

Topics of all U.S. Inaugural Addresses from 1789-2021

- A Text Analysis with R (Studio) -

Motivation

The United States of America is one of the most powerful and influential countries in the world. It is led by the U.S. President who is the head of state, head of government, and Commander-in-Chief of the armed forces. [1] Analyzing the presidents' inaugural addresses allows drawing conclusions regarding the political, social, and economic development of the U.S. over the last centuries.

Research Questions

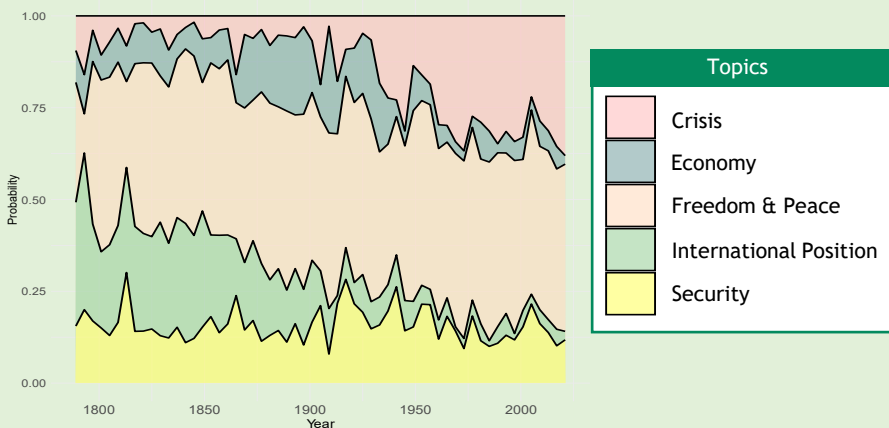
1. How do topical priorities change over time?
2. How does the framing of the terms *freedom* and *peace* change over the centuries?
3. Can thematic differences between the speeches of democratic presidents and those of republican presidents be identified?

Conclusion

- **Freedom & peace** has maintained its position as topic with the **highest relative distribution** throughout the years.
- **Its context has expanded to a worldwide focus**, while simultaneously remaining closely tied to the notion of **country/nation**.
- **Differences in word choice between the two political parties exist** and they are very prominent, indeed.

Topic Modelling (RQ 1)

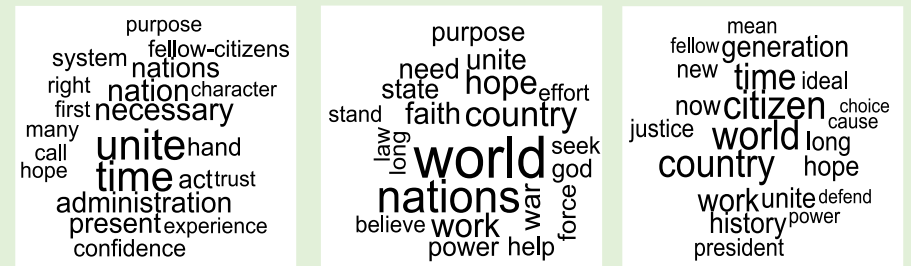
Method: Seeded LDA (Latent Dirichlet Allocation) Topic Modelling with own dictionary to ensure valid results. Identification of an optimal topic number k based on the comparison of divergence rates. Plotted in a stacked area chart.



Interpretation: While the percentage share of the topic **crisis** has generally increased over the decades, the topic **international position** has shrunk. No trend can be observed for **security**, **economy**, and **freedom & peace** which additionally don't fluctuate as much as the other topics. The latter proportionally forms the biggest topic over the majority of presidential turns.

Context Change over Time - Topic: Freedom & Peace (RQ 2)

