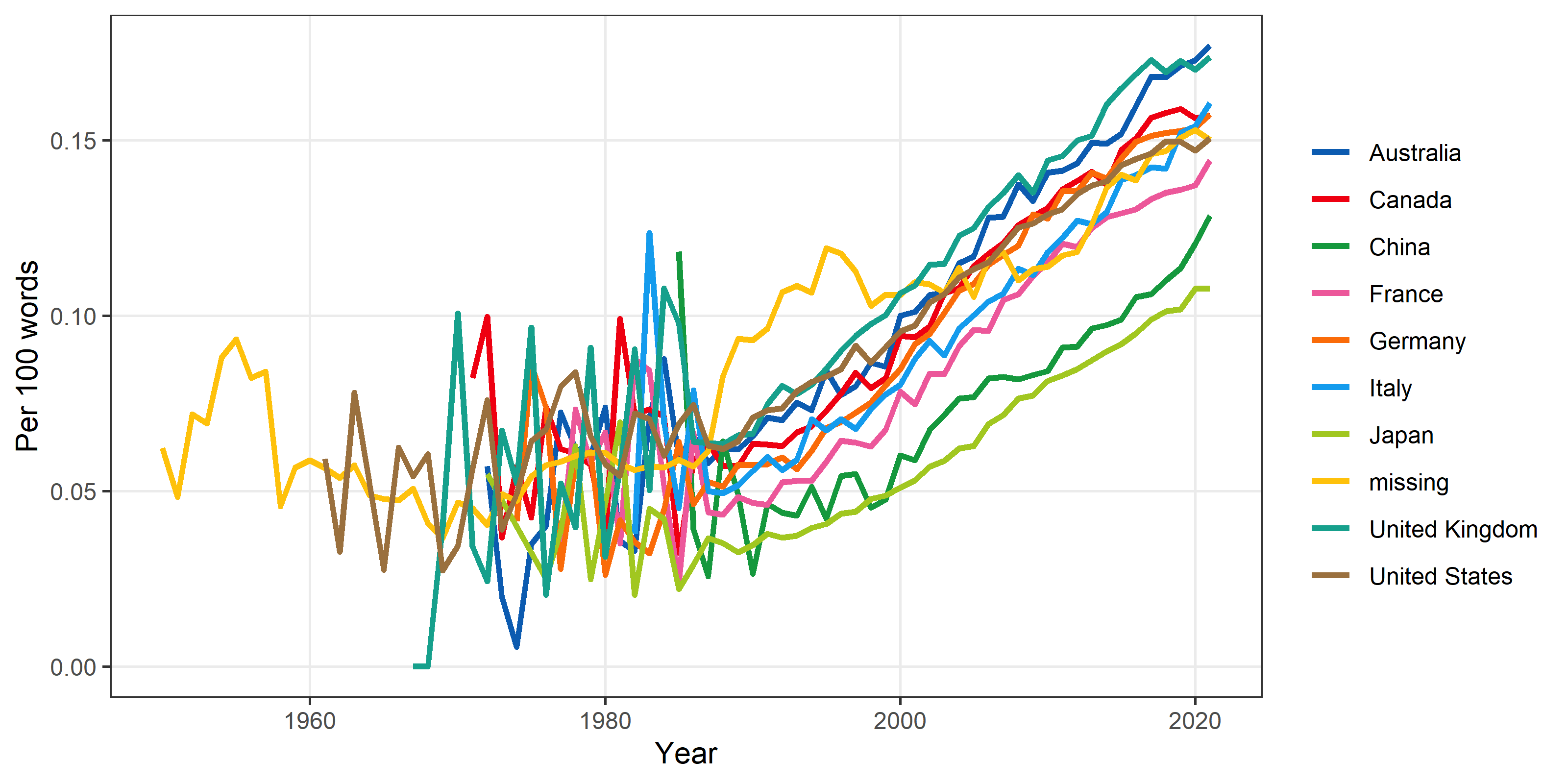
## Importance



