Charles University

Faculty of Social Sciences
Institute of Economic Studies

MASTER'S THESIS

Reflection of the Czechoslovak Monetary Policy in the thoughts and writings of German-speaking Economist that worked at the Charles-Ferdinand University during 1918 – 1938

Author: **Mgr. Hana Zatřepálková**

Study program: **Economics and Finance**

Supervisor: **PhDr. Jaromír Baxa, Ph.D.**

Academic Year: **2023/2024**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently; using only the listed resources and literature, and the thesis has not been used to obtain a different or the same degree.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, January 3, 2024 Hana Zatřepálková

Acknowledgments

The author is grateful especially to the thesis supervisor, PhDr. Jaromír Baxa, Ph.D., for not only providing insightful guidance but also for his patience, personal understanding, and invaluable motivation and encouragement. Additionally, the author extends heartfelt thanks to her family for their unwavering support.

Abstract

The thesis focuses on German-speaking economists, specifically Oskar Engländer, Alfred Amonn, and Hugo Müller, who taught economically oriented subjects at the German University in Prague during the interwar period. Through a combination of historical research and textual analysis of their works, the thesis explores the extent to which these economists expressed interest in the monetary policy of Czechoslovakia and its notable figures, Alois Rašín and Karel Engliš. Additionally, it investigates the predominant themes in their works during this period and provides insights into the potential applications of computer-based textual analysis in the field of the history of economic thought.

JEL Classification	B31, B49, E52, N01, N14,
Keywords	Alfred Amonn, Hugo Müller, Oskar Engländer, Czechoslovakia, 1918 – 1938, Karel Engliš, Alois Rašín, monetary policy, text minig analysis
Title	Reflection of the Czechoslovak Monetary Policy in the thoughts and writings of German-speaking Economist that worked at the Charles-Ferdinand University during 1918 – 1938

Abstrakt

Práce se zaměřuje na německy mluvící ekonomy, konkrétně Oskara Engländra, Alfreda Amonna a Huga Müllera, kteří vyučovali ekonomicky orientované předměty na Německé univerzitě v Praze v meziválečném období. Skrze kombinaci historického výzkumu a textové analýzy děl těchto ekonomů se práce zabývá otázkou, do jaké míry projevovali zájem o měnovou politiku Československa a její významné představitele, Aloise Rašína a Karla Engliše. Zároveň zkoumá, jaká témata převažovala v jejich pracích během tohoto období, a poskytuje náhled na možnosti využití počítačové textové analýzy v oblasti historie ekonomického myšlení.

Klasifikace	F12, F21, F23, H25, H71, H87
Klíčová slova	Alfred Amonn, Hugo Müller, Oskar Engländer, Československo 1918 - 1938, Karel Engliš, Alois Rašín, měnová politika, text miningová analýza
Název práce	Reflexe československé měnové politiky v díle a myšlenkách německy mluvících ekonomů vyučujících na Univerzitě Karlo-Ferdinandově v Praze v letech 1918 – 1938

Contents

List of Tables

List of Figures

Acronyms

AUK Archive of Charles Univerzity [Archiv Univerzity Karlovy]

CAS Czech Academy of Sciences [Akademie věd České republiky]

DTM Document-term matrix

Idf Inverse document frequency

LDA Latent Dirichlet Allocation

NACH National Archive Chodovec [Národní archive Chodovec]

OCR Optical Character Recognition

STM Structural Topic Models

Tf Term frequency

Master's Thesis Proposal

Author:	Mgr. Hana Zatřepálková
Supervisor:	PhDr. Jaromír Baxa, Ph.D.
Defense Planned:	January 2024

Proposed Topic:
Reflection of the Czechoslovak Monetary Policy in the thoughts and writings of German-speaking Economist that worked at the Charles-Ferdinand University during 1918 – 1938

Motivtion:
The first half of the twentieth century belongs to the important periods of modern economic history. The First World War, the collapse of the Central Powers, emergence of the successor states and their need to accommodate to the new economic situation, the Great Depression and finally the economy of World War II – that all contributed significantly to the development of economic thinking. It is the period of formation economic theories that influenced next generations of economists.

Similar holds also for the Czech economic area although there are few specifics that should be taken into consideration. After the disintegration of Austria-Hungary, significant minority of German speaking economists stayed and worked there. The thoughts and work of this minority have not been paid enough attention yet. The thesis aims to fill this gap and it focuses particularly on the reflection of economic problems of Czechoslovakia in the work of German speaking economists.

Hypotheses:
1. What was the attitude of German speaking economists to the monetary politics of Czechoslovak First Republic? (H1: The particular economist did pay attention to the Czech monetary politics and exerted positive/negative attitude towards it.)
2. Can we distinguish between “Rašín’s” and “Engliš” conceptions in the thoughts of German speaking economists? (H2: The particular economist took interest in the person of Rašín/Engliš and evaluated their contributions positively/negatively.)
3. What is the reason for their attitudes and could this influence their lecturing? (H3: The particular economist can be associated with a particular school of economic thought which shaped his attitudes.)

Methodology:
The research questions should be answered via detailed analysis of the texts written by German speaking economist in the period 1918 – 1938. The analysis

should contain both – qualitative and quantitave parts. In the first stage these texts are to be scanned and digitalised via open source technology (such as Tesseract Open Source OCR Engine). The main tool of the quantitative analysis should be text mining techniques (such as cluster or sentiment analysis) in R-software. The qualitative analysis should compare 3 texts written by German speaking economists concerning Czech monetary politics and explain them in the context of existing theories and contemporary documents.

Expected Contribution:
As already mentioned, there is no literature that would be interested in the contribution of German speaking economists to the overall Czech economic thinking. The thesis want to fill this gap and to open this topic and so leave a space for further research..

Outline:
1. Introduction
2. Monetary politics of Czechoslovak republic in 1918 – 1938
3. Tradition of German speaking economists at the Charles-Ferdinand University
4. Three Comments on the Monetary Policy of the Czechoslovak Government - Comparative analysis
5. The schools of thoughts of German speaking economists – Text mining analysis
6. Conclusion

Core Bibliography:

1. ENGLÄNDER, Oskar. Die Devalvation der tschechoslowakischen Krone im Jahre 1934 . Jena: G. Fischer, 1936
2. ENGLÄNDER, Oskar. Wechselkurse, internationaler Handel, Währungs-politik. Brünn: Rudolf M. Rohrer, 1933
3. AMONN, Alfred. Währungsreform in der Tschechoslovakei und in Sowjet-Russland. München: Duncker & Humblot, 1924
4. KRAUS, Oskar. Die Werttheorien : Geschichte und Kritik. Wien: Rudolf M. Rohrer, 1937
5. MÜLLER, Hugo. Zwei tschechische Schriften über Währungsreform. Schmollers Jahrbuch für Gesetzgebung, Verwaltung und Volkswirtschaft. 1926, (50,2.1926). ISSN 0036-6234
6. SCHRANIL, Rudolf. Gutachten zur tschechoslowakischen Verwaltungsreform. Praha: Prager Juristischen Zeitschrift, 1927
7. SPITALER, Armin Egon Hilbert. Die Ursachen der Weltwirtschaftskrise und die Möglichkeiten ihrer Milderung. Reichenberg: Gebrüder Stiepel Gesellschaft, 1931
8. HOLMAN, Robert. Dějiny ekonomického myšlení. 2.vydání. Praha: C. H. Beck, 2001. ISBN 80-7179-631-X
9. BARTH, Boris (ed.), Josef FALTUS (ed.), Jan KŘEN (ed.) a Eduard KUBŮ (ed.). Konkurence i partnerství: Německé a československé hospodářství v letech 1918 - 1945. Praha: Karolinum, 1999. ISBN 80-7184-708-9
10. HOLT, Robert T. (ed.) a John E. TURNER (ed.). The Methodology of Comparative Research. New York: The Free Press, 1970. ISBN Není.
11. MAHONEY, James (ed.) and Dietrich RUESCHEMEYER (ed.). Comparative Historical Analysis in the Social Science. New York: Cambridge University Press, 2003. ISBN 0 521 01645 2.
12. SILGE, Julia and David ROBINSON. Text Mining with R: A Tidy Approach. Sebastopol: O'Reilly, 2017. ISBN 978-1491981658

1 Introduction

The first half of the twentieth century with its two World Wars and the Great Depression belongs to the most turbulent eras in modern history that dramatically changed the power structures of the world. Related to this was the need for accommodating the everchanging economic situation that contributed significantly to the development of economic thinking. This era can be therefore seen as a period of the formation of economic theories that influenced the next generations of economists. In the Central European context, the collapse of the Central powers after the Great War led directly to the emergence of the successor states and so enabled the establishment of an independent Czechoslovakia. After centuries of perceived oppression, Czechs and Slovaks were free to decide their fate in their national state but at the same time, a significant German-speaking minority[1] with its experts and eminent persons (which also included economists and professors) staying and working in Czechoslovakia experienced this situation from a very different point of view.

This thesis closely examines a specific group of German-speaking economists who were affiliated with the German University in Prague between 1918 and 1938. Given the limited attention to the ideas and work of this group, the thesis aims to address this gap by focusing on their reflections on Czechoslovak monetary policy. The study also explores the theme of isolation versus engagement within the German-speaking community concerning economic events in interwar Czechoslovakia. The main research questions revolve around the attitudes of German-speaking economists toward Czechoslovak monetary policy, their opinions on Alois Rašín and Karel Engliš (the two most influential Czechoslovak finance ministers with differing views on the role of currency and monetary policy for economic prosperity), and the identification of general topics and patterns within a particular school of economic thought present in their writings.

From a methodological approaches, this thesis integrates both qualitative and quantitative methodological principles. It involves in-depth historical research, the collection of literature, articles, and archival materials. Additionally, modern technologies, including OCR software and methods of text mining analysis (such as word frequency analysis, sentiment analysis, or topic modeling), are employed. This multidisciplinary approach aligns with the growing significance of such projects in modern science. Furthermore, the thesis explores the potential application of

[1] As described later, this so-called minority was in fact second largest and accounted for ca. 30 % of the total population of Czechoslovakia.

computer-based textual analysis in studying the history of economic thought, delving into the aspect of human auditing of computer-based models.

The thesis encompasses both archival work and a review of historical and contemporary literature, along with an exploration of text-mining analysis methods. First, it delves into the development of interwar Czechoslovak monetary policy, placing particular emphasis on the roles of the two most influential finance ministers, Alois Rašín and Karel Engliš. This part of the thesis draws extensively from specialized literature, with noteworthy contributions from Doležalová (2007), who examines interwar Czechoslovak fiscal policy in relation to prominent figures, and Kubů and Pátek (2000), who explore the alignment of economic progress in interwar Czechoslovakia with reality. Then, the evolution of the teaching of economic subjects at the University in Prague[2] is presented and short biographies of selected German-speaking economists who taught here are offered. This section relies on publications addressing the history of Charles University (Bečvář 1998; Havránek 1997) and, with regard to German-speaking economists, it is based on archival materials from the Archives of Charles University and the National Archives. As far as the author is aware, this thesis stands as the first publication to present comprehensively these German-speaking economists, as existing anthologies, such as Doležalová (2018), Vencovský (1997), Sojka (2002) or Yagi (2013), typically exclude those working in Prague during the interwar period, both in the Czech, Austrian, and German contexts. Broadly speaking, a prevailing perspective has traditionally leaned toward a national (nationalistic) approach when categorizing individuals among Czech economists, rather than adopting a regional (patriotic) viewpoint. Third, the thesis seeks to uncover to what extent the German-speaking economists reflected the developments of the Czechoslovak economy and of their colleagues writing dominantly in the Czech language to analyze potential links between these two communities. Therefore, we compare three specific texts authored by the selected German-speaking economists, which directly address the monetary policy of the Czechoslovak government in the interwar period. These texts include "Die tschechoslowakische Währung und Währungspolitik" (Amonn 1924), "Die Devaluation der tschechoslowakischen Krone

[2] When referring to the University in Prague in this text, it is essential to acknowledge that the name and structure of this institution underwent several changes over the years. The Charles-Ferdinand University originated in the aftermath of the Thirty Years' War, resulting from the merger of Charles University and the Jesuit College. In 1882, the university underwent a division into Czech and German sections, maintaining this configuration until 1921. At that juncture, the German segment officially became an independent entity known as the German University in Prague, where German-speaking economists continued to teach. Concurrently, the Czech segment reverted to its original name, Charles University. Throughout this discussion, unless explicitly stated otherwise, the focus is directed towards the German branch of these institutions.

im Jahre 1934" (Engländer 1936), and "Zwei tschechische Schriften über Währungsreform" (Müller 1926).

The quantitative part of this thesis explores the potential applications of text-mining approaches such as word frequency analysis, sentiment analysis, and topic modeling (Silge and Robinson, 2022) in the field of the history of economic thought. The results indicate that the text mining analysis aligns with the reading experience, albeit with certain techniques, like topic modeling, showing less efficiency. The amalgamation of approaches affirms the German economists' interest in Czechoslovak monetary policy. However, their attention to prominent Czechoslovak economists in their writings appears relatively limited. It's worth noting that additional materials reveal that, despite seeming indifference in their works, these economists were not entirely disinterested in developments of the Czechoslovak economy and the work of Czech-speaking economists and it can be asserted that, aside from a few reservations, they generally held a positive perception of their Czechoslovak colleagues. Therefore, the conclusion can be made that the German-speaking economists were not isolated from the developments in Czechoslovakia and their work constitute a valid contribution to the history of economic thought in the Czech Republic.

This thesis is structured as follows. Chapters 2 and 3 provide the historical framework and biographies of selected German-speaking economists working in Prague. Chapter 4 analyses their three texts devoted on the developments of the Czechoslovak economic policy. Chapters 5 to 8 present the text-mining analysis. Initially, Chapter 5 provides a description of how the texts were obtained and subsequently digitized for further text-mining analysis, accompanied by a comprehensive table presenting their basic characteristics. Chapter 6 describes the methods used for the analysis, and Chapters 7 and 8 presents the results. Short conclusions close the thesis.

2 Monetary Policy of the Czechoslovak Republic in 1918 – 1938

The collapse of the Austro-Hungarian Empire into the successor states had a great impact on their economic issues. Newly constituted Czechoslovakia inherited from the monarchy many problems among which there were for example vast differences between its western and eastern parts. Whereas the Czech lands accommodated almost 90 % of industrial capacity, Slovakia and Carpathian Ruthenia belonged to the rural areas and the economic coordination of both parts of the republic became extremely complicated, partly thanks to the insufficient transportation network. Additionally, during the first months after the Great War, industrial production was devasted and the lack of basic amenities such as food or clothes caused social unrest and riots. Therefore the quick economic stabilization and establishment of an integrated national economic system was necessary to the further development of an independent democracy in Czechoslovakia. The key aspects of such processes were primarily the establishment of own currency, nostrification of shares, and land reform (See Novotný & Šouša 1996, pg. 80 – 81).

Fiscal and monetary policy are the two main tools governments use to influence economic activity in a country (Novotný & Šouša 1996, pg. 24). The independent Czechoslovak Crown achieved stability within three years after the end of the Great War, which can be considered a great success, especially in comparison to the neighboring states that did not manage to do so up until 2 – 4 years later (Kubů & Pátek 2000, pg. 12). Monetary and fiscal policy was the responsibility of the Ministry of Finance until the mid-1920s. On 1 April 1926, however, the National Bank of Czechoslovakia began its activities and the management of the currency was entrusted to it. Instead of the so-called state notes, banknotes began to be issued. The Czechoslovak Crown maintained its stability, when about 3 dollars were exchangeable for 100 Czechoslovak Crowns, until 1934. The Great Depression, which hit Czechoslovakia with a slight delay, affected its exports very negatively, and as a result, the first and second devaluation of the Czechoslovak Crown were approved by parliament in 1934 and 1936 respectively as part of the crisis measures. This measure, although it had its opponents, especially on the part of the management of the Živnostenská banka, proved to be effective, and only after 1934 exports increased by 24 % compared with their greatest decline (Kubů & Pátek 2000, pg. 20 - 21).

Another heritage of the Austro-Hungarian Empire causing major problems was the national composition of the population. The constitution of Czechoslovakia in the interwar period was justified by two influential theories: the theory of

Czechoslovakism[3] and the right to self-determination[4]. A combination of those two meant that other nationalities except for Czechs and Slovaks were considered only minorities whose rights should be respected but whose national state was somewhere else. That led to the situation where the German-speaking population in the territory of Czechoslovakia was labeled as a minority although it accounted for more than 30 % of the whole population and was second after the Czechs (Němec 2010, pg. 15). Naturally, this transition was not easily embraced, as the German-speaking citizens had enjoyed their privileged position for centuries. The shift not only impacted social status but also had repercussions on economic relations. Even before the war, Albín Bráf foresaw: "The struggle for the preservation of (Czech) national identity will be decided on the economic field" (Kárník 2017, pg. 204). Despite Czech economic expansion before the Great War, it is estimated that around 70% of the capital in the newly founded Czechoslovakia was still under German administration: "Only exceptionally did it happen that a German worker worked in a factory controlled by a Czech businessman or bank, but in almost half of the cases, a Czech worker worked in enterprises controlled by German capital" (Kárník 2017, pg. 206). Post-revolution, this dynamic was expected to change. Nostrification and land reform were not merely tools to preserve stability but were also crucial in ensuring dominance and success in the Czech national economic revolution.[5]

The complexity of economic and social relations in interwar Czechoslovakia encourages further in-depth research. Especially, the German perspective has often been overlooked in the past, including their view of Czechoslovak monetary policy. Before delving into the background of the German-speaking academic elite interested in economics, it is essential to provide a more detailed description of the main monetary policy events during the interwar period. Subsequent paragraphs will be devoted to exploring the ideas and actions of the two most influential figures in Czechoslovak monetary policy during the First Republic period.

[3] Czechoslovakism is a theory widely used in the 20th century that Czechs and Slovaks constitute one nation. During the interwar period, it became an official ideology in newly constituted Czechoslovakia (Žaloudek 1999, pg. 54).

[4] The right to self-determination is a principle that grants every nation the right to its state. It is the right of a nation to decide freely whether to establish its state or to form a federation with other nations (Žaloudek 1999, pg. 362).

[5] For a deeper understanding of the problem of Czech-German economic struggles see for example Jančík& Kubů (2011).

2.1 Rašín's Monetary Policy of 1919 - 1922

Alois Rašín was born on the 18th of October 1867 in Nechanice. After graduating from the Law Faculty of Charles University in Prague in 1891, he commenced his career in law. During the Great War, he became a member of the resistance movement and was imprisoned for treason (Doležalová 2007, pg. 126). While in prison, he authored Národní hospodářství (1922), his only theoretical economic work. The slogan "Work and Save", the name of his first speech as Czechoslovak Finance Minister, sums up the basic principles of his economic philosophy (Lacina 1992, pg. 41). Rašín was a man of action, preferring practical applications over theory. He actively participated in the establishment of Czechoslovakia during the coup, contributing to the formulation of many economic laws. Serving as the first Czechoslovak Minister of Finance from the 14th of November 1918 until the 8th of July 1919, Rašín successfully implemented the monetary separation from Austria-Hungary in March 1919, playing a vital role in the subsequent stabilization of the Czechoslovak currency (Doležalová 2007, pg. 47). However, despite the success of his pragmatic approach and actions, which often bore fruit with delay, not everyone appreciated them and mainly his deflationary policy of 1922 aroused heated debates among both the lay and professional public, as it as it clashed with the post-war crisis and meant for example also a weakening of the export industry (see e. g. Čechurová 1996). In January 1923, he became the target of an assassination attempt, succumbing to injuries a few weeks later (Doležalová 2007, pg. 47).

As mentioned earlier, Rašín did not place much emphasis on theoretical economics and is generally regarded as more of a politician than an economist[6]. Nonetheless, he demonstrated an ability to surround himself with capable advisors[7], adeptly navigate complex situations, and make effective decisions (Lacina 1992, pg. 41). Monetary policy and a robust currency were crucial aspects for him in maintaining a stable and independent economy. Identifying as a metalist, Rašín considered backing the currency with precious metals essential for restoring and strengthening confidence in money (Holmann 2001, pg. 507-508). Rašín's most significant contribution to the monetary policy of the newly established Czechoslovakia was undoubtedly the establishment of the Czechoslovak currency in the spring of 1919. On Rašín's proposal, the National Assembly approved a law during a secret meeting on 25 February 1919. The subsequent day saw the closure of borders, interruption of postal connections and international rail transport, and the halt of deposits and remittances (Staněk 1995, pg. 160). As a result of the law, all circulating Austro-Hungarian banknotes were withdrawn between 3 and 12 March 1919, marked, and half of the withdrawn quantity was reintroduced into circulation. The remaining half was withheld, and 1%

[6] Holmann (2001, pg. 507) calls him "a pragmatic eclectic"

[7] Namely J. Preiss, V. Pospíšil, K. Leopold or, A Novák (Lacina 1992, pg. 44).

government bonds were issued in their place, designated for repaying the extraordinary levy on property (Ulrich 2001, pg. 31). Banknotes of denominations 10, 20, 100, and 1000 Koruna, totaling almost 7,158 million, were marked. Coins and lower notes remained valid until 14 October 1919 without stamps. The stamped currency, described as the first Czech state notes, was gradually replaced by new ones starting in July 1919 (Staněk 1995, pg. 160).

Rašín aimed to halt inflation simultaneously with the creation of the currency in 1919. According to his perspective, a stable and robust currency was a top priority, reflecting an economically strong, stable, and reliable state. However, the economic and social situation was not yet favorable, requiring him to postpone his plans. Over the next three years, the political and economic landscape remained far from stable, with the position of Minister of Finance changing hands among five individuals. Disputes arose over the budget deficit, and the deflationary monetary policy, which resumed at the end of 1921, further exacerbated the challenges (Vencovský 1997, pg. 189). Even without the ministerial post, Rašín continued to exert influence as a member of parliament and presented his plan to introduce deflation, involving a significant appreciation of the Czechoslovak Crown, in a speech at the beginning of 1922 (Lacina 1992, pg. 45). The deflationary policy intended to be applied through the exchange rate, and there was consideration for a substantial increase to around 20 Swiss francs per 100 Czechoslovak crowns (Vencovský 1997, pg. 189). Eventually, from the end of 1921 until the close of 1922, the exchange rate rose from 5 to 16 francs per 100 crowns, maintaining this level until 1934. However, the sharp appreciation of the Czechoslovak Crown had its downsides, manifesting in the autumn of 1922 with reduced exports, downward pressure on wages, and increased unemployment. A general deterioration in the social situation fueled anti-Rašín sentiment among the population, culminating in his assassination in early 1923 (Geršlová 1999, pg. 26).

2.2 Engliš's Perspective on the Czechoslovak Monetary Policy

Another influential figure in the interwar economics of Czechoslovakia and a counterpart to Rašín was Professor and politician Karel Engliš. Born on the 17th of August 1880 in Hrabyně near Opava, he pursued law studies at the University of Prague and Munich, under the guidance of the renowned Albín Bráf. Bráf considered a "founding father" of Czech economic terminology, recommended Engliš, his promising student, for a position in the Provincial Statistical Office (Zemský statistický úřad). In 1910, Engliš was appointed associate professor at the Technical University in Brno, swiftly rising to a full professor a year later (Holman 2001, pg. 508). By 1919, he had also assumed a professorship at Masaryk University in Brno, where he served as the first rector. Following the establishment of Czechoslovakia, Engliš held frequent appointments as Minister of Finance, serving in a total of six First Republic governments until 1931. From 1934 onwards, he took on the role of

Governor of the Czechoslovak National Bank. In 1947, he briefly became the rector of Charles University. After the February coup in Czechoslovakia, he faced persecution by the communist regime, leading him to go into seclusion. Engliš passed away in his birthplace of Hrabyně on the 13th of June 1961.

If Rašín was primarily a practitioner whose theoretical philosophy must be gleaned between the lines of his actions and reflections, Engliš was already a recognized academic at the time of the establishment of independent Czechoslovakia. As the author of several theoretical treatises, he developed his economic teleological theory, forming the basis for a later-founded school of economic thought (Holman 2001, pg. 508). Despite their mutual respect and lack of open personal disputes, the economic thinking of Engliš and Rašín were complete opposites. Engliš, one of Rašín's first critics, openly challenged the monetary reform of 1919. He perceived it as an application of the quantitative theory of money that did not consider the reform measures' impact on overall price developments. Later, he disagreed with Rašín's deflationary policy, attributing it as the main cause of the financial crises in the early '20s, and advocated for currency stabilization instead. It wasn't until 1925, during another deflationary crisis, that his opinion gained widespread acceptance after the government was compelled to abandon Rašín's advocated monetary policy (Vencovský 1993, pg. 16-19). Engliš continued to exert influence on Czechoslovak economic policy, serving as Minister of Finance, influential expert, professor, and later as the Governor of the National Bank. At the onset of the Great Depression in 1929, Engliš, a nominalist, criticized the National Bank's policy, specifically denouncing the premature merger of the Czechoslovak crown with gold and the liberalization of the foreign exchange system. He anticipated the need for currency devaluations and the abandonment of the gold standard (Vencovský 1993, pg. 20).

In early 1934, in his capacity as a "mere" professor, he compiled a summary of his articles and polemics related to the economic crisis, critiquing the monetary policy of the National Bank. Shortly thereafter, he actively advocated for a one-sixth devaluation of the Czechoslovak crown. This move proved to be an effective measure in alleviating the crisis, leading to his subsequent appointment as the Governor of the National Bank, a position he held until February 1939. In his role as Governor, he played a crucial part in ensuring that the Czechoslovak crown became one of the most sought-after European currencies. His efforts focused on stabilizing the currency and upholding the purity of the National Bank's issuing policy, which strictly prohibited the granting of long-term loans to commercial banks and any loans to the state budget (Vencovský 1993, pg. 20, 34-35).

3 German-speaking Economists at Charles-Ferdinand[8] University

The upcoming chapter presents a concise overview of the history of teaching economic subjects at universities in Prague, with a specific emphasis on the German University in Prague. This historical context is essential for gaining a deeper understanding of the background of the docents and professors whose ideas are explored in more detail within the following text. The final section of this chapter introduces the biographies of the selected economists who delivered lectures at the Faculty of Law of the German University in Prague from 1918 to 1938.

3.1 The Tradition of German-speaking Economists at Charles-Ferdinand University

3.1.1 Economics Courses at Charles-Ferdinand University in Prague until 1882

The first academic institution at Prague University to offer courses in economic subjects was the Faculty of Law. By the end of the eighteenth century, initial indications of the inclusion of political science in full professorships had already emerged. In 1810, the establishment of the professorship of statistics marked the emergence of another independent discipline – one that can be considered among the earliest attempts to depict *the economic landscape of the world*. It is worth noting, however, that the statistics of the early nineteenth century should not be equated with contemporary statistics. Instead, they were primarily focused on interpreting a form of *history of the present*. Their objective was to provide the most accurate information available about the geographical, demographic, and economic conditions of the region, employing limited numerical data and mathematical methods. At the outset of the nineteenth century, the curriculum was rigidly defined by the state, with statistics being a foundational first-year subject. Other courses related to economic theory, finance, and economics included commercial law, fee law, and accounting (Havránek 1997, pg. 41-42).

[8] Charles-Ferdinand University was established in the aftermath of the Thirty Years' War through the amalgamation of Charles University and the Jesuit College. In 1882, the university underwent a division into Czech and German sections. This configuration persisted until 1921, at which point the German segment became an independent entity known as the German University in Prague, where German-speaking economists continued to teach. Meanwhile, the Czech portion underwent a renaming, reverting to its original name of Charles University.

One of the most notable economists in the early nineteenth century was Václav Gustav, who held the position of professor of political science at the Faculty of Law. His most significant contribution lies in his work concerning commercial legislation. Given the stringent state regulations governing education during that era, Václav Gustav was compelled to align with the declining Austrian mercantilism in his teachings. However, alternative sources suggest that he displayed an appreciation for economic liberalism inspired by the principles of Adam Smith (Havránek 1997, pg. 48).

The transformative year of 1848, along with the March reforms of universities, ushered in a significant change not only in the realm of economics education. Beyond the introduction of the German language for instruction, a change made as early as 1784, the reform also marked the inclusion of the Czech language within academic circles. Liberalization had other implications as well; it facilitated stronger connections with German and other foreign universities. Starting in the 1850s, young and driven professionals from abroad, particularly from Germany, joined Karl-Ferdinand University as educators, leading to a transformation that endowed the university with a distinctively German character (Havránek 1997, pg. 104). Post-revolution, the designation of *Law Studies* was revised to *Studies in Law and State Sciences* (Havránek 1997, pg. 126). During this period, political economy and national economy became recognized state science disciplines, both experiencing significant growth. Their rapid development hinted at the future high quality and stature of this field in Prague. Notably, until 1882, the University of Prague boasted prominent figures in these fields, including Professor Franz Makowicz from Cracow, Professor Leopold Hasner Jr. in political economy, and Petr Mischler. Furthermore, Eberhard Novák, a statistics professor, and Karl Thomas Richter, a disciple of Lorenz v. Stein, joined the law faculty after 1861. After the departure of these two professors, their roles were assumed by two noteworthy figures: Emil Sax and Karl Theodor v. Inama-Sternegg (Havránek 1997, pg. 136).

3.1.2 Political Economy at the German Charles-Ferdinand University in Prague until 1918

The development of Czech national consciousness and the promotion of Czech as an academic language resulted in the split of Charles-Ferdinand University into two separate institutions: one with instruction in German and the other in Czech. On April 11, 1881, the Emperor issued a decree stating: "*The Charles-Ferdinand University in Prague will henceforth be organized such that one university in Prague will use German as the language of instruction, and the other will employ Czech, and both will continue to be known as Karl-Ferdinand University*" (Kvirenc & Kunstová 2006, pg. 185). However, the relevant law formally took effect the following February, and the actual division of individual faculties occurred gradually. The Faculty of Law was among the first to undergo this transition (Havránek 1997, pg. 185).

The German University in Prague played a pivotal role in upholding German culture and advancing scientific knowledge in Bohemia until its dissolution in 1945. However, the university's position was marked by challenges, given that Prague had a predominantly Czech population at the time of its establishment, causing the German minority to become increasingly isolated. A significant issue was the reluctance of German professors to learn or speak Czech[9], resulting in minimal communication between the Czech and German academic communities (Havránek 1997, pg. 305).

During this period, National Economics and State Studies at the Faculty of Law of the German Charles-Ferdinand University experienced significant advancements. Despite a shift towards a more purely scientific orientation in research, the faculty achieved outstanding results. Alongside the influence from the preceding era, exemplified by the renowned Inama statistical school, the 1880s saw the emergence of notable figures who played pivotal roles in German economics education in Bohemia, even during the First Republic. Among these luminaries were Heinrich Rauchberg and Robert Zuckerkandl. Additionally, key figures who, while not active in the academic life of the First Republic, made substantial contributions to the field, included Emil Sax, a representative of the Austrian School of National Economy, and Friedrich von Wieser (Havránek 1997, pg. 322).

3.1.3 The Status of the German University after the Establishment of the Czechoslovak Republic

Even before the outbreak of World War I, the German division of Charles-Ferdinand University faced challenges in the Czech environment. However, after the First World War, the university encountered an even more complex situation. The matter of how to address the existence of two universities in Prague in the newly established republic became a subject of concern for both academics and politicians from the very inception of the republic. Various parliamentary proposals were put forward to address this situation, including the possibility of relocating the German university to one of the border towns. Notably, Liberec, which served as the primary center for Germans in Czech territory at that time, received significant consideration. This proposal found favor not only among certain Czech representatives but also among some members of the German student and professorial communities, including the then-rector of the German University, August Naegle. However, there were opposing viewpoints, and ultimately, the German University remained in Prague (Bečvář 1998, pg. 20-23, 181-182).

The formal relationship between the two universities was ultimately established through Act No. 135 Coll. of 19 February 1920, which addressed the interconnection of the Prague universities. This legislation, often referred to as Lex Mareš, was named after its proponent, František Mareš, a nationalist politician and professor at the

[9] The instruction of the Czech language at the German part of the university did not begin until 1906.

Faculty of Medicine. The initial paragraph of this law held great significance: "The Czech University is recognized as the rightful successor to the ancient Charles University. The names of the two universities in Prague, established under the Act of 28 February 1882, No. 24 of the Laws of 1882, namely 'The Czech Charles-Ferdinand University' and 'The German Charles-Ferdinand University' are hereby abolished. The name 'Charles University' is reinstated for the Czech University. A special law will determine the name of the German University. The law also outlined the ownership relations of the two universities, although these relations were not fully clarified until later years. In essence, it can be concluded that not only did Charles University inherit the cultural and symbolic legacy, but it also largely acquired the material legacy (Bečvář 1998, pg. 23). In 1930, the German University also pursued acknowledgment of its lineage to the "*Ancient Learning of Charles*," but this appeal was denied by the Senate. Tensions persisted between the universities, occasionally leading to minor disputes among their representatives. Nevertheless, amid these tensions, there were instances of amicable gestures, such as mutual attendance at funerals or the exchange of letters of condolence (Bečvář 1998, pg. 39-41).

In terms of student enrollment, the German university was approximately half the size of the Czech university. Charles University counted approximately six thousand students, while the German University had around three thousand. During the First Republic, there was a notable redistribution of students among the faculties of the German University, particularly within the faculties of medicine and law. In the early 1920s, the medical faculty of the German University had 1600 students, but by the late 1930s, this number had risen to 2400. In contrast, the Faculty of Law experienced the opposite trend, with only a quarter of its 1920s enrollment remaining by the 1930s. The primary reason for the substantial decline in interest and enrollment in law studies at the German University stemmed from the limited prospects for subsequent employment as a German-speaking lawyer within Czechoslovakia. The dissolution of the Austro-Hungarian Empire significantly curtailed these opportunities due to the language barrier (Bečvář 1998, pg. 184).

On the brink of the outbreak of the Second World War, the German University aligned itself with Nazi ideology. Consequently, numerous professors of Jewish origin were compelled to depart from the university following the Munich Agreement. Simultaneously, there was a substantial exodus of Jewish students, who had previously constituted a significant portion of the university's enrollment. Following the closure of the Czech universities in 1939, the assets of Charles University were transferred to the German University. As a result, many Jewish students and professors, who were unable to emigrate, were forced to discontinue their association with the university. Tragically, a considerable number of them eventually found themselves interned in concentration camps. After the conclusion of World War II, the German University was dissolved by the decree of the President of the Republic (Bečvář 1998, pg. 185).

3.2 Economists Lecturing at the German University in Prague during the First Republic Era

Economic subjects, particularly political economy and national economy, continued to be part of the curriculum at both law faculties of Prague universities. Furthermore, these doctrines were also taught at the technical universities in Prague, and in 1919 Vysoká škola obchodní (College of Commerce) was founded (Bečvář 1998, pg. 86). During the First Republic era, the subjects and state examinations for aspiring lawyers were rigorously defined by law, with the study of national economy as a mandatory component. This requirement applied to the German University as well, which, as a fully-fledged state university in Czechoslovakia, adhered to the established regulations (Havelka & Placht 1934, pg. 1291-1293).

Law students in the First Republic period were obliged to complete theoretical state examinations in history and law, jurisprudence, and state science to earn their doctorate. The examinations in jurisprudence and state science were conducted at the conclusion of their studies. Attending lectures on national economy or financial science was a prerequisite for passing these exams. Specifically, students were required to engage with subjects such as the science of national economy and national economic policy, financial science with a focus on Czechoslovak financial legislation, and Czechoslovak commercial and exchange law. These compulsory lectures also constituted examination subjects. National economy and financial science were central to the state science examination, while commercial and exchange law was a component of the judicial examination. Furthermore, the law faculties were mandated to offer elective lectures on state accounting or financial law (Havelka & Placht 1934, pg. 1291-1293).

