



University of
Zurich ^{UZH}

News Across the Röstigraben

How News Reporting Differs Across the French and German Swiss Language Regions. A Computational Analysis of the Online Publication "20 Minuten".

Anne-Sophie Skarabis, 20-725-487

`anne-sophie.skarabis@uzh.ch`

Fall Semester 2023

Multilingual Computational Methods for Text-based Media

University of Zurich, Department of Communication and Media Research (IKMZ)

Morley Weston, MA

Zurich, January 11, 2024

GitHub Code Repository

Please find the code for my final project in Computational Methods at the following GitHub repository:

github.com/anne-tanne/askara-final-project-computational-methods

Abstract

This paper analyses the differences and similarities in reporting between the German and French editions of "20 Minuten", a leading Swiss newspaper. Using natural language processing and data analysis, 1,954 online articles from 10 to 24 November 2023 were analysed. The study focuses on regional and international reporting and the sentiment in the headlines. The results indicate a strong focus on the German-speaking regions in the German edition, while the French edition has a less pronounced focus on French-speaking Switzerland. In addition, the French edition has a broader international perspective. Both editions show a tendency towards negative headlines. This sheds light on how linguistic and regional preferences influence news coverage in a multilingual country like Switzerland.

Contents

1	Introduction	1
2	Theoretical Foundations	1
2.1	Media Consumption and Integration Patterns	2
2.2	Domestic Reporting in Switzerland	2
2.3	International Reporting in Switzerland	3
2.3.1	News Factor Theory	3
2.4	The Role of Sentiment in Headlines: News Factors and Audience Attention	4
3	Methods	4
3.1	Location Analysis	5
3.2	Sentiment Analysis	5
4	Results	6
4.1	Regional Coverage Analysis – RQ1	6
4.2	Country Analysis, RQ2	7
4.3	Sentiment Variations Across Headlines, RQ3	9
4.3.1	Sentiment Distribution Within Editions	10
5	Discussion	10
5.1	Interpretation of Regional Focus in Media Coverage	10
5.2	Patterns of International Coverage	11
5.3	Sentiment Analysis and Its Implications	11
6	Limitations	11

1 Introduction

This paper aims to provide insights into the potential disparities and similarities in news reporting between the German- and French-language editions of "20 Minuten".

"20 Minuten"¹ is a free daily newspaper in Switzerland available in German, French and Italian. With a total reach of around 3.25 million people (WEMF AG, 2023), "20 Minuten" is the most widely used private medium in Switzerland and the press market leader in Switzerland (Swiss Federal Department of Foreign Affairs, 2023).

To achieve the objectives of this paper, the study takes a quantitative approach using natural language processing and data analysis tools. The following sections discuss the theoretical background, methodological approach, results and implications for understanding the differences in Swiss news reporting.

2 Theoretical Foundations

To understand the Swiss media landscape, it is important to recognise one of its characteristic features: its regional segmentation. "Given that media markets are languagemarkets, media markets are even smaller in countries with different language communities" (Puppis, 2009, p. 10). Switzerland's media market is hence segmented into three submarkets based on its three dominant language regions: German-speaking Switzerland, French-speaking Romandie and Italian-speaking Ticino (Beier et al., 2020, p. 1056).

This paper focuses on the media markets of German-speaking Switzerland and Romandie, which account for the largest share of media usage. The German-speaking market with 5.2 million adults and the Romandie market with 1.8 million adults are the primary focus, as they have a larger audience compared to the Italian-speaking Ticino, which has only 300,000 adults (Udris et al., 2023, p. 5). These two markets provide an interesting context for this study: the online editions of "20 Minuten", which share the same website across both regions, provides the basis for this research. In contrast, the Italian-language edition "20 Minuti" differs as its contents are integrated into the *tio.ch* portal.

¹"20 Minutes" in Romandie. For the purposes of this paper, all references to the newspaper, regardless of the language edition, will be uniformly referred to as "20 Minuten".

2.1 Media Consumption and Integration Patterns

The Reuters Institute News Report 2023 shows that the public broadcaster SRF/ RTS and the private outlet "20 Minuten" dominate media consumption in both language regions. This dominance is apparent in both traditional media formats and online platforms, with "20 Minuten" leading online consumption. Consumption patterns show similar trends in both regions.

The report also shows that media consumption in Switzerland is influenced by the exposure to broadcasting stations from neighboring countries. In both the German and French regions, private and public channels from neighbouring Germany and France rank third and fourth in popularity. As stated by Signer et al. (2011, p. 422): "Each language region constitutes a (...) small media system (...) with a (...) giant neighbor (...) sharing the same language. As a consequence, the market shares of overspilling foreign (...) channels in all three regions are higher than 60 percent."