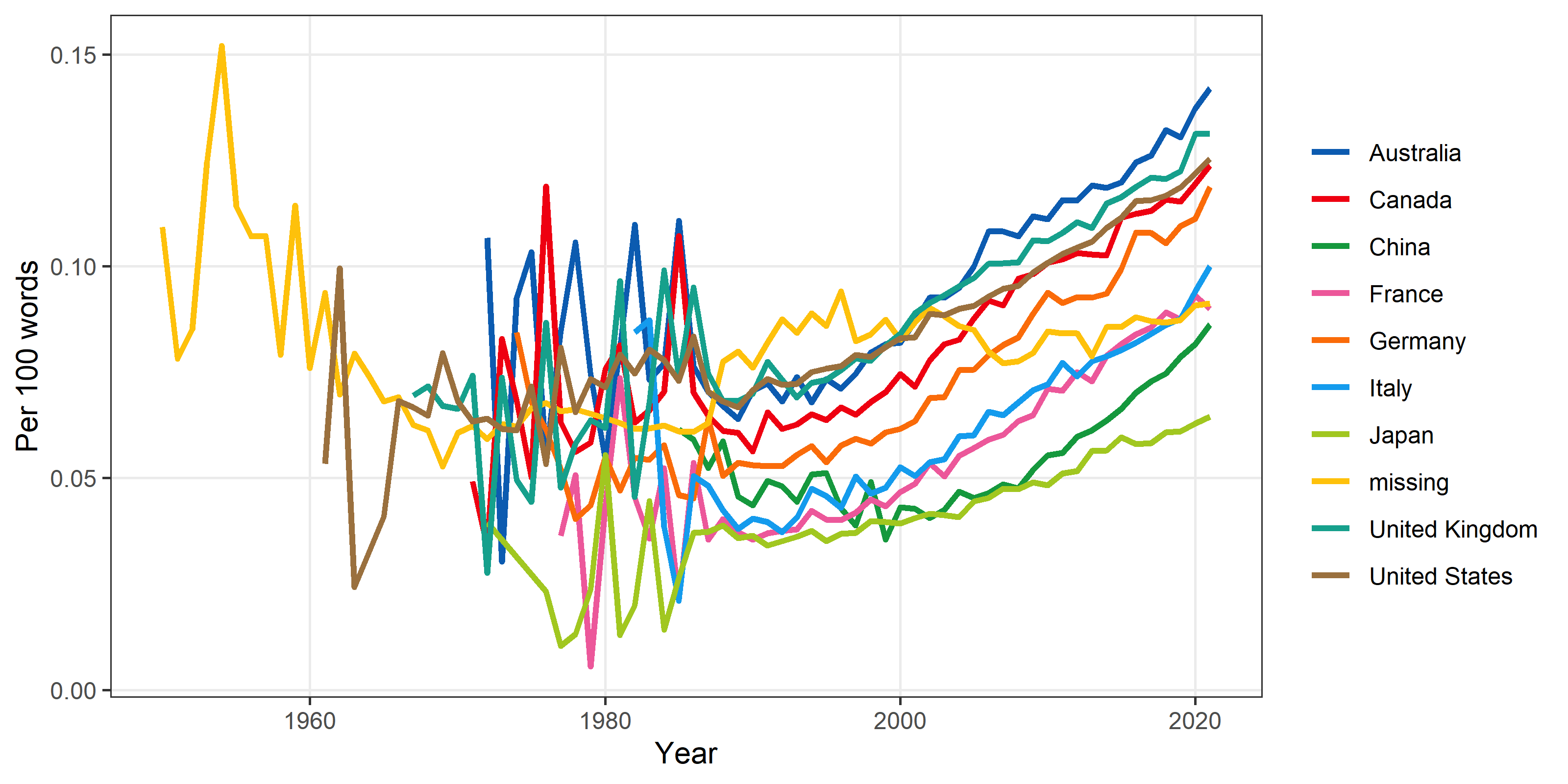
## Novelty