Table 3.1: Key Figures in the Instruction of Economics at the German University in Prague

Name		Born		Died		Period of academic activity in Prague		Anticipated professional orientation
Zuckerkandl	Robert	3.12.1856	Ráb	25.5.1926	Prague	1894	1926	Austrian school of marginal utility
Rauchberg	Heinrich	12.4.1860	Vienna	26.9.1938	Prague	1896	1838	Statistics
Frankl	Otto	4.10.1855	Prague	26.3.1923	Prague	1883	1923	Commercial law
Engländer	Oskar	31.10.1876	Pardubice	31.12.1936	Prague	1918	1936	Austrian school of marginal utility
Amonn	Alfred	1.6.1883	Bruneck	2.11.1962	Bern	1920	1929	"Schumpeterian school of thought"
Weiss	Franz X.	18.4.1885	Vienna	19.3.1956	Oprington	1926	1938	Austrian school (Böhm-Bawerk)
Müller	Hugo	30.10.1882	Ústí nad Labem	23.4.1942	Praha	1926	1939	Theory of money and bank credit
Worliczek	Camillo	18.7.1892	Moravský Krumlov	?1951	?	1929	1945	Agricultural policy
Marschner	Robert	4.7.1865	Prague	8.8.1934	Šluknov	1920	1930	Insurance
Totomianz	Vahan	2.2.1875	Astrakhan	9.5.1964	France	1922	1923	Cooperativism
Flusser	Gustav	18.3.1885	Rakovník	18.6.1940	Buchenwald	1927	1939	Business science
Spitaler	Armin	11.6.1898	Prague	29.10.1964	Köln am Rhein	1932	1941	Taxes
Schranil	Rudolf	21.1.1885	Mikulášovice	22.7.1957	Brühl	1921	1941	Finance
Mayer	Theodor	24.8.1883	Neukirchen	26.11.1972	Salzburg	1925	1930	Economic history

The group of professors who taught economics at the German University in Prague exhibited considerable diversity. Within this group, a distinction can be made between the generation that had already been instructing at the university before the outbreak of the First World War. This earlier generation featured notable figures such as Robert Zuckerkandl, who held the position of professor in political economy and economics, Heinrich Rauchberg, an expert in statistics and civic science, and Otto Frankl, a professor specializing in commercial law. The subsequent generation of economists and national economists at the law faculty of the German University included individuals like Oskar Engländer, Alfred Amonn, Franz Xaver Weiss, Hugo Müller, and, lastly, Camillo Worliczek and Egon Armin Spitaler.

Furthermore, there was a separate group of associate professors and professors who specialized in particular subjects related to economics. This category encompassed Vahan Totomianz, who held the rank of professor in cooperative studies, Robert Marschner, with expertise in insurance, Gustav Flusser, responsible for teaching merchant arithmetic and business science, Rudolf Schranil, serving as a professor of

financial law, and Theodor Mayer, who was a professor specializing in economic history. Basic information about all the figures mentioned can be found in Table 3.1.

This thesis primarily focuses on the contributions of Oskar Engländer, Alfred Amonn, and Hugo Müller, considering the selected time frame, their scientific focus, and overall impact. The biographical information about these individuals is discussed in the subsequent paragraphs, which, if not specifically stated, draws from corresponding funds from the Archive of Charles University and National Archive Chodovec (see Archive materials on pg. 79). Biographical profiles of the other notable figures are provided in Appendix A.

3.2.1 Alfred Amonn

Alfred Amonn was born on 1 June 1883 in Bruneck. Following his studies in law and economics in Innsbruck and Vienna, he embarked on his academic journey. In 1910, he assumed the position of an extraordinary professor in Freiburg. In 1912, he attained the status of a full professor at the University of Czernowitz, following in the footsteps of his classmate Josef Alois Schumpeter[10]. Subsequently, in 1920, he transitioned to the Faculty of Law at the German University in Prague, where he was appointed as a full professor of political economy on 30 April (Winterberger 1983). Besides his academic activities at the German University in Prague, a rare mention in the daily press suggests that he paid attention to the Czechoslovak monetary policy, when he gave a lecture in Teplice on the currency policy of Czechoslovakia at the beginning of February 1923, where he stated that Finance Minister Rašín had prevented inflation by stamping banknotes and the timely financial separation of Czechoslovakia from Austria. Another task of Czechoslovak financial policy should have been to stabilize the crown and create a gold reserve (Anonym 1923).

Figure 3.1: Alfred Amonn

[10] Amonn's fellow students at the University of Vienna included Joseph A. Schumpeter, Ludwig Mises, Emil Lederer, Otto Bauer, Hans Kelsen, and Felix Somary. Amonn had a particularly close relationship with Schumpeter, who, on several occasions, recommended him as his replacement or substitute. For an in-depth exploration of the close relationship between Amonn and Schumpeter, refer to Gehrke (2017).

Source: National Archive Chodovec, PŘ 1931–1940, box 4380, sign. A306/10–Alfréd Amonn

Upon his relocation to Prague, Alfred Amonn faced persistent housing challenges. He was initially provided with a small temporary apartment spanning a mere 40 square meters. However, this living space proved inadequate for his family, which consisted of seven members, including four children[11] and three adults. The constrained living conditions significantly impeded Amonn's academic endeavors. Given his inability to secure a more suitable residence within three years of his appointment as a full professor in Prague, Amonn sought an interest-free loan to join a housing cooperative with plans to construct apartments in Letná. After several delays and negotiations regarding the loan amount, Amonn was eventually approved for the loan. Nonetheless, before he could fully resolve the housing situation, he departed from Prague for good.

Alfred Amonn's time in Prague also involved several trips abroad for various academic purposes. In Easter 1921, he embarked on a study tour to Berlin and Vienna. In September 1922, he participated in the Jubilee Congress for Social Policy in Eisenach. Subsequently, in November 1923, the Ministry sanctioned his lectures on the protection of workers, which he was tasked with delivering at the law faculty of the German University. In the summer of 1924, Amonn accepted an invitation from the Rockefeller Foundation to engage in studies at American universities during the summer recess. He returned to Prague for the winter semester of 1924/1925. During this term, he also served as a substitute professor teaching economics at the German Technical University in Prague. In the spring of 1926, Alfred Amonn received an invitation from the Imperial University of Tokyo, which extended an offer for him to lecture there for a period of two to three years. This invitation was again linked to the fact that his former colleague, Josef Alois Schumpeter, had to refuse this offer as he had been called to the University of Bonn in that year (compare Winterberger 1983 & Gehrke 2017). The professorial staff of the law faculty at the German University unanimously approved Amonn's journey to Japan, granting him unpaid leave until the end of September 1928. Nevertheless, Amonn extended his stay until June 1929. During this time, he was offered a lectureship in Bern, Switzerland.

Professor Amonn did not have the luxury of much contemplation regarding his decision to accept the offer in Bern. Consequently, he formally requested a temporary leave of absence from the German University in Prague to pursue teaching in Switzerland. In a letter dated 29 October 1929, he elaborated on his predicament to the dean. Amonn expressed his strong desire to remain in Prague due to the personal connections he had forged but clarified that the existing circumstances, coupled with the limited prospects for adequate housing and a higher standard of living on his present salary, were not particularly conducive to supporting a family of seven.

[11] Alfred Amonn had four sons: Kurt (*1917), Herbert (*1919), Egon (*1921) and Jan (*1924). His wife's name was Anna and was born on 11 July 1889 in Innsbruck.

However, he maintained that if conditions were to change, he would prefer to remain in Prague. In response, the faculty made efforts to accommodate him. He was granted a paid leave of absence, contingent upon the understanding that the funds would be disbursed should he return to the faculty. Additional measures were also taken to enhance his financial situation in Prague. Nevertheless, Amonn ultimately decided to remain in Bern. On 5 April 1930, he confirmed his acceptance of the Swiss offer and consequently requested to be released from the civil service of the Czechoslovak Republic[12]. He passed away in Bern on 2 November 1962.

3.2.2 Oskar Engländer

Oskar Engländer was born on 31 October 1876 in Pardubice. In 1894, he commenced his law studies at the German University in Prague, where he attended lectures delivered by Professors Zuckerkandl and Wieser. These individuals, in particular, had a profound influence on him and significantly shaped his future academic pursuits. Engländer successfully completed his studies at the Faculty of Law in 1899. Following graduation, he briefly served as a trainee in the judicial service and also worked as an assistant at the Institute of State Studies at the German University. Subsequently, he dedicated a year to voluntary military service. Starting in 1900, Engländer held a position in the administration of the state railways, from which he departed in 1921, having achieved the rank of chief councilor. He then assumed the role of director at the private railway Liberec-Jablonec-Tanvald, a position he held until the railway's nationalization (Schranil 1936, pg. 345).

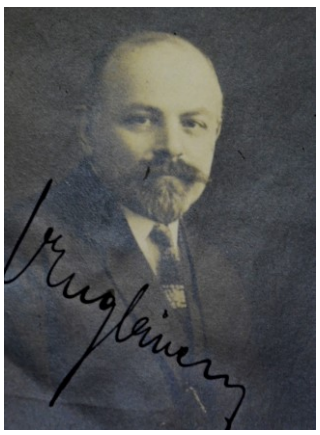


Figure 3.2: Oskar Engländer

Source: National Archive Chodovec, PŘ 1931–1940, box 5589, sign. E387/5–Oskar Engländer

In June 1917, Oskar Engländer applied for the venia docendi in the field of transport, specifically focusing on the economics of transport and transport law, at the law

[12] However, his attachment to Czechoslovakia did not completely disappear even after his relocation. For instance, during the winter semester of 1932-33, Judr. Václav Chytil, later the secretary to the deputy governor of the National Bank of Czechoslovakia, Karel Engliš, and a recipient of financial support from the Rašín Scholarship Fund, studied under Amonn in Bern (Anonym 1932).

faculty of the German University. For his habilitation thesis, he submitted "Eisenbahnschlüsse," a work he had completed in 1914. The assessment of Engländer's eligibility and the feasibility of his habilitation revolved around a contentious question: whether the field of transport law was suitable for university-level instruction. The assessment committee, which included Rauchberg, Zuckerkandl, and Frankl, weighed in on this matter. Zuckerkandl provided an approximately three-page evaluation, concluding that the subject matter within the field of transport studies was sufficiently extensive to warrant granting a venia docendi. On 15 March 1918, Engländer's habilitation and examination lecture were successful, resulting in his appointment as a private associate professor of transport engineering.

A year later, Oskar Engländer sought an extension of his teaching license to include national economy and national economic policy. In his application, he cited other scholarly works he had produced in the interim. The professorial board unanimously approved this request, and in June 1919, the ministry granted him the additional license. Subsequently, in March 1922, he was appointed as an extraordinary professor, and by 30 April 1927, he had achieved the status of a full professor. In this role, Engländer partly assumed the responsibilities of Professor Zuckerkandl, whose passing in 1926 led to Engländer teaching his courses as a substitute. Notably, the media of the era responded to his appointment as a full professor, with the *Socialdemokrat* newspaper raising concerns about his simultaneous public positions, including chief councilor of the State Railways and full professor at the German University, and expressing fears of potential conflicts of interest.

Oskar Engländer was the parent of two children: a son named Hans, who graduated from law school in 1933, and a daughter named Lilly, who completed her education at the German Real Gymnasium in 1934 and pursued further studies at the Faculty of Science of the German University. In 1932, Engländer began experiencing vision problems that ultimately resulted in total blindness. Remarkably, this disability did not hinder his academic and scientific diligence.

Engländer's economic thinking bore the influence of his mentors, particularly Wieser and Zuckerkandl, who could be classified within the analytical Austrian school of marginal utility. However, according to Schranil (1936, pg. 346-347), this influence waned, especially after Wieser departed from the University, which should be evident in Engländer's notable work from 1921, "Die Bestimmungsgründe des Preises," in which he espoused views distinct from the staunch adherents of the Vienna School of economics. Oskar Engländer passed away on 31 December 1936.

3.2.3 Hugo Müller

Hugo Müller was born on 30 October 1882 in Ústí nad Labem into a Catholic family, to parents Edmund and Flora (née Heller). On 25 November 1915, he married Theodora Bermann, and together they had two children: a son named Bedřich (born in 1915) and a daughter named Helena (born in 1919). The family frequently vacationed

in Yugoslavia, and Hugo himself often traveled to Austria, especially to Vienna, in his capacity as secretary to the director of the Hypoteční banka.

Figure 3.3: Hugo Müller

Source: National Archive Chodovec, PŘ 1941–1951, box 7466, sign. M2723/1–Hugo Müller

On the 20th of October 1924, Hugo Müller assumed the role of deputy director of the Hypoteční banka in Prague. Two years later, on 22 March 1926, he submitted an application for a venia docendi at the Faculty of Law of the German University. He aspired to offer advanced courses covering topics such as money, the theory of money, bank credit, exchange rate issues, and national economic theories. His habilitation thesis, "Wechselkurs und Gütterpreise," was duly presented. On 30 November 1926, Hugo Müller delivered an examination lecture on "Pfandbrief- und Pfandbriefskredit" before the professorial board, which deemed the lecture satisfactory. On the same day, he was appointed as a private lecturer in political economy at the German University in Prague. Hugo Müller passed away on the 23 th of April 1942 in Prague due to arteriosclerosis.

4 Three Comments on the Monetary Policy of the Czechoslovak Government

The following chapter introduces the three most important articles written by Alfred Amonn, Oskar Engländer, and Hugo Müller concerning the monetary policy of the First Republic of Czechoslovakia. It then proceeds to compare these articles, aiming to address the following research questions:

Firstly, this chapter delves into the overarching stance of the three German-speaking economists towards the monetary policies of the Czechoslovak First Republic. The mere existence of their articles serves as a direct manifestation of their interest in the monetary policies implemented by the Czechoslovak Government. The critical question emerges: were their assessments predominantly favorable or critical? Secondly, the examination explores the relationships of Alfred Amonn, Hugo Müller, and Oskar Engländer with Rašín and Engliš, two pivotal figures in Czechoslovak monetary politics. Did these German-speaking economists show interest in and make direct references to Rašín and Engliš in their writings? Can distinctions be drawn between Rašín's and Engliš's concepts in their works? If so, how did they evaluate the contributions of these Czech figures – negatively or positively? Finally, the analysis addresses the broader inquiry of the specific topics and theoretical orientations evident in the texts of Amonn, Müller, and Engländer. It aims to unravel how these elements may influence their perspectives on Czechoslovak monetary policy and its key representatives.

The first article, "Die Tschechoslowakische Währung und Währungspolitik," authored by Alfred Amonn and published in 1924 in Munich, constitutes a chapter in the fourth volume of the book series titled "Geschichte der Stabilisierungsversuche." This volume, entitled "Währungsreform in der Tschechoslowakei und Sowjet Russland," encompasses Amonn's contribution and features Michael von Bernatzky's article, "Der Zusammenbruch der russischen Währung und die Aussichten auf ihre Wiederherstellung." Amonn's work serves as an annotated account detailing the unfolding of the 1919 currency reform and tracking the subsequent developments in currency, exchange rates, and money supply up to February 1924. Following a concise overview of the immediate context surrounding the creation of Czechoslovak state notes, Amonn delves into the pre-World War I era, elucidating the evolution of Austria-Hungary's money supply. This historical backdrop sets the stage for understanding the constraints on the money supply during the monetary separation, which, as Amonn (1924) asserts, gradually eased to reach approximately 2 billion crowns. Amonn then delves into the establishment of the "Banking Office of the

Ministry of Finance of the Czechoslovak Republic," outlining its activities and competencies. The narrative further encompasses the dynamics of cash and cash equivalents, Lombard loans, escompte, and elastic and non-elastic means of payment. Additionally, Amonn provides insights into interest rates, quantities of precious metals and foreign exchange, and the intricate relationships between these elements, unraveling the reasons for their fluctuations.

Based on his astute observations, he unequivocally concludes: "It is evident from all these considerations that, until this juncture, there has been no explicit, intentional, or purposeful restriction[13] of the quantity of means of payment in Czechoslovakia. The prevailing monetary policy of Czechoslovakia, in actuality, has been driven by entirely different factors inclined towards restriction. Concerning the quantity of means of payment, the focus has hitherto leaned more towards fostering stability than the imposition of constraints"[14] (Amonn 1924). Amonn consistently downplays the contributions of governmental authorities, emphasizing that favorable developments often find their roots in the broader economic, financial, and monetary landscape. He contends that relatively minor interventions could catalyze these favorable shifts.

The second article, "Zwei Tschechische Schriften über Währungsreform" by Hugo Müller, was published in "Schmollers Jahrbuch für Gesetzgebung, Verwaltung und Volkswirtschaft im Deutschen Reiche" in 1926. This article assesses and partially compares two publications concerning monetary policy in Czechoslovakia: "Můj finanční plán," written in 1920 by the first finance minister of Czechoslovakia, Alois Rašín, and "Nástin měnového plánu stabilizačního po jeho stránkách metodické, technické a státofinanční," published in 1924 by the former Minister for Trade in Cisleithania, Josef Fořt. In the first two paragraphs, Müller recognizes Rašín's exceptional personality as a political leader and commends his intuition and his ability to make quick, accurate, and functional decisions. Rašín's sometimes almost violent energy radiates from "Můj finanční plan," which deservedly evokes respect. According to Müller (1926): "at this point, however, the praise must stop, as some unpleasant things stand in the way."[15] Although Müller admires Rašín's decisiveness and

[13] Amonn rejects the use of the word "deflation" and "deflationary" to describe the monetary policy of a state as an unfortunate neologism that is used ambiguously in public debate and is not suitable for scientific purposes. He therefore proposes and uses the term "restriction" or "contraction".

[14] „Aus allem dem geht hervor, dass bisher eine tatsächliche und beabsichtigte oder gewollte, währungspolitisch bedeutsame Restriktion der Zahlungsmittelmenge in der Tschechoslowakei überhaupt nicht erfolgt ist und die tatsächliche Währungspolitik der Tschechoslowakei von ganz Anderen restriktionistischen Gesichtspunkten beherrscht war.1 Was die Zahlungsmittelmenge anbelangt, war sie bisher viel eher auf Stabilität gerichtet als auf Restriktion."

[15] „Bei diesem Punkte muß freilich das Lob halt machen, eil sich manches Unerfreuliche in der Weg stellt."

directness, he adds that it sends a shudder down the spine and criticizes, in particular, the lack of reflection on the possible negative consequences of Rašín's decisions. As an example, Müller cites the decision to apply the territorial principle to the monetary separation from Austria-Hungary, which Rašín cited as the only feasible solution, without any consideration of its downsides and how they could be mitigated. Müller also evaluates negatively the quality of Rašín's extensive theoretical explanations, which, according to him, can at best be considered "clever popular economics."

Müller devotes only a third of his article to Rašín's book, with the remaining portion discussing the ideas of Josef Fořt. Müller acknowledges Fořt's economic essay as of high quality, attempts to explain it, identifies its mistakes, and compares Fořt's demands with the actual Czechoslovak monetary policy. Based on his analysis, Müller characterizes Fořt as a proponent of stabilization, an opponent of deflation[16], and a staunch metalist. Of particular interest to Müller is Fořt's belief that the stabilization of the Czechoslovak Crown was premature, as the natural exchange rate had not yet been achieved. Müller goes on to explain why he tends to disagree with Fořt's perspective.

The third article, "Die Devaluation der tschechoslowakischen Krone im Jahre 1934", by Oskar Engländer was published in 1936 in the professional journal „Weltwirtschaftliches Archiv, Zeitschrift für Allgemeine und Spezielle Weltwirtschaftslehre". The article is divided into three parts. The introduction delves into the historical developments of monetary policy in Austria-Hungary and subsequently in Czechoslovakia during the 1920s. Engländer deems this section necessary, as the devaluation of 1934 was "not a stand-alone event" but rather "a certain stage in the history of the Czechoslovak currency." After a brief explanation of the monetary system of pre-war Austria-Hungary, characterized as a "perfectly organized monetary system," he proceeds to elucidate the consequences of the Great War and the establishment of the new Czechoslovak currency post-war. Engländer asserts that all the measures implemented directed Czechoslovak monetary policy toward a specific path – deflation or, at the very least, stabilization. In this context, he mentions Alois Rašín as the only personality in his entire article, linking Rašín's death to the ultimate fate of this policy. Following this, Engländer continues with a description of the process of introducing the nominal gold currency in 1929, concluding the first part of his article.

The second and third parts then focus on the devaluation of 1934 itself. While the second part delves into its causes and implementation, the third part outlines its effects. Engländer naturally attributes the primary reasons for the shift in Czechoslovak monetary policy to changes in world monetary policy and the global economic crisis. The development of this crisis in Czechoslovakia and the potential, as well as the

[16] Interestingly, unlike Amonn, Müller does not oppose the use of the term "deflation" and uses it freely without any remarks.

implemented defense measures, constitute a significant portion of the second chapter. The conclusion of the second chapter addresses the means of implementing the devaluation. Engländer notes that this step faced opposition, especially from the former party of the late minister Rašín, although its proponents often included industrialists who would have benefited from the devaluation.

In the final chapter, Engländer extensively examines the real and perceived effects of devaluation. He concludes the chapter by asserting that the devaluation had only a short-term positive impact on boosting exports, after which the economy reverted to stagnation. Engländer highlights that the situation had already begun to show slight improvement two months before the devaluation. However, describing the devaluation as a failure would be inaccurate, as it is conceivable that without it, the economy might have experienced a more pronounced decline. He emphasizes the need for cautious use of devaluation as a monetary policy instrument, stating that the primary objective of monetary policy should be the stability of currencies, whether in the form of gold currencies or otherwise.

All three articles provide evidence that Alfred Amonn, Hugo Müller, and Oskar Engländer paid attention to developments in the Czechoslovak economy and monetary policy. Regarding their overall attitude toward Czechoslovak monetary policy, while they do not express overt negativity, a certain reticence and reserve in assessing its achievements can be discerned in the aforementioned works. Amonn points out that the favorable development is mainly attributed to the generally positive market conditions. Engländer emphasizes that although the devaluation may not be considered a failure, it lacked lasting effects. Müller criticizes Rašín's inability to at least attempt to reflect on the adverse consequences of his reforms.

The level of interest in personalities within Czechoslovak monetary policy varies significantly. Amonn deems it unimportant to mention the figures behind the 1919 monetary reform. Engländer, however, refers solely to Rašín in the context of advocating deflation and due to his death. In contrast, Müller takes a direct interest in the writings of two Czech personalities with political influence, essentially providing a review in which, despite noting some mistakes, he appreciates Rašín's strong spirit and intuition and Fořt’s scientific expertise. Interestingly, none of them directly mention Karel Engliš. In the first two articles, this omission is understandable, given the focus on the era of the initial monetary reform and deflationary policies associated mainly with Alois Rašín. Karel Engliš, despite being among the first critics and having served as Minister of Finance, appears not to have garnered their attention. Oskar Engländer also omits any mention of Karel Engliš in his article, even though it specifically addresses the devaluation proposed and advocated by Engliš for the Czechoslovak Crown. The sole reference can be found in the bibliography, where Engländer cites Engliš's article "Rok devalvace," published in the 1935 economic section of "Přítomnost" on page 65.

From a theoretical standpoint, it is noteworthy that Amonn, Müller, and Engländer adopt distinct positions regarding the term "deflation." Engländer and Müller do not find it necessary to challenge the use of this term in describing the trajectory of Czechoslovak monetary policy in the 1920s, while Amonn explicitly deems the term inaccurate, amateurish, and unsuitable for professional discourse. Instead, he opts for the terms "restriction" or "contraction." The demand for a higher level of terminological expertise is also reflected in the style with which Amonn presents the Czechoslovak monetary reform. By tracing the development of selected variables, he sheds light on the interrelationships between them and seeks internal connections. Engländer, on the other hand, adopts a more narrative approach, detailing the chronological development of the variables and situating them within the context of contemporary events. In Müller's article, he reviews two writings by Czechoslovak politicians. While he provides a relatively brief assessment of Rašín's book, he meticulously examines Fořt's contribution, describing both his theory and personality, praising the text's scholarly merits, and attempting to identify some of the erroneous reasoning in relation to the theories of other world economists.

5 Data

The sources of the following text-mining analysis are the books, articles and lectures written by Oskar Engländer, Alfred Amonn and Hugo Müller between the years 1918 and 1938. Some texts, mainly articles published in *Jahrbücher für Nationalökonomie und Statistik* were accessed digitally via the service DigiZeitschriften.de, others had to be borrowed from various libraries in Prague (such as the Czech National Library, Library of the Institute of History of the CAS or Specialised Library of Czech National Bank) and consequently manually photocopied by digital camera. In both cases[17] the obtained scans had to be converted to text via text recognition software. To achieve the most accurate conversion, the scans were first manually edited in a bitmap graphics editor, which included changing the brightness, contrast and cropping.

The source texts were written not only in Latin but also in Fractur - a type of Neo-Gothic fractured script used mainly in Germany from the 16th century until 1941. That narrowed down the possibilities of text recognition software selection because Fractur is a very specific, old-fashioned type of script whose characters are in many cases quite different from Latin. Therefore the open-source optical character recognition (OCR) software Tesseract was chosen to convert the visual materials to texts. First of all, the Tesseract-ocr-all package was installed on the Ubuntu Linux version 18.04. This installed the Tesseract OCR library in version 4.1.1-rc2-21-gf4ef and the files *deu.traineddata* and *deu_frak.traineddata* from the freely available set of trained models at https://github.com/tesseract-ocr/tessdata for German language and German language with Fraktur were added consequently. The text files were then manually checked and edited in MS Word. The editions included corrections of misspellings and unification of the footnote system, where the footnotes were in all cases moved to the end of each text.

For the purpose of analysis, the texts were then one by one translated via DeepL Translate software from German to English (UK) and the translations were copied and saved as .txt files in UTF-8 format. Table 5.1 lists all the analyzed texts with their basic characteristics:

[17] Although some articles can be accessed online, their text is only scanned as pictures without text recognition and with no possibility to search through the text of the articles digitally.

Table 5.1: Text Corpus Characteristics

	Author	Title	Year	Type	# of words	# of sen.	# of par.	Journal/Book/Series
1	Amonn	Die Hauptprobleme der Socialisierung	1920	Book	36198	1178	300	-
2	Amonn	Die gegenwartige Wirtschaftskrise und ihre Überwindung	1921	Article	2117	70	8	Die Wirtschaft
3	Amonn	Die Valutenkrise in der Tschechoslowakei	1922	Article	1287	31	7	Die Wirtschaft
4	Amonn	Der Ausweis des Bankamtes des Finanzministeriums	1923	Article	4348	153	14	Die Wirtschaft
5	Amonn	Die tschechoslowakische Währung und Währungspolitik	1924	Chapter	6897	448	245	Währungsreform in der Tschechoslowakei und Sowjet Russland
6	Amonn	Ricardo als Begründer der theoretischen Nationalökonomie	1924	Book	54213	1822	313	-
7	Amonn	Zur Frage der steuerlichen Lastenverteilung	1925	Article	15175	522	87	Jahrbücher für Nationalökonomie und Statistik
8	Amonn	Die Begriffe Volkseinkommen und Volksvermögen und ihre Bedeutung für die Volkswirtschaftslehre	1926	Article	2023	70	10	Volkseinkommen und Volksvermögen. Begriffskritische Untersuchungen
9	Amonn	Die Einnahmen aus Privat- und offentlichwirtschaftlichen Unternehmungen	1926	Chapter	20060	605	108	Handbuch der Finanzwissenschaft, Bd. 1
10	Amonn	Emil Lederers Grundzuge der ökonomischen Theorie	1926	Article	10509	432	76	Weltwirtschaftliches Archiv, Zeitschrift für Allgemeine und Spezielle Weltwirtschaftslehre
11	Amonn	Der Unternehmergewinn	1928	Chapter	6608	218	37	Einkommensbildung, Allgemeine Prinzipien, Lohn, Zins, Grundrente, Unternehmergewinn, Spezialprobleme
12	Amonn	Gegenwartsaufgaben der National-Ökonomie	1928	Article	12127	373	46	Zeitschrift für die gesamte Staatswissenschaft
13	Amonn	Das Lohnproblem	1930	Book	20279	557	114	Wirtschaftsprobleme der Gegenwart

14	Amonn	Grundfragen der Konjunkturtheorie und Krisenpolitik	1937	Chapter	7092	275	56	Oskar Engländer. Festschrift zur Feier des 60. Geburtstages, gewidmet von Fachkollegen, Freunden und Schülern
15	Amonn	Zwei offene Fragen der allgemeinen Preistheorie	1937	Article	14354	435	84	Zeitschrift für Nationalökonomie
16	Amonn	Keynes Allgemeine Theorie der Beschäftigung	1938	Article	11356	428	93	Jahrbücher für Nationalökonomie und Statistik
17	Amonn	Volkswirtschaftliche Grundbegriffe und Grundprobleme	1938	Book	66115	2129	498	-
18	Engländer	Das Geld ohne Eigenwert und die Preislehre	1922	Article	19192	793	89	Jahrbücher für Nationalökonomie und Statistik
19	Engländer	Zur Theorie und Praxis der Personenfahrpreise	1925	Article	4344	173	21	Zeitschrift für die gesamte Staatswissenschaft
20	Engländer	Standort	1926	Entry	10749	345	86	Handwörterbuch der Staatswissenschaften
21	Engländer	Unternehmergewinnn	1926	Entry	8929	280	30	Handwörterbuch der Staatswissenschaften
22	Engländer	Kapitalzins und wirtschaftliche Entwicklung	1928	Article	15609	560	52	Zeitschrift für die gesamte Staatswissenschaft
23	Engländer	Volkswirtschaftliche Probleme der Gegenwart	1929	Article	4283	210	32	Zeitschrift "Hochschulwissen"
24	Engländer	Das Seelische und die Volkswirtschaftslehre	1935	Article	13156	510	68	Jahrbücher für Nationalökonomie und Statistik
25	Engländer	Die Devalvation der tschechoslowakischen Krone im Jahre 1934	1936	Article	12179	579	125	Weltwirtschaftliches Archiv - Zeitschrift für Allgemeine und Spezielle Weltwirtschaftslehre
26	Müller	Die Geldvermehrung und die Sicherung der Wirtschaftsführnug	1921	Article	2491	86	28	Die Wirtschaft
27	Müller	Wechselkurs und Gütterpreise	1926	Book	50402	1490	653	-
28	Müller	Zwei tschechische Schriften uber Währungsreform	1926	Article	6257	247	87	Schmollers Jahrbuch für Gesetzgebung

29	Müller	Geld und Verteilung	1929	Lecture	15717	517	208	Sammlung Gemeinnütziger Vorträge
30	Müller	Kapital als dritter Produktionsfaktor	1933	Chapter	9118	404	107	Oskar Engländer. Festschrift zur Feier des 60. Geburtstages, gewidmet von Fachkollegen, Freunden und Schülern

Note: The digitized works by Hugo Müller and Oskar Engländer are freely accessible under this link: https://drive.google.com/drive/folders/1jM2LbRau4wQ1e52fS7DGEKxkOYkTj0JM?usp=sharing.

Alfred Amonn's work is still under copyright protection, which extends for 70 years after the author's death.

6 Text Mining Analysis

To analyse the text corpus described in the previous chapter and to address the stated research questions following three methods of text mining analysis are used: word frequency analysis, sentiment analysis and topic modeling. This chapter describes the methodology of these techniques and their implementation in R – software. Prior to the analysis itself, however, a few data transformations have to be made to be able to effectively proceed with the chosen methods.

First of all a specific data format has to be chosen. The basic format in computer science referring to textual values is so-called strings. A string is a sequence of characters and it is the most common way in which software including R primarily uploads text values. In text mining analysis, the term "corpus" is often used. In linguistics a textual corpus is defined as "*an internally structured, unified and usually indexed and comprehensive large collection of electronically stored and processed linguistic data, usually in text form, organized with a view to using for a specific purpose*" (Čermák 1995, pg. 119). Concerning a computer science, a corpus can be defined as an object containing "*raw strings annotated with additional metadata and details*" (Silge & Robinson 2022). In our case, a text corpus is a collection of texts written by Amonn, Müller and Engländer annotated with metadata such as author, year of publication, or type similar to in Table 5.1. The corpus can be also represented by a more structured format called Document-term matrix (DTM), which is frequently used in a text analysis as it can easily describe the count or tf-idf statistic of each term for a collection of documents. The structure of DTM looks as follows: "*Each row represents a document, each column represents a term, and the cell values are the counts of the occurrences of the term for the particular document*" (Engel & Bailey 2022). Finally, we are going to look more in detail at a tidy data approach - the style of data (including text) formatting and cleaning that is very useful for text mining analysis particularly in R because it enables it to work with a lot of popular packages.

The Tidy dataset described by Wickham & Grolemund (2017) is a table that has to fulfil the following three rules: *1. Each variable must have its own column. 2. Each observation must have its own row. 3. Each value must have its own cell.* From these rules, Silge & Robinson (2022) define a Tidy text format as a table with one token per row, where a token is a text unit that provides meaning such as a word, sentence, paragraph, or a grouping of consecutive words called n-gram that is of interest for the given analysis. Tokenization is then called the process, where the text is split into tokens. In R, the tidytext package with the unnest_tokens() function provides the functionality of converting the text into the tidy data format. Using this procedure, punctuation is automatically stripped and unless specified the tokens are converted to lowercase. In our case, the tidy data approach is used in most cases as it is flexible and

allows for effective usage of various text mining and visualization techniques in R. We can also easily set our token to a desirable value and it is also possible to merge tidy data table with metadata (and get tidy text corpus) table or covert tidytext to DTM.

Table 6.1 The Most Frequent Words in the Text Corpus

word	count
the	40694
of	26456
in	12436
and	11151
to	10746
a	9742
is	8985
that	5970
be	5476
it	5202
this	4653
for	4649
as	4607
which	3722
not	3585
or	3461
by	3330
on	2796
with	2732
are	2730

After achieving the preferred data format, other transformations are usually needed. Table 6.1 shows the most frequent words used in our whole tidy data corpus. We can see that there are only words that either do not bear any specific meaning and have only a lexical function such as prepositions, conjunctions, articles, or auxiliary verbs or their meaning is of lesser value such as verbs go, consider, indicate, and so on. Those words are called stopwords and we can define them as "*highly common words that are considered to provide non-relevant information about the content of a text*" (Engel & Bailey 2022). Generally[18], we want to omit them in our analysis, because they are simply redundant, we do not need them as they only distract from the significant words and bring zero value in abstracting general topics or sentiment. In the package tidytext in R, there exists a list of 1,149 English stopwords taken from different lexicons that serve our purposes just well. Combining our tidy data and a tibble[19] stop_words

[18] There can be problematic cases concerning mainly names or phrases such as „Doctor Who", To be or not to be", and „Take five", the omission of which could be undesirable depending on the type of analysis. By inspecting the given list of stopwords, we, however, did not come across this problem in our analysis. Moreover, term frequency and inverse document frequency (Tf-Idf) could be used without stopwords implementation as alternative tools for comparison (see Chapter 6.1).

[19] A "tibble" is a type of data frame in the R programming language, particularly associated with the tidyverse ecosystem.

through function anti_join, we cleaned our text of stopwords and went from 463,174 to 161,679 rows.

Word stemming is another transformation considered when performing text analysis. It is a procedure that reduces words to their stems for example reducing *stabilization, stability* and *stabilized* to the stem *stabil*. In our whole tidy text corpus, it reduces the unique count of words from 9,641 to 5,911. Table 6.2 compares 5 most frequent stems with the words they come from.

Table 6.2 Five most frequent stems

stem	count	words
product	3370	production, product, productive, products, productivity, productions, productively
price	3163	price, prices, priced
econom	2309	economic, economy, economics, economically, economies, economical, economise, economists, economist, economizing, economizes, economize, economist's, economy's, economising, economises
mean	1759	meaning, means, meanings, meaningless, meantime, meant, meaningful
capit	1721	capitalism, capital, capitalist, capitalistically, capitalists, capitalised, capitals, capitalless, capita, capitalization, capitalistic, capitalizing, capitalist's, capitale

In the course of the analysis itself, we then encounter certain issues that require us to go back and perform further corrections and transfromations of our dataset. Those trnaformations include mainly removing numbers and word substitutions. The numerical values that appear in tables or formulas in the text are not relevant to this analysis and constitute unnecessary noise that carries no meaning. Only the values denoting years have been retained, as they carry meaning that can be related across all texts. Furthermore, it seemed advisable to completely delete "i. e.", which did not appear in stop words, and to replace possessive adjectives with nouns (from Rašín's to Rašín), as stemming could not cope with them and evaluated them as two words with different stems. In addition, inaccuracies were found created by the translation of the texts, especially for the designation of the Czechoslovak currency (koruna), which in the original texts also appeared in different forms (koruna, krone or kč) and the translator translated also it in different ways as: crown, koruna, krone, kč, or kčč. Therefore all terms denoting the Czechoslovak currency were replaced by the English term crown.