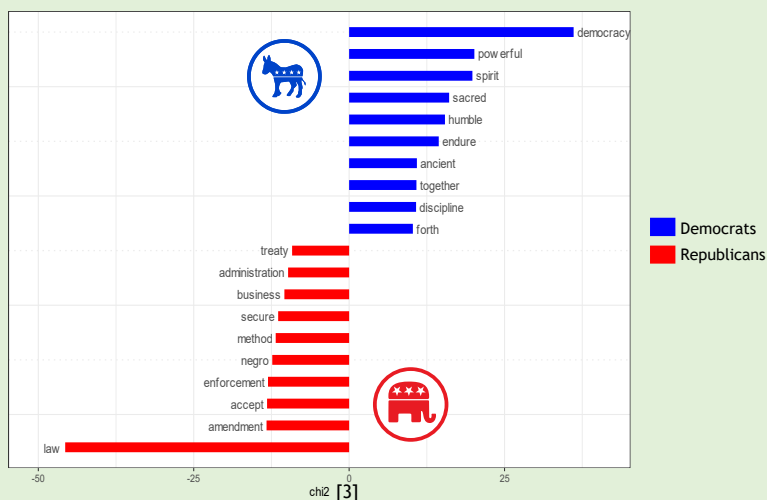
Method: Formation of one subset per century. Definition of a window of ten words as context for instances which fit into the topic of *freedom & peace*. Usage of a Feature Co-Occurrence Matrix to quantify common appearance rates of words. Display in word clouds. The bigger the word in the cloud, the higher its occurrence and relevance. [2]



Interpretation: The need for peace was expressed especially in the 19th and 20th century. *Freedom & peace* had its greatest international focus in the 20th century (here also strongly linked to the topic of war). The general focus shifted from *nation* to *world*, indicating international interests and relations. Yet, *country* is a persisting constant.

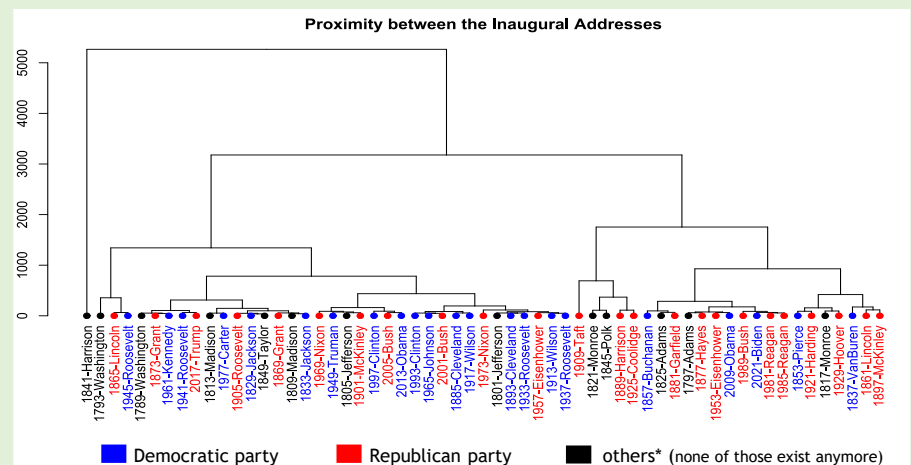
Party Differences (RQ 3)

Method: Juxtaposing democratic and republican addresses using a preprocessed corpus subset which consist of those 46 speeches only (22x democratic, 24x republican). Creation of a keyness-statistic.



Interpretation: **Democrats** place greater emphasis on transmitting values (such as **democracy**, **discipline**) and emotionally loaded expressions (like **humble**, **together**). Compared to this, **Republicans** incorporate much more impersonal, factual speech, as well as administrative terms (i.e., **administration**, **amendment**).

Additionally: Comparison of all 59 presidential speeches on the word level. Creation of a distance matrix and depiction of hierarchical clustering in a dendrogram.



Interpretation: Proximity between speeches of the same party is only slightly recognizable. It should be noted, however, that there is no trend visible in the process of time concerning the allocation of presidents to groups. More recent as well as earlier presidents occasionally have similar speeches which might indicate that one president drew inspiration from another.

Annotations

- [1] cf. whitehouse.gov
- [2] The data preprocessing is based on lemmatization. Exclusion of the tokens *nation* and *nations* was done purposely due to a difference in meaning between the singular and plural form. The word clouds are based on the tokens 'equal', 'free', 'freedom', 'liberty', 'peace', and 'peaceful'
- [3] The chi2 value: A number which indicates to which degree observations deviate from the expected results. It is positive "if the observed value in the target [the upper group] exceeds its expected value". (quanteda.io)

Data Basis

The quanteda package 'data_corpus_inaugural' which consists of 59 presidential inaugural addresses of the last 46 U.S. presidents (1789-2021)

Lara Backer & Luisa Reis
Seminar *Introduction to Computer-Based Text Analysis*
Summer Term 2023
Prof. Mirco Schönfeld

