Task 3: Text analysis (PELEGRI)

Patrick Pelegri-O'Day

3/6/2022

Overview

This report visualizes the most common words used by two senior UN officials at the press conference for the release of the most recent International Panel on Climate Change (IPCC) report: UN Secretary General Antonio Guterres and United Nations Environment Programme director Inger Andersen. Sentiment analysis of the speeches is also presented. The speeches were recorded live on YouTube, and the speech text was obtained from YouTube's transcription feature.

Data Citation: IPCC Press Conference - Climate Change 2022: Impacts, Adaptation, and Vulnerability. Recorded 2022-02-28. https://www.youtube.com/watch?v=JpK7eeYRhjQ

Citation for NRC lexicon: Crowdsourcing a Word-Emotion Association Lexicon, Saif Mohammad and Peter Turney, Computational Intelligence, 29 (3), 436-465, 2013.

Setup

Read in data

Initial wrangling

Remove stop words and numbers
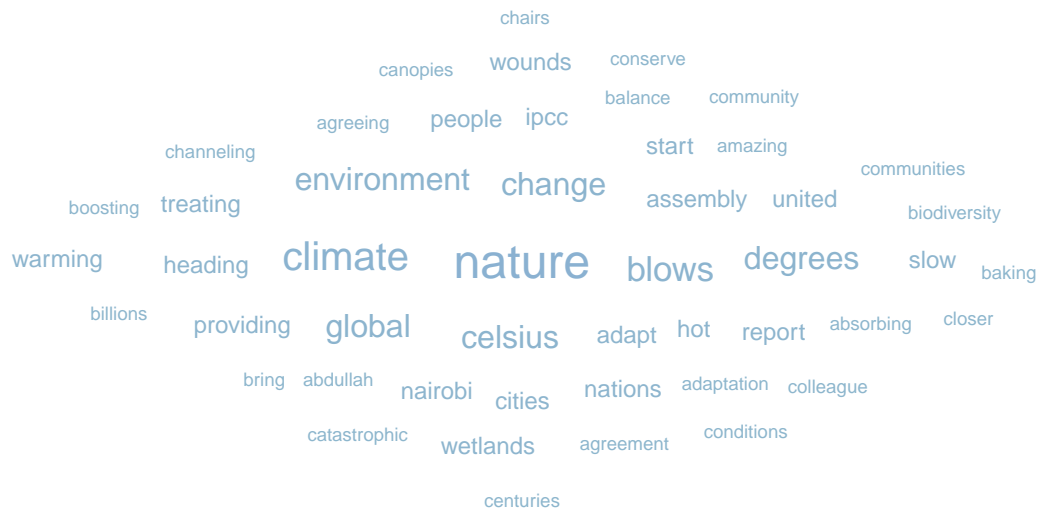
Count word occurrence