6.1 Word Frequency Analysis, Term Frequency-Inverse Document Frequency, N-Grams and Co-occurrences

Word frequency analysis is a fundamental technique that tallies the raw occurrences of words within a document. By setting tokens to word values, we count these occurrences in our tidytext, which has been cleansed of stopwords and numerical values (excluding years), and undergone stemming, along with other mentioned modifications. This allows for a comparison of values and the creation of illustrative visualizations, such as word clouds or bar charts, for each author, providing foundational insights into the content of the analyzed texts. Additionally, this initial word count analysis can uncover potential shortcomings in the text, such as typographical errors or synonymous expressions.

As an alternative to word frequency analysis, which counts words with absolute numbers, the term frequency-inverse document frequency (tf-idf) statistic can be employed to "*measure how important a word is to a document in a collection (or corpus) of documents*" (Silge & Robinson 2022). Term frequency (tf) is a relative number defined as "*how often the word appears relative to the total number of words in a text*" (Engel & Bailey 2022):

$$tf(term) = \frac{n_{word}}{n_{total\ words\ in\ a\ document}}$$

Inverted document frequency (idf) "*decreases the weight for commonly used words and increases the weight for words that are not used very much in a collection of documents*" (Silge & Robinson 2022) and can be defined as:

$$idf(term) = \ln\left(\frac{n_{documents}}{n_{documents\ containing\ term}}\right)$$

The tf-idf statistic is then calculated by multiplying tf and idf scores. In R, the tidytext package includes the bind_tf_idf() function that can automatically calculate these statistics.

Another technique closely related to the frequency and occurrence of words in a text involves assessing the proximity of two words standing together in the text. This can be addressed through n-grams, which are sequences of words (phrases) appearing closely together in the text (e.g., bigram - two words next to each other, such as "interest rate"[20]). Alternatively, we can study the co-occurrence of two or more words,

[20] Note that n-grams in a text can overlap. For example, when looking for all bigrams in a sentence like "The interest rate is low," we get the following bigrams: "the interest," "interest rate," "rate is," "is low." To obtain meaningful bigrams, the stopwords routine must be implemented, or the calculation of tf-idf statistics should filter the most meaningful bigrams.

which, unlike n-grams, do not have to be right next to each other in the text. For n-gram analysis, we can easily implement it by setting the token to n-gram using the unnest_tokens() function. We then proceed similarly to token = word: either stopwords are addressed, and n-gram frequency analysis is conducted, or tf-idf statistics can be calculated. In this case, we do not stem our words, as it makes no sense and would complicate the results, making them difficult to read and interpret. N-grams can provide better context for our frequency analysis, reveal concealed text structures, and be very useful in sentiment analysis.

We are also interested in pairs of words that frequently appear together in a text unit, even if they do not occur directly next to each other. In our analysis, it makes the most sense to analyze the co-occurrence of words in each paragraph, but it is also possible to arbitrarily set sections, such as every ten lines or so. To obtain truly meaningful pairs of co-occurring words in one paragraph, the Pearson correlation coefficient φ is calculated for each such pair:

$$\varphi = \frac{n_{11}n_{00} - n_{10}n_{01}}{\sqrt{n_{1\cdot}n_{0\cdot}n_{0\cdot}n_{1\cdot}}}$$

where variables are defined in the following table:

Table 6.3 Variables for correlation coefficient of co-occurring words

	Has word Y	No word Y	Total
Has word X	n_{11}	n_{10}	$n_{1\cdot}$
No word X	n_{01}	n_{00}	$n_{0\cdot}$
Total	$n_{\cdot 1}$	$n_{\cdot 0}$	n

Source: Silge & Robinson 2022

This coefficient indicates how often the co-occurring pair of words appear together relative to how often they appear separately (Silge & Robinson 2022).

Note that for manual calculations of co-occurrences and correlations, the tidy data format with its row structure would not be very useful, and a rather wide matrix structure is needed. Fortunately, R contains the widyr package with a function called pairwise_count() that automatically transforms our tidytext to a matrix and allows us to compute co-occurrences and correlations easily.

6.2 Sentiment Analysis

Sentiment analysis is a technique anchored in the intersections of linguistics and computer science that aims to automatically evaluate the opinions, thoughts and impressions of the author of a given text providing information about the level of positivity or negativity of the author to a certain topic. Its growing popularity can be attributed to its widespread use, whether in marketing or science, and the ever-increasing amount of user-generated content on the internet and social networks as potential material suitable for analysis.

Sentiment analysis can be performed at different levels, such as Document Level, Sentence Level, Phrase Level, or Aspect Level, depending on the definition of its basic unit (document, sentence, sequence of words, words). The two most common methods used in sentiment analysis are the Lexicon-Based Approach and the Machine Learning Approach.

The Lexicon-Based approach involves the use of predefined lexicons that assign a score to words (or tokens) determining whether the word has a positive, negative, or neutral connotation. Words are assigned scores such as -1 for negative words, +1 for positive words, and 0 for neutral words. In practice, the researched document is split into tokens of single words, and the polarity of each token is calculated and aggregated. This approach is feasible and suitable for sentiment analysis performed at the sentence or feature level. Since it does not require a training dataset, it can be considered an unsupervised approach. However, it is a domain-based technique that does not take into account the context-sensitivity of some words. For example, the word "low" has a different sentiment in phrases like "low salary" versus "low debt."

The Machine Learning approach, on the other hand, can work better with context and even detect irony or sarcasm. It can be used in supervised or unsupervised forms, with supervised models being more commonly employed for their more accurate results. The application of these algorithms to actual data requires training on a training dataset (Wankhade, Rao, & Kulkarni 2022).

In our analysis, we aim to discover the sentiment associated with terms related to Czechoslovakia and the individuals Rašín and Engliš. The sentiment analysis is conducted at the feature level using a lexicon-based approach that assigns values to individual words and aggregates them over the paragraphs in which the listed words (or n-grams) occur. In R-software, the tidytext package provides several lexicons, and for our analysis, the AFINN lexicon by Finn Årup Nielsen was chosen. This lexicon assigns scores to words ranging between -5 and 5, with negative scores indicating negative sentiment and positive scores indicating positive sentiment (Silge & Robinson 2022). To address the context insensitivity of the Lexicon-based approach, corrections are made for sentiment scores of words preceded by certain negative words, such as "not," "no," "never," "without," and others, using bigrams.

6.3 Latent Dirichlet Allocation and Structural Topic Models

An alternative approach to frequency and co-occurrence analysis when trying to classify an unknown text via computer-based techniques offers topic modeling: “Topic models find patterns of words that appear together and group them into topics. The researcher decides on the number of topics and the algorithms then discover the main topics of the texts without prior information, training sets or human annotations” (Meyer & Pushmann 2019).

The main method in topic modeling includes Latent Dirichlet Allocation (LDA) and its extension Structural Topic Modeling (STM). LDA, or Latent Dirichlet Allocation, is a probabilistic model used for handling discrete data, particularly in the context of document analysis. In this Bayesian mixture model, the assumption is made that topics are independent of each other. Essentially, LDA provides a framework for understanding how documents within a dataset are generated. The model involves assigning an arbitrary number of topics (denoted as K), where each topic represents a probability distribution over a fixed vocabulary. From this perspective, every document is treated as a composite of words, with each word associated with one of the K topics. Moreover, LDA adheres to the "bag of words" approach, treating each word within a document as an independent entity, divorced from its sequential or contextual relationships. The similarity of the topics can be thereafter assessed by creating dendrograms. A dendrogram is a tree diagram frequently used in hierarchical clustering to illustrate the arrangement of clusters formed during the analysis. In the context of data visualization, dendrograms are often employed to represent relationships and similarities between data points (Meyer & Pushmann 2019).

Structural Topic Models (STMs) represent a widely embraced enhancement of conventional LDA models. Distinguishing themselves through versatility, STMs provide an innovative extension by accommodating metadata information for each document in the topic modeling process. This inclusion of metadata enables a richer analysis, where external information can influence the topic modeling outcomes. Additionally, STMs introduce an alternative initialization mechanism known as "Spectral," offering a different approach to the initialization step compared to standard LDA models. Notably, STMs allow the integration of covariates into priors, allowing for a more nuanced exploration of the relationships between document content and additional contextual information. This adaptability makes STM a valuable tool for researchers seeking a comprehensive understanding of topics within a corpus while considering associated metadata and covariates (Meyer & Pushmann 2019). In R, the quanteda package provides a robust platform for implementing both Structural Topic Models (STMs) and Latent Dirichlet Allocation (LDA) in text analysis.

7 Text Mining Analysis of the Commentaries on the Czechoslovak Monetary Policy

This chapter unveils the results derived from the computer-based textual analysis methods, as detailed in Chapter 6[21], applied to three articles of Alfred Amonn, Hugo Müller, and Oskar Engländer, which were also discussed in Chapter 4. These outcomes are subsequently compared with the earlier findings obtained through the manual reading of the texts. Moving forward, Chapter 8 broadens the scope of text analysis to include all digitized texts specified in Table 5.1.

7.1 Word Frequency Analysis

The initial word frequency analysis of Amonn's, Engländer's, and Müller's articles on Czechoslovak monetary policy is visually presented in Figure 7.1 using wordcloud graphics, highlighting the 100 most frequently used stems in each article. These visual representations yield valuable insights. Notably, Czechoslovak personalities are prominently featured in Hugo Müller's article, with Rašín mentioned 17 times and Fořt 39 times. In Müller's text, the most prevalent meaningful words are related to rate (64x), price (58x), exchange (54x), and stabil (53x), emphasizing the central theme of price and exchange rate stability. In Amonn's and Engländer's texts, the most common stems include words like crown, million, and designations of years or months within the top 100 frequencies. The indication of years places Amonn's focus in the twenties, while Engländer's centers around the thirties. Furthermore, the inclusion of the Czech currency (crown) with numerical values (million) suggests numerous references to monetary sums.

[21] From the methods explained in Chapter 6, topic modeling methods like Latent Dirichlet Allocation (DLA) and Structural Topic Models (STM), employing a machine learning approach, necessitate a larger sample of texts. Consequently, they could only be applied to the entire data corpus. Their utility in our computer-based text analysis was assessed by comparing their results with those obtained from previous methods. As DLA and STM did not yield consistent results and do not provide additional information value, their outcomes are only briefly presented on example of Alfred Amonn in Appendix C: Latent Dirichlet Allocation and Structural Topic Models.

Figure 7.1 Three Comments on the Czechoslovak Monetary Policy: Wordclouds

(a) Amonn Alfred: Die Tschechoslowakische Währung und Währungspolitik

(b) Oskar Engländer: Die Devalvation der Tschechoslowakischen Krone im Jahre 1934

(c) Hugo Müller: Zwei Tschechische Schriften über Währungsreform

Despite the relevance of these entries for general text characterization, excluding months, years, and numerical terms (million, millions, billion, billions, tausend, tausends) provides a more focused analysis of Amonn's and Engländer's main themes. The refined results are presented in Table 7.1[22], which depicts the 20 most frequent stems of words for each article after eliminating both arbitrary and self-designated stopwords. Amonn's article reveals stems that appear to be associated with general banking operations, including terms such as payment, bank, cash, lombard, and amount. In contrast, Engländer's article demonstrates stems that lean towards foreign trade and international market operations, featuring terms such as export, import, exchange, foreign, domestic, and country. Notably, Amonn's stem "restric" alludes to the deflationary policy of the Czechoslovak government in the 1920s, while Engländer's "devalue" pertains to the act of devaluation in 1934.

Results for Müller's article did not change after excluding mentions of months, years, and numerals, as his work did not encompass any tables with statistical values. The topic of his article, based on the stems and alongside the focus on Fořt and Rašín, appears to refer similarly to international transactions and operations. Normative judgments may be present in the article, as indicated by stems like "stabil," "maintain," "view," or "premature." However, several stems seem to originate from words forming a phrase (or n-gram), such as exchange rate, foreign currency, means of payment, or banking office. Therefore, expanding the analysis to bigrams can provide additional context, enabling a more comprehensive assessment of the topics addressed by economists in their articles. Additionally, values from the tf-idf analysis on both word and bigram levels could offer a more meaningful basis for comparisons among the authors.

[22] For comparison, further visualisation of the refined results in a form of wordclouds can be found in Figure B.0.1 in Appendix B: Supplementary Graphics and Tables.

Table 7.1 Word Frequency Analysis - Three Comments on the Monetary Policy of the Czechoslovak Government

Amonn: Die tschechoslowakische Wahrung und Wahrungspolitik		Engländer: Die Devalvation der tschechoslowakischen Krone im Jahre 1934		Müller: Zwei tschechische Schriften uber Wahrungsreform	
stem	**count**	**stem**	**count**	**Stem**	**count**
crown	110	crown	154	rate	63
amount	88	price	148	price	58
payment	76	increas	136	exchang	54
mean	68	devalu	123	stabil	53
bank	56	foreign	88	fořt	39
currenc	55	currenc	71	maintain	20
exchang	50	gold	71	polici	20
foreign	42	export	62	view	19
time	38	exchang	55	chang	18
offic	35	rate	52	time	18
rate	34	import	50	currenc	17
increas	29	czechoslovak	45	mean	17
balanc	28	bank	44	rašín	17
cash	28	result	44	artifici	15
level	27	countri	41	econom	15
restrict	26	domest	41	gold	14
economi	24	econom	39	monetari	14
lombard	24	abroad	38	czechoslovak	12
note	24	balanc	38	czechoslovakia	12
separ	20	effect	34	forc	12

7.2 Term Frequency - Inverse Document Frequency Analysis & Analysis of Bigrams

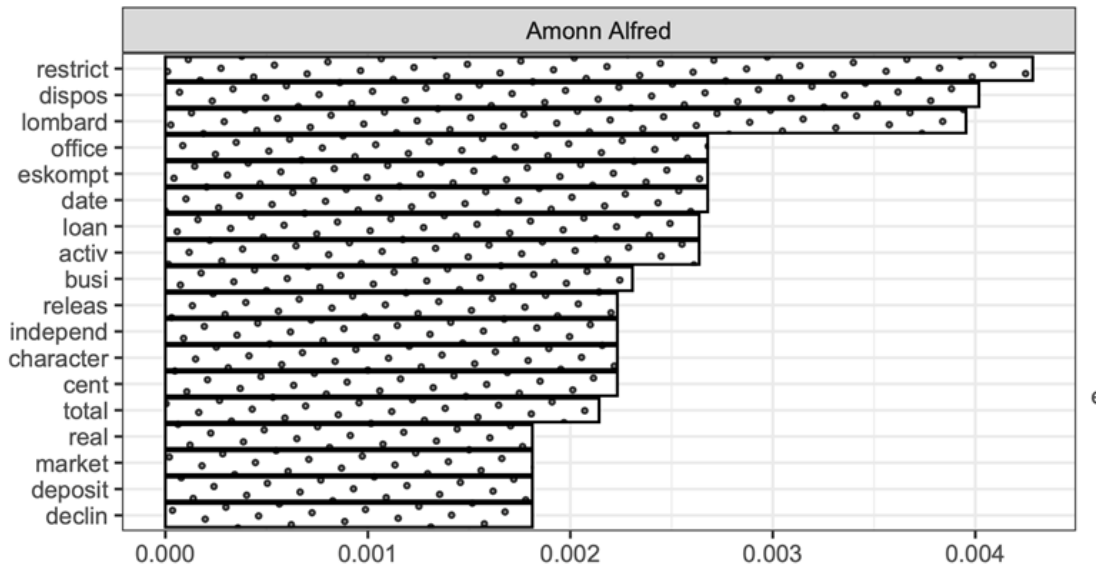
The term frequency-inverse document frequency (tf-idf) analysis of stems and its extension to bigrams enriches our analysis by providing additional context. It also offers an opportunity to validate our previous findings. The application of tf-idf analysis to a small dataset did not effectively eliminate all necessary stopwords. Given that our objective is to compare words that are uniquely significant to each article in relation to others, we employ all previously utilized techniques, including filtering for arbitrary and self-designated stopwords (involving months, years, numbers, and abbreviations like "p," "pp"). The results are depicted in the bar charts in Figure 7.2[23]

Comparing the results to the previous absolute counts reveals significant differences. The tf-idf statistics pinpoint the most impactful words in a given article compared to others, providing insights into what sets it apart. Analyzing Hugo Müller's results in Figure 7.2, we note that the term with the highest tf-idf statistics is "Fořt" – underscoring his distinct focus on the former minister for trade, a feature not shared by the articles of Engländer and Amonn. In addition to "Fořt," other high-tf-idf words include "book," "person," "view," "theory," and "Rašín." This pattern indicates that Müller's article stands out for its emphasis on the perspectives and contributions of Rašín and Fořt, essentially framing his work as a comprehensive review. An intriguing observation is the presence of the word "premature" and the reference to Keynes in Müller's article. The term "premature" is linked to Fořt's perspective on stabilizing the Czechoslovak currency and the government's pursuit of a fixed, stable exchange rate, even if it leads to internal price instability. Müller, in this context, invokes Keynes and his suggestion to prioritize sacrificing exchange rate stability over price stability if necessary.

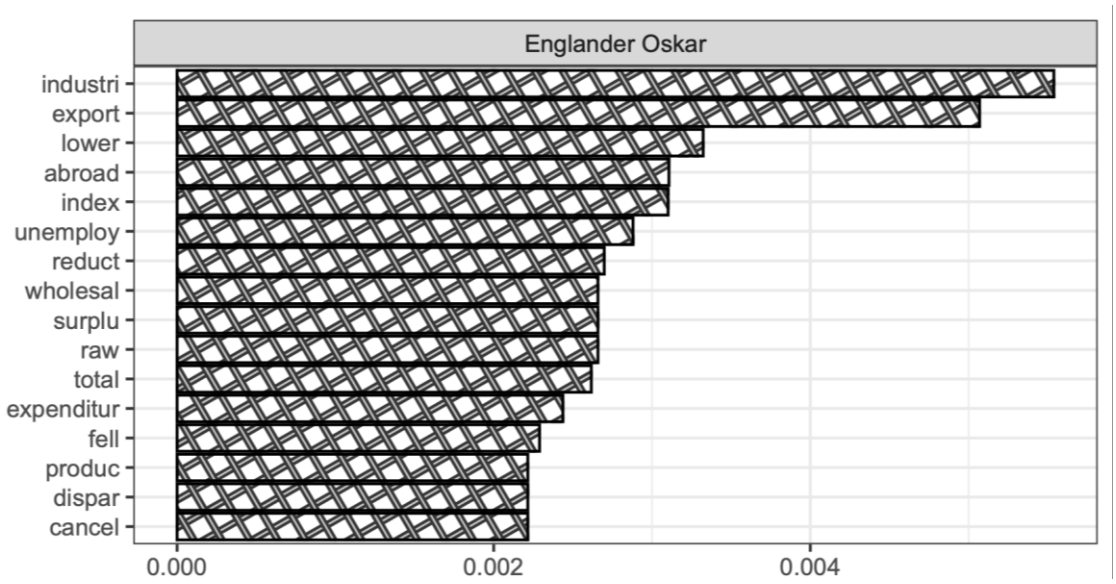
Regarding Amonn and Engländer, the tf-idf analysis not only affirmed but also nuanced our earlier findings. In Amonn's case, the term "restriction" prominently emerges as a distinctive feature that sets the article apart from the others. Upon closer examination, it becomes evident that Amonn deliberately avoids employing the terms "deflation" and "deflatory" in a professional context, opting instead for "restriction" or "contraction." The focal point of the article revolves around the currency policy of the Czechoslovak government in the 1920s, a topic closely linked with the adjective "deflatory." Other high-tf-idf words like "lombard," "eskompte," "loan," and "deposit" further underscore his concentration on banking operations.

[23] For comparison, we present the 15 words with the highest tf-idf statistics. If the lowest tf-idf statistic is the same for multiple words, the number of bars in a graph exceeds 15. Detailed statistics can be found in Table B.1 in Appendix B: Supplementary Graphics and Tables.

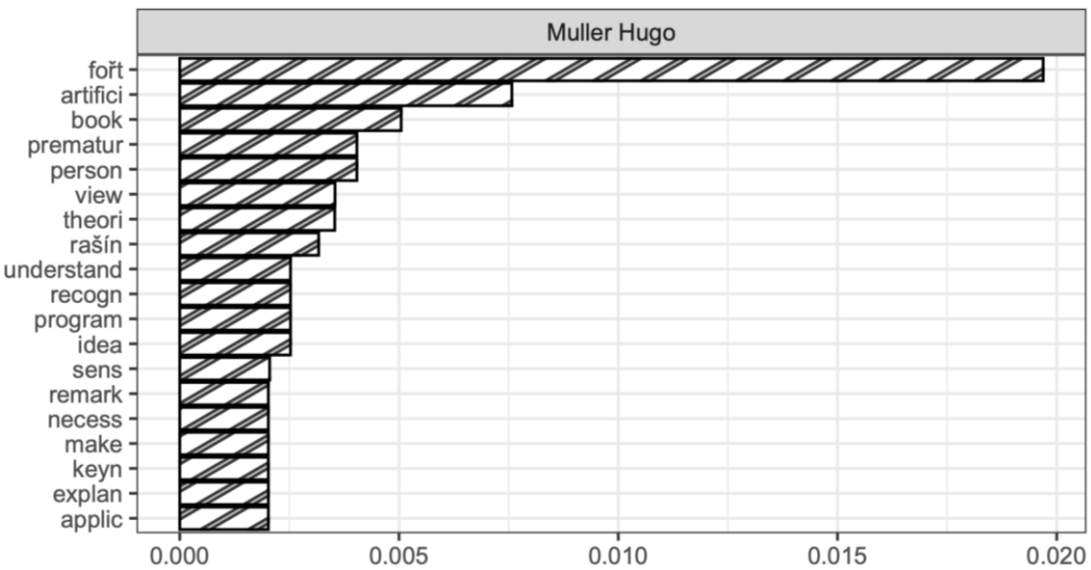
Figure 7.2 Tf – Idf Analysis of Stems



(a) Amonn Alfred: Die Tschechoslowakische Währung und Währungspolitik



b) Oskar Engländer: Die Devalvation der Tschechoslowakischen Krone im Jahre 1934



(c) Hugo Müller: Zwei Tschechische Schriften über Währungsreform

Examining the results of the tf-idf analysis of stems for Engländer's article, the term with the highest tf-idf statistics is "industry." Paired with terms such as “low,” “unemployment,” “disparity,” “fell,” or “reduct,” it suggests that the defining theme in Engländer’s article revolves around events related to the Great Depression.

The tf-idf analysis of stems provided additional context by highlighting the relatively unique themes that appeared in the articles. In the following lines, a frequency analysis of bigrams and their tf-idf analysis are conducted. Table 7.2 depicts the 20 most frequent bigrams found in the analyzed articles for each author. Regarding Amonn’s article, significant bigrams are associated with the first monetary reform, such as "currency separation," "banking office," or "property levy." Other bigrams can be linked to the operations conducted by the banking office for exchange rate management. In Engländer’s case, the bigrams affirm his emphasis on foreign trade. Additional topics include the gold standard ("gold content," "gold currency") and the presence of a national or central bank. For Müller, the bigram frequency is relatively low, with the only notable focus on the exchange rate.

Concerning the tf-idf analysis of bigrams[24], it does not provide significantly new information for either Amonn or Engländer. The primary theme distinguishing Amonn's article, "Die tschechoslowakische Währung und Währungspolitik," according to bigram analysis, is currency separation. Other important bigrams are related to the operations of the banking office. Engländer's bigrams revolve around the reduction of gold content during devaluation and touch on the economic crisis, its symptoms, and consequences in relation to foreign trade. For Müller, two relatively new phrases emerge: "personal principle” vs. “territorial principle." These phrases are linked to Müller’s main critique of Rašín, emphasizing Rašín's failure to discuss alternative approaches or potential negative consequences of his final choices (illustrated, for instance, by the adaptation of territorial over personal principle during currency separation) and how these could be mitigated. “Fořt’s concerns” in bigrams indicate his opinion on the premature stabilization of the crown.

[24] The results can be found in bar-charts in Figure B.0.2 in Appendix B: Supplementary Graphics and Tables.

Table 7.2 Frequency Analysis of Bigrams

Amonn: Die tschechoslowakische Währung und Währungspolitik		Engländer: Die Devalvation der tschechoslowakischen Krone im Jahre 1934		Müller: Zwei tschechische Schriften uber Währungsreform	
bigram	**count**	**bigram**	**count**	**bigram**	**count**
banking office	34	exchange rate	36	exchange rate	52
foreign exchange	20	purchasing power	18	credit policy	6
currency separation	19	foreign trade	16	monetary policy	5
exchange rate	18	foreign exchange	13	czechoslovak monetary	4
national economy	15	price increase	13	economic situation	4
austro hungarian	11	price level	13	gold currency	4
lombard loans	10	czechoslovak republic	12	personal principle	4
precious metals	9	gold content	12	chosen exchange	3
property levy	9	gold currency	12	coin parity	3
credit balances	8	monetary policy	12	financial plan	3
crown rate	7	raw materials	11	fořt's concerns	3
foreign currency	7	czechoslovak crown	9	free currency	3
monetary policy	7	domestic prices	9	price determination	3
banking office's	6	foreign currency	9	territorial principle	3
cover assets	6	power disparity	9	2nd ed	2
disposable means	6	totalled crown	9	art iii	2
rate level	6	austro hungarian	8	artificial exchange	2
real cover	6	national bank	8	artificial rate	2
total amount	6	central bank	7	austria hungary	2
business transactions	5	asset balance	6	austro hungarian	2

7.3 Co-occurrences

We now explore the co-occurrences of words at the paragraph level, displaying the most correlated pairs in a correlation diagram for each article. In these diagrams, connected word pairs are represented by lines, with the darkness of the line indicating the strength of the correlation. A sufficient threshold has to be selected to ensure clarity of the visualizations. The threshold value indicates that pairs with a correlation lower than the chosen threshold are not displayed in the diagram.

The diagrams of Amonn's and Engländer's articles depicted in Figure 7.3 are commented on and compared in the following lines. The threshold for these diagrams was set at $\varphi \geq 0.4$[25]. Concerning Amonn, there is one large cluster of interconnected and highly correlated words, accompanied by one standalone pair of words "Austro" and "Hungarian." This cluster of linked words provides evidence that Amonn builds his interpretation on coherent theoretical concepts interlinked among each other. The most correlated pairs in this group - "national economy," "foreign exchange," "money market," and "currency separation" – are consecutively linked through different notions to construct one coherent system.

For Engländer, the situation is a bit different. Setting the same threshold $\varphi \geq 0.4$, we can see 5 isolated clusters, which, rather than the development of a comprehensive theoretical concept, indicate a sole interpretation of individual topics.[26] Exept for 3 bigger clusters

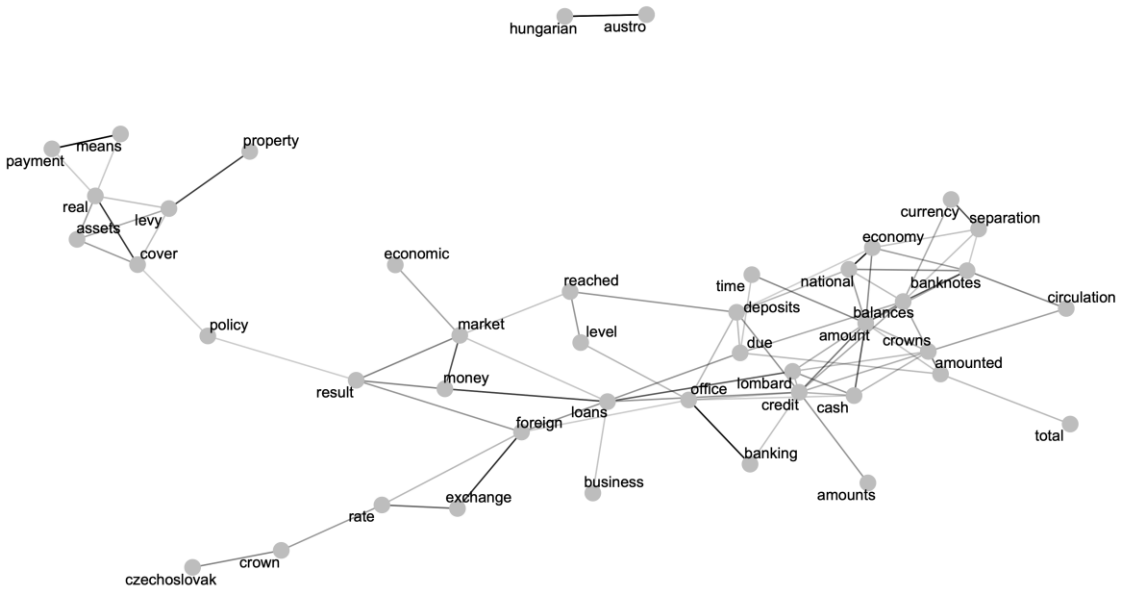
[25] There is no established rule of thumb for determining the threshold when tracking correlations in text analysis. In this study, the author chose an absolute threshold value that aimed for a clear diagram while incorporating as much relevant information as possible. This approach allows for a partial comparison between Ammann and Engländer, offering insights into the variations when adjusting the threshold. Müller's text, characterized by a low degree of correlation, significantly differs from the previous two texts and is, therefore, discussed in isolation. Another approach involves comparing relative values, such as setting the number of edges (i. e. links between two words displayed in diagram) to a specific value. For a number of edges equal to 75, the threshold corresponds to $\varphi \geq 0.4$ for Amonn, $\varphi \geq 0.44$ for Engländer, and $\varphi \geq 0.14$ for Müller. The diagrams with these values, available in Figure B.0.4 in Appendix B: Supplementary Graphics and Tables, align with and support the interpretations presented in the main text of the thesis.

[26] This situation becomes more evident when the threshold is set higher, allowing only Figure pairs of truly highly correlated words to be displayed in the diagram. In this case, raising the correlation coefficient by 5% sets the threshold at $\varphi \geq 0.45$. As can be seen in Figure B.0.3 in Appendix B: Supplementary Graphics and Tables, even with the higher threshold, Amonn constructs an interconnected system, whereas Engländer's diagram is represented by nine isolated clusters.

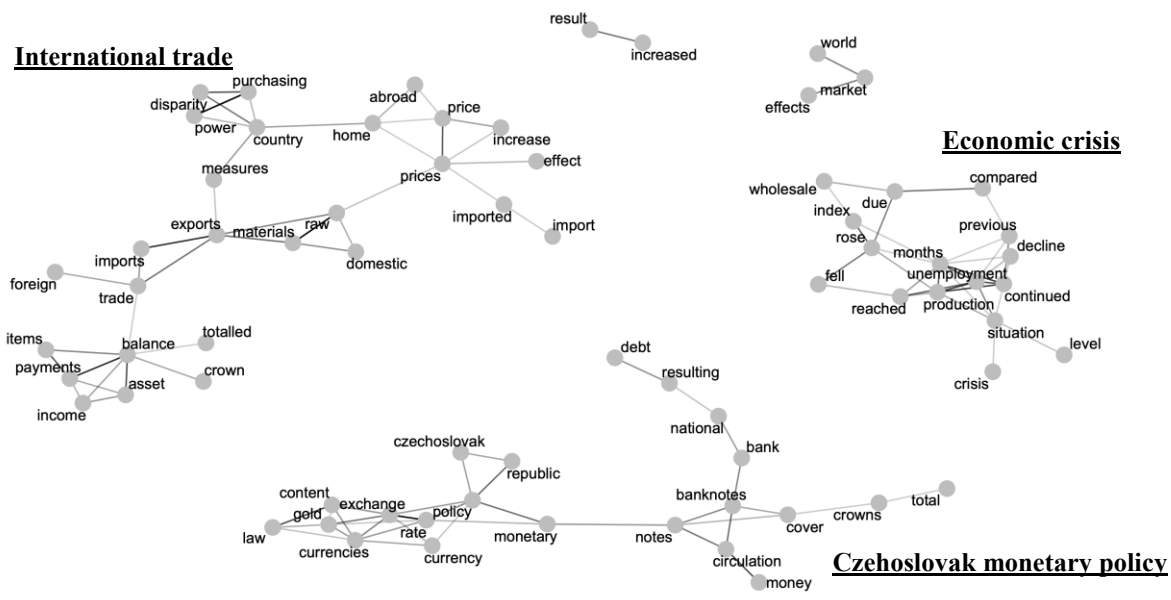
representing topics concerning International trade, Czechoslovak monetary policy and Economic crisis, we can see one pair and one trinity of highly correlated words that are “increased” – “result” and “effect” – “market” – “world”.

Regarding Müller’s article, the overall level of correlation between word pairs is significantly lower than in the previous two articles. To achieve a reasonable visualization of correlated pairs of words, the threshold was set to $\varphi \geq 0.25$. As seen in Figure 7.4, there is one large cluster of points corresponding to Müller's interpretations of Fořt’s theory, and two smaller groups of points, with one of them referring to frequent mentions of Fořt and his assessment of Czechoslovak monetary policy.

Figure 7.3 Pairs of words that show at least 0.4 correlation of appearing within the same paragraph

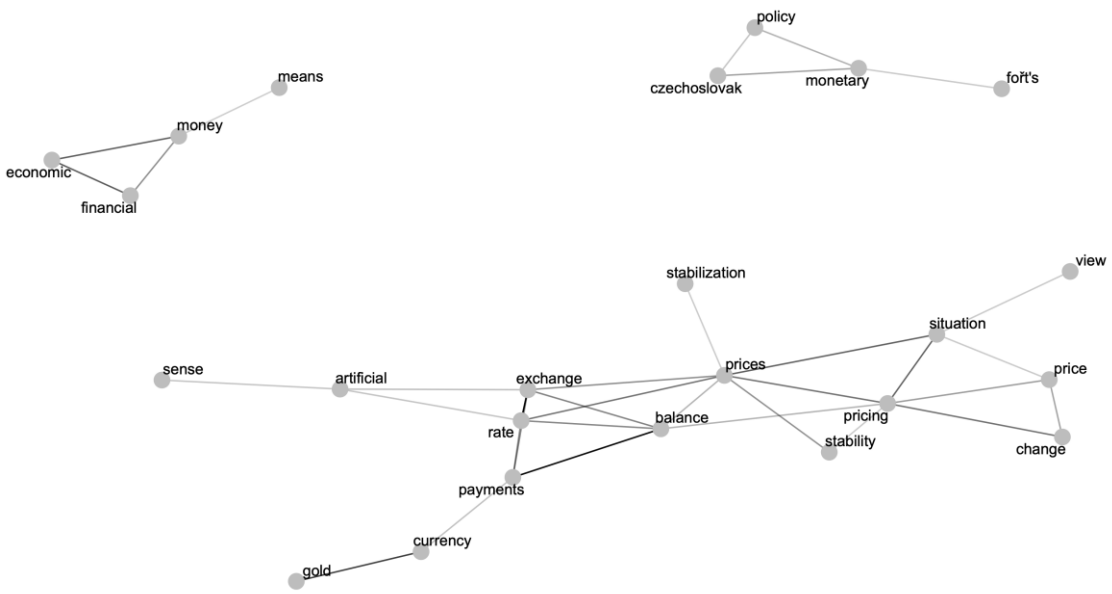


(a) Amonn Alfred: Die Tschechoslowakische Währung und Währungspolitik



b) Oskar Engländer: Die Devalvation der Tschechoslowakischen Krone im Jahre 1934

Figure 7.4 Pairs of words in Hugo Müller's "Zwei Tschechische Schriften über Währungsreform" that show at least 0.25 correlation of appearing within the same paragraph



7.4 Sentiment Analysis

With sentiment analysis, our aim is to discern the sentiments expressed regarding references to Rašín or Engliš in the three selected articles of German-speaking economists and evaluate the possible application of sentiment analysis to historical scientific texts. Since Engliš is not mentioned in any of the analyzed texts, our initial focus will be on understanding the sentiment associated with Rašín in Müller's and Engländer's articles. Given that Amonn's article does not mention either Engliš or Rašín, we will at least examine his general attitude toward Czechoslovakia and evaluate the sentiment surrounding the words "Czechoslovakia," "Czechoslovakian," "Czech," and "Czecholovak" in his article.