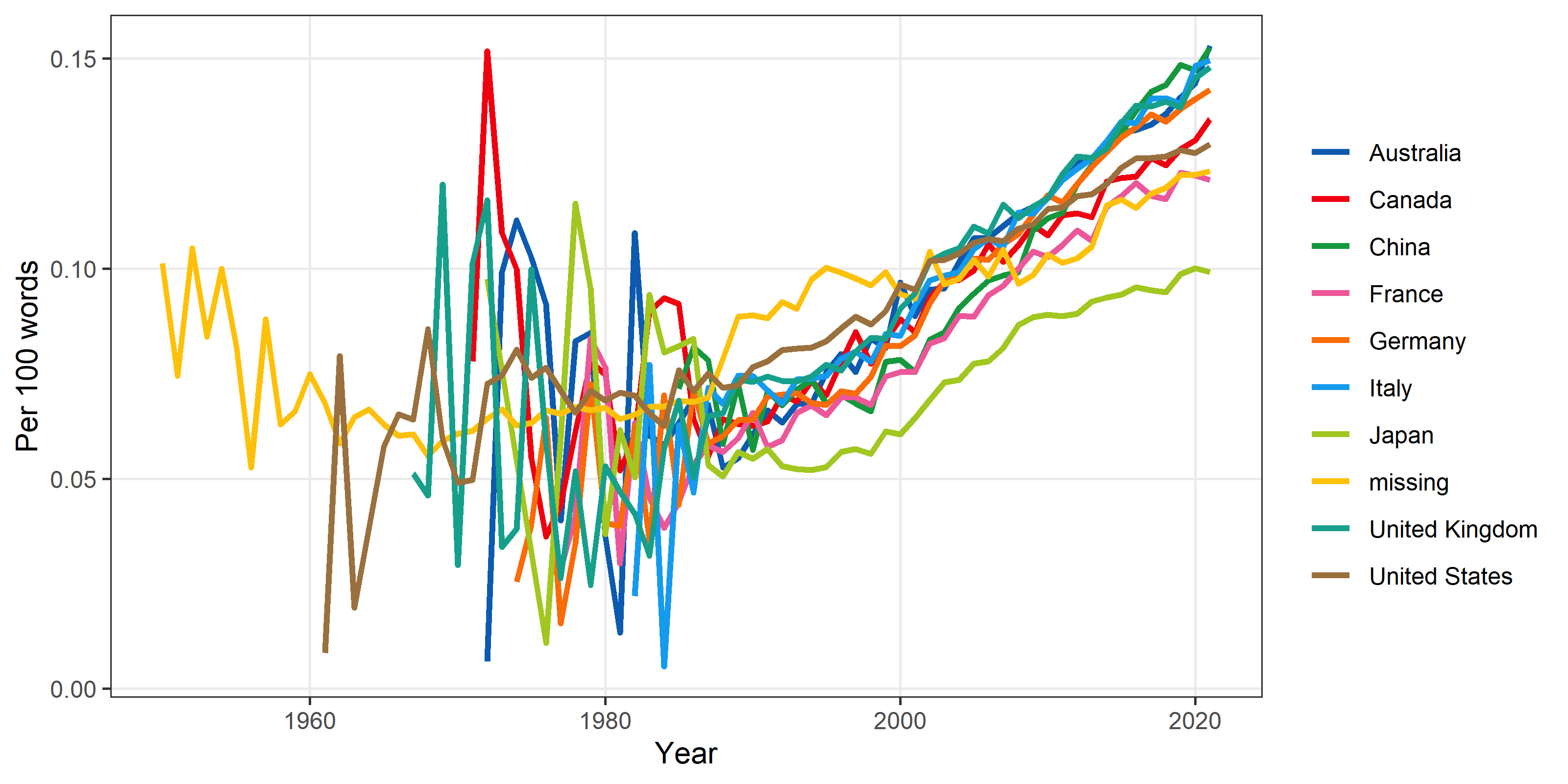
## Rigor



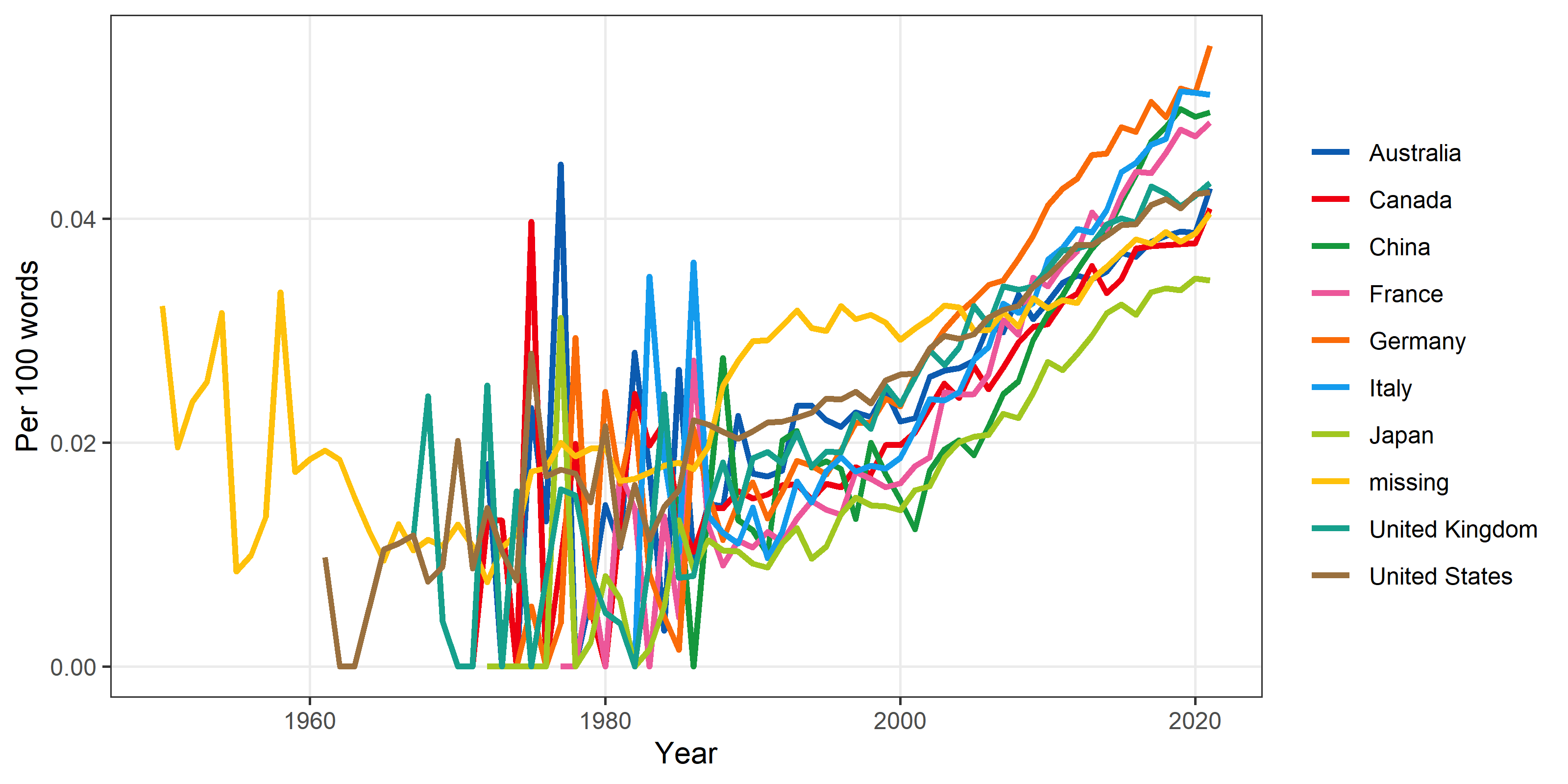