2.2 Domestic Reporting in Switzerland

Domestic reporting plays a crucial role in creating a close link between the public and the democratic nation state (Gerhards, 1998, p. 271). Strong domestic reporting minimises the problems of divergence between the public sphere and jurisdiction. Nevertheless, Switzerland is confronted with the problem of geographical segmentation and fragmentation of media consumption, with certain regions being isolated or focusing more on media from neighbouring countries. This situation raises questions about the cohesion and integrative role of domestic reporting in Switzerland. The Swiss Broadcasting Corporation SRG SSR is therefore mandated by its licence (Article 3.4) to promote knowledge about the different parts of the country (BAKOM, 2018, p. 2) in order to ensure integration at regional level for a national public (Beier et al., 2020, p. 1056).

Still, this objective is only achieved to a limited extent: SRG SSR programming tends to focus more on foreign countries and national events, with Italian-speaking Switzerland often being underrepresented in their broadcasts (Publicom AG, 2015). The 2019 "Jahrbuch Qualität der Medien" confirms these findings: The report shows that German-speaking Switzerland receives a high proportion of coverage in Romandie and Ticino (33% and 36% respectively), while the media in the German-speaking part dedicate

considerably less coverage to French-speaking areas (13%) and Ticino (5%). Conversely, Ticino's media devotes 15% of its coverage to Romandie, which in turn allocates only 4% to Ticino. This indicates an uneven distribution of domestic reporting, which may impact the integrative role of media in Swiss democracy.

These differences between the language regions lead to the hypothesis that such patterns of presentation will also be found in the language editions of "20 Minuten", forming the basis for the first research question:

RQ1: Is there a recognisable difference in regional geographical focus that is reflected in the frequency and distribution of place names in the articles of the language-specific editions of "20 Minuten"?

2.3 International Reporting in Switzerland

Foreign reporting is important for several reasons: From a nation-state perspective, it aids in learning from others, enhances risk perception, and helps shape a collective identity through external contrasts. From a cosmopolitan viewpoint, it highlights transnational networks, fosters a global collective identity, encourages appreciation for cultural diversity, and promotes dialogue across national backgrounds (Brüggemann & Königslöw, 2013; Imhof, 2011). Swiss media demonstrates a traditionally intensive and diverse foreign news reportage, indicating a cosmopolitan editorial stance (Brüggemann & Königslöw, 2013). This international focus is also influenced by Switzerland's proximity to its major neighbouring countries (Beier et al., 2020, p. 1061).

2.3.1 News Factor Theory

International reporting often features certain nations more prominently. This can be related back to news factor theory (Galtung & Ruge, 1965), where the factor "Elite Nations" increases the likelihood of a news story being chosen for publication.

Generally, the international reporting landscape within Swiss media is characterised by a focus on a select group of countries, with prominent coverage of the USA, France, Germany and the UK reflecting factors like geopolitical proximity and relevance to Swiss interests (Wilke et al., 2012, p. 314). There too is a concentration of reporting on nations with significant political and economic impact (such as Israel

or Palestine). This pattern forms the basis for the following research question:

RQ2: Which countries are predominantly covered in the German and French editions of "20 Minuten"?

Are there differences in the international geographical focus as reflected in the place names of the articles in the language-specific editions of "20 Minuten"?

2.4 The Role of Sentiment in Headlines: News Factors and Audience Attention

Headlines serve as the initial connection between news and readers. Emotions in headlines, as suggested by Galtung and Ruge (1965), are likely to attract the audience's attention, especially in the context of negativity. This is coherent with negativity bias theory, which states that negative events or emotions have a stronger influence on a person's psychological state and processes than positive ones (Rozin & Royzman, 2001). In news reporting, this bias means that headlines with negative emotions are more likely to engage readers and shape their perception of news value and importance (Shoemaker & Reese, 2014)².

This leads us to the following research question:

RQ3: How do sentiment trends in "20 Minuten" headlines vary among the language-specific editions?

3 Methods

A total of 1,954 articles from the 20 Minuten online portals were collected for this study (German = 894, French = 1060). The time frame for this collection was from 10 November to 24 November 2023. The articles were collected as full census by scraping the News sitemaps of the respective "20 Minuten" websites.