In Figure 7.5, the bar charts depict the sentiment surrounding Rašín. The dark bars belong to the paragraphs where the word Rašín (or Rašín's) can be found. For Müller's article, the sum of sentiment surrounding Rašín in each paragraph is as follows:

$$1\text{-}3+1+3+5+8+1+11+1 = 28$$

and can, therefore, be regarded as predominantly positive. Now, let us examine the 4th paragraph, where the sentiment is negative (-3). For demonstration, we will show how the sentiment in this paragraph is determined. The words that contribute to the sentiment in this paragraph are as follows: disregards (-2), **help (2),** certain (1), **shy (-1)**, violence (-3), loved (3), harms (-2), interests (1), protect (1), **want (1),** protect (1), forced (-1). Their sentiment values from the AFINN package are in parentheses, and the words in bold are preceded by negatives such as "not," "no," "never," "without," "cannot," and so on, so their final value will be multiplied by -1. The sentiment of a paragraph is, therefore, the sum of the mentioned values:

$$-2\mathbf{-2}+1\mathbf{+1}-3+3-2+1+1\mathbf{-1}+1-1 = -3.$$

In context, the sentiment evaluation of this paragraph is correct. The paragraph itself is preceded by a standalone sentence beginning with: "At this point, of course, the praise must stop…" and in the paragraph, Rašín is described as "a man of the revolution who did not shy away from violence" and that "he did not take into account the other harms associated with the means he had in mind." It is, therefore, a paragraph where Rašín's personality is criticized the most in Müller's article. Other paragraphs contain a more balanced mixture of positive and negative sentiments. Another negative sentiment can be connected with the second paragraph that describes what Rašín had to fight against after the war (without mentioning his name directly) and it includes words such as "problems," "war," "fight," or "shattered."

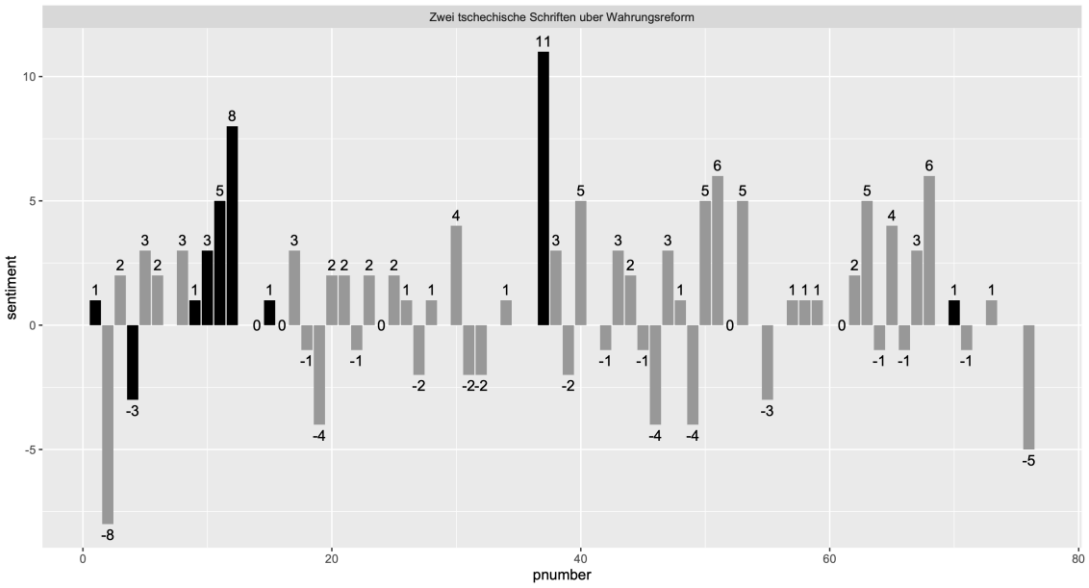
We observe a notably positive sentiment value for paragraph number 37, reaching a score of 11. In this context, Rašín is acknowledged as the architect of an artificially increased exchange rate for the crown during his era, coupled with endeavors to align prices accordingly. The subsequent portion of the paragraph maintains a positive tone, asserting: "*The very fact that it could be sustained for several years* ***without interruption*** *is a clear indication that the adjustment of prices has been* ***successful****, and it can no longer be deemed an artificial rate.*"

The final paragraphs include remarks and references. The last negative reference initiates with the sentence, "The protection of tenants is not placed in the right context by Fořt," and proceeds to critique certain explanations provided by Fořt. Preceding the remarks and references, the last paragraph maintains a positive sentiment with a value of 6. In this article, the overall contribution of Fořt's publication is summarized as positive.

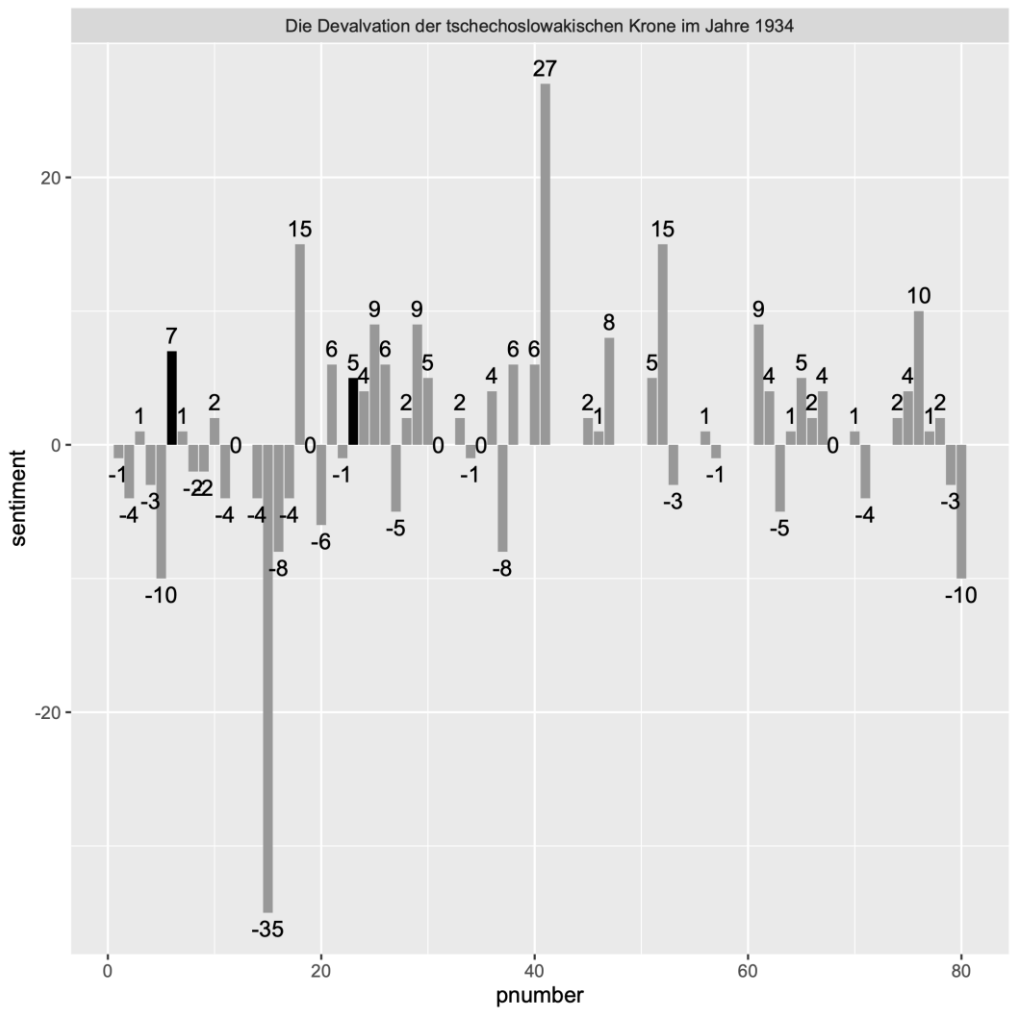
Rašín is briefly mentioned in Engländer's article, and the sentiment surrounding him appears to be positive. The first paragraph where Rašín is mentioned begins with this sentence: "*Alongside these more emotional considerations, however, there were now legitimate economic considerations, albeit perhaps not with complete clarity*." The article then proceeds to clarify these "*legitimate economic considerations*" but ends with a reference to Rašín in the following manner: "*In this way, the monetary policy of the Czechoslovak Republic was steered in a very specific direction towards deflation or at least stabilization, a monetary policy that was then sealed by the death of Rašín, the first Minister of Finance of the Republic*." Therefore, Rašín is directly mentioned only in connection with his death, which could be considered rather negative (the word "death" itself has a sentiment value of -2). However, the prevalence of positive words such as "clarity," "increased," "justified," "granting," "reached," or "growth" in the preceding lines surrounds the mention of Rašín with a rather positive sentiment.

In this article, two significant outliers stand out: a negative value of -35 in paragraph number 15 and a positive value of 27 in paragraph number 41. Paragraph 15 describes the beginning of the Great Depression and, with words such as "weak," "unemployment," "isolated," "lowest," "worsening," "crisis," and so on, correctly conveys a negative sentiment for this paragraph. On the other hand, paragraph 41 begins with the word: "However, the general direction of price movements has a significant influence on the success of devaluation." The word "increase" is present multiple times in this paragraph, ensuring its positive sentiment.

Figure 7.5 Sentiment surrounding Rašín



(a) Hugo Müller: Zwei Tschechische Schriften über Währungsreform



(b) Oskar Engländer: Die Devalvation der Tschechoslowakischen Krone im Jahre 1934

Now, let us examine Figure 7.6, where the bar chart illustrates the sentiments surrounding words related to Czechoslovakia in Amonn's article. The dark bars represent paragraphs where words like "Czechoslovakia," "Czechoslovakian," "Czech," and "Czechoslovak" are mentioned. The total sentiment for each paragraph is calculated as follows:

$$3\text{-}4\text{-}7+4\text{-}2+9+18+2+5\text{-}9=19,$$

indicating an overall positive sentiment. Our focus will be on analyzing outliers, specifically paragraph 31 with a sentiment value of 18 and the last paragraph with a sentiment value of -9.

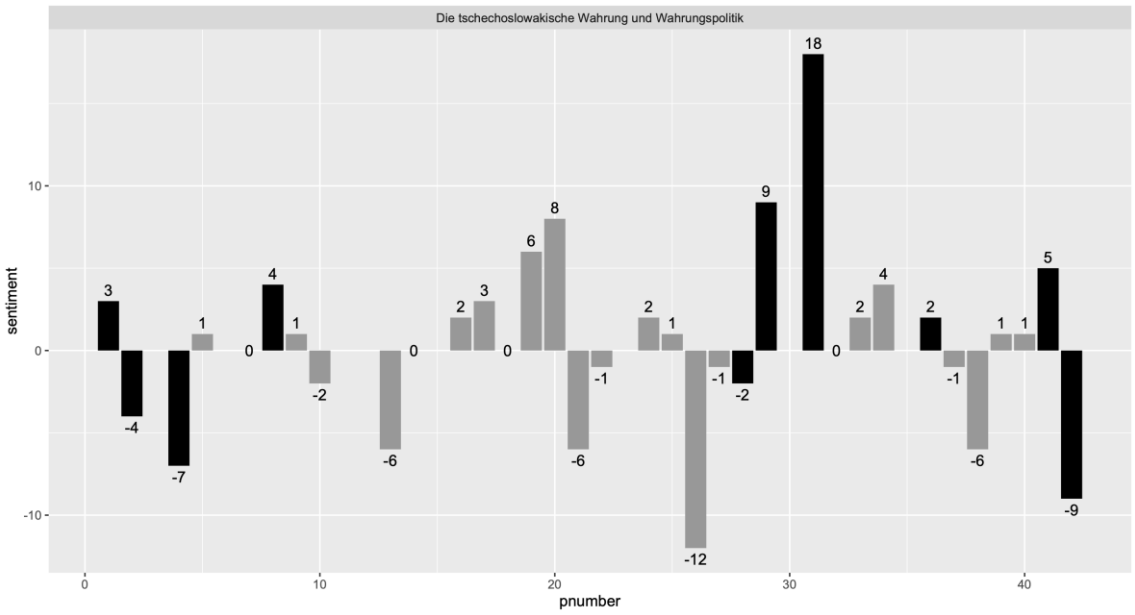
In paragraph 31, Amonn extensively describes the development of the exchange rate of the crown on world markets since September 17, 1921, when it was first listed in New York. Since periods of growth have prevailed, the words "growth" or "rise," having a positive sentiment value, also dominate. On the other hand, in this same paragraph, Amonn also highlights that positive development was due to the positive economic situation rather than because of activities of Czechoslovak banking authorities[27].

The last paragraph is a footnote where Amonn criticizes the inadequate and, for professional debate, inappropriate use of the word "deflation" and proposes to use "restriction" instead—a word which, in itself, has a value of -2 in the used dictionary of sentiment.

The comparison of sentiment analysis results with the insights obtained in Chapter 4 reveals a correspondence with the reader's experience. The analysis effectively assigns a general sentiment to paragraphs, maintaining accuracy in reflecting positivity when discussing possitive events. Nonetheless, caution is warranted when drawing conclusions about the author's opinions solely through sentiment analysis, especially in the realm of professional literature. The presence of positive events in the text does not automatically indicate an uncritical or all-encompassing perspective on the subject or individuals discussed

[27] Examples like these illustrate the necessity of a critical approach and the indispensability of the reader's experience with the text when analyzing the sentiment of technical texts. The sentiment of the paragraph has been correctly identified, as the entire paragraph is about the favorable development of the Crown exchange rate. Thus, it is only fair to say that, in relation to Czechoslovakia, Amonn is rather positive in his attitude towards the country, writing about positive events. Without reading it, however, we would miss certain nuances of restraint and critical evaluation in relation to the Czechoslovak authorities, which were not so common among Czechoslovak economists in the optics of events at that time.

Figure 7.6 Sentiment surrounding Czechoslovakia-related words in Amonn's article



8 Text Mining Analysis of the Whole Text Corpus by German-speaking Economists

In this section, various techniques of textual analysis, including word frequency analysis, TF-IDF analysis, bigram analysis, and co-occurrences, are applied to all the texts listed in Table 5.1. The analyses are conducted and compared at the authorial level to identify the most common topics addressed by economists and to assess the extent to which Czechoslovakia and Czechoslovak Finance Ministers Rašín and Engliš are mentioned. Subsequently, the texts are examined to identify any new occurrences of Rašín or Engliš, and these instances are subjected to sentiment analysis.

8.1 Word Frequency Analysis of the Whole Text Corpus

In Table 8.1, the 20 most frequent stems in the corpus for each economist are presented, and they are further visualized in a form of wordclounds in Figure 8.1, depicting the 100 most frequent stems for each economist. Both in the table and in the wordclouds, it is evident that the most common words differ, particularly for Hugo Müller. Stems derived from words like "exchange," "rate," "money," "foreign," "Franc," "Crown," "country," "domestic," and "currency" collectively indicate a narrow focus on monetary and currency issues in Müller's case. On the other hand, the most common words for Amonn and Engländer encompass a broader range of economic theory interests.

When comparing Amonn and Engländer, it becomes apparent that Amonn exhibits a dominance of concepts related to labor economics (such as labor, wage, entrepreneur, employment), while Engländer's prevalent words, such as price, product, quantity, and margin, reveal a focus on microeconomic topics. For more detailed insights, however, additional approaches such as Tf-Idf analysis and Analysis of Bigrams are considered in the following part.

Table 8.1 Word Frequency Analysis – Whole corpus

Amonn Alfred		Engländer Oskar		Müller Hugo	
stem	**count**	**stem**	**count**	**stem**	**count**
product	2623	price	1040	price	880
econom	1928	product	601	exchang	842
labor	1634	monei	462	rate	663
mean	1363	capit	457	demand	510
price	1243	quantiti	385	monei	492
capit	1165	cost	340	foreign	366
quantiti	1133	labour	335	purchas	330
economi	1085	increas	318	franc	252
wage	903	type	296	determin	241
demand	790	economi	285	crown	240
nation	789	unit	250	amount	238
increas	744	econom	230	chang	231
exchang	732	capitalist	215	suppli	208
incom	685	ratio	210	mean	196
sens	649	determin	200	currenc	193
time	621	mean	200	domest	179
determin	537	result	199	increas	177
social	537	theori	196	countri	172
land	535	entrepreneur	192	power	171
question	528	margin	189	result	170

Figure 8.1 The Word Frequncy Analysis - Whole Corpus: Wordclouds

(a) Amonn Alfred

(b) Oskar Engländer

(c) Hugo Müller:

8.2 Term Frequency-Inverse Document analysis & Analysis of Bigrams (Whole Text Corpus)

Performing Tf–idf analysis provides us with additional insightful information on the dominant topics in each economist's writings throughout the analyzed corpus, revealing distinctive patterns when comparing these three scholars. The results are illustrated in bar charts in Figure 8.2. For Alfred Amonn, the labor topic once again stands out as the most distinguishing factor, affirming our previous observations from word frequency analysis. Amonn, unlike Engländer and Müller, also delves more deeply into the subjects of socialization and taxation, as evidenced by the tf-idf analysis results.

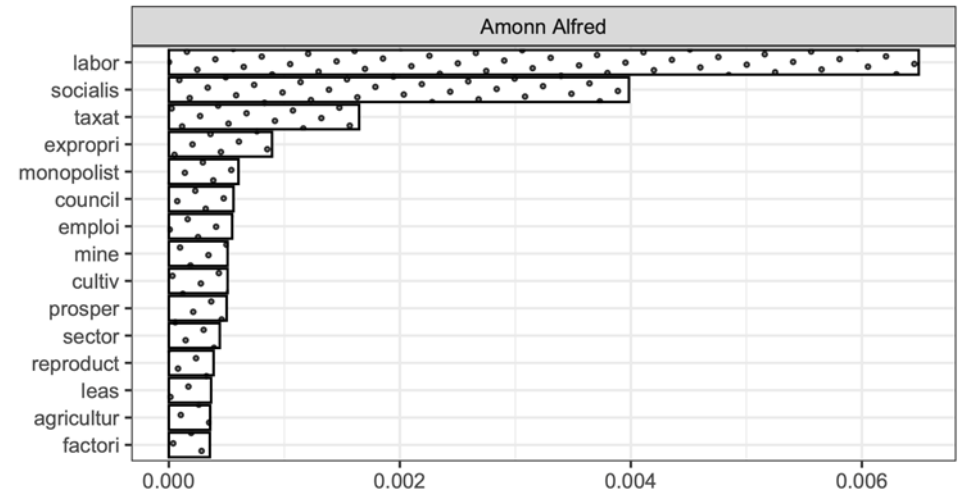
Concerning Oskar Engländer, the tf-idf analysis highlights the topic of transportation, as reflected in words such as "fare," "freight," "passenger," "kilometer," "journey," "velocity," "ticket," and "railway." This focus is unique to Engländer and aligns with his biographical data. Given his background as a former administrative worker at state railways, subsequent director at the private railway Liberec-Jablonec-Tanvald, and later a lecturer in the economics of transport (his first academic role), such outcomes were expected.

For Hugo Müller, the tf-idf statistics don't immediately reveal a strong connection to currency and monetary issues as a distinctive trait of his writings. This is likely because Engländer and Amonn also frequently addressed these topics. However, what stands out is the mention of Fořt, indicating that neither Engländer nor Amonn devoted any attention to him, and Mises. Ludwig von Mises, an economist of the Austrian school, made a groundbreaking contribution in 1912 with the publication of "The Theory of Money and Credit." This work marked the first instance of integrating monetary theory into the core framework of economic analysis, offering novel perspectives on the nature of money and its functions in the economy. This aligns with Müller's known focus on topics related to money and currency, raising the possibility of speculating whether he could be assigned to the Misesian branch of the Austrian school of economics.

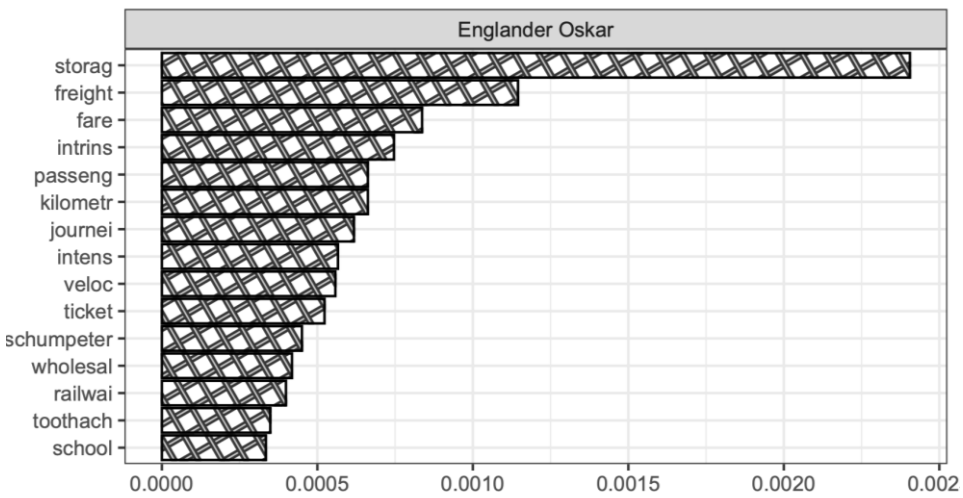
The results of the bigram[28] analysis largely confirm our previous findings and don't provide significant new insights. It's worth noting, however, the relatively frequent mention of Adam Smith in Amonn's texts. This observation indicates that Amonn constructed his economic theories within a broader context, considering the development of the history of economic thought.

[28] The results are also depicted in Table B.3 and Figure B.0.5 in Appendix B: Supplementary Graphics and Tables.

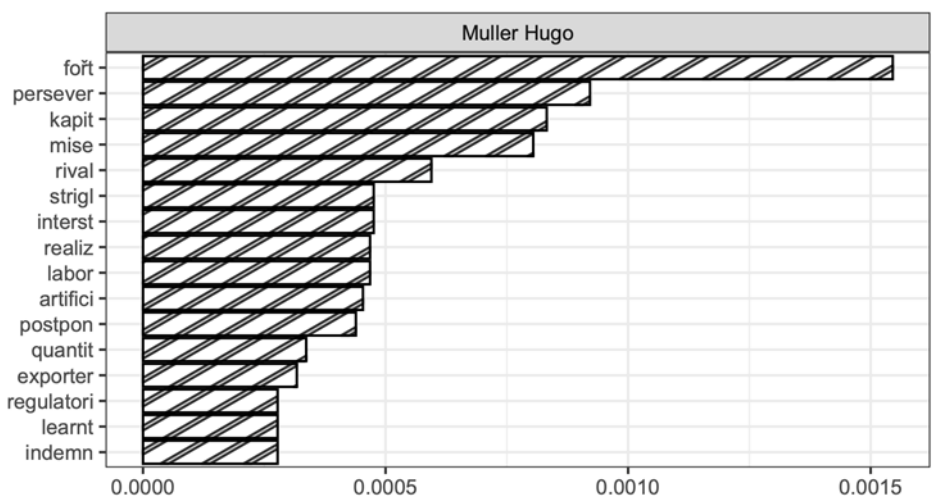
Figure 8.2 Tf - Idf Analysis - Whole Text Corpus



(a) Amonn Alfred



(b) Oskar Engländer



(c) Hugo Müller

8.3 Co-occurrences

The analysis of word pairs' co-occurrences within the same paragraph offers an additional perspective on the overall textual structure for each author. Utilizing pairwise correlation calculations, we visually represent the most correlated pairs in diagrams specific to each economist. Lines connecting these correlated word pairs vary in darkness, indicating the strength of the correlation. For Alfred Amonn and Oskar Engländer, we achieved clear yet detailed visualizations by setting the threshold to $\varphi \geq 0.57$, while for Hugo Müller, a slightly lower threshold of $\varphi \geq 0.4$ was applied[29]. The diagrams exhibit larger clusters, smaller bundles, and isolated word pairs. It is important to note that the distance between these points is arbitrary and does not convey information about the proximity in meaning among individual groups.

In Figure 7.28, Amonn's most correlated word pairs throughout the entire text corpus are depicted. The prominent cluster resembles the core of Amonn's theoretical foundation, featuring interconnected points related to contemporaries such as Karl Diehl, John Maynard Keynes, and Emil Lederer. Interestingly, these surnames are not linked to their first names, suggesting Amonn references them based on their work rather than as personalities. Stand-alone word pairs referring to Adam Smith and Otto Bauer are also present. Keynes occupies a central position, linking to words like "marginal," "Economics," "Classical," or "entrepreneurial," leading to a monetary theme centered around the crown. Czechoslovakia forms another cluster along with terms like territory, Austro-Hungarian, and crisis. Additionally, a smaller cluster relates to taxation. Other correlated words include opposites (active–passive, static–dynamic), common bigrams (raw materials, precious metals), historical references (mercantilists–guild–constitution), and animal kingdom examples (kill –deer - kill – beaver). Although the topic labor is not immediately evident, it could be implicitly represented. Notably, Josef Schumpeter is missing among the names.

For Engländer, the diagram is challenging to interpret, possibly due to the extensive mention of various topics without forming a complex theoretical model. However, interconnected bundles suggest associations with transportation, a notable aspect of his work. References to Schumpeter, Sombart and Bendixen are evident. Surprisingly, Czechoslovakia does not feature prominently among words correlated with others above $\varphi=0.57$, but at lower-level (at $\varphi=0.51$) correlation exists at least for the word czechoslovak.

[29] For comparison also the approach of setting the fixed numer of edges was applied (see Footnote 25) and can be found in **Chyba! Nenalezen zdroj odkazů.**

Müller's diagram indicates references to Böhm-Bawerk's book, connections to Amonn, and to Rašín's book "My Financial Plan," and associations with Fořt and Czechoslovakia. The most numerous interconnected points form a sequential group related to an economic model, encompassing production, labor, transport, and land.

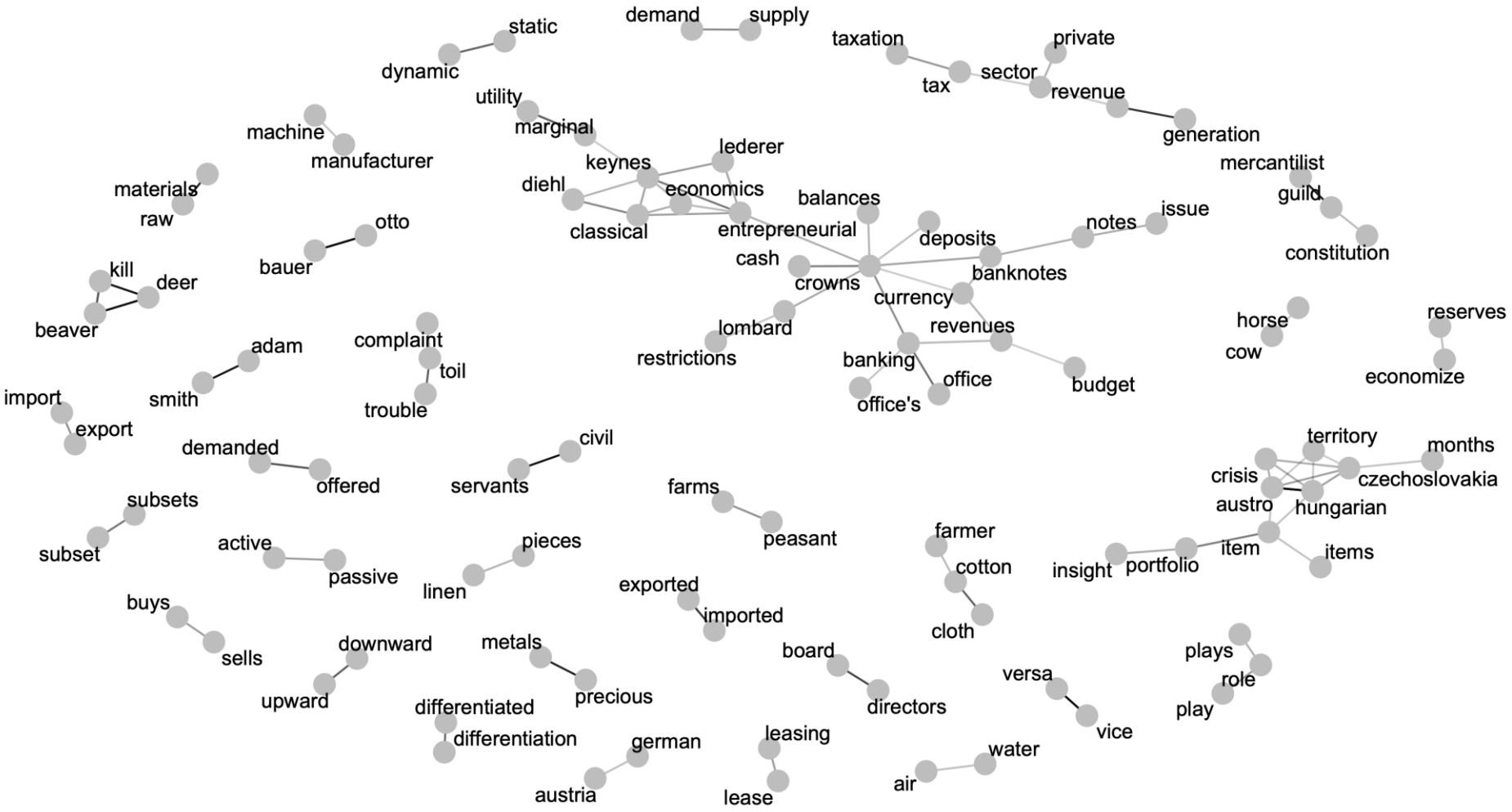


Figure 8.3 Alfred Amonn - Co-occurrences

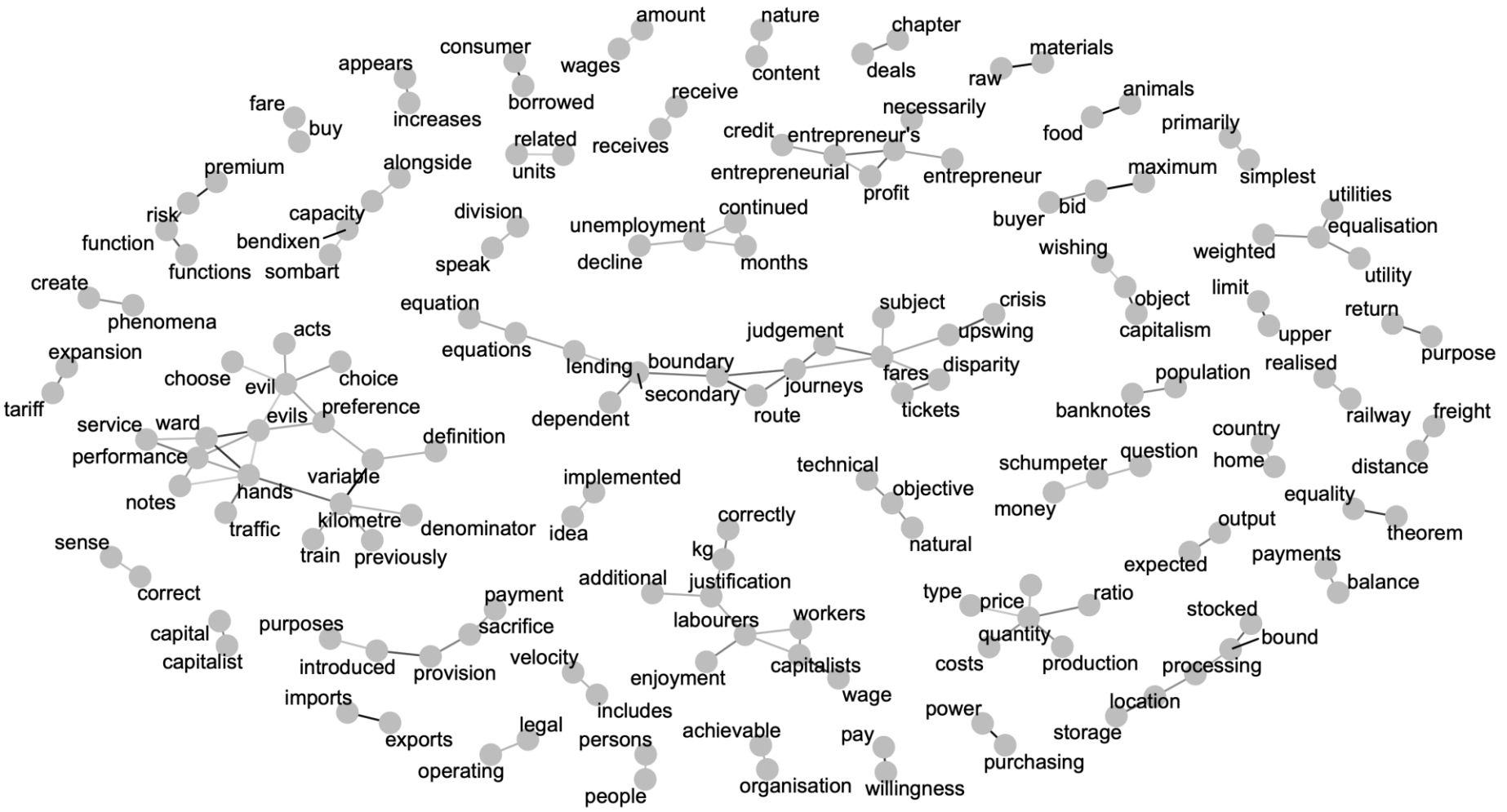


Figure 8.4 Oskar Engländer - Co-occurrences

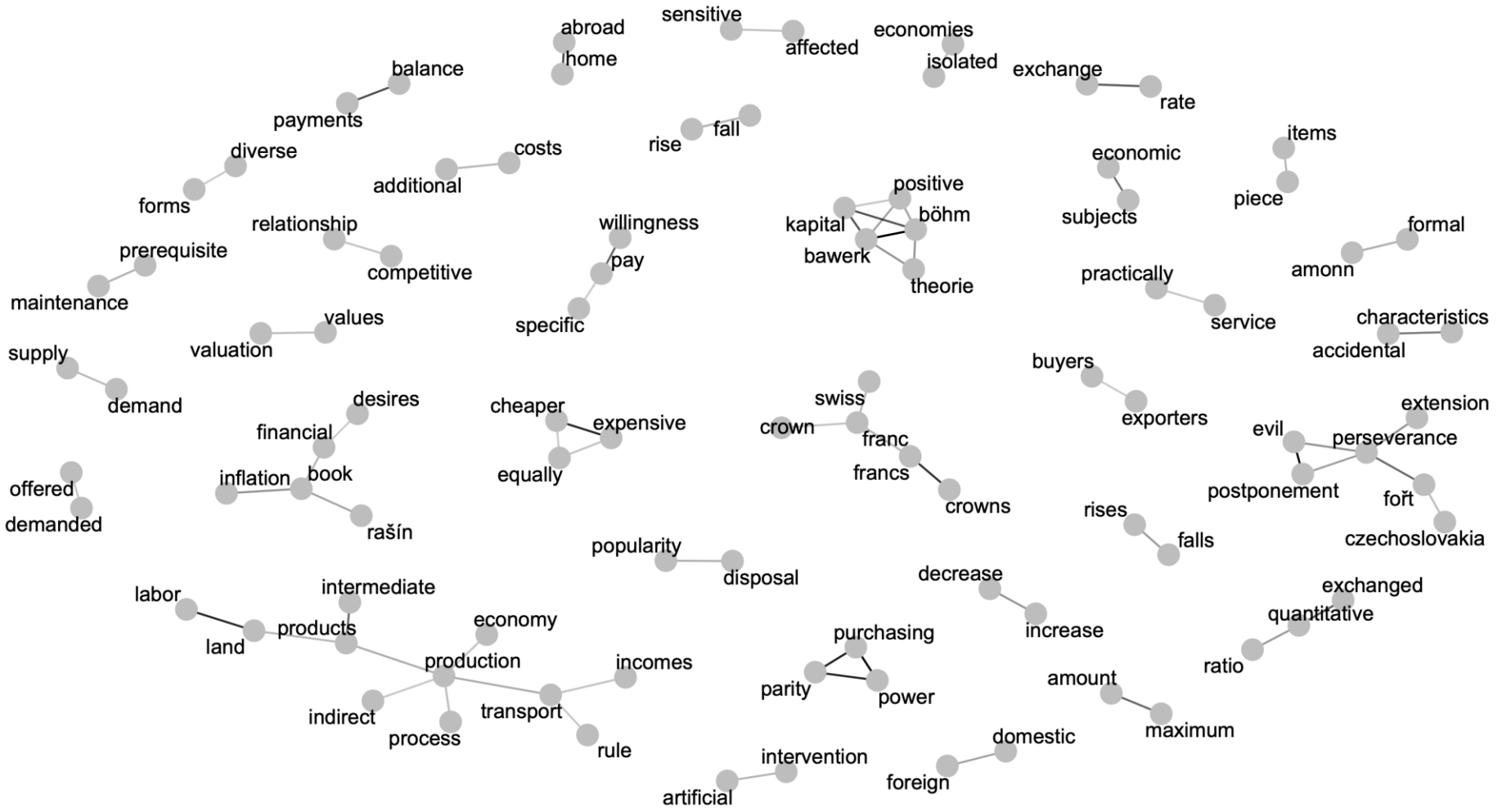
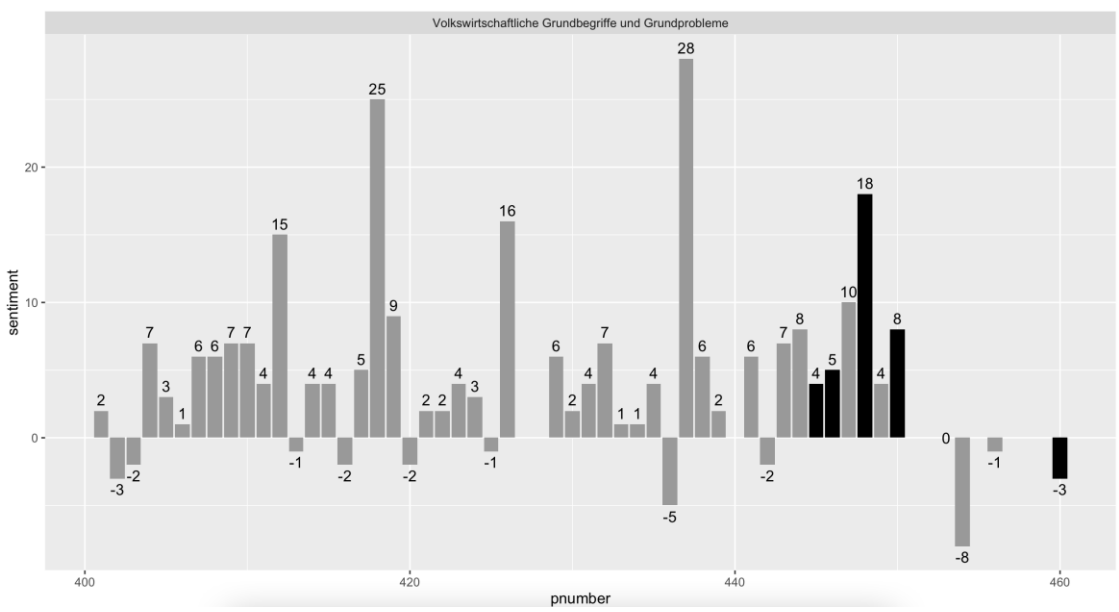


Figure 8.5 Hugo Müller - Co-occurences

8.4 Sentiment Analysis

In this section, sentiment analysis pertaining to Rašín and Engliš in the texts of German-speaking economists is intended. However, during the search for additional mentions of them, only one extra text concerning Engliš[30] was discovered. Rašín is not mentioned in any further texts. A few mentions of Engliš appear in Amonn (1938): "Volkswirtschaftliche Grundbegriffe und Grundprobleme." Engliš appears in this book only in footnotes, mainly in connection to shares in the national product. Nevertheless, Amonn's interpretation of Engliš, albeit confined to footnotes, is quite extensive, and the sentiment surrounding Engliš is predominantly positive, as shown in Figure 8.6,. The figure displays the last few paragraphs of the said book, indicating the paragraphs in which Engliš appears in dark color. Amonn introduces Engliš's statement with the following words: "*An interesting attempt to define a concept, because of the peculiarity of its consideration and method, and at the same time admirable because of its clarity and logical consistency, is that of ENGLIŠ in his 'Teleological Theory of the State Economy'.*" A positive assessment of Engliš's work continues in the next paragraphs. The only instance of negative sentiment is associated with the last paragraph, where Amonn clarifies some of his own definitions in opposition to terms defined by Engliš. Amonn is not criticizing him; he only states that he uses some terms in a different way.