Create word clouds

Figure 1. 50 most frequently used words in Inger Andersen's speech for the press conference of Part 2 of the IPCC's 6th Climate Assessment.
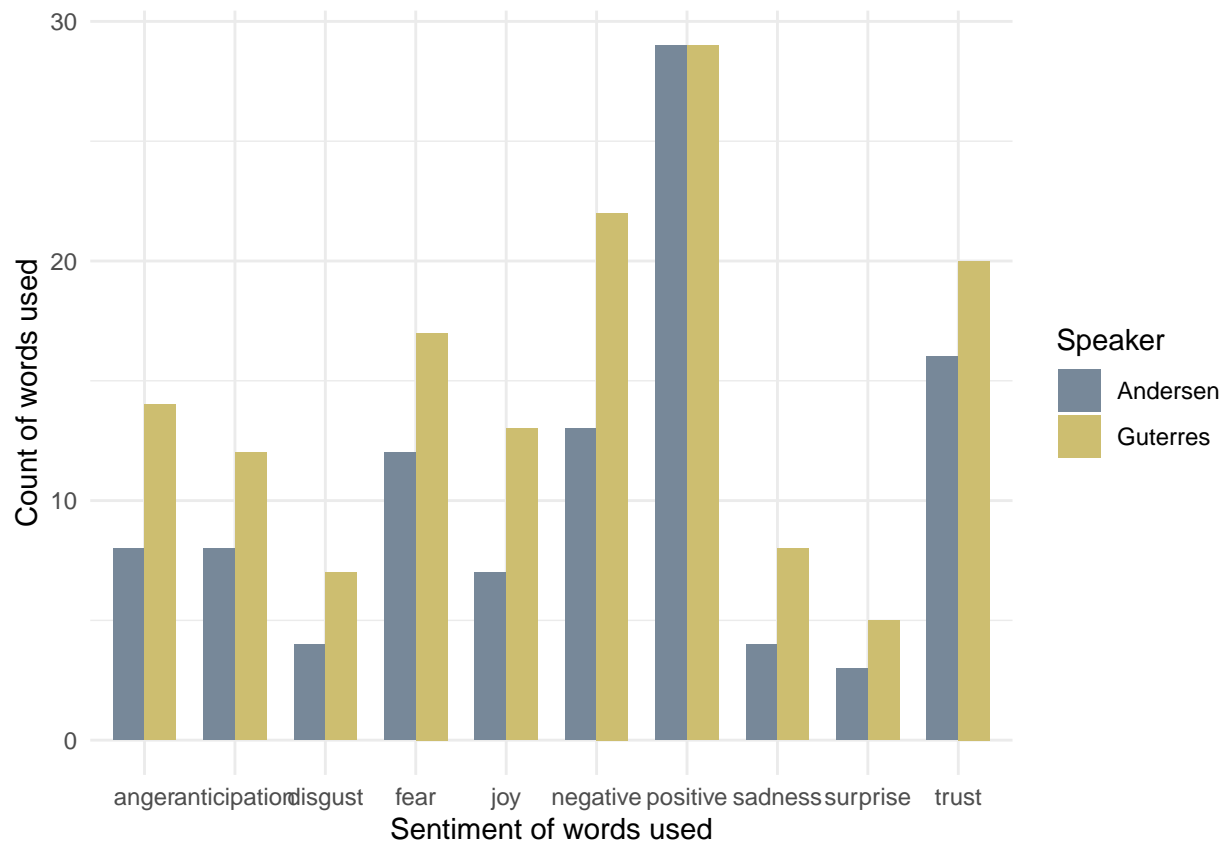
Figure 2. 50 most frequently used words in António Guterres's speech for the press conference of Part 2 of the IPCC's 6th Climate Assessment.

Takeaways from comparing most frequently used words

- Guterres's speech seems to be more focused on energy; some of the most frequently used words are "energy", "coal", "fuel", and "emissions."

- By contrast, Andersen's speech is more focused on nature, using words like "nature", "environment", "biodiversity", and "conserve."

- Guterres's speech also emphasizes large sociopolitical systems, while Andersen's speech focuses on personal, firsthand language. We see this through Guterres's discussion of "banks", "economy", "security", "finance", and "multilateral".

- Andersen, on the other hand, uses an extended metaphor of global warming as an assailant on humanity's unified body through words like "blows" and "wounds", and includes more human-scale, evocative, and softer language such as "communities", "agreeing", "channeling", "boosting," and "absorbing".

Sentiment analysis

Figure 3. Sentiment analysis of Andersen's and Guterres's speeches using NRC sentiment lexicon.