## Scale



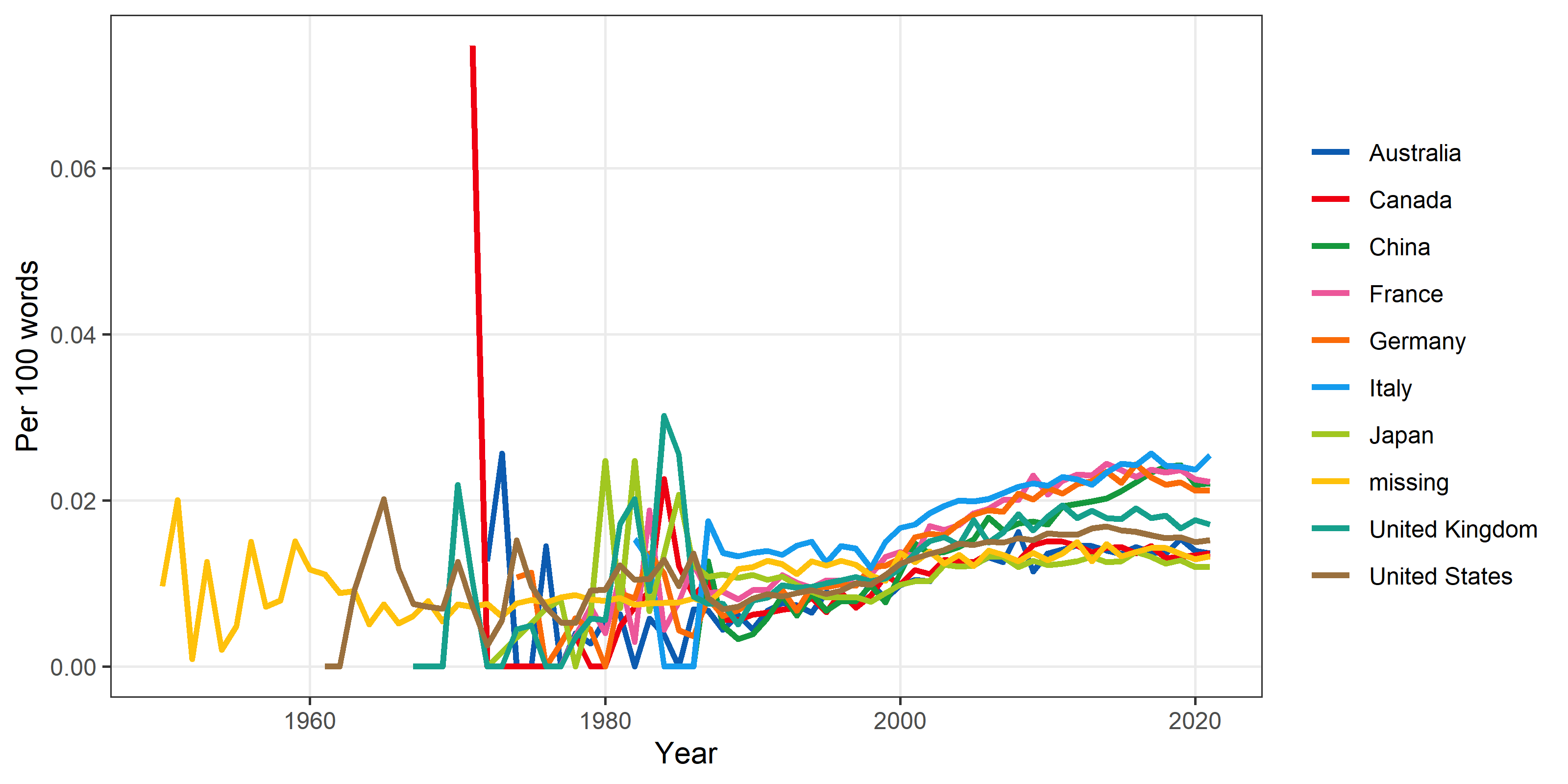