Figure 8.6 Sentiment Analysis of Engliš in Amonn's book „Volkswirtschaftliche Grundbegriffe und Grundprobleme"



[30] The mention of Engliš in Engländer's article, "Die Devalvation der Tschechoslowakischen Krone im Jahre 1934," is not discussed here, as it was only a reference to Engliš's article in *Přítomnost* in the list of references as stated in Chapter 4.

9 Conclusion

This thesis explores the life and work of German-speaking economists in interwar Czechoslovakia. Its structure combines historical research based on archival materials, literature, and contemporary publications with a quantitatively focused text mining analysis of selected German-speaking economists' texts. The main objective was to uncover the relations of the German-speaking economists to the Czechoslovak and its authorities, and examine the potential applications of text mining in the history of economic thought.

Qualitative and quantitative research has yielded new insights. The focus on the three economists is intriguing. Alfred Amonn exhibited a clear connection to Czechoslovakia. Contemporary publications indicate his advocacy for Rašín, and in his academic writings, he engaged with Engliš, expressing a positive view of Engliš's work. Despite not residing in Czechoslovakia, Amonn seemed to have an internal connection to the country. The discovery of Hugo Müller is notable, as he had not been extensively discussed before. His work can also mostly be characterised by a positive view of Czechoslovak economists. Oskar Engländer appears to have a more explanatory focus, particularly regarding transportation, In his work we have not found any extensive references to Czech economists.

The evaluation of textual analysis provides interesting perspectives and serves as a basis for further exploration of economists' work. Methods such as Word Frequency, TF-IDF, Bigrams, and Co-occurrences prove useful for analyzing lengthy or voluminous texts, offering insights and orientation for reading. Sentiment analysis shows promise but requires further validation. Techniques like DLA and STM did not significantly outperform standard Frequency Analysis in quick analyses.

The challenge remains in the digitization of works, a time-consuming process that, once accomplished, enhances the utility of text analysis. The thesis is pioneering, offering avenues for further research. Beyond the three economists analyzed, there is potential to explore additional figures teaching at the German University, as listed in the Appendix A. The text corpus could be expanded to include works from before 1918 and after 1938.

This work lays a solid foundation for subsequent research, providing a fresh perspective on significant economic figures in interwar Czechoslovakia.

Bibliography

Literature

Amonn, A. (1924): Die tschechoslowakische Währung und Währungspolitik. Währungsreform in der Tschechoslowakei und in Sowjet-Russland. München: Duncker & Humblot.

Amonn, A. (1938): Volkswirtschaftliche Grundbegriffe und Grundprobleme. Jena: Gustav Fischer.

Anonym (1923): Deutsches Lob für Dr. Rašín. Prager Presse. 31(1).

Anonym (1932): Der Rašín-Fonds 1932, 48(13).

Bečvář, J. et al. (1998): Dějiny Univerzity Karlovy. IV, 1918-1990. Praha: Univerzita Karlova, ISBN 80-7184-539-6.

Čechurová, J. (1996): Alois Rašín. Praha: Národohospodářský ústav Josefa Hlávky, ISBN 80-270-8131-9.

Čermák, F. (1995): Jazykový korpus: Prostředek a zdroj poznání. Slovo a slovenost, 56(2): pp. 119 – 140.

Doležalová, A. (2007): Rašín, Engliš a ti druzí: československé státní rozpočty v letech 1918-1938. Praha: Oeconomica, ISBN 978-80-245-1356-0.

Doležalová, A. (2018): A History of Czech Economic Thought. London and New York: Routledge. ISBN 978-1-138-91416-2

Engel C. & Bailey S. (2022): Text Analysis with R. https://cengel.github.io/R-text-analysis/

Engländer, O. (1936): Die Devalvation der tschechoslowakischen Krone im Jahre 1934. Weltwirtschaftliches Archiv - Zeitschrift für Allgemeine und Spezielle Weltwirtschaftslehre. 43. Bd., pp. 220-254.

Gehrke, K. (2017): Schumpeter's substitute in Graz (and elsewhere): Alfred Amonn Paper prepared for the 21th Annual Conference of the European Society for the History of Economic Thought at the University of Antwerp, Belgium. https://medialibrary.uantwerpen.be/oldcontent/container42730/files/Gehrke-367381.pdf

Geršlová, J. (1999): Zrození nové doby. Valašské Meziříčí: Vysoká škola báňská - Technická univerzita Ostrava, ISBN 80-7078-729-5.

Havránek, J. et al. (1997): Dějiny Univerzity Karlovy. III, 1802-1918. Praha: Univerzita Karlova, ISBN 80-7184-320-2.

Holman, R. (2001): Dějiny ekonomického myšlení. Praha: C. H. Beck, ISBN: 80-7179-631-X

Jančík, D. & E. Kubů (eds) (2011): Nacionalismus zvaný hospodářský: Střety a zápasy o nacionální emancipaci/převahu v českých zemích (1859 – 1945). Praha: Dokořán, IBSN 978-80-7363-351-6.

Kubů, E. & J. Pátek (2000): Mýtus a realita hospodářské vyspělosti Československa mezi světovými válkami. Praha: Karolinum, ISBN 80-7184-716-X.

Kvirenc, J. & E. Kunstová (2006): České dějiny do roku 1914: historie v dokumentech. Liberec: Dialog, ISBN 80-86761-43-6.

Lacina, V. (1922): Alois Rašín. Praha: Mladá fronta, ISBN 80-204-0303-5.

Meyer, C. & Puschmann, C. (2019): Advancing Text Mining with R and quanteda. Methods Bites - Blog of the MZES Social Science Data Lab, https://www.mzes.uni-mannheim.de/socialsciencedatalab/author/cornelius-puschmann/

Müller, H. (1926): Zwei tschechische Schriften über Währungsreform. Schmollers Jahrbuch für Gesetzgebung. Vol. 50.

Němec, M. (2010): Erziehung zum Staatsbürger?: deutsche Sekundarschulen in der Tschechoslowakei, 1918-1938. Veröffentlichungen zur Kultur und Geschichte im östlichen Europa, Bd. 33. Essen: Klartext Verlag, ISBN 978-3-8375-0065-3.

Novotný, J. & I. Šouša (1996): Banka ve znamení zeleného čtyřlístku: Agrární banka 1911-1938 (1948).Praha: Karolinum, ISBN 80-7184-240-0.

Placht, O. & F. Havelka (1934): Příručka školské a osvětové správy: pro potřebu služby školských a osvětových úřadů a orgánů. Praha: Státní nakladatelství.

Schranil, R. (1936): Oskar Engländer: Festschrift zur Feier des 60. Geburtstages gewidmet von Fachkollegen, Freunden und Schüllern. Brno: Verlag Rudolf M. Rohrer.

Silge J. & Robinson D. (2022): Welcome to Text Mining with R. https://www.tidytextmining.com/

Sojka, M. (2002): KDO BYL KDO – světoví a čeští ekonomové. Praha: Libri. ISBN: 80-7277-055-1.

Sojka, M. (2010): Dějiny ekonomických teorií. Praha: Havlíček Brain Team, ISBN: 978-80-87109-21-2.

Staněk, J. (1995): Peníze v českých zemích. Praha: P.A.S., ISBN 80-901893-1-8.

Ulrich, M. (2001): Bankovnictví. Praha: Credit, ISBN 80-213-0815-X.

Vencovský, F. (1997): Dějiny českého ekonomického myšlení do roku 1948. Brno: Masarykova univerzita. ISBN 80-210-1624-8

Vencovský, F.: (1993): Karel Engliš. Brno: Nadace Universitas Masarykiana, ISBN 80-7028-006-9.

Vencovský, F.: (1997): Dějiny českého ekonomického myšlení do roku 1948. Brno: Masarykova univerzita, ISBN 80-210-1624-8.

Wankhade, M., Rao, A.C.S. & Kulkarni, C. (2022): A survey on sentiment analysis methods, applications, and challenges. Artif Intell Rev 55, pp.5731–5780. https://doi.org/10.1007/s10462-022-10144-1

Wickham H. & Grolemund G. (2017): R for Data Science. https://r4ds.had.co.nz/index.html

Winterberger, G. (1983): „Alfred Amonn und Joseph Alois Schumpeter: zum 100. Geburtstag zweier grosser Ökonomen“. Schweizer Monatshefte: Zeitschrift für Politik, Wirtschaft, Kultur. 63(5): pp. 387-396. DOI: http://dx.doi.org/10.5169/seals-164023.

Yagi, K. (2013). Austrian and German Economic Thought: From Subjectivism to Social Evolution. Great Britain: Taylor & Francis.

Žaloudek, K. (1999): Encyklopedie politiky. Praha: Libri, ISBN 80-85983-75-3.

Archive materials

AUK: Alfred Amonn - Archive of Charles University, Fund of the Faculty of Law of the German University in Prague, personal files: Alfred Amonn

AUK: Hugo Müller - Archive of Charles University, Fund of the Faculty of Law of the German University in Prague, personal files: Hugo Müller

AUK: Oskar Engländer - Archive of Charles University, Fund of the Faculty of Law of the German University in Prague, personal files: Oskar Engländer

NACH: Alfréd Amon - National Archive Chodovec, Fund of Police Directorate Prague II - general registry - 1931 – 1940 [Policejní ředitelství Praha II - všeobecná spisovna – 1931 – 1940], box 4380, sign. A306/10–Alfréd Amonn

NACH: Hugo Müller - National Archive Chodovec, Fund of Police Directorate Prague II - general registry - 1941 – 1950 [Policejní ředitelství Praha II - všeobecná spisovna – 1941 – 1950], box 7466, sign- M2723/1–Hugo Müller

NACH: Oskar Engländer - National Archive Chodovec, Fund of Police Directorate Prague II - general registry - 1931 – 1940 [Policejní ředitelství Praha II - všeobecná spisovna – 1931 – 1940], box 5589, sign. E387/5–Oskar Engländer

Appendix A: A biographies of German economists teaching at the German University in Prague

přeložit

Robert Zuckerkandl

Robert Zuckerkandl was born on December 3, 1856, in the Hungarian town of Győr. He came from a Jewish family and had three brothers who successfully established themselves in their respective fields. One was an anatomist, the second a urologist at the University of Vienna, and the third a prominent figure in German industry. Robert Zuckerkandl himself studied law at the University of Vienna. After completing his studies, he joined the law firm of Dr. Heinrich Jackues, who was a member of the Imperial Council at that time. In September 1886, following his habilitation under Carl Menger, he became a private lecturer in political economy at the University of Vienna. In 1894, he was appointed to the German University in Prague, succeeding Emil Sax, as an associate professor of political economy and collaborator of Friedrich von Wieser. Two years later, he was promoted to a full professor at the same university (Engländer 1926; Schulak & Unterköfler 2011, pg. 57) From 1899-1900 and 1915-1916, he served as the dean of the Law Faculty at the German University. On July 27, 1911, Robert Zuckerkandl was honored with the title of Court Counselor. During the First Republic period, he frequently undertook study trips during the long vacations, especially to Germany. [31] (AUK, PF NU, Zuckerkandl)

As early as 1889, Zuckerkandl published one of his most significant theoretical works, "Zur Theorie des Preises". This work, in fact, constituted his only comprehensive monograph. In this piece, Zuckerkandl "*dedicated considerable space to the historical development of price theory.*" This contribution positioned his work as a bridge between two very distinct schools of economic thought: the German historical school, which appreciated it not only for its historical perspective but also for its "*smooth, clear, and straightforward presentation of ideas*," and the Austrian school of marginal

[31] During the summer of 1921, Zuckerkandl embarked on a study trip abroad, for which he received a state grant of three thousand crowns. Similarly, the following year, he traveled through Germany and Austria, this time with a support of two thousand crowns. In July 1924, he participated in the congress of the Association for Social Policy in Stuttgart, for which he received fifteen hundred crowns.

utility, principally represented by Carl Menger, whom Zuckerkandl studied under. (Engländer 1926)

In the 1890s, he published several articles, particularly in the journal "Jahrbücher für Nationalökonomie und Statistik," which is still in publication today. Before the outbreak of the First World War, he also released several smaller works, mainly addressing the tax and banking situation following the Austro-Hungarian Compromise. However, it can be generally stated that the core of his theoretical publishing activities fell within the period before he joined the German University in Prague. Nevertheless, as Oskar Engländer writes, this apparent publication hiatus was deceptive. "*In reality, he was intensely engaged in a major work meant to conclude his lifelong endeavors. He spoke of this work repeatedly, initially in general terms, later more specifically. The work was supposed to delve into economic production from entirely new perspectives.*" (Engländer 1926)

Robert Zuckerkandl passed away on May 28, 1926, at the age of 70. At the time of his death, he held the position of the chairman of the examination board for political science, the chairman of the examination board for teachers in higher commercial schools, and was also a full member of the German Society of Sciences and Arts for the Czechoslovak Republic in Prague. (AUK, PF NU, Zuckerkandl)

After Zuckerkandl's death, his wife Therese requested permission to relocate from Prague's Smíchov to Jena in Thuringia, where her family, including her adopted daughter Helene, resided at that time. She sought to maintain the full widow's pension during the move. In her application, she stated, "*I was born on December 20, 1861, in Gleiwitz, Upper Silesia, making me 65 years old. All my living relatives are in Jena, Thuringia, including my mother Mrs. Agnes Kern, my sister Mrs. Professor Straub, and the adopted daughter of my late husband, Mrs. Dr. Lang, and her children. In Prague and throughout the territory of the Czechoslovak Republic, I have no relatives, not even a close friend. Therefore, I am entirely alone here and have no one to assist me in case of illness. ... My request for my pension to be sent uncut abroad is primarily justified by the fact that I am a sick, weak lady who requires constant medical care and treatment. I have lost sight in one eye due to cataracts and will have to undergo surgery shortly. ...*" Her request was deemed legitimate, facilitated through Karel Engliš, who knew and respected Robert Zuckerkandl as a "*colleague from the academic sphere*." (AUK, PF NU, Zuckerkandl)

Upon moving to Jena, Therese had a house built by the architect Walter Gropius. The house, where Mrs. Zuckerkandl lived with her adopted daughter's family from 1929, stood in Weinbergstraße and soon became a social hub in the city (STOLPERSTEINE in Jena 2007). However, approximately six years later, in 1933, Therese Zuckerkandl initially approached the rector of the German University with a request not to reduce

her widow's pension and to waive consular fees. Her request was forwarded to the dean of the Law Faculty, Fritz Sander, who advised her to directly contact the president of the republic through the faculty. Ultimately, her request was rejected, citing that "*Mrs. Zuckerkandl, in addition to her pension of Kč 2475 per month, owns a house and land worth 40,000.-Km. Neither the house nor the land is burdened with debts. The said person takes care of herself.*" (AUK, PF NU, Zuckerkandl)

Therese Zuckerkandl took her own life after receiving a deportation order to a concentration camp in 1942 (STOLPERSTEINE in Jena 2007).

Heinrich Rauchberg

Heinrich Rauchberg was born on April 12, 1860, in Vienna. He grew up in his hometown, completed his secondary education, and continued his studies at the University of Vienna, graduating from the Faculty of Law in 1883. Influential throughout his studies and later in gaining his initial work experience was the statistician and economist Inama-Sternegg. From 1884, Rauchberg worked under Inama-Sternegg's guidance at the Central Austrian Statistical Commission. Initially focusing on banking and finance, his interests later expanded to include social issues. In 1890, he was tasked with conducting the census in Austria-Hungary. In 1891, he qualified as a lecturer in statistics at the University of Vienna, and five years later, in 1896, he was appointed as a full professor of statistics at the German University in Prague. At that time, two men taught economics at the Prague university: Robert Zuckerkandl and Friedrich von Wieser. Oskar Engländer, a former student of Rauchberg and later a professor of political economy, recalled , "*Rauchberg quickly distinguished himself among them (referring to Zuckerkandl and Wieser) and could conduct lectures and seminars with lively enthusiasm, bridging the gap between theory and practical application, approaching students adeptly."* Engländer also reminisced about industrial excursions organized by Rauchberg. These excursions, scheduled during the summer semester from 1921 to 1928, involved discussions on economic and technical issues within the industries, as well as labor relations related to working hours and payment methods (Boháč 1930; AUK, PF NU, Rauchberg; Heinrich Rauchberg, Österreichisches Biographisches Lexikon 1815-1950).

Another significant achievement by Rauchberg at the German University in Prague was the establishment of the Political Science Institute. In 1898, the institute found its home in the Clamm-Gallas Palace, later the headquarters of the Ministry of Finance. The number of specialized books in the institute gradually expanded, allowing the institute to continuously advance its scientific activities across all branches of political science, including finance and economics. (AUK, PF NU, Rauchberg)

In 1906, he was awarded the title of Court Counselor. Although his statistical work is arguably his greatest life's achievement, Rauchberg demonstrated his versatility and thorough interest in other legal and political science fields. In addition to statistics, he dedicated himself to administrative law, and in later years, particularly after the establishment of independent Czechoslovakia, he focused his attention on civic education, international law, and minority protection. For the purpose of studying and acquiring modern publications related to international law, he frequently traveled abroad. In 1925, he announced that his son Herbert had completed his studies, and therefore, he was no longer eligible for receiving child allowances. He had to repay the annual excess amount of 2,307.60 Kč in twelve monthly installments. After the death of Professor Zuckerkandl in 1926, Rauchberg was appointed as the chairman of the political science examination commission at the German University. He continued teaching statistics, international law, and civic science in an unchanged manner even after his retirement in 1930. In that year, Rauchberg celebrated his seventieth birthday. He passed away in Prague on September 26, 1938. (Boháč 1930; AUK, Rauchberg; Heinrich Rauchberg, Österreichisches Biographisches Lexikon 1815-1950)

Heinrich Rauchberg was unquestionably an inspiring figure who influenced many of his successors. Testifying to his impact on Czech statistics is Boháč's work (1930), published on the occasion of Rauchberg's seventieth birthday: *"...I cannot refrain from expressing my personal gratitude. He profoundly affected my life. The reading of 'Der nationale Besitzstand in Böhmen' decided my future scientific activities, transforming me from a linguist into a statistician*." The document provides a detailed description and partial evaluation of Rauchberg's work up to 1930. Boháč (1930) notes the nationalist, pro-German mindset that influenced Rauchberg's interpretation of statistical data, especially before 1918. However, he also appreciates Rauchberg's remarkable precision, the quality of his work and analytical thinking, as well as his engaging presentation of statistical data and his willingness to adapt to the new conditions of the First Republic.

Otto Frankl

Otto Frankl belongs to the last representatives of the old generation who continued teaching at the German University in Prague even after the dissolution of Austria-Hungary. Frankl was born on October 4, 1855, in Prague. He studied at universities in Prague, Göttingen, and Leipzig, and in 1883, he began teaching as a private lecturer, becoming a full professor at the German University in Prague in 1891. Otto Frankl was more of a jurist than an economist, focusing particularly on mining law, commercial and negotiable law, as well as bankruptcy law. During his tenure at the Prague university, he was also appointed as a Court Counselor. Additionally, jointly with professors Amonn and Rauchberg, he conducted one-hour economic lectures on

modern large-scale enterprises from the summer semester of 1921[32] to the summer semester of 1922, likely associated with the excursions organized by Rauchberg. This reflects Otto Frankl's interest in national economy. Among his works exclusively related to economics, we can mention the article "Zeit- und Mengenkosten in der Industrie Teildr. aus 'Die Anwendbarkeit der Zeit- und Mengenkosten(rechnung) in der Papierfabrikation," published in 1937, presumably posthumously. Otto Frankl passed away on March 26, 1923, in Prague (Singer 2002).

Franz X. Weiss

Franz Xaver Weiss was born on April 18, 1885, in Vienna, into a Jewish family. In 1909, he completed his law studies at the University of Vienna and shortly thereafter began working at the Vienna Commercial Company. Simultaneously, Weiss engaged in scientific work, specifically writing articles and analyses that focused particularly on the works of Böhm-Bawerk and his interest theory. In 1925, he was appointed as a full professor of political economy at the German Technical School in Prague, where he habilitated with his extensive critique of David Ricardo (Schulak & Unterköfler 2011, p. 130; Blumesberger et al. 2002, p. 1449).

In early May 1926, Weiss submitted a request to the German University in Prague to be appointed as a private lecturer in political economy: "*As theoretical economics is my main field of work, it would mean a lot to me if I could be honored and allowed to teach at the German University in Prague, a place where many famous and outstanding representatives of the Austrian school have worked.*" The request was initially approved by Professor Peterka, who found it to be formally in order. However, he noted that the presented lecture schedule was too vague and did not meet the requirements. Despite this, Weiss's request underwent further evaluation. Alfred Amonn, the expert chosen, assessed the expertise of Weiss's habilitation thesis "Produktionsumwege und Kapitalzins." Based on Amonn's positive evaluation and recommendation, the faculty decided to waive the colloquium and lecture exams, and Franz X. Weiss was habilitated as a private lecturer in political economy in October.

On October 21, 1930, Weiss was appointed as a full professor of national economy and financial science at the Faculty of Law of the German University. Simultaneously, he was relieved of his duties as a full professor at the technical school. A year later, he applied to the state for a subsidy to enrich the economic library, which was in a dismal state at that time. However, due to the state's financial situation during the economic

[32] This course was already conducted in the summer semester of 1914 with his then-colleague Professor Spiethoff. The course was free of charge and limited to 50 participants. The timing of the lecture was not fixed; it depended on an agreement with the students.

crisis, the request was rejected. In 1932, he was entrusted with teaching labor law, a subject he took over from Amonn, who had left Prague in 1929 (AUK, PF, Weiss).

Professor Weiss had two children who did not have Czech citizenship until the end of February 1931, yet he had been receiving family allowances for them since 1926 without authorization. Monthly deductions were ordered from his salary until the overpayment was fully repaid. However, Weiss, well-versed in legal regulations, appealed persistently until the authorities acknowledged that it was a procedural error on their part. They conceded that when family allowances were granted to him in 1926, the mistake was on their part. All the deducted salary deductions had to be refunded to Weiss (AUK, PF, Weiss).

On October 5, 1938, Professor Weiss flew to Bern for a quick visit and was supposed to return two days later. However, he couldn't find a return ticket within the specified timeframe, so he wrote a letter of apology to the dean. Weiss believed that the shortage of tickets was caused by the flow of refugees who, after the acute danger had passed, were returning to Prague. He feared that his delay might be interpreted as an attempt to emigrate. The dean surely knows that Weiss could have left long ago if he had wanted to but could not abandon Prague and his work in times of peril (AUK, PF, Weiss). It is not known whether Weiss ever returned to Prague or if this letter served as a formal document of his loyalty and his intention to come back. By November 1938, he was already in Britain. His journey there was labeled as a business trip for scientific reasons in Prague, and from January 1939, he was granted paid leave, which was revoked in March 1940, stating that, according to the latest information, he was in London and probably a Jewish migrant (AUK, PF, Weiss) He passed away on March 19, 1956, in Orpington (Blumesberger et al., 2002, p. 1449).

Camillo Worliczek

Camillo Worliczek was born on July 18, 1892. He simultaneously studied at the Agricultural University in Vienna and law in Vienna and Innsbruck, completing his studies in 1919. From 1929, he served as an unpaid associate professor of law at the German University. He lectured on Czechoslovak agricultural laws, agrarian policy, and later also on cooperatives. In 1937, he was appointed an associate professor at the Technical University in Brno. However, he continued to lecture in Prague, adding courses on business and transport policy and economics. He taught at both universities during the Second World War. "At the beginning of 1940, he was entrusted with the commissioner's administration of the agricultural university in Brno." He remained in Brno until 1945 and passed away in 1951 (Šišma 2004, pg. 134)

Robert Marschner

Robert Marschner was born on July 4, 1865, in Prague. On June 27, 1889, he graduated from the Faculty of Law at the German University. From 1903 onwards, he taught social law and insurance at the German Technical University in Prague, where he became an associate professor in 1915. At the end of April 1920, he applied for venia docendi at the Faculty of Law of the German University for the same field. His habilitation thesis was titled "Die Stellung des Unfallversicherungsgesetzes zu Privaten Schadenersatzrecht" from 1915. Heinrich Rauchberg prepared the assessment of the habilitation thesis and overall academic capabilities. In his assessment, Rauchberg considered Marschner's habilitation thesis sufficient, taking into account his other works and, most importantly, his extensive teaching experience. Therefore, he recommended granting venia docendi to Marschner without the need for a colloquium and lecture for examination. For the academic year 1922/1923, Marschner wanted to skip lectures due to excessive workload – he was also a member of the Supreme Administrative Court. However, the schedule of the German University includes his lectures until the winter semester of 1923/1924, when he took a one-year break. He continued teaching until 1930, although in the last years, after moving to Šluknov, he lectured according to an extraordinary schedule for the first two weeks of each semester every day. On June 30, 1930, Marschner wrote a letter to the dean's office, announcing that he had suffered a concussion after the last lecture of the summer semester. Due to his current health condition, he was uncertain if he could continue teaching. His final lectures took place in the winter semester of 1930/1931. Even after leaving the academic sphere, he wanted to stay in touch and had the rector's speeches and lecture schedules sent to him. [Archive of the University of Cologne, Faculty of Arts, Marschner] He passed away in Šluknov on September 8, 1934 (Marschner, Robert Anton, Österreichisches Biographisches Lexikon 1815-1950).

Vahan Totomianz

Vahan Fomic Totomianz was born on February 2, 1875, in Astrakhan. He was one of the leading figures involved in cooperative movements. Teaching in Moscow, he became a professor at the Russian Institute for Cooperation in Prague. In the winter semester of 1922/1923, he conducted lectures on cooperatives at the Faculty of Law of the German University in Prague. However, he couldn't complete the lectures in this semester (he is not listed in the schedule for this semester), and he also conducted lectures in the summer semester of 1923 (AUK, PF, Totomianz). From 1925 onwards, he led lectures on cooperation (Genossenschaftswesen) at the Berlin School of Commerce (Berlin im Jahr 1875, Lexikon von A-Z zur Berlingeschichte und Gegenwart). He passed away in France on May 9, 1964 (Vahan Fomich Totomianz, Find a Grave).

Gustav Flusser

Gustav Flusser was born on March 18, 1885, in Rakovník. In 1902, his entire family moved to Prague. Flusser attended lectures on mathematics and physics at the German Technical High School and the German Technical University in Prague. Additionally, he attended lectures on philosophy and German literature. In 1908, he passed the teaching exams and, a year later, began teaching at the German Business Academy in Prague, where he taught algebra, calculus, and Czech language. From 1927, he also worked as a substitute professor at the Faculty of Law of the German University, delivering lectures on bookkeeping and correspondence, later expanding to commercial arithmetic and business studies. Despite his unsuccessful attempts to have his lectures recognized as an honored lectureship, he was only rewarded with dormitory fees (Kollegiengeld). On average, his lectures attracted 110 students per semester. In late 1931 and 1932, he sought a reduction in workload due to his busy schedule, but his request was denied. In 1932, he became the director of the German Business Academy (Koeltzsch 2007; AUK, PF, Flusser] In early September 1939, Gustav Flusser was arrested. He was initially transferred from the Prague prison to the Dachau concentration camp and later from Dachau to Buchenwald, where he died on June 18, 1940 (Koeltzsch 2007).

Armin Spitaler

Armin Spitaler was born in 1898 in Prague. After studying law in Prague and Vienna, he initially worked as a candidate in the state financial service. In 1922, he moved to Liberec, where he worked in the Chamber of Commerce (Handels- und Gewerbekammer). In 1936, he received a scholarship from the Rockefeller Foundation for studies in New York (Prof. Dr. Armin Spitaler, Institut für Steuerrecht an der Universität zu Köln). From the summer semester of 1932, he taught at the Faculty of Law of the German University as a private lecturer in the theory of public bonds and conducted lectures on taxes. From 1939, he also lectured on economics. In 1941, he became an associate professor for finance and tax law. However, a year later, he had to join the German army, and in 1945, he became a prisoner of war. After his release, he professionally established himself at the University of Cologne, where in 1953, he became a full professor of international finance and tax law. He died of a heart attack in 1963 (Prof. Dr. Armin Spitaler, Institut für Steuerrecht an der Universität zu Köln).

Rudolf Schranil

Rudolf Schranil was born on January 21, 1885, in Mikulášovice in the Šluknov region. He graduated from the Faculty of Law at the German Charles-Ferdinand University in Prague in 1909. During his studies, he became a collaborator at the State Financial

Office in Prague, later working as a financial counselor at the Ministry of Finance in Vienna. From 1918, he taught public law at the University of Vienna, and in 1921, he was called to Prague, where he became an associate professor of financial and administrative law. In 1927, he was appointed a full professor (Rudolf Schranil. Die Martin-Luther-Universität Halle-Wittenberg). In 1936, Rudolf Schranil became the dean of the Faculty of Law and played a key role in the decision to appoint Hans Kelsen, a Prague native of Jewish origin and an expert in legal sciences, to the Charles University. As the dean, Schranil had a deciding vote in case of a tie and (although counted among "reliable nationalists") voted in favor of Kelsen's appointment. "*After the university's fascistization, he was forced to leave Prague.*" (Bečvář 1998, pg. 198) He died on July 22, 1957, in Brühl, Germany (Rudolf Schranil. Die Martin-Luther-Universität Halle-Wittenberg)

Theodor Mayer

Theodor Mayer was born on August 24, 1883, in the village of Neukirchen an der Enknach. He studied at the University of Vienna, particularly at the Institute for Austrian Historical Research. In 1923, he became an associate professor at the German University in Prague and taught economic history at the Faculty of Law . He remained in Prague until 1930. He passed away on November 26, 1972, in Salzburg (Mayer, Theodor, Deutsche Biographie; Mayer, Theodor, Gemeinsamer Verbundkatalog).

Others

Oskar Kraus never taught at the Faculty of Law of the German University. However, he was one of the significant philosophers during the First Republic era, and his philosophical ideas also touched upon economic science. Born on July 24, 1872, into a Jewish family, he later converted to Protestantism. At the German University in Prague, he studied both law and philosophy, with a particular focus on philosophy. Later, he became a professor of philosophy at the Faculty of Philosophy of the German University in Prague. Oskar Kraus was a highly influential figure who bridged Czech and German academic societies with Western thought. He formulated his own a priori theory of values, opposing Marxist theory, and applied it to economics. He is renowned for his criticism of Einstein, the compilation of Bretan's work, and collaboration with Masaryk. During the early days of the occupation, he was interned in a concentration camp. After his release, he emigrated to Britain in 1939, where he passed away on September 26, 1942, due to cancer (Vytěsněná elita: 2. část, *Židovská obec Brno*).

Literature and Other Resources

AUK, PF, Archive of Charles University, Fund of the Faculty of Law of the German University in Prague, personal files: Robert Zuckerkandl

AUK, PF, Archive of Charles University, Fund of the Faculty of Law of the German University in Prague, personal files: Franz Xaver Weiss

AUK, PF, Archive of Charles University, Fund of the Faculty of Law of the German University in Prague, personal files: Vahan Totomianz

AUK, PF, Archive of Charles University, Fund of the Faculty of Law of the German University in Prague, personal files: Heinrich Rauchberg

AUK, PF, Archive of Charles University, Fund of the Faculty of Law of the German University in Prague, personal files: Robert Marschner

AUK, PF, Archive of Charles University, Fund of the Faculty of Law of the German University in Prague, personal files: Gustav Flusser

BEČVÁŘ, Jindřich et al. *Dějiny Univerzity Karlovy. IV, 1918-1990*. Vyd. 1. Praha: Univerzita Karlova, ©1998. 671 s. ISBN 80-7184-539-6.