Takeaways from NRC sentiment analysis

- Guterres's speech scored higher on all measurements except positivity — both speeches were equally positive.
- Andersen's speech had a higher ratio of positive language to negative language
- Andersen's speech used more neutral language as measured by the NRC sentiment lexicon: Guterres's speech had more instances of language in almost every category, including positive sentiments such as joy and negative sentiments such as fear.
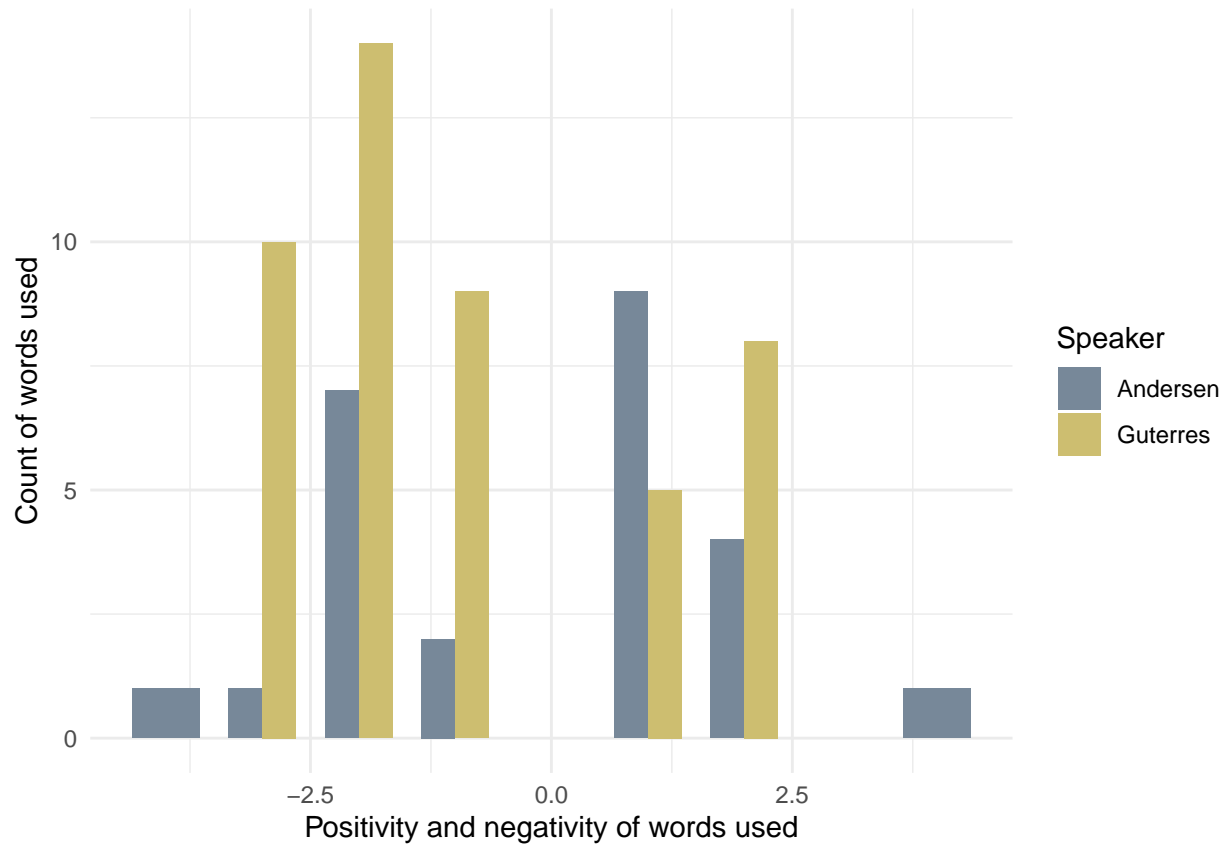
Setup for AFINN sentiment comparison

Figure 4. Sentiment analysis of Andersen's and Guterres's speeches using AFINN sentiment lexicon.

Takeaways from AFINN sentiment analysis

- Using AFINN lexicon sentiment analysis, we more clearly see the trend that Andersen's speech is more positive than Guterres's speech
- We see that Guterres's speech is significantly more negative than positive in language used. Andersen's speech appears slightly more positive than negative.