## Utility



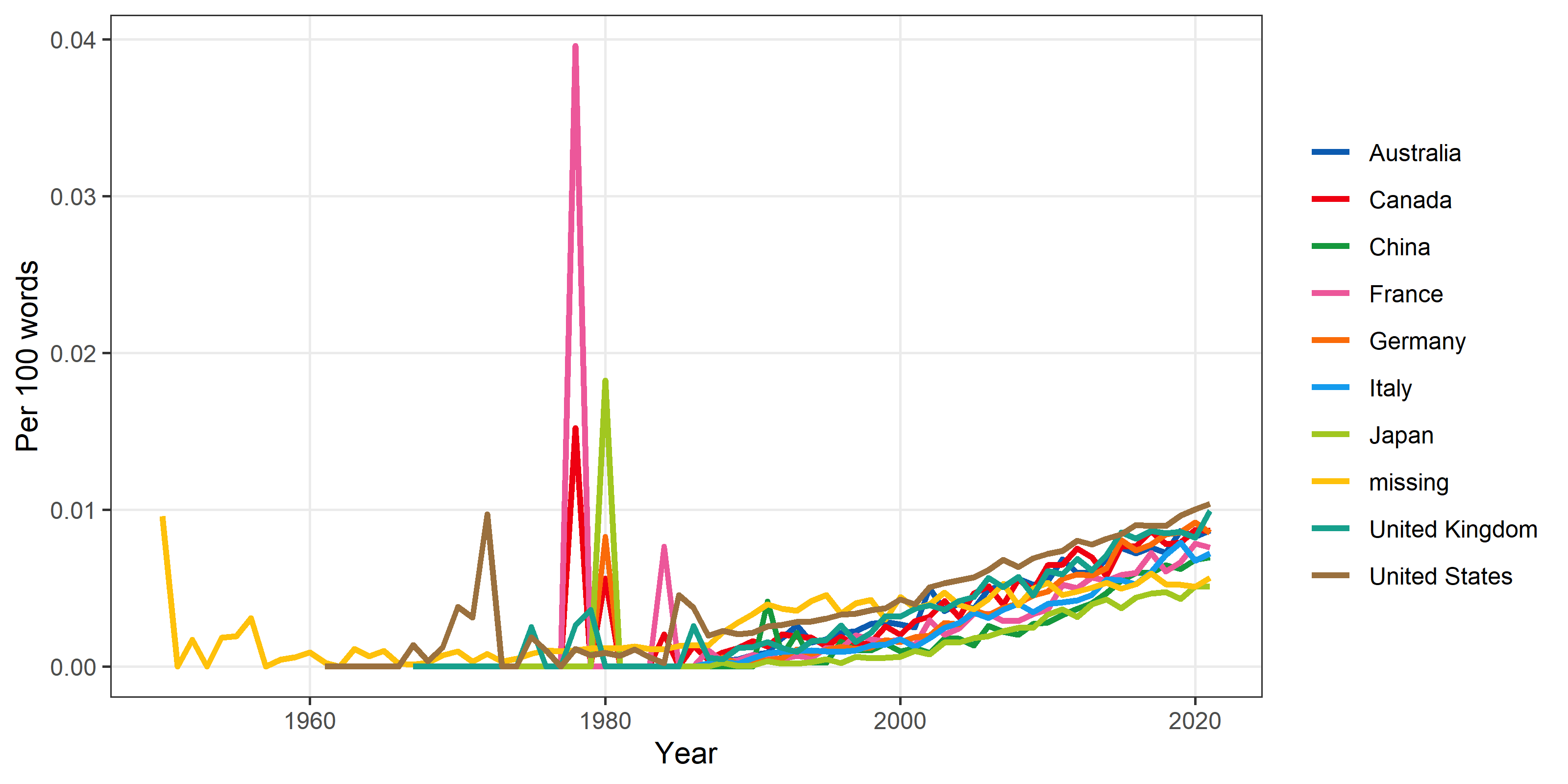
## Quality



## Attitude



## Problem



## Additional hype words