Berlin im Jahr 1875. *Lexikon von A-Z zur Berlingeschichte und Gegenwart* [online]. [qoat.2015-12-28 Accessed from: http://www.luise-berlin.de/kalender/jahr/1875.htm

BLUMESBERGER, Susanne, Michael DOPPELHOFER a Gabriele MAUTHE. *Handbuch österreichischer Autorinnen und Autoren jüdischer Herkunft: 18. bis 20. Jahrhundert* [online]. Mnichov: K G Saur, 2002, 2008-03-19 [quot.2015-12-28]. ISBN 3-598-11545-8. Accessed from: https://books.google.cz/books/about/Handbuch_%C3%B6sterreichischer_Autorinnen_un.html?id=MOhbAAAAMAAJ&redir_esc=y

BOHÁČ, Antonín. *Prof. Dr. Heinrich Rauchberg k sedmdesátým narozeninám: Professeur Dr. Heinrich Rauchberg son 70e anniversaire*. Praha: nákl. vlastním, 1930.

ENGLÄNDER, Oskar. Robert Zuckerkandl. Warnsdorf: Ed. Strache, 1926.

EUGEN MARIA SCHULAK AND HERBERT UNTERKÖFLER a TRANSLATED BY ARLENE OOST-ZINNER. The Austrian school of economics: a history of its

ideas, ambassadors, and institutions [online]. Auburn, Ala: Ludwig von Mises Institute, 2011 [quot..2015-12-26]. ISBN 978-161-0161-343. Accessed from: https://books.google.cz/books?id=hElf69BAElMC&printsec=frontcover&hl=cs&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

KOELTZSCH, Ines. Gustav Flusser. Biographische Spuren eines deutschen Juden in Prag vor dem Zweiten Weltkrieg. *Flusser Studies* [online]. 2007, **4**(5): 1-13 [quot..2015-12-29]. ISSN 1661-5719. Accessed from: http://www.flusserstudies.net/sites/www.flusserstudies.net/files/media/attachments/gustav-flusser.pdf

Marschner, Robert Anton. *Österreichisches Biographisches Lexikon 1815-1950 Online-Edition* [online]. Vídeň: Verlag der Österreichischen Akademie der Wissenschaften [quot..2015-12-27]. Accessed from: http://www.biographien.ac.at/oebl/oebl_M/Marschner_Robert-Anton_1865_1934.xml

Mayer, Theodor. *Deutsche Biographie* [online]. Německo, 2015 [quot..2015-12-31]. Dostupné z: http://www.deutsche-biographie.de/sfz59557.html

Mayer, Theodor. *Gemeinsamer Verbudkalatog* [online]. Německo, 2015 [quot..2015-12-31]. Accessed from: http://gso.gbv.de/DB=2.1/PPNSET?PPN=079374220&COOKIE=U999,K999,D2.1,E36887b65-10,I0,B9994++++++,SY,A%5C9008+J,,1,,U,,4,,7,,M,H13-15,,17-23,,30,,50,,60-61,,73-78,,88-90,NGAST,R77.240.103.224,FN

Prof. Dr. Armin Spitaler. *Institut für Steuerrecht an der Universität zu Köln* [online]. Německo, 2014, 2015-12-23 [quot..2015-12-31]. Accessed from: http://www.steuerrecht.uni-koeln.de/8751.html

Rauchberg, Heinrich. *Österreichisches Biographisches Lexikon 1815-1950 Online-Edition* [online]. Vídeň: Verlag der Österreichischen Akademie der Wissenschaften [quot..2015-12-27]. Accessed from: http://www.biographien.ac.at/oebl/oebl_R/Rauchberg_Heinrich_1860_1938.xml

Rudolf Schranil. *Die Martin-Luther-Universität Halle-Wittenberg* [online]. 2015 [quot..2015-12-31]. Accessed from: http://www.catalogus-professorum-halensis.de/schranilrudolf.html

SCHRANIL, Rudolf (ed.). *Oskar Engländer : Festschrift zur Feier des 60. Geburtstages gewidmet von Fachkollegen, Freunden und Schüllern*. Brno: Verlag Rudolf M. Rohrer, 1936.

SINGER, Isidore. Frankl, Otto. Jewish Encyclopedia: The unedited full-text of the 1906 Jewish Encyclopedia [online]. 2002, 2011 [quot..2015-12-27]. Accessed from: http://www.jewishencyclopedia.com/articles/6319-frankl-otto

ŠIŠMA, Pavel. *UČITELÉ NA NĚMECKÉ TECHNICE V BRNĚ 1849–1945* [online]. Praha: Společnost pro dějiny věd a techniky, 2004 [quot..2015-12-29]. Práce z dějin techniky a přírodních věd: Svazek 2: Svazek 2. Accessed from: http://www.ntm.cz/data/veda-a-vyzkum/publikace/sisma_ucitele_tech_brno.pdf

Therese Zuckerkandl und Dr. Helene Langer. *STOLPERSTEINE in Jena* [online]. Jena, 2007 [quot..2015-12-26]. Accessed from: http://www3.jena.de/stolpersteine/zuckerkandl.htm

Vahan Fomich Totomianz. *Find a Grave* [online]. [quot..2015-12-28]. Accessed from: http://www.findagrave.com/cgi-bin/fg.cgi?page=gr&GRid=95680237

Vytěsněná elita: 2. část. *Židovská obec Brno* [online]. Brno, 2010 [quot..2015-12-28]. Accessed from: http://zob.cz/files2/Vyt%C4%9Bsn%C4%9Bn%C3%A1%20elita2.d%C3%ADl.pdf

WINTERBERGER, Gerhard. Alfred Amonn und Joseph Alois Schumpeter : zum 100. Geburtstag zweier grosser Ökonomen. *Schweizer Monatshefte : Zeitschrift für Politik, Wirtschaft, Kultur* [online]. 1983, 2015-05-26, **63**(5): 387-396 [quot..2015-12-27]. DOI: http://dx.doi.org/10.5169/seals-164023. Accessed from: http://retro.seals.ch/digbib/view?pid=smh-002:1983:63::1236

Appendix B: Supplementary Graphics and Tables

Figure B.0.1 Sensitivity Analysis - Three Comments on the Czechoslovak Monetary Policy: Wordclouds

(a) Amon Alfred: Die Tschechoslowakische Währung und Währungspolitik

(b) Oskar Engländer: Die Devalvation der Tschechoslowakischen Krone im Jahre 1934

(c) Hugo Müller: Zwei Tschechische Schriften über Währungsreform

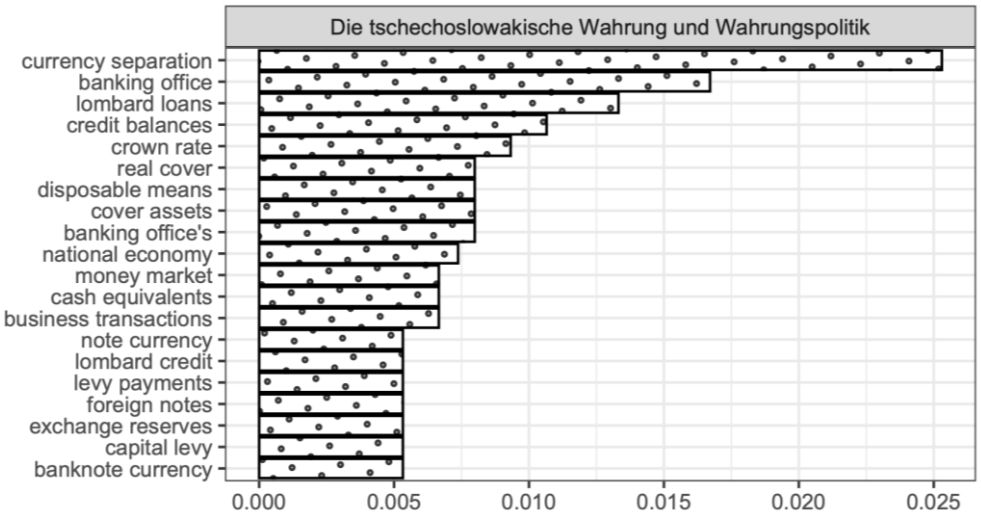
Table B.1 Tf - idf Statistics

Alfred Amonn				
stem	**n**	**tf**	**idf**	**tf_idf**
restrict	26	0.0106	0.405	0.00428
dispos	9	0.00366	1.10	0.00402
lombard	24	0.00975	0.405	0.00395
date	6	0.00244	1.10	0.00268
eskompt	6	0.00244	1.10	0.00268
office	6	0.00244	1.10	0.00268
activ	16	0.00650	0.405	0.00264
loan	16	0.00650	0.405	0.00264
busi	14	0.00569	0.405	0.00231
cent	5	0.00203	1.10	0.00223
character	5	0.00203	1.10	0.00223
independ	5	0.00203	1.10	0.00223
releas	5	0.00203	1.10	0.00223
total	13	0.00528	0.405	0.00214
declin	11	0.00447	0.405	0.00181
deposit	11	0.00447	0.405	0.00181
market	11	0.00447	0.405	0.00181
real	11	0.00447	0.405	0.00181

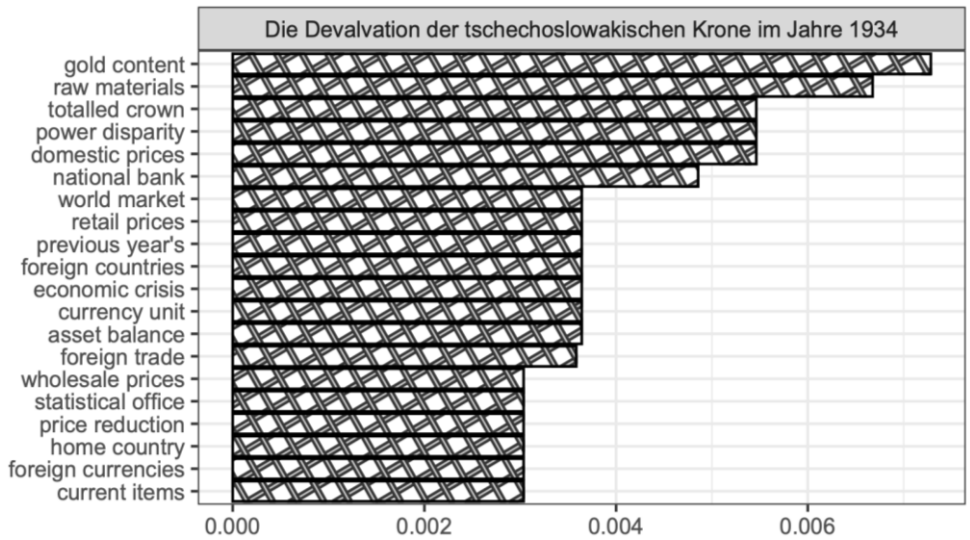
Oskar Engländer				
stem	**n**	**tf**	**idf**	**tf_idf**
industri	25	0.00504	1.10	0.00554
export	62	0.0125	0.405	0.00507
lower	15	0.00302	1.10	0.00332
abroad	38	0.00766	0.405	0.00311
index	14	0.00282	1.10	0.00310
unemploy	13	0.00262	1.10	0.00288
reduct	33	0.00665	0.405	0.00270
raw	12	0.00242	1.10	0.00266
surplu	12	0.00242	1.10	0.00266
wholesal	12	0.00242	1.10	0.00266
total	32	0.00645	0.405	0.00262
expenditur	11	0.00222	1.10	0.00244
fell	28	0.00565	0.405	0.00229
cancel	10	0.00202	1.10	0.00222
dispar	10	0.00202	1.10	0.00222
produc	10	0.00202	1.10	0.00222

Hugo Müller				
stem	**n**	**tf**	**idf**	**tf_idf**
fořt	39	0.0179	1.10	0.0196
artifici	15	0.00688	1.10	0.00756
book	10	0.00459	1.10	0.00504
person	8	0.00367	1.10	0.00403
prematur	8	0.00367	1.10	0.00403
view	19	0.00871	0.405	0.00353
theori	7	0.00321	1.10	0.00353
rašín	17	0.00779	0.405	0.00316
idea	5	0.00229	1.10	0.00252
program	5	0.00229	1.10	0.00252
recogn	5	0.00229	1.10	0.00252
understand	5	0.00229	1.10	0.00252
sens	11	0.00504	0.405	0.00204
applic	4	0.00183	1.10	0.00201
explan	4	0.00183	1.10	0.00201
keyn	4	0.00183	1.10	0.00201
make	4	0.00183	1.10	0.00201
necess	4	0.00183	1.10	0.00201
remark	4	0.00183	1.10	0.00201

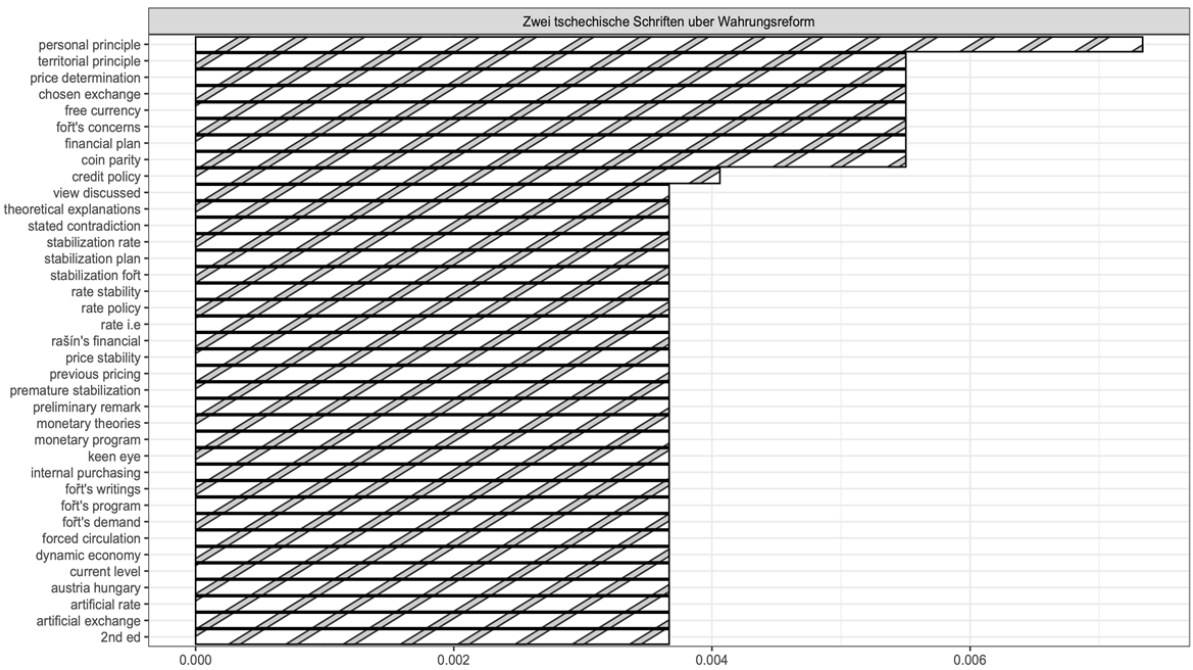
Figure B.0.2 Tf - Idf Analysis of Bigrams



(a) Amonn Alfred: Die Tschechoslowakische Währung und Währungspolitik

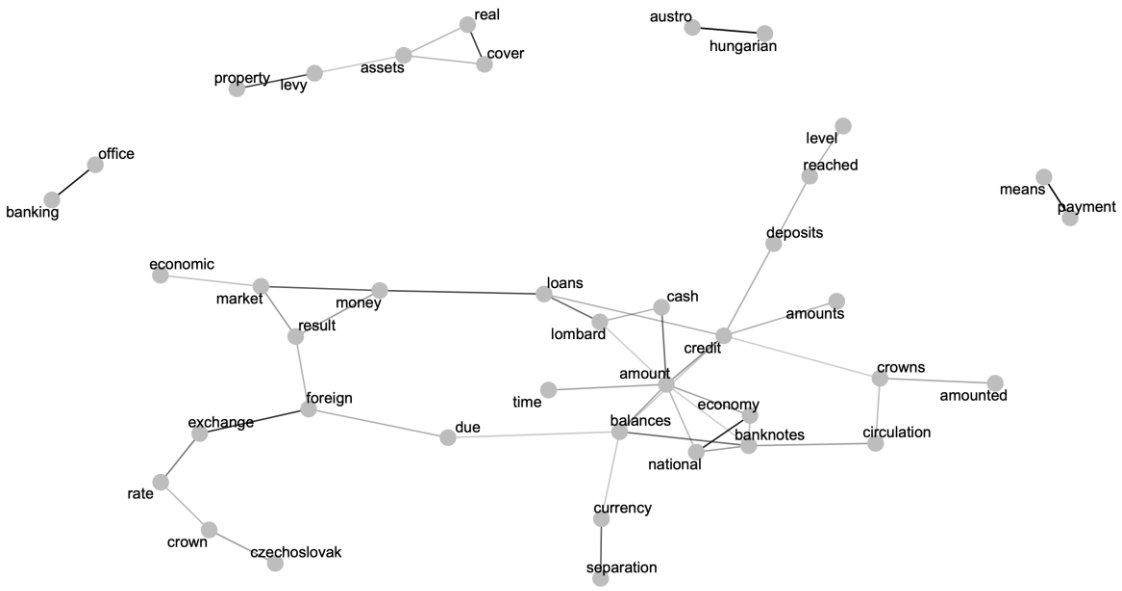


b) Oskar Engländer: Die Devalvation der Tschechoslowakischen Krone im Jahre 1934

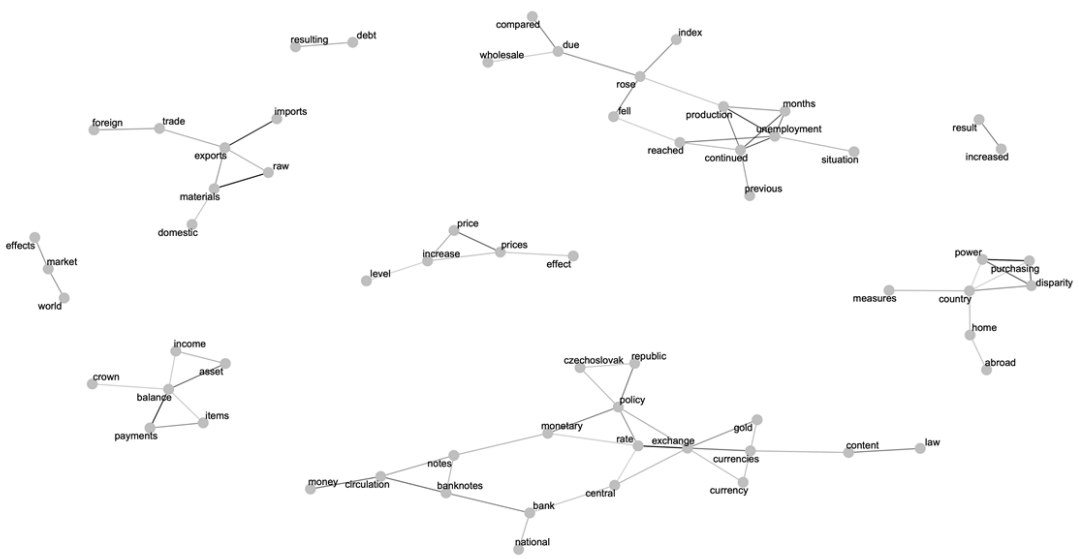


(c) Hugo Müller: Zwei Tschechische Schriften über Währungsreform

Figure B.0.3 Pairs of words that show at least 0.45 correlation of appearing within the same paragraph

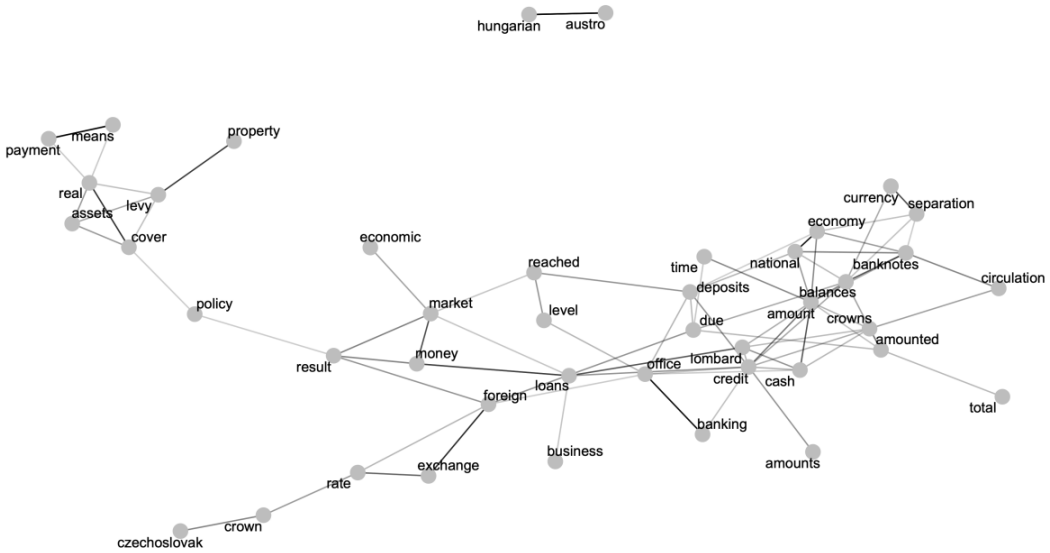


(a) Amonn Alfred: Die Tschechoslowakische Währung und Währungspolitik

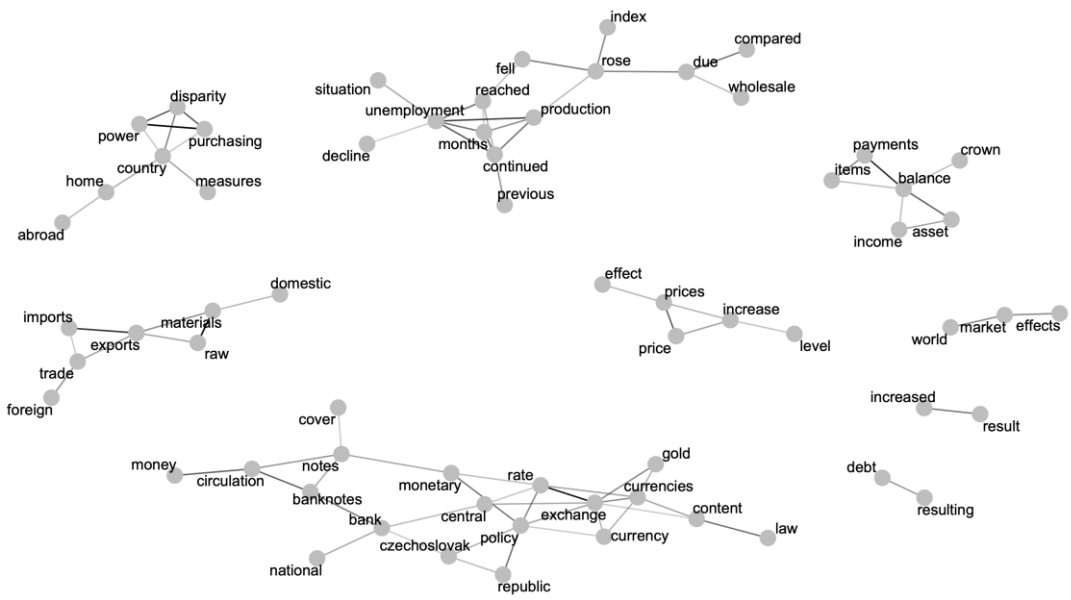


b) Oskar Engländer: Die Devalvation der Tschechoslowakischen Krone im Jahre 1934

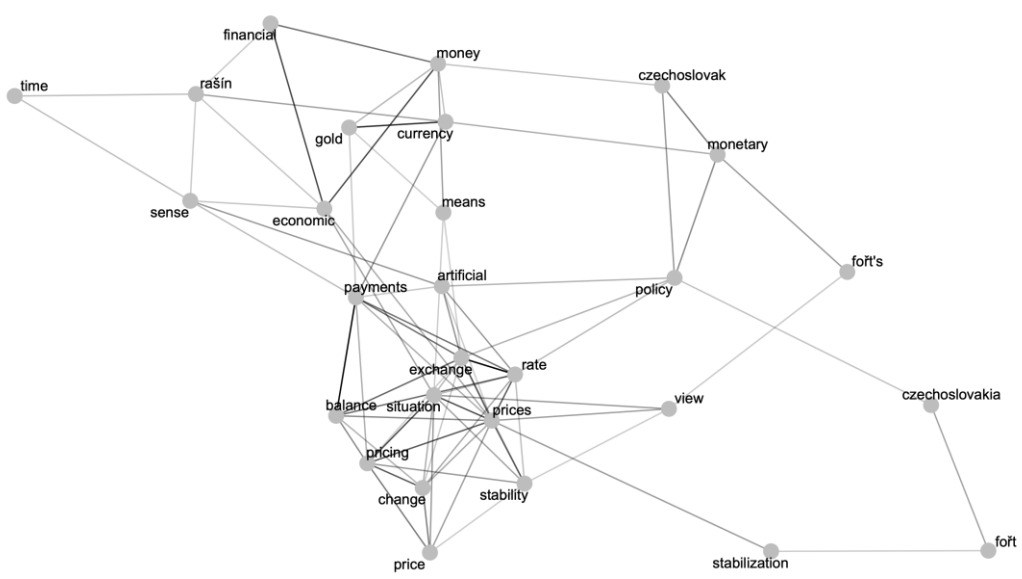
Figure B.0.4 Co-occurrence, the number of edges = 75



(a) Amonn, φ≥0.4



(b) Engländer, φ≥0.44



(c) Müller, φ≥0.14

Teble B.2 Tf - Idf Statistics - Whosle Text Corpus

Alfred Amonn				
stem	**n**	**tf**	**idf**	**tf_idf**
labor	1634	0.0160	0.405	0.00649
socialis	370	0.00363	1.10	0.00398
taxat	153	0.00150	1.10	0.00165
expropri	83	0.000813	1.10	0.000894
monopolist	56	0.000549	1.10	0.000603
council	52	0.000510	1.10	0.000560
emploi	138	0.00135	0.405	0.000548
cultiv	128	0.00125	0.405	0.000509
mine	128	0.00125	0.405	0.000509
prosper	126	0.00123	0.405	0.000501
sector	111	0.00109	0.405	0.000441
reproduct	36	0.000353	1.10	0.000388
leas	34	0.000333	1.10	0.000366
agricultur	90	0.000882	0.405	0.000358
factori	33	0.000323	1.10	0.000355

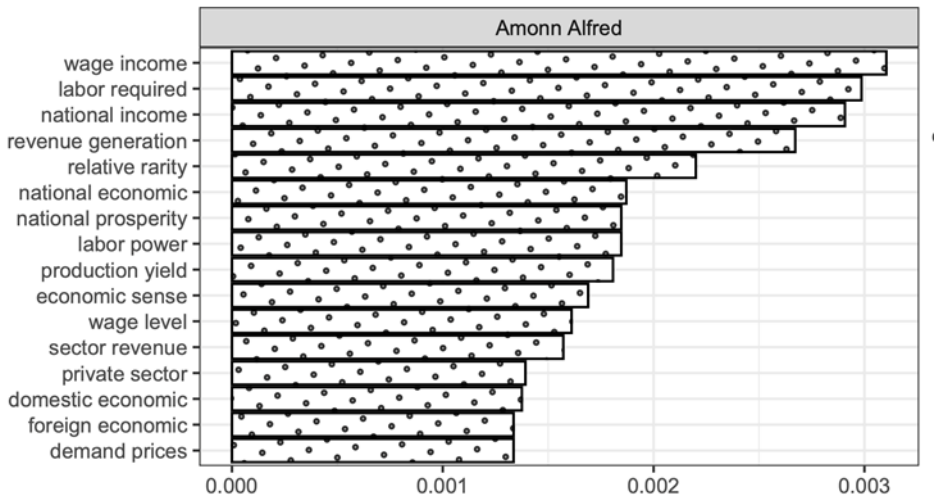
Oskar Engländer				
stem	**n**	**tf**	**idf**	**tf_idf**
storag	69	0.00219	1.10	0.00241
freight	89	0.00282	0.405	0.00115
fare	24	0.000762	1.10	0.000837
intrins	58	0.00184	0.405	0.000746
kilometr	19	0.000603	1.10	0.000662
passeng	19	0.000603	1.10	0.000662
journei	48	0.00152	0.405	0.000618
intens	44	0.00140	0.405	0.000566
veloc	16	0.000508	1.10	0.000558
ticket	15	0.000476	1.10	0.000523
schumpeter	35	0.00111	0.405	0.000450
wholesal	12	0.000381	1.10	0.000418
railwai	31	0.000984	0.405	0.000399
toothach	10	0.000317	1.10	0.000349
school	26	0.000825	0.405	0.000335

Hugo Müller				
stem	**n**	**tf**	**idf**	**tf_idf**
fořt	39	0.00141	1.10	0.00155
persever	63	0.00227	0.405	0.000921
kapit	21	0.000758	1.10	0.000832
mise	55	0.00198	0.405	0.000804
rival	15	0.000541	1.10	0.000594
interst	12	0.000433	1.10	0.000476
strigl	12	0.000433	1.10	0.000476
labor	32	0.00115	0.405	0.000468
realiz	32	0.00115	0.405	0.000468
artifici	31	0.00112	0.405	0.000453
postpon	30	0.00108	0.405	0.000439
quantit	23	0.000830	0.405	0.000336
exporter	8	0.000289	1.10	0.000317
indemn	7	0.000253	1.10	0.000277
learnt	7	0.000253	1.10	0.000277
regulatori	7	0.000253	1.10	0.000277

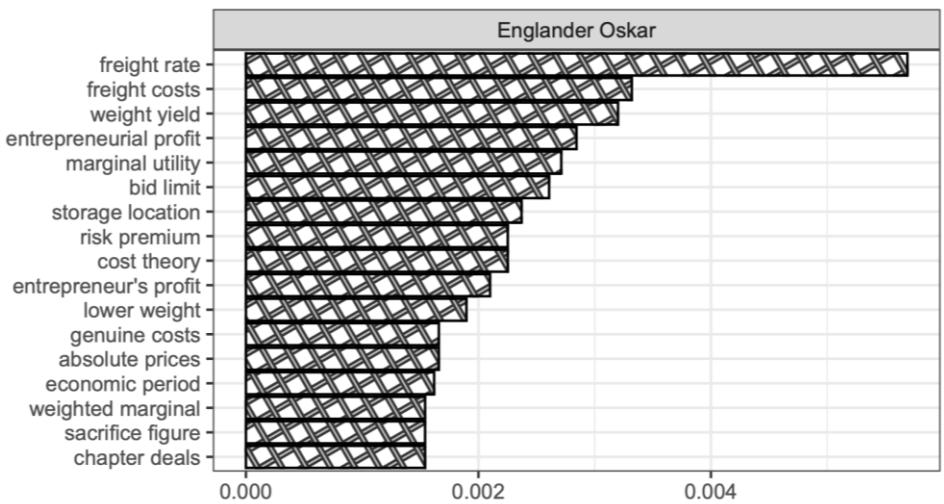
Table B.3 Frequency Analysis of Bigrams - Whole Text Corpus

Amonn		Engländer		Müller	
bigram	**count**	**bigram**	**count**	**bigram**	**count**
national economy	322	national economy	80	exchange rate	424
national economic	129	entrepreneurial profit	65	purchasing power	148
private sector	96	price ratio	63	power parity	98
marginal utility	84	marginal utility	62	foreign exchange	84
economic activity	80	entrepreneur's profit	48	maximum amount	82
wage income	79	freight rate	48	foreign money	61
labor required	76	purchasing power	39	exchange values	39
national income	74	economic period	37	economic subjects	35
national product	74	exchange rate	37	static rate	30
banking office	71	money supply	36	foreign currency	27
real wage	69	productive capital	33	price change	25
revenue generation	68	labour wages	32	domestic price	21
adam smith	59	price structure	31	money supply	21
relative rarity	56	maximum bid	29	price level	21
free competition	54	freight costs	28	competitive relationship	20
economic agents	52	lower limit	28	exchange rates	20
natural price	49	upper limit	28	domestic money	19
social economy	48	weight yield	27	concrete money	18
labor power	47	money capital	26	home country	17
national prosperity	47	bid limit	22	indirect exchange	17

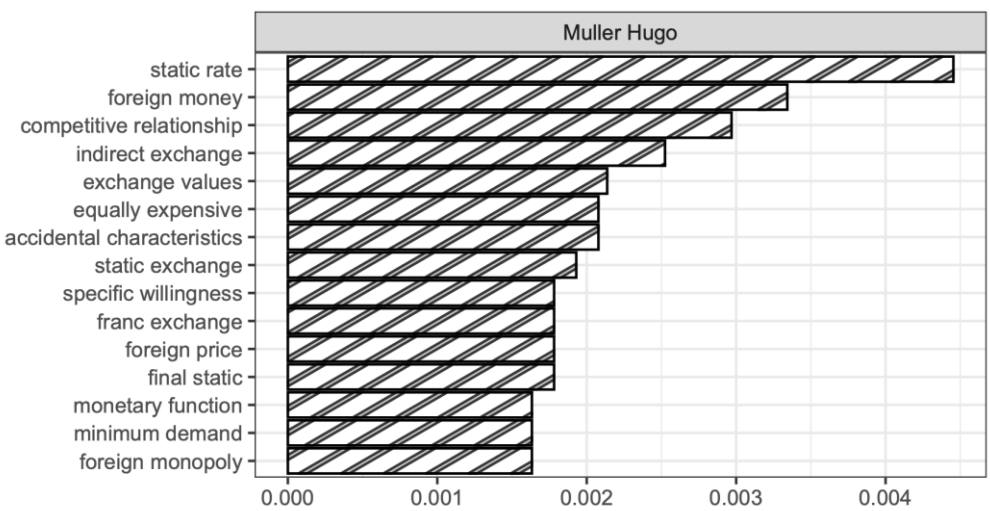
Figure B.0.5 Tf - Idf Analysis of Bigrams - Whole Text Corpus



(a) Amonn Alfred

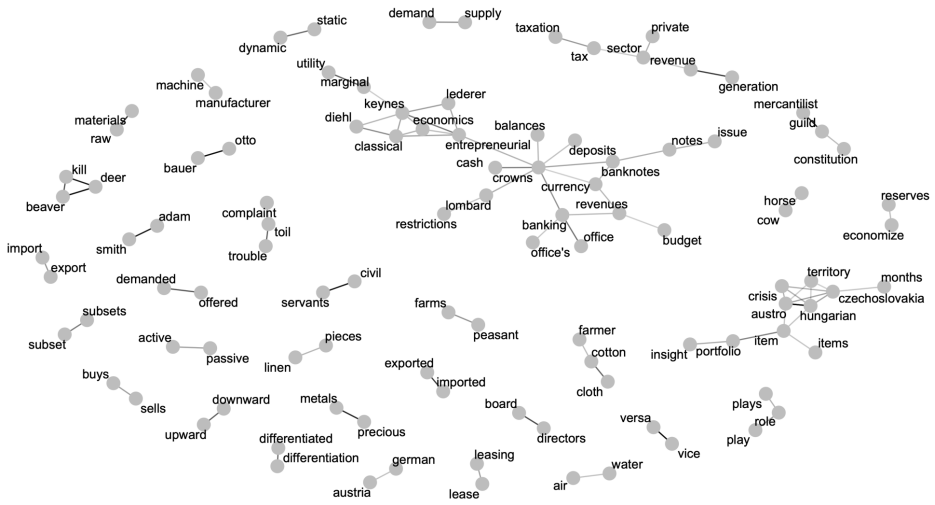


(b) Oskar Engländer

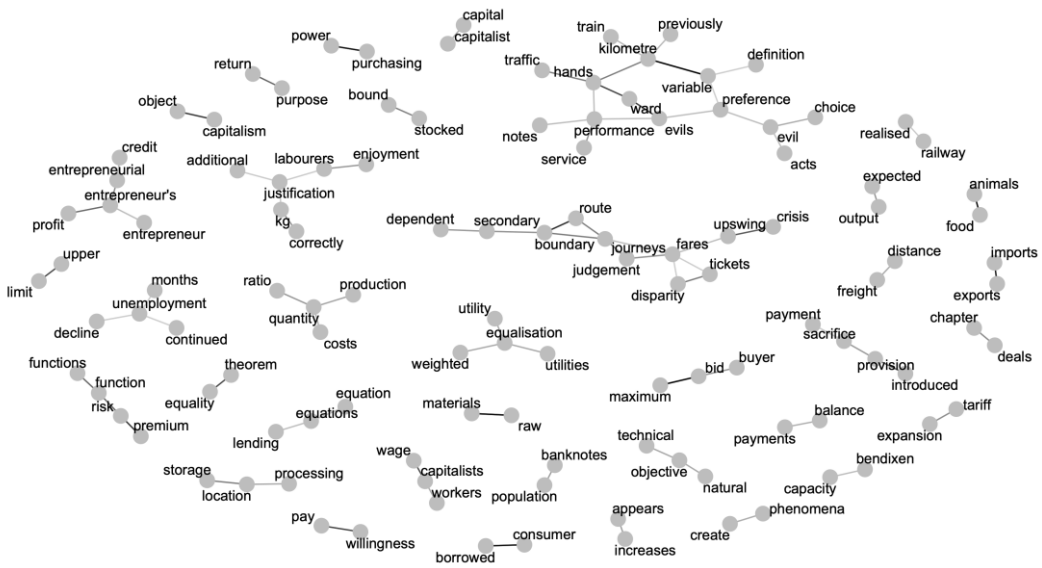


(c) Hugo Müller

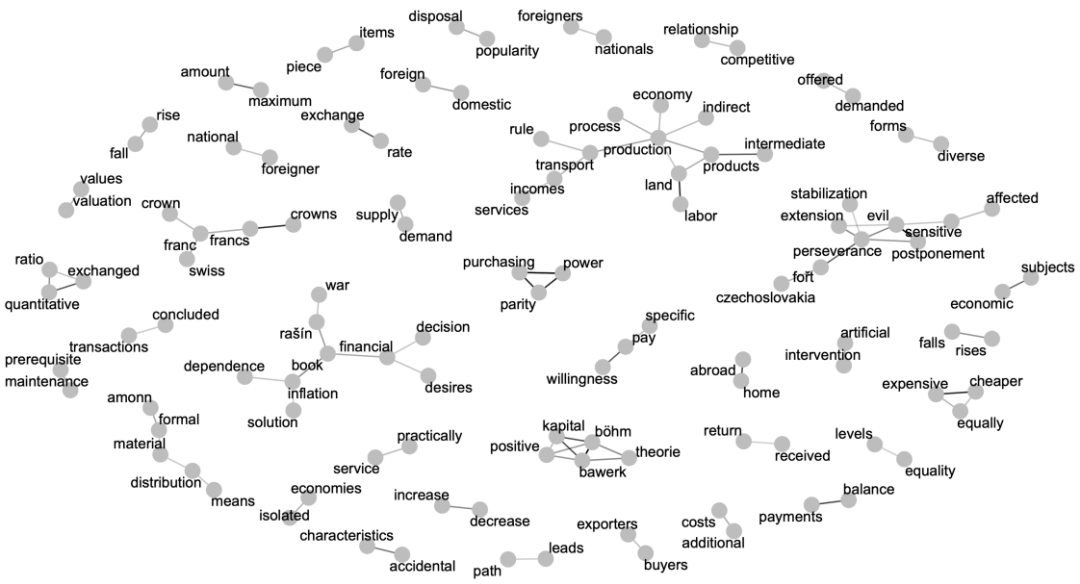
Figure B.0.6 Co-occurrence - number of Edges= 85



(a) Amonn, φ≥0.57



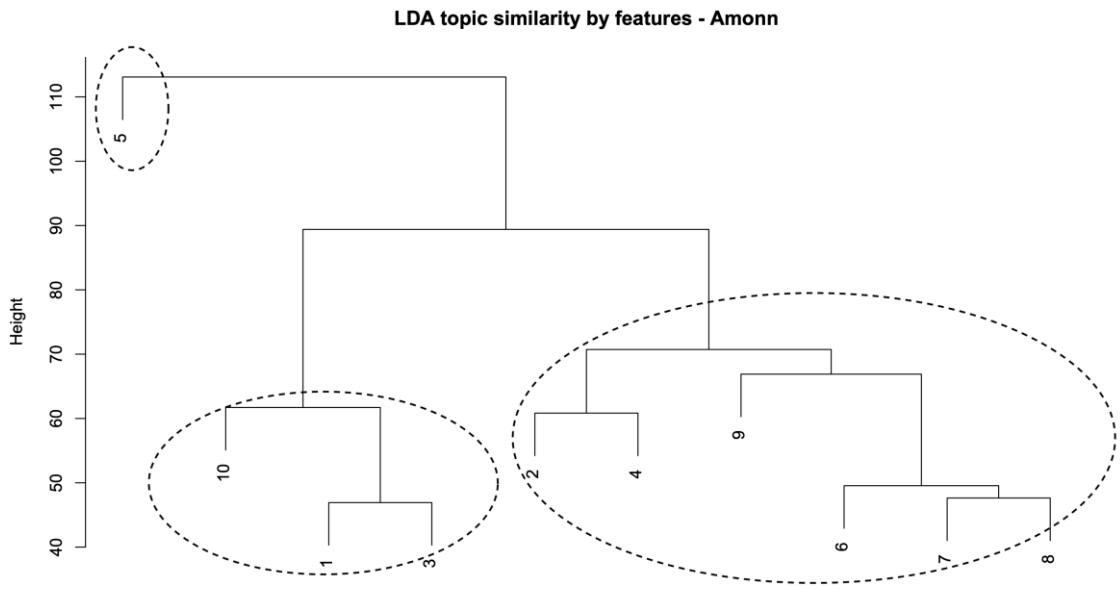
(b) Engländer, φ≥0.592



(c) Müller, φ≥0.379

Appendix C: Latent Dirichlet Allocation and Structural Topic Models – Results for Amonn

Firstly, the model was set to generate 10 topics. The results were not interpretable. The following dendrogam was used to detect the similarity of topics:



Analysis repeated with 3 topics but it was again difficult to interpret.

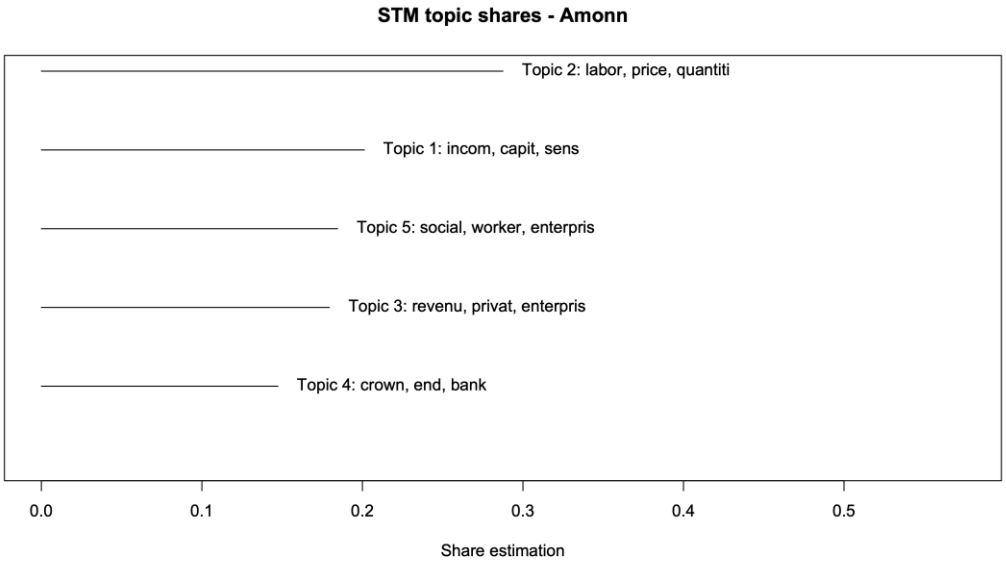
Amonn LDA

T1	T2	T3
labor	incom	enterpris
quantiti	capit	privat
price	sens	social
wage	concept	public
demand	money	crown
capit	question	bank
exchang	countri	amount
produc	tax	end
land	view	manag
determin	social	worker

So we try 5 topics with another model (STM):

Amonn STM

T1	T2	T3	T4	T5
incom	labor	revenu	crown	social
capit	price	privat	end	worker
sens	quantiti	enterpris	bank	enterpris
money	wage	public	amount	form
countri	demand	profit	offic	question
concept	capit	social	payment	work
tax	exchang	entrepreneur	currenc	industri
quantiti	produc	view	foreign	labour
labor	land	oper	exchang	larg
price	determin	communiti	note	expropri



We generated the graphics showing the share of each topic in corpus. The topic labour stand out, similarly as in the previous analysis. However, the topics can be interpreted only with difficulty using the generated list of words.

Appendix D: R - script

```
# # # Text Mining Analysis of the Commentaries on the Czechoslovak Monetary Policy

# Libraries needed

library(tm)

library(tidyverse)

library(tidytext)

library(readtext)

library(magrittr)

library("wordcloud")

library("RColorBrewer")

library("SnowballC")

library("textdata")

library(ggplot2)


# Loading data and creation of tibble

metadata1 <- read.csv("/localPathToFiles/data-text-analysis/metadata.csv", header=TRUE, stringsAsFactors=FALSE, sep = ';')

str(metadata1)


file_names <- list.files("/localPathToFiles/data-text-analysis/txt2")


file_contents = c()

for (file in file_names) {
```

```
  file_contents = append(file_contents, paste(readLines(
    paste("/localPathToFiles/data-text-analysis/txt2/", file, sep = "")
  ), collapse="\n"))
}

data_whole <-
metadata1 %>%
arrange(Author, Year, Title) %>% # sort metadata
bind_cols(file_names) %>% # combine with texts
rename(Files = ...4) %>% # combine with texts
bind_cols(file_contents) %>% # combine with texts
rename(text = ...5) %>% # combine with texts
as_tibble() # convert to tibble for better screen viewing

glimpse(data_whole)

#####################################################################
# 01 - most frequent words, word clouds
#####################################################################

# tokenize: split by words (lowercase, remove punctuation)
data_words <- data_whole %>%
  filter(
    Title == "Die tschechoslowakische Wahrung und Wahrungspolitik"
```

```
  | Title == "Die Devalvation der tschechoslowakischen Krone im Jahre 1934"
  | Title == "Zwei tschechische Schriften uber Wahrungsreform"
 ) %>%
 unnest_tokens(word, text, to_lower = TRUE, strip_punct = TRUE, drop = FALSE)

# remove stop words (and custom stop words)
my_stop_words <- tribble(
 ~word, ~lexicon,
 "millions", "CUSTOM",
 "million", "CUSTOM",
 "thousands", "CUSTOM",
 "thousand", "CUSTOM",
 "billions", "CUSTOM",
 "billion", "CUSTOM",
 "january", "CUSTOM",
 "february", "CUSTOM",
 "march", "CUSTOM",
 "april", "CUSTOM",
 "may", "CUSTOM",
 "june", "CUSTOM",
 "july", "CUSTOM",
 "august", "CUSTOM",
 "september", "CUSTOM",
 "october", "CUSTOM",
```

```
  "november", "CUSTOM",
  "december", "CUSTOM"
)

tidy_data_words <- data_words %>%
  anti_join(stop_words) %>%
  anti_join(my_stop_words) %>%
  filter(
    !grepl('^[0-9]+$', word)
    & !grepl('^i\\.e$', word)
  )

# replacements (particular words)
tidy_data_words$word <- replace(tidy_data_words$word,
tidy_data_words$word == "kr", "crown")
tidy_data_words$word <- replace(tidy_data_words$word,
tidy_data_words$word == "koruna", "crown")
tidy_data_words$word <- replace(tidy_data_words$word,
tidy_data_words$word == "krone", "crown")
tidy_data_words$word <- replace(tidy_data_words$word,
tidy_data_words$word == "kčč", "crown")
tidy_data_words$word <- replace(tidy_data_words$word,
tidy_data_words$word == "kč", "crown")
tidy_data_words$word <- replace(tidy_data_words$word,
tidy_data_words$word == "metallist", "metalist")
tidy_data_words$word <- replace(tidy_data_words$word,
tidy_data_words$word == "metallism", "metalism")
```

```
# stem words (using wordStem from SnowballC package)
tidy_data_words_s <- tidy_data_words %>%
  mutate(stem = wordStem(word))

# then replace all words (dynamically) ending with apostrophe (using regex = regular expression). E.g. rašín' -> rašín, fořt' -> fořt, author' -> author
tidy_data_words_s$stem <- gsub("^(.+)\\'$", "\\1", tidy_data_words_s$stem)

# top 20 most frequent words (after stemming only); uncomment particular title and continue below
tidy_data_words_s %>%
  filter(
    Title == "Die tschechoslowakische Wahrung und Wahrungspolitik"
#    Title == "Die Devalvation der tschechoslowakischen Krone im Jahre 1934"
#    Title == "Zwei tschechische Schriften uber Wahrungsreform"
  ) %>%
  group_by(stem) %>%
  summarise(
    w_count = n()
  ) %>%
  arrange(desc(w_count)) %>%
  print(n=20)

dtm <- TermDocumentMatrix(
```

```
  tidy_data_words_s %>%
  filter(
    Title == "Die tschechoslowakische Wahrung und Wahrungspolitik"
#    Title == "Die Devalvation der tschechoslowakischen Krone im Jahre 1934"
#    Title == "Zwei tschechische Schriften uber Wahrungsreform"
  ) %>%
  pull(stem)
)
m <- as.matrix(dtm)
v <- sort(rowSums(m),decreasing=TRUE)
d <- data.frame(word = names(v),freq=v)

set.seed(123)
wordcloud(words = d$word, freq = d$freq, min.freq = 1, max.words = 100,
random.order=FALSE, rot.per=0.35, colors=brewer.pal(8, "Greys")[seq(4,9)])

#####################################################################
# 02 - tf-idf
#####################################################################

install.packages("remotes")
remotes::install_github("coolbutuseless/ggpattern")
library("ggpattern")
```

```
data_words_tfidf <- data_whole %>%
  filter(
    Title == "Die tschechoslowakische Wahrung und Wahrungspolitik"
    | Title == "Die Devalvation der tschechoslowakischen Krone im Jahre 1934"
    | Title == "Zwei tschechische Schriften uber Wahrungsreform"
  ) %>%
  unnest_tokens(word, text, to_lower = TRUE, strip_punct = TRUE, drop = FALSE)


# manual stop words removal
my_stop_words <- tribble(
  ~word, ~lexicon,
  "millions", "CUSTOM",
  "million", "CUSTOM",
  "thousands", "CUSTOM",
  "thousand", "CUSTOM",
  "billions", "CUSTOM",
  "billion", "CUSTOM",
  "january", "CUSTOM",
  "february", "CUSTOM",
  "march", "CUSTOM",
  "april", "CUSTOM",
  "may", "CUSTOM",
  "june", "CUSTOM",
  "july", "CUSTOM",
```

```
  "august", "CUSTOM",
  "september", "CUSTOM",
  "october", "CUSTOM",
  "november", "CUSTOM",
  "december", "CUSTOM",
  "year", "CUSTOM",
  "per", "CUSTOM",
  "cent", "CUSTOM",
  "p", "CUSTOM",
  "pp", "CUSTOM",
)

# remove numbers and remove "i.e"
tidy_data_words <- data_words_tfidf %>%
  anti_join(stop_words) %>%
  anti_join(my_stop_words) %>%
  filter(
    !grepl('^[0-9]+$', word)
    & !grepl('^i\\.e$', word)
  )

# replacements (particular words)
tidy_data_words$word <- replace(tidy_data_words$word,
tidy_data_words$word == "kr", "crown")
```

```
tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "koruna", "crown")

tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "krone", "crown")

tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "kčč", "crown")

tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "kč", "crown")

tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "metallist", "metalist")

tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "metallism", "metalism")


tidy_data_words_s <- tidy_data_words %>%

 mutate(stem = wordStem(word))


# then replace all words (dynamically) ending with apostrophe (using regex = regular expression). E.g. rašín' -> rašín, fořt' -> fořt, author' -> author

tidy_data_words_s$stem <- gsub("^(.+)\\'$", "\\1", tidy_data_words_s$stem)


# aggregate for author, count words by same "stem" values

tidy_data_words_count <- tidy_data_words_s %>%

 count(Author, stem, sort = TRUE)


plot_words_s <- tidy_data_words_count %>%

 bind_tf_idf(stem, Author, n) %>%

 mutate(
```

```
    Author = factor(
      Author,
      levels = c("Amonn Alfred", "Englander Oskar", "Muller Hugo")
    )
  )

# output top 15 words
plot_words_s %>%
  group_by(Author) %>%
  slice_max(tf_idf, n = 15) %>%
  ungroup() %>%
  mutate(stem = reorder(stem, tf_idf)) %>%
  print(n=60)

# create diagram
plot_words_s %>%
  group_by(Author) %>%
  slice_max(tf_idf, n = 15) %>%
  ungroup() %>%
  mutate(stem = reorder(stem, tf_idf)) %>%
  ggplot(aes(tf_idf, stem)) +
  scale_color_grey(start=0.8, end=0.2) +
  geom_col_pattern(
    aes(
```

```
      pattern = (
        ifelse(Author == "Amonn Alfred", "stripe",
          ifelse(Author == "Englander Oskar", "polygon_tiling", "none")
        )
      )
    ),
    fill    = 'white',
    colour  = 'black'
  ) +
  theme_bw() +
  theme(legend.position = 'none') +
  labs(x = "tf-idf", y = NULL) +
  facet_wrap(~Author, ncol = 2, scales = "free")

# additional manual removal
mystopwords <- tibble(stem = c("pp"))
tidy_data_words_s <- anti_join(tidy_data_words_s, mystopwords, by = "stem")

plot_words_s <- tidy_data_words_s %>%
  count(Author, stem, sort = TRUE) %>%
  bind_tf_idf(stem, Author, n) %>%
  mutate(
    Author = factor(
      Author,
```

```
      levels = c("Amonn Alfred", "Englander Oskar", "Muller Hugo")
    )
  )

plot_words_s %>%
  group_by(Author) %>%
  slice_max(tf_idf, n = 15) %>%
  ungroup() %>%
  mutate(stem = reorder(stem, tf_idf)) %>%
  ggplot(aes(tf_idf, drlib::reorder_within(stem, tf_idf, Author))) +
  scale_color_grey(start=0.8, end=0.2) +
  geom_col_pattern(
    aes(
      pattern = (
        ifelse(Author == "Amonn Alfred", "stripe",
          ifelse(Author == "Englander Oskar", "polygon_tiling", "none")
        )
      )
    ),
    fill    = 'white',
    colour  = 'black'
  ) +
  theme_bw() +
  theme(legend.position = 'none') +
```

```
  labs(x = "tf-idf", y = NULL) +
  facet_wrap(~Author, ncol = 2, scales = "free") +
  drlib::scale_y_reordered()



######################################################################
# 03 - n-grams (bigrams)
######################################################################

library(dplyr)
library(tidyr)
library(forcats)

data_bigrams <- data_whole %>%
  filter(
    Title == "Die tschechoslowakische Wahrung und Wahrungspolitik"
    | Title == "Die Devalvation der tschechoslowakischen Krone im Jahre 1934"
    | Title == "Zwei tschechische Schriften uber Wahrungsreform"
  ) %>%
  unnest_tokens(bigram, text, token = "ngrams", n = 2) %>%
  filter(!is.na(bigram))

bigrams_separated <- data_bigrams %>%
  separate(bigram, c("word1", "word2"), sep = " ")
```

```
# replacements (particular words)

bigrams_separated$word1 <- replace(bigrams_separated$word1,
bigrams_separated$word1 == "kr", "crown")

bigrams_separated$word1 <- replace(bigrams_separated$word1,
bigrams_separated$word1 == "koruna", "crown")

bigrams_separated$word1 <- replace(bigrams_separated$word1,
bigrams_separated$word1 == "krone", "crown")

bigrams_separated$word1 <- replace(bigrams_separated$word1,
bigrams_separated$word1 == "kčč", "crown")

bigrams_separated$word1 <- replace(bigrams_separated$word1,
bigrams_separated$word1 == "kč", "crown")

bigrams_separated$word1 <- replace(bigrams_separated$word1,
bigrams_separated$word1 == "metallist", "metalist")

bigrams_separated$word1 <- replace(bigrams_separated$word1,
bigrams_separated$word1 == "metallism", "metalism")


bigrams_separated$word2 <- replace(bigrams_separated$word2,
bigrams_separated$word2 == "kr", "crown")

bigrams_separated$word2 <- replace(bigrams_separated$word2,
bigrams_separated$word2 == "koruna", "crown")

bigrams_separated$word2 <- replace(bigrams_separated$word2,
bigrams_separated$word2 == "krone", "crown")

bigrams_separated$word2 <- replace(bigrams_separated$word2,
bigrams_separated$word2 == "kčč", "crown")

bigrams_separated$word2 <- replace(bigrams_separated$word2,
bigrams_separated$word2 == "kč", "crown")

bigrams_separated$word2 <- replace(bigrams_separated$word2,
bigrams_separated$word2 == "metallist", "metalist")
```

```
bigrams_separated$word2 <- replace(bigrams_separated$word2,
bigrams_separated$word2 == "metallism", "metalism")


my_stop_words <- tribble(
  ~word, ~lexicon,
  "millions", "CUSTOM",
  "million", "CUSTOM",
  "thousands", "CUSTOM",
  "thousand", "CUSTOM",
  "billions", "CUSTOM",
  "billion", "CUSTOM",
  "january", "CUSTOM",
  "february", "CUSTOM",
  "march", "CUSTOM",
  "april", "CUSTOM",
  "may", "CUSTOM",
  "june", "CUSTOM",
  "july", "CUSTOM",
  "august", "CUSTOM",
  "september", "CUSTOM",
  "october", "CUSTOM",
  "november", "CUSTOM",
  "december", "CUSTOM"
)
```

```
bigrams_filtered <- bigrams_separated %>%
  filter(!word1 %in% stop_words$word) %>%
  filter(!word2 %in% stop_words$word) %>%
  filter(!grepl('^[0-9]+$', word1)) %>%
  filter(!grepl('^[0-9]+$', word2)) %>%
  filter(!word1 %in% my_stop_words$word) %>%
  filter(!word2 %in% my_stop_words$word)

# new bigram counts
bigrams_filtered %>%
  filter(
    Title == "Die tschechoslowakische Wahrung und Wahrungspolitik"
#    Title == "Die Devalvation der tschechoslowakischen Krone im Jahre 1934"
#    Title == "Zwei tschechische Schriften uber Wahrungsreform"
  ) %>%
  unite(bigram, word1, word2, sep = " ") %>%
  count(bigram, sort = TRUE) %>%
  print(n=20)

bigrams_united <- bigrams_filtered %>%
  unite(bigram, word1, word2, sep = " ")

bigram_tf_idf <- bigrams_united %>%
  count(Title, bigram) %>%
```

```
  bind_tf_idf(bigram, Title, n) %>%
  arrange(desc(tf_idf))


bigram_tf_idf %>%
  group_by(Title) %>%
  slice_max(tf_idf, n = 15) %>%
  ungroup() %>%
  ggplot(aes(tf_idf, fct_reorder(bigram, tf_idf))) +
  scale_color_grey(start=0.8, end=0.2) +
  geom_col_pattern(
    aes(
      pattern = (
        ifelse(Title == "Die tschechoslowakische Wahrung und Wahrungspolitik",
"stripe",
          ifelse(Title == "Die Devalvation der tschechoslowakischen Krone im
Jahre 1934", "polygon_tiling", "none")
        )
      )
    ),
    fill    = 'white',
    colour  = 'black'
  ) +
  theme_bw() +
  theme(legend.position = 'none') +
  labs(x = "tf-idf", y = NULL) +
```

```
  facet_wrap(~Title, ncol = 2, scales = "free")


bigram_tf_idf %>%
  filter(
    Title == "Zwei tschechische Schriften uber Wahrungsreform"
  ) %>%
  group_by(Title) %>%
  slice_max(tf_idf, n = 15) %>%
  ungroup() %>%
  ggplot(aes(tf_idf, fct_reorder(bigram, tf_idf))) +
  scale_color_grey(start=0.8, end=0.2) +
  geom_col_pattern(
    pattern = "stripe",
    fill    = 'white',
    colour  = 'black'
  ) +
  theme_bw() +
  theme(legend.position = 'none') +
  labs(x = "tf-idf", y = NULL) +
  facet_wrap(~Title, ncol = 1, scales = "free")


#####################################################################
# 04 - co-occurrences (correlation)
```

```
#####################################################################

library(widyr)
library(igraph)
library(ggraph)
remotes::install_github("dgrtwo/drlib")

data_words_corr <- data_whole %>%
  filter(
    Title == "Die tschechoslowakische Wahrung und Wahrungspolitik"
    | Title == "Die Devalvation der tschechoslowakischen Krone im Jahre 1934"
    | Title == "Zwei tschechische Schriften uber Wahrungsreform"
  ) %>%
  unnest_tokens(paragraph, text, token = "paragraphs", to_lower = TRUE)
%>%
  filter(nchar(paragraph) > 50) %>%
  group_by(Author, Title) %>%
  mutate(pnumber = row_number()) %>%
  ungroup() %>%
  unnest_tokens(word, paragraph, to_lower = TRUE, strip_punct = TRUE)

# manual stop words removal
my_stop_words <- tribble(
  ~word, ~lexicon,
  "millions", "CUSTOM",
```

```
  "million", "CUSTOM",
  "thousands", "CUSTOM",
  "thousand", "CUSTOM",
  "billions", "CUSTOM",
  "billion", "CUSTOM",
  "january", "CUSTOM",
  "february", "CUSTOM",
  "march", "CUSTOM",
  "april", "CUSTOM",
  "may", "CUSTOM",
  "june", "CUSTOM",
  "july", "CUSTOM",
  "august", "CUSTOM",
  "september", "CUSTOM",
  "october", "CUSTOM",
  "november", "CUSTOM",
  "december", "CUSTOM",
  "year", "CUSTOM",
  "per", "CUSTOM",
  "cent", "CUSTOM",
  "p", "CUSTOM",
)
```

```
# remove numbers and remove "i.e"
tidy_data_words <- data_words_corr %>%
 anti_join(stop_words) %>%
 anti_join(my_stop_words) %>%
 filter(
  !grepl('^[0-9]+$', word)
  & !grepl('^i\\.e$', word)
 )

# replacements (particular words)
tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "kr", "crown")
tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "koruna", "crown")
tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "krone", "crown")
tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "kčč", "crown")
tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "kč", "crown")
tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "metallist", "metalist")
tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "metallism", "metalism")

word_cors <- tidy_data_words %>%
 filter(
```

```
    Title == "Die tschechoslowakische Wahrung und Wahrungspolitik"
#   Title == "Die Devalvation der tschechoslowakischen Krone im Jahre 1934"
#   Title == "Zwei tschechische Schriften uber Wahrungsreform"
  ) %>%
  group_by(word) %>%
  filter(n() >= 10) %>% # threshold: at least in 10 different pairs (should be adjusted based on dataset size)
  pairwise_cor(word, pnumber, sort = TRUE) # do include upper triangle of matrix

word_cors %>%
  filter(item1 %in% c("engliš", "rašín", "fořt", "czechoslovakia", "czechoslovak")) %>%
  group_by(item1) %>%
  slice_max(correlation, n = 6) %>%
  ungroup() %>%
  mutate(item2 = reorder(item2, correlation)) %>%
  ggplot(aes(drlib::reorder_within(item2, correlation, within = item1), correlation)) +
  geom_bar(stat = "identity") +
  facet_wrap(~ item1, scales = "free") +
  drlib::scale_x_reordered() +
  coord_flip()

# correlation relations - each text/author has own threshold (to have clear chart)
```

```
set.seed(123)

word_cors %>%

 filter(correlation > .45) %>%

 graph_from_data_frame() %>%

 ggraph(layout = "fr") +

 geom_edge_link(aes(edge_alpha = correlation), show.legend = FALSE) +

 geom_node_point(color = "gray", size = 5) +

 geom_node_text(aes(label = name), repel = TRUE) +

 theme_void()



#####################################################################

# 05 - sentiment

#####################################################################


# bigrams with paragraph number

data_bigrams_sent <- data_whole %>%

 filter(

  Title == "Die tschechoslowakische Wahrung und Wahrungspolitik"

  | Title == "Die Devalvation der tschechoslowakischen Krone im Jahre 1934"

  | Title == "Zwei tschechische Schriften uber Wahrungsreform"

 ) %>%

 unnest_tokens(paragraph, text, token = "paragraphs", to_lower = TRUE)
%>%

 filter(nchar(paragraph) > 50) %>%
```

```
  group_by(Author, Title) %>%
  mutate(pnumber = row_number()) %>%
  ungroup() %>%
  unnest_tokens(bigram, paragraph, token = "ngrams", n = 2) %>%
  filter(!is.na(bigram))

bigrams_separated_sent <- data_bigrams_sent %>%
  separate(bigram, c("word1", "word2"), sep = " ")

# replacements (particular words)
bigrams_separated_sent$word1 <- replace(bigrams_separated_sent$word1, bigrams_separated_sent$word1 == "kr", "crown")
bigrams_separated_sent$word1 <- replace(bigrams_separated_sent$word1, bigrams_separated_sent$word1 == "koruna", "crown")
bigrams_separated_sent$word1 <- replace(bigrams_separated_sent$word1, bigrams_separated_sent$word1 == "krone", "crown")
bigrams_separated_sent$word1 <- replace(bigrams_separated_sent$word1, bigrams_separated_sent$word1 == "kčč", "crown")
bigrams_separated_sent$word1 <- replace(bigrams_separated_sent$word1, bigrams_separated_sent$word1 == "kč", "crown")
bigrams_separated_sent$word1 <- replace(bigrams_separated_sent$word1, bigrams_separated_sent$word1 == "metallist", "metalist")
bigrams_separated_sent$word1 <- replace(bigrams_separated_sent$word1, bigrams_separated_sent$word1 == "metallism", "metalism")

bigrams_separated_sent$word2 <- replace(bigrams_separated_sent$word2, bigrams_separated_sent$word2 == "kr", "crown")
```

```
bigrams_separated_sent$word2 <- replace(bigrams_separated_sent$word2,
bigrams_separated_sent$word2 == "koruna", "crown")

bigrams_separated_sent$word2 <- replace(bigrams_separated_sent$word2,
bigrams_separated_sent$word2 == "krone", "crown")

bigrams_separated_sent$word2 <- replace(bigrams_separated_sent$word2,
bigrams_separated_sent$word2 == "kčč", "crown")

bigrams_separated_sent$word2 <- replace(bigrams_separated_sent$word2,
bigrams_separated_sent$word2 == "kč", "crown")

bigrams_separated_sent$word2 <- replace(bigrams_separated_sent$word2,
bigrams_separated_sent$word2 == "metallist", "metalist")

bigrams_separated_sent$word2 <- replace(bigrams_separated_sent$word2,
bigrams_separated_sent$word2 == "metallism", "metalism")


negation_words <- c(
  "not",
  "no",
  "never",
  "without",
  "cannot",
  "none",
  "nowhere",
  "neither",
  "nor",
  "isn't",
  "don't",
  "doesn't",
  "won't",
```

```
  "haven't",
  "hasn't"
)

not_words <- bigrams_separated_sent %>%
  filter(word1 %in% negation_words) %>%
  inner_join(get_sentiments("afinn"), by = c(word2 = "word")) %>%
  mutate(value = -2*value)

data_not_words_sent <- not_words %>%
  rename(word = word2) %>%
  select(-word1)

data_words_sent <- data_whole %>%
  filter(
    Title == "Die tschechoslowakische Wahrung und Wahrungspolitik"
    | Title == "Die Devalvation der tschechoslowakischen Krone im Jahre 1934"
    | Title == "Zwei tschechische Schriften uber Wahrungsreform"
  ) %>%
  unnest_tokens(paragraph, text, token = "paragraphs", to_lower = TRUE)
%>%
  filter(nchar(paragraph) > 50) %>%
  group_by(Author, Title) %>%
  mutate(pnumber = row_number()) %>%
  ungroup() %>%
```

```
  unnest_tokens(word, paragraph, to_lower = TRUE, strip_punct = TRUE) %>%
  inner_join(get_sentiments("afinn"))

data_sentiment_afinn <- data_words_sent %>%
  bind_rows(data_not_words_sent) %>%
  group_by(Author, Title, Year, pnumber) %>%
  summarise(sentiment = sum(value)) %>%
  mutate(method = "AFINN") %>%
  ungroup() %>%
  arrange(pnumber)

ggplot(data_sentiment_afinn, aes(pnumber, sentiment, fill = Title)) +
  geom_col(show.legend = FALSE) +
  facet_wrap(~Title, ncol = 2, scales = "free_x")

# Sentiment related to particular words: "engliš", "rašín", "fořt", "czechoslovakia", "czechoslovak"

data_sent_particular <- data_whole %>%
  filter(
    Title == "Zwei tschechische Schriften uber Wahrungsreform"
  ) %>%
  unnest_tokens(paragraph, text, token = "paragraphs", to_lower = TRUE) %>%
  filter(nchar(paragraph) > 50) %>%
```

```
  group_by(Author, Title) %>%
  mutate(pnumber = row_number()) %>%
  ungroup()

word_matched_paragraphs <- data_sent_particular %>%
  filter(grepl("rašín", paragraph)) %>%
  select(pnumber)

ggplot(
    data_sentiment_afinn %>%
      filter(
        Title == "Zwei tschechische Schriften uber Wahrungsreform"
      ) %>%
      mutate(fill = ifelse(pnumber %in% word_matched_paragraphs$pnumber, "matched", "notmatched")),
    aes(pnumber, sentiment)
  ) +
  geom_col(aes(pnumber, fill = fill), show.legend = FALSE) +
  geom_text(aes(label=sentiment, vjust=0.5-sign(sentiment))) +
  facet_wrap(~Title, ncol = 2, scales = "free_x") +
  scale_fill_manual(values=c("matched" = "#000000", "notmatched" = "#999999"))

data_sent_particular <- data_whole %>%
  filter(
```

```
    Title == "Die Devalvation der tschechoslowakischen Krone im Jahre 1934"
  ) %>%
  unnest_tokens(paragraph, text, token = "paragraphs", to_lower = TRUE) %>%
  filter(nchar(paragraph) > 50) %>%
  group_by(Author, Title) %>%
  mutate(pnumber = row_number()) %>%
  ungroup()

word_matched_paragraphs <- data_sent_particular %>%
  filter(grepl("rašín", paragraph)) %>%
  select(pnumber)

ggplot(
    data_sentiment_afinn %>%
      filter(
        Title == "Die Devalvation der tschechoslowakischen Krone im Jahre 1934"
      ) %>%
      mutate(fill = ifelse(pnumber %in% word_matched_paragraphs$pnumber, "matched", "notmatched")),
    aes(pnumber, sentiment)
  ) +
  geom_col(aes(pnumber, fill = fill), show.legend = FALSE) +
  geom_text(aes(label=sentiment, vjust=0.5-sign(sentiment))) +
  facet_wrap(~Title, ncol = 2, scales = "free_x") +
```

```
  scale_fill_manual(values=c("matched" = "#000000", "notmatched" = "#999999"))


data_sent_particular <- data_whole %>%
  filter(
    Title == "Die tschechoslowakische Wahrung und Wahrungspolitik"
    | Title == "Die Devalvation der tschechoslowakischen Krone im Jahre 1934"
    | Title == "Zwei tschechische Schriften uber Wahrungsreform"
  ) %>%
  unnest_tokens(paragraph, text, token = "paragraphs", to_lower = TRUE) %>%
  filter(nchar(paragraph) > 50) %>%
  group_by(Author, Title) %>%
  mutate(pnumber = row_number()) %>%
  ungroup()


word_matched_paragraphs <- data_sent_particular %>%
  filter(
    grepl("czechoslovakia", paragraph)
    | grepl("czech", paragraph)
    | grepl("czechoslovak", paragraph)
    | grepl("czechoslovakian", paragraph)
  ) %>%
  select(Title, pnumber) %>%
  mutate(match = 1)
```

```
ggplot(
   data_sentiment_afinn %>%
   filter(
#     Title == "Die tschechoslowakische Wahrung und Wahrungspolitik"
#        Title == "Die Devalvation der tschechoslowakischen Krone im Jahre 1934"
     Title == "Zwei tschechische Schriften uber Wahrungsreform"
   ) %>%
   merge(word_matched_paragraphs,
      by=c("Title","pnumber"), all.x=TRUE) %>%
   mutate(fill = ifelse(match == 1, "matched", "notmatched")),
   aes(pnumber, sentiment)
 ) +
 geom_col(aes(pnumber, fill = fill), show.legend = FALSE) +
 geom_text(aes(label=sentiment, vjust=0.5-sign(sentiment))) +
 facet_wrap(~Title, ncol = 2, scales = "free_x") +
 scale_fill_manual(values=c("matched" = "#000000", "notmatched" = "#999999"))


data_sent_particular <- data_whole %>%
 filter(
   Title == "Die tschechoslowakische Wahrung und Wahrungspolitik"
   | Title == "Die Devalvation der tschechoslowakischen Krone im Jahre 1934"
   | Title == "Zwei tschechische Schriften uber Wahrungsreform"
```

```
  ) %>%
  unnest_tokens(paragraph, text, token = "paragraphs", to_lower = TRUE)
%>%
  filter(nchar(paragraph) > 50) %>%
  group_by(Author, Title) %>%
  mutate(pnumber = row_number()) %>%
  ungroup()

word_matched_paragraphs <- data_sent_particular %>%
  filter(
    grepl("metalist", paragraph)
    | grepl("metalism", paragraph)
    | grepl("metallist", paragraph)
    | grepl("metallism", paragraph)
  ) %>%
  select(Title, pnumber) %>%
  mutate(match = 1)

ggplot(
    data_sentiment_afinn %>%
    filter(
#      Title == "Die tschechoslowakische Wahrung und Wahrungspolitik"
#         Title == "Die Devalvation der tschechoslowakischen Krone im Jahre
1934"
      Title == "Zwei tschechische Schriften uber Wahrungsreform"
```

```
  ) %>%
  merge(word_matched_paragraphs,
      by=c("Title","pnumber"), all.x=TRUE) %>%
  mutate(fill = ifelse(match == 1, "matched", "notmatched")),
  aes(pnumber, sentiment)
 ) +
 geom_col(aes(pnumber, fill = fill), show.legend = FALSE) +
 geom_text(aes(label=sentiment, vjust=0.5-sign(sentiment))) +
 facet_wrap(~Title, ncol = 2, scales = "free_x") +
 scale_fill_manual(values=c("matched" = "#000000", "notmatched" =
"#999999"))

#####################################################################
06 - STM/LDA
#####################################################################

library(quanteda) # For NLP
library(topicmodels) # For topicmodels (and LDA)
library(stm) # For structural topic models (STM)

mycorpus <- corpus(
 data_whole %>%
 filter(Author == "Amonn Alfred")
#  filter(Author == "Englander Oskar")
#  filter(Author == "Muller Hugo")
```

```
)
mycorpus.stats <- summary(mycorpus)

token <-
  tokens(
    mycorpus,
    split_hyphens = TRUE,
    remove_numbers = TRUE,
    remove_punct = TRUE,
    remove_symbols = TRUE,
    remove_url = TRUE,
    include_docvars = TRUE
  )

mydfm <- dfm(
  token,
  tolower = TRUE,
  stem = TRUE,
  remove = stopwords("english")
)

mydfm.trim <-
  dfm_trim(
    mydfm,
```

```
  min_docfreq = 0.075,
  # min 7.5%
  max_docfreq = 0.90,
  # max 90%
  docfreq_type = "prop"
 )

topic.count <- 10

dfm2topicmodels <- convert(mydfm.trim, to = "topicmodels")
lda.model <- LDA(dfm2topicmodels, topic.count)

# print model words as table
as.data.frame(terms(lda.model, 10))

lda.similarity <- as.data.frame(lda.model@beta) %>%
 scale() %>%
 dist(method = "euclidean") %>%
 hclust(method = "ward.D2")

par(mar = c(0, 4, 4, 2))

# dendrogram
plot(lda.similarity,
```

```
    main = "LDA topic similarity by features - Muller",
    xlab = "",
    sub = "")

# STM

topic.count <- 5

dfm2stm <- convert(mydfm.trim, to = "stm")

model.stm <- stm(
  dfm2stm$documents,
  dfm2stm$vocab,
  K = topic.count,
  data = dfm2stm$meta,
  init.type = "Spectral"
)

# print topic terms
as.data.frame(t(labelTopics(model.stm, n = 10)$prob))

plot(
  model.stm,
  type = "summary",
```

```
  text.cex = 1,
  main = "STM topic shares - Muller",
  xlab = "Share estimation"
)
```

Text Mining Analysis of the Whole Text Corpus by German-speaking Economists

```
# Libraries needed
library(tm)
library(tidyverse)
library(tidytext)
library(readtext)
library(magrittr)
library("wordcloud")
library("RColorBrewer")
library("SnowballC")
library("textdata")
library(ggplot2)


# Loading data and creration of tibble
metadata1 <- read.csv("/localPathToFiles/data-text-analysis/metadata.csv", header=TRUE, stringsAsFactors=FALSE, sep = ';')
str(metadata1)


file_names <- list.files("/localPathToFiles/data-text-analysis/txt2")
```

```
file_contents = c()
for (file in file_names) {
  file_contents = append(file_contents, paste(readLines(
    paste("/localPathToFiles/data-text-analysis/txt2/", file, sep = "")
  ), collapse="\n"))
}

data_whole <-
  metadata1 %>%
  arrange(Author, Year, Title) %>% # sort metadata
  bind_cols(file_names) %>% # combine with texts
  rename(Files = ...4) %>% # combine with texts
  bind_cols(file_contents) %>% # combine with texts
  rename(text = ...5) %>% # combine with texts
  as_tibble() # convert to tibble for better screen viewing

glimpse(data_whole)

#################################################################
# 01 - most frequent words, word clouds
#################################################################

# tokenize: split by words (lowercase, remove punctuation)
```

```
data_words <- data_whole %>%
 unnest_tokens(word, text, to_lower = TRUE, strip_punct = TRUE, drop = FALSE)


# remove stop words (and custom stop words)
my_stop_words <- tribble(
 ~word, ~lexicon,
 "millions", "CUSTOM",
 "million", "CUSTOM",
 "thousands", "CUSTOM",
 "thousand", "CUSTOM",
 "billions", "CUSTOM",
 "billion", "CUSTOM",
 "january", "CUSTOM",
 "february", "CUSTOM",
 "march", "CUSTOM",
 "april", "CUSTOM",
 "may", "CUSTOM",
 "june", "CUSTOM",
 "july", "CUSTOM",
 "august", "CUSTOM",
 "september", "CUSTOM",
 "october", "CUSTOM",
 "november", "CUSTOM",
 "december", "CUSTOM"
```

```
)

tidy_data_words <- data_words %>%
 anti_join(stop_words) %>%
 anti_join(my_stop_words) %>%
 filter(
  !grepl('^[0-9]+$', word)
  & !grepl('^i\\.e$', word)
 )

# replacements (particular words)
tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "kr", "crown")
tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "koruna", "crown")
tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "krone", "crown")
tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "kčč", "crown")
tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "kč", "crown")
tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "metallist", "metalist")
tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "metallism", "metalism")

# stem words (using wordStem from SnowballC package)
```

```
tidy_data_words_s <- tidy_data_words %>%
  mutate(stem = wordStem(word))


# then replace all words (dynamically) ending with apostrophe (using regex = regular expression). E.g. rašín' -> rašín, fořt' -> fořt, author' -> author
tidy_data_words_s$stem <- gsub("^(.+)\\'$", "\\1", tidy_data_words_s$stem)


# 3 tables: 20 most frequent words (after stemming only)
tidy_data_words_s %>%
  filter(
#    Author == "Amonn Alfred"
#    Author == "Englander Oskar"
    Author == "Muller Hugo"
  ) %>%
  group_by(stem) %>%
  summarise(
    w_count = n()
  ) %>%
  arrange(desc(w_count)) %>%
  print(n=20)


dtm <- TermDocumentMatrix(
  tidy_data_words_s %>%
  filter(
#   Author == "Amonn Alfred"
```

```
#    Author == "Englander Oskar"
    Author == "Muller Hugo"
  ) %>%
  pull(stem)
)
m <- as.matrix(dtm)
v <- sort(rowSums(m),decreasing=TRUE)
d <- data.frame(word = names(v),freq=v)

set.seed(123)
wordcloud(words = d$word, freq = d$freq, min.freq = 1, max.words = 100, random.order=FALSE, rot.per=0.35, colors=brewer.pal(8, "Greys")[seq(4,9)])

#####################################################################
# 02 - tf-idf
#####################################################################

install.packages("remotes")
remotes::install_github("coolbutuseless/ggpattern")
library("ggpattern")

data_words_tfidf <- data_whole %>%
  unnest_tokens(word, text, to_lower = TRUE, strip_punct = TRUE, drop = FALSE)
```

```
# manual stop words removal
my_stop_words <- tribble(
  ~word, ~lexicon,
  "millions", "CUSTOM",
  "million", "CUSTOM",
  "thousands", "CUSTOM",
  "thousand", "CUSTOM",
  "billions", "CUSTOM",
  "billion", "CUSTOM",
  "january", "CUSTOM",
  "february", "CUSTOM",
  "march", "CUSTOM",
  "april", "CUSTOM",
  "may", "CUSTOM",
  "june", "CUSTOM",
  "july", "CUSTOM",
  "august", "CUSTOM",
  "september", "CUSTOM",
  "october", "CUSTOM",
  "november", "CUSTOM",
  "december", "CUSTOM",
  "year", "CUSTOM",
  "per", "CUSTOM",
  "cent", "CUSTOM",
```

```
  "p", "CUSTOM",
)

# remove numbers and remove "i.e"
tidy_data_words <- data_words_tfidf %>%
  anti_join(stop_words) %>%
  anti_join(my_stop_words) %>%
  filter(
    !grepl('^[0-9]+$', word)
    & !grepl('^i\\.e$', word)
  )

# replacements (particular words)
tidy_data_words$word <- replace(tidy_data_words$word,
tidy_data_words$word == "kr", "crown")
tidy_data_words$word <- replace(tidy_data_words$word,
tidy_data_words$word == "koruna", "crown")
tidy_data_words$word <- replace(tidy_data_words$word,
tidy_data_words$word == "krone", "crown")
tidy_data_words$word <- replace(tidy_data_words$word,
tidy_data_words$word == "kčč", "crown")
tidy_data_words$word <- replace(tidy_data_words$word,
tidy_data_words$word == "kč", "crown")
tidy_data_words$word <- replace(tidy_data_words$word,
tidy_data_words$word == "metallist", "metalist")
tidy_data_words$word <- replace(tidy_data_words$word,
tidy_data_words$word == "metallism", "metalism")
```

```
tidy_data_words_s <- tidy_data_words %>%
  mutate(stem = wordStem(word))

# then replace all words (dynamically) ending with apostrophe (using regex = regular expression). E.g. rašín' -> rašín, fořt' -> fořt, author' -> author
tidy_data_words_s$stem <- gsub("^(.+)\\'$", "\\1", tidy_data_words_s$stem)

# aggregate for author, count words by same "stem" values
tidy_data_words_count <- tidy_data_words_s %>%
  count(Author, stem, sort = TRUE)

plot_words_s <- tidy_data_words_count %>%
  bind_tf_idf(stem, Author, n) %>%
  mutate(
    Author = factor(
      Author,
      levels = c("Amonn Alfred", "Englander Oskar", "Muller Hugo")
    )
  )

# first - remove some other words
plot_words_s %>%
  group_by(Author) %>%
  slice_max(tf_idf, n = 15) %>%
```

```
  ungroup() %>%
  mutate(stem = reorder(stem, tf_idf)) %>%
  ggplot(aes(tf_idf, stem)) +
  scale_color_grey(start=0.8, end=0.2) +
  geom_col_pattern(
    aes(
      pattern = (
        ifelse(Author == "Amonn Alfred", "stripe",
          ifelse(Author == "Englander Oskar", "polygon_tiling", "none")
        )
      )
    ),
    fill    = 'white',
    colour  = 'black'
  ) +
  theme_bw() +
  theme(legend.position = 'none') +
  labs(x = "tf-idf", y = NULL) +
  facet_wrap(~Author, ncol = 2, scales = "free")


# manual removal
mystopwords <- tibble(stem = c("pp","ibidem","ki","ka","ma","pi","pa","p1","pn","0.25","aa","ai","kpa"))
tidy_data_words_s <- anti_join(tidy_data_words_s, mystopwords, by = "stem")
```

```
tidy_data_words_count <- tidy_data_words_s %>%
  count(Author, stem, sort = TRUE)


plot_words_s <- tidy_data_words_count %>%
  bind_tf_idf(stem, Author, n) %>%
  mutate(
    Author = factor(
      Author,
      levels = c("Amonn Alfred", "Englander Oskar", "Muller Hugo")
    )
  )


plot_words_s %>%
  group_by(Author) %>%
  slice_max(tf_idf, n = 15) %>%
  ungroup() %>%
  mutate(stem = reorder(stem, tf_idf)) %>%
  print(n=60)


plot_words_s %>%
  group_by(Author) %>%
  slice_max(tf_idf, n = 15) %>%
  ungroup() %>%
  mutate(stem = reorder(stem, tf_idf)) %>%
```

```
  ggplot(aes(tf_idf, drlib::reorder_within(stem, tf_idf, Author))) +
  scale_color_grey(start=0.8, end=0.2) +
  geom_col_pattern(
   aes(
    pattern = (
     ifelse(Author == "Amonn Alfred", "stripe",
      ifelse(Author == "Englander Oskar", "polygon_tiling", "none")
     )
    )
   ),
   fill    = 'white',
   colour  = 'black'
  ) +
  theme_bw() +
  theme(legend.position = 'none') +
  labs(x = "tf-idf", y = NULL) +
  facet_wrap(~Author, ncol = 2, scales = "free") +
  drlib::scale_y_reordered()

#####################################################################
# 03 - n-grams (bigrams)
#####################################################################

library(dplyr)
```

```
library(tidyr)

library(forcats)


data_bigrams <- data_whole %>%

  unnest_tokens(bigram, text, token = "ngrams", n = 2) %>%

  filter(!is.na(bigram))


bigrams_separated <- data_bigrams %>%

  separate(bigram, c("word1", "word2"), sep = " ")


# replacements (particular words)

bigrams_separated$word1 <- replace(bigrams_separated$word1,
bigrams_separated$word1 == "kr", "crown")

bigrams_separated$word1 <- replace(bigrams_separated$word1,
bigrams_separated$word1 == "koruna", "crown")

bigrams_separated$word1 <- replace(bigrams_separated$word1,
bigrams_separated$word1 == "krone", "crown")

bigrams_separated$word1 <- replace(bigrams_separated$word1,
bigrams_separated$word1 == "kčč", "crown")

bigrams_separated$word1 <- replace(bigrams_separated$word1,
bigrams_separated$word1 == "kč", "crown")

bigrams_separated$word1 <- replace(bigrams_separated$word1,
bigrams_separated$word1 == "metallist", "metalist")

bigrams_separated$word1 <- replace(bigrams_separated$word1,
bigrams_separated$word1 == "metallism", "metalism")


bigrams_separated$word2 <- replace(bigrams_separated$word2,
bigrams_separated$word2 == "kr", "crown")
```

```
bigrams_separated$word2 <- replace(bigrams_separated$word2,
bigrams_separated$word2 == "koruna", "crown")

bigrams_separated$word2 <- replace(bigrams_separated$word2,
bigrams_separated$word2 == "krone", "crown")

bigrams_separated$word2 <- replace(bigrams_separated$word2,
bigrams_separated$word2 == "kčč", "crown")

bigrams_separated$word2 <- replace(bigrams_separated$word2,
bigrams_separated$word2 == "kč", "crown")

bigrams_separated$word2 <- replace(bigrams_separated$word2,
bigrams_separated$word2 == "metallist", "metalist")

bigrams_separated$word2 <- replace(bigrams_separated$word2,
bigrams_separated$word2 == "metallism", "metalism")


my_stop_words <- tribble(

 ~word, ~lexicon,

 "millions", "CUSTOM",

 "million", "CUSTOM",

 "thousands", "CUSTOM",

 "thousand", "CUSTOM",

 "billions", "CUSTOM",

 "billion", "CUSTOM",

 "january", "CUSTOM",

 "february", "CUSTOM",

 "march", "CUSTOM",

 "april", "CUSTOM",

 "may", "CUSTOM",

 "june", "CUSTOM",
```

```
  "july", "CUSTOM",
  "august", "CUSTOM",
  "september", "CUSTOM",
  "october", "CUSTOM",
  "november", "CUSTOM",
  "december", "CUSTOM",
  "op", "CUSTOM",
  "cit", "CUSTOM"
)

bigrams_filtered <- bigrams_separated %>%
  filter(!word1 %in% stop_words$word) %>%
  filter(!word2 %in% stop_words$word) %>%
  filter(!grepl('^[0-9]+$', word1)) %>%
  filter(!grepl('^[0-9]+$', word2)) %>%
  filter(!word1 %in% my_stop_words$word) %>%
  filter(!word2 %in% my_stop_words$word)

# new bigram counts:
bigrams_filtered %>%
  filter(
#    Author == "Amonn Alfred"
#    Author == "Englander Oskar"
    Author == "Muller Hugo"
```

```
  ) %>%
  unite(bigram, word1, word2, sep = " ") %>%
  count(bigram, sort = TRUE) %>%
  print(n=20)

bigrams_united <- bigrams_filtered %>%
  unite(bigram, word1, word2, sep = " ")

bigram_tf_idf <- bigrams_united %>%
  count(Author, bigram) %>%
  bind_tf_idf(bigram, Author, n) %>%
  arrange(desc(tf_idf))

bigram_tf_idf %>%
  group_by(Author) %>%
  slice_max(tf_idf, n = 15) %>%
  ungroup() %>%
  ggplot(aes(tf_idf, fct_reorder(bigram, tf_idf))) +
  scale_color_grey(start=0.8, end=0.2) +
  geom_col_pattern(
    aes(
      pattern = (
        ifelse(Author == "Amonn Alfred", "stripe",
          ifelse(Author == "Englander Oskar", "polygon_tiling", "none")
```

```
      )
    )
  ),
  fill    = 'white',
  colour  = 'black'
) +
theme_bw() +
theme(legend.position = 'none') +
labs(x = "tf-idf", y = NULL) +
facet_wrap(~Author, ncol = 2, scales = "free")


#####################################################################
# 04 - co-occurrences (correlation)
#####################################################################

library(widyr)
library(igraph)
library(ggraph)

remotes::install_github("dgrtwo/drlib")

data_words_corr <- data_whole %>%
  unnest_tokens(paragraph, text, token = "paragraphs", to_lower = TRUE)
%>%
```

```
  filter(nchar(paragraph) > 50) %>%
  group_by(Author, Title) %>%
  mutate(pnumber = row_number()) %>%
  ungroup() %>%
  unnest_tokens(word, paragraph, to_lower = TRUE, strip_punct = TRUE)

# manual stop words removal
my_stop_words <- tribble(
  ~word, ~lexicon,
  "millions", "CUSTOM",
  "million", "CUSTOM",
  "thousands", "CUSTOM",
  "thousand", "CUSTOM",
  "billions", "CUSTOM",
  "billion", "CUSTOM",
  "january", "CUSTOM",
  "february", "CUSTOM",
  "march", "CUSTOM",
  "april", "CUSTOM",
  "may", "CUSTOM",
  "june", "CUSTOM",
  "july", "CUSTOM",
  "august", "CUSTOM",
  "september", "CUSTOM",
```

```
"october", "CUSTOM",
"november", "CUSTOM",
"december", "CUSTOM",
"year", "CUSTOM",
"per", "CUSTOM",
"cent", "CUSTOM",
"p", "CUSTOM",
"a1", "CUSTOM",
"a2", "CUSTOM",
"op", "CUSTOM",
"cit", "CUSTOM",
"loc", "CUSTOM",
"p1", "CUSTOM",
"pn", "CUSTOM",
"cf", "CUSTOM",
"2nd", "CUSTOM",
"ff", "CUSTOM",
"die", "CUSTOM",
"4th", "CUSTOM",
"und", "CUSTOM",
"pp", "CUSTOM",
"jena", "CUSTOM",
"des", "CUSTOM",
"der", "CUSTOM",
```

```
  "ed", "CUSTOM",
  "vol", "CUSTOM",
  "ka", "CUSTOM",
  "ki", "CUSTOM",
  "pa", "CUSTOM",
  "pi ", "CUSTOM",
)

# remove numbers and remove "i.e"
tidy_data_words <- data_words_corr %>%
  anti_join(stop_words) %>%
  anti_join(my_stop_words) %>%
  filter(
    !grepl('^[0-9]+$', word)
    & !grepl('^i\\.e$', word)
  )

# replacements (particular words)
tidy_data_words$word <- replace(tidy_data_words$word,
tidy_data_words$word == "kr", "crown")
tidy_data_words$word <- replace(tidy_data_words$word,
tidy_data_words$word == "koruna", "crown")
tidy_data_words$word <- replace(tidy_data_words$word,
tidy_data_words$word == "krone", "crown")
tidy_data_words$word <- replace(tidy_data_words$word,
tidy_data_words$word == "kčč", "crown")
```

```
tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "kč", "crown")

tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "metallist", "metalist")

tidy_data_words$word <- replace(tidy_data_words$word, tidy_data_words$word == "metallism", "metalism")


word_cors <- tidy_data_words %>%

 filter(

#   Author == "Amonn Alfred"

   Author == "Englander Oskar"

#   Author == "Muller Hugo"

 ) %>%

 group_by(word) %>%

 filter(n() >= 10) %>% # threshold (can be adjusted based on dataset size): at least in 10 different pairs (number of relations)

 pairwise_cor(word, pnumber, sort = TRUE) # do include upper triangle of matrix


word_cors %>%

 filter(item1 %in% c("engliš", "rašín", "fořt", "czechoslovakia", "czechoslovak")) %>%

 group_by(item1) %>%

 slice_max(correlation, n = 6) %>%

 ungroup() %>%

 mutate(item2 = reorder(item2, correlation)) %>%
```

```
  ggplot(aes(drlib::reorder_within(item2, correlation, within = item1), correlation)) +
  geom_bar(stat = "identity") +
  facet_wrap(~ item1, scales = "free") +
  drlib::scale_x_reordered() +
  coord_flip()

# correlation relations - each text/author has own threshold (to have clear chart)
set.seed(123)
word_cors %>%
  filter(correlation > .51) %>%
  graph_from_data_frame() %>%
  ggraph(layout = "fr") +
  geom_edge_link(aes(edge_alpha = correlation), show.legend = FALSE) +
  geom_node_point(color = "gray", size = 5) +
  geom_node_text(aes(label = name), repel = TRUE) +
  theme_void()

######################################################################
# 05 - sentiment
######################################################################

# bigrams with paragraph number
data_bigrams_sent <- data_whole %>%
```

```
  unnest_tokens(paragraph, text, token = "paragraphs", to_lower = TRUE)
%>%

  filter(nchar(paragraph) > 50) %>%

  group_by(Author, Title) %>%

  mutate(pnumber = row_number()) %>%

  ungroup() %>%

  unnest_tokens(bigram, paragraph, token = "ngrams", n = 2) %>%

  filter(!is.na(bigram))


bigrams_separated_sent <- data_bigrams_sent %>%

  separate(bigram, c("word1", "word2"), sep = " ")


# replacements (particular words)

bigrams_separated_sent$word1 <- replace(bigrams_separated_sent$word1,
bigrams_separated_sent$word1 == "kr", "crown")

bigrams_separated_sent$word1 <- replace(bigrams_separated_sent$word1,
bigrams_separated_sent$word1 == "koruna", "crown")

bigrams_separated_sent$word1 <- replace(bigrams_separated_sent$word1,
bigrams_separated_sent$word1 == "krone", "crown")

bigrams_separated_sent$word1 <- replace(bigrams_separated_sent$word1,
bigrams_separated_sent$word1 == "kčč", "crown")

bigrams_separated_sent$word1 <- replace(bigrams_separated_sent$word1,
bigrams_separated_sent$word1 == "kč", "crown")

bigrams_separated_sent$word1 <- replace(bigrams_separated_sent$word1,
bigrams_separated_sent$word1 == "metallist", "metalist")

bigrams_separated_sent$word1 <- replace(bigrams_separated_sent$word1,
bigrams_separated_sent$word1 == "metallism", "metalism")
```

```
bigrams_separated_sent$word2 <- replace(bigrams_separated_sent$word2,
bigrams_separated_sent$word2 == "kr", "crown")

bigrams_separated_sent$word2 <- replace(bigrams_separated_sent$word2,
bigrams_separated_sent$word2 == "koruna", "crown")

bigrams_separated_sent$word2 <- replace(bigrams_separated_sent$word2,
bigrams_separated_sent$word2 == "krone", "crown")

bigrams_separated_sent$word2 <- replace(bigrams_separated_sent$word2,
bigrams_separated_sent$word2 == "kčč", "crown")

bigrams_separated_sent$word2 <- replace(bigrams_separated_sent$word2,
bigrams_separated_sent$word2 == "kč", "crown")

bigrams_separated_sent$word2 <- replace(bigrams_separated_sent$word2,
bigrams_separated_sent$word2 == "metallist", "metalist")

bigrams_separated_sent$word2 <- replace(bigrams_separated_sent$word2,
bigrams_separated_sent$word2 == "metallism", "metalism")


negation_words <- c(

 "not",

 "no",

 "never",

 "without",

 "cannot",

 "none",

 "nowhere",

 "neither",

 "nor",

 "isn't",

 "don't",
```

```
  "doesn't",
  "won't",
  "haven't",
  "hasn't"
)

not_words <- bigrams_separated_sent %>%
  filter(word1 %in% negation_words) %>%
  inner_join(get_sentiments("afinn"), by = c(word2 = "word")) %>%
  mutate(value = -2*value)

data_not_words_sent <- not_words %>%
  rename(word = word2) %>%
  select(-word1)

data_words_sent <- data_whole %>%
  unnest_tokens(paragraph, text, token = "paragraphs", to_lower = TRUE)
%>%
  filter(nchar(paragraph) > 50) %>%
  group_by(Author, Title) %>%
  mutate(pnumber = row_number()) %>%
  ungroup() %>%
  unnest_tokens(word, paragraph, to_lower = TRUE, strip_punct = TRUE)
%>%
  inner_join(get_sentiments("afinn"))
```

```
data_sentiment_afinn <- data_words_sent %>%
  bind_rows(data_not_words_sent) %>%
  group_by(Author, Title, Year, pnumber) %>%
  summarise(sentiment = sum(value)) %>%
  mutate(method = "AFINN") %>%
  ungroup() %>%
  arrange(pnumber)

data_sentiment_afinn %>%
  filter(Author == "Muller Hugo" & Title == "Zwei tschechische Schriften uber Wahrungsreform")

# skipped
ggplot(data_sentiment_afinn %>%
  filter(Author == "Muller Hugo" & Title == "Zwei tschechische Schriften uber Wahrungsreform"), aes(pnumber, sentiment, fill = Title)) +
  geom_col(show.legend = FALSE) +
  facet_wrap(~Title, ncol = 3)

data_sentiment_afinn %>%
  group_by(Author, Title) %>%
  mutate(sum = sum(sentiment)) %>%
  mutate(count = n()) %>%
  mutate(avg = sum / count) %>%
```

```
  select(-Year,-pnumber,-sentiment,-method) %>%
  distinct() %>%
  print(n=40)

data_matching_englis <- data_whole %>%
  filter(grepl("engliš", tolower(text)))

data_matching_rasin <- data_whole %>%
  filter(grepl("rašín", tolower(text)))

data_matching_fort <- data_whole %>%
  filter(grepl("fořt", tolower(text)))

data_matching_czech <- data_whole %>%
  filter(
    grepl("czechoslovakia", tolower(text))
    | grepl("czechoslovak", tolower(text))
  )

word_matched_paragraphs <- data_whole %>%
  filter(
    Title == "Volkswirtschaftliche Grundbegriffe und Grundprobleme"
  ) %>%
  unnest_tokens(paragraph, text, token = "paragraphs", to_lower = TRUE)
%>%
```

```
  filter(nchar(paragraph) > 50) %>%
  group_by(Author, Title) %>%
  mutate(pnumber = row_number()) %>%
  ungroup() %>%
  filter(grepl("engliš", paragraph)) %>%
  select(pnumber)


ggplot(
    data_sentiment_afinn %>%
      inner_join(data_matching_englis, by = c(Author = "Author", Title = "Title")) %>%
      filter(
        Title == "Volkswirtschaftliche Grundbegriffe und Grundprobleme"
      ) %>%
      filter(pnumber > 400) %>%
      mutate(fill = ifelse(pnumber %in% word_matched_paragraphs$pnumber, "matched", "notmatched")),
    aes(pnumber, sentiment)
  ) +
  geom_col(aes(pnumber, fill = fill), show.legend = FALSE) +
  geom_text(aes(label=sentiment, vjust=0.5-sign(sentiment))) +
  facet_wrap(~Title, ncol = 2, scales = "free_x") +
  scale_fill_manual(values=c("matched" = "#000000", "notmatched" = "#999999"))


word_matched_paragraphs %>
```

```
    inner_join(data_matching_englis, by = c(Author = "Author", Title =
"Title")) %>%
    filter(
      Title == "Volkswirtschaftliche Grundbegriffe und Grundprobleme"
    ) %>%
    filter(n > 40)



#####################################################################
06 STM/LDA ##########################################################
#####################################################################

library(quanteda) # For NLP
library(topicmodels) # For topicmodels (and LDA)
library(stm) # For structural topic models (STM)

mycorpus <- corpus(
  data_whole %>%
  filter(Author == "Amonn Alfred")
#  filter(Author == "Englander Oskar")
#  filter(Author == "Muller Hugo")
)
mycorpus.stats <- summary(mycorpus)

token <-
```

```
  tokens(
    mycorpus,
    split_hyphens = TRUE,
    remove_numbers = TRUE,
    remove_punct = TRUE,
    remove_symbols = TRUE,
    remove_url = TRUE,
    include_docvars = TRUE
  )

merged_stopwords <- stopwords("english") %>%
  append(c(
    "pp",
    "ibidem",
    "ki",
    "ka",
    "ma",
    "pi",
    "pa",
    "p1",
    "pn",
    "0.25",
    "aa",
    "ai",
```

"kpa",

"p",

"millions",

"million",

"thousands",

"thousand",

"billions",

"billion",

"january",

"february",

"march",

"april",

"may",

"june",

"july",

"august",

"september",

"october",

"november",

"december",

"year",

"per",

"cent",

"a1",

```
  "a2",
  "op",
  "cit",
  "loc",
  "cf",
  "2nd",
  "ff",
  "die",
  "4th",
  "und",
  "pp",
  "jena",
  "des",
  "der",
  "ed",
  "vol",
  "pi "
 ))

mydfm <- dfm(
  token,
  tolower = TRUE,
  stem = TRUE,
  remove = merged_stopwords
```

```
)

mydfm.trim <-
  dfm_trim(
    mydfm,
    min_docfreq = 0.075,
    # min 7.5%
    max_docfreq = 0.90,
    # max 90%
    docfreq_type = "prop"
  )

topic.count <- 10

dfm2topicmodels <- convert(mydfm.trim, to = "topicmodels")
lda.model <- LDA(dfm2topicmodels, topic.count)

# print model words as table
as.data.frame(terms(lda.model, 10))

lda.similarity <- as.data.frame(lda.model@beta) %>%
  scale() %>%
  dist(method = "euclidean") %>%
  hclust(method = "ward.D2")
```

```
par(mar = c(0, 4, 4, 2))


# dendrogram
plot(lda.similarity,
     main = "LDA topic similarity by features - Muller",
     xlab = "",
     sub = "")


# less topics (targetted)


mycorpus <- corpus(
  data_whole %>%
#  filter(Author == "Amonn Alfred")
#  filter(Author == "Englander Oskar")
  filter(Author == "Muller Hugo")
)
mycorpus.stats <- summary(mycorpus)


token <-
  tokens(
    mycorpus,
    split_hyphens = TRUE,
    remove_numbers = TRUE,
```

```
    remove_punct = TRUE,
    remove_symbols = TRUE,
    remove_url = TRUE,
    include_docvars = TRUE
  )

merged_stopwords <- stopwords("english") %>%
  append(c(
    "pp",
    "ibidem",
    "ki",
    "ka",
    "ma",
    "pi",
    "pa",
    "p1",
    "pn",
    "0.25",
    "aa",
    "ai",
    "kpa",
    "p",
    "millions",
    "million",
```

```
"thousands",
"thousand",
"billions",
"billion",
"january",
"february",
"march",
"april",
"may",
"june",
"july",
"august",
"september",
"october",
"november",
"december",
"year",
"per",
"cent",
"a1",
"a2",
"op",
"cit",
"loc",
```

```
    "cf",
    "2nd",
    "ff",
    "die",
    "4th",
    "und",
    "pp",
    "jena",
    "des",
    "der",
    "ed",
    "vol",
    "pi "
  ))

mydfm <- dfm(
  token,
  tolower = TRUE,
  stem = TRUE,
  remove = merged_stopwords
)

mydfm.trim <-
  dfm_trim(
```

```
    mydfm,
    min_docfreq = 0.075,
    # min 7.5%
    max_docfreq = 0.90,
    # max 90%
    docfreq_type = "prop"
  )

topic.count <- 4

dfm2topicmodels <- convert(mydfm.trim, to = "topicmodels")
lda.model <- LDA(dfm2topicmodels, topic.count)

# print model words as table
as.data.frame(terms(lda.model, 10))

lda.similarity <- as.data.frame(lda.model@beta) %>%
  scale() %>%
  dist(method = "euclidean") %>%
  hclust(method = "ward.D2")

par(mar = c(0, 4, 4, 2))

# less topics based on dendogram (mark with circles common topics)
```

```
topic.count <- 3
plot(lda.similarity,
     main = "LDA topic similarity by features - Amonn",
     xlab = "",
     sub = "")

# STM

mycorpus <- corpus(
  data_whole %>%
#  filter(Author == "Amonn Alfred")
#  filter(Author == "Englander Oskar")
  filter(Author == "Muller Hugo")
)
mycorpus.stats <- summary(mycorpus)

token <-
  tokens(
    mycorpus,
    split_hyphens = TRUE,
    remove_numbers = TRUE,
    remove_punct = TRUE,
    remove_symbols = TRUE,
    remove_url = TRUE,
```

```
  include_docvars = TRUE
 )

merged_stopwords <- stopwords("english") %>%
 append(c(
  "pp",
  "ibidem",
  "ki",
  "ka",
  "ma",
  "pi",
  "pa",
  "p1",
  "pn",
  "0.25",
  "aa",
  "ai",
  "kpa",
  "p",
  "millions",
  "million",
  "thousands",
  "thousand",
  "billions",
```

```
"billion",
"january",
"february",
"march",
"april",
"may",
"june",
"july",
"august",
"september",
"october",
"november",
"december",
"year",
"per",
"cent",
"a1",
"a2",
"op",
"cit",
"loc",
"cf",
"2nd",
"ff",
```

```
    "die",
    "4th",
    "und",
    "pp",
    "jena",
    "des",
    "der",
    "ed",
    "vol",
    "pi "
  ))

mydfm <- dfm(
  token,
  tolower = TRUE,
  stem = TRUE,
  remove = merged_stopwords
)

mydfm.trim <-
  dfm_trim(
    mydfm,
    min_docfreq = 0.075,
    # min 7.5%
```

```
  max_docfreq = 0.90,
  # max 90%
  docfreq_type = "prop"
 )

topic.count <- 5

dfm2stm <- convert(mydfm.trim, to = "stm")

model.stm <- stm(
 dfm2stm$documents,
 dfm2stm$vocab,
 K = topic.count,
 data = dfm2stm$meta,
 init.type = "Spectral"
)

# print topic terms
as.data.frame(t(labelTopics(model.stm, n = 10)$prob))

plot(
 model.stm,
 type = "summary",
 text.cex = 1,
```

```
  main = "STM topic shares - Amonn",
  xlab = "Share estimation"
)
```