To facilitate data collection, a series of automated scripts were developed using the R programming language. The data extraction process included steps to clean up file names to account for multilingual content and to systematically organise the retrieved data for subsequent analysis. The end result were two

²Shoemaker and Reese (2014, p. 55) note that this tendency towards negativity bias is particularly pronounced in men, while women are often more receptive to positive news

data frames, each containing the headline, the header (typically 1-3 words located above the headline) and the content of the articles from the respective language editions. The HTML files were also saved as part of the data collection process to provide a complete record of the online articles for future reference and analysis.

3.1 Location Analysis

The location analysis of the study was carried out using Python, where an analysis of the geographical references in the articles of both "20 Minuten" editions was conducted. The main goal was to identify and quantify mentions of countries and cities to gain insight into the geographical focus of the coverage.

The `spaCy` NLP library was used to identify geographical entities in the article texts. For this purpose, language models for German (`de_core_news_sm`) and French (`fr_core_news_sm`) were used. Recognised locations were then matched against a dataset of geographical units. The scope of the location survey included all countries worldwide, while only Swiss cities were surveyed. The *Position-Stack API* was used to convert the place names (in the case of countries, their capitals) into latitude and longitude coordinates.

The results of the location analysis were aggregated and the frequency of mentions of countries and cities was calculated. These were then merged with the corresponding geographical data and exported to `.csv` files for further analysis and visualisation.

3.2 Sentiment Analysis

To analyse the variance in sentiment trends between the German and French articles, a sentiment analysis was conducted using the *John Snow Labs SparkNLP* library. The reason for this choice was the availability of comparable sentiment analysis models for German and French.

The process began with the setup of `PySpark` and `SparkNLP`. The article headlines, which were initially stored in `.csv` format, were transferred to `Spark DataFrames` to meet the requirements of the `SparkNLP` processing environment.

The sentiment analysis pipeline was created using Spark's NLP tools. The pipeline started with the `DocumentAssembler`, which converted the article headings into a suitable format for NLP pro-

cessing. The `BertSentenceEmbeddings` model, in particular the `labse` variant, was then used to generate embeddings from these headings. These embeddings serve as numerical representations of the text and capture their linguistic nuances. The core of the pipeline was the `ClassifierDLModel` (`classifierdl_bert_sentiment`) trained for sentiment analysis. This model had the task of classifying each headline as either positive or negative.

Post-processing, each headline was assigned a sentiment label, adding a column containing the sentiment information to the dataset.

4 Results

4.1 Regional Coverage Analysis – RQ1

In the German edition of "20 Minuten", there was a clear focus on cities in German-speaking Switzerland. The vast majority (89.4%) of city mentions related to this region, which indicates a strong regional focus. The Romandie and Ticino regions, on the other hand, received significantly less attention (9.26% and 1.32% respectively).

In the French edition, the focus was clearly on the Romandie (58.7% of city mentions), but German-speaking regions still accounted for 39.6% of city mentions, indicating a more pronounced concentration on areas outside the primary language group. Similar to the German edition, the French edition allocated only about 1.7% of its coverage to the Ticino.

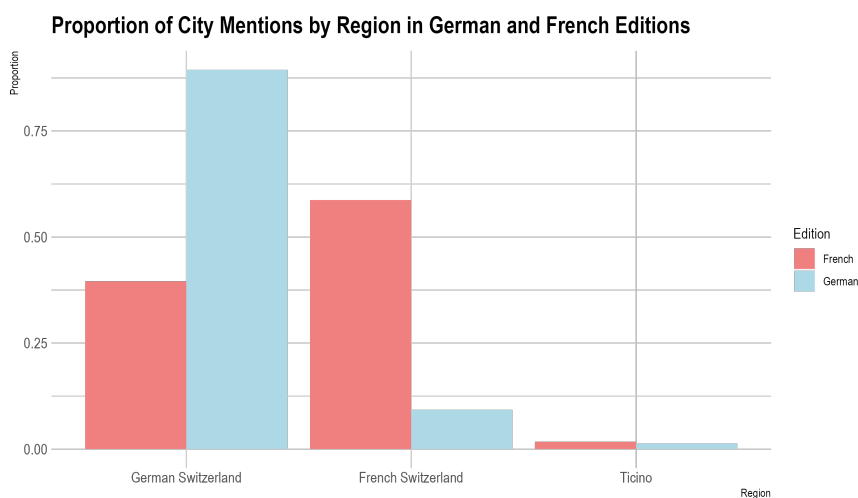


Figure 1: Proportion of City Mentions by Region in German and French Editions

The application of a chi-square test to compare the distribution of city mentions per region between the German and French editions yielded statistically significant results. The p-value ($< 2.2e - 16$) suggests significantly different editorial focuses. This finding is consistent with the hypothesis set out alongside RQ1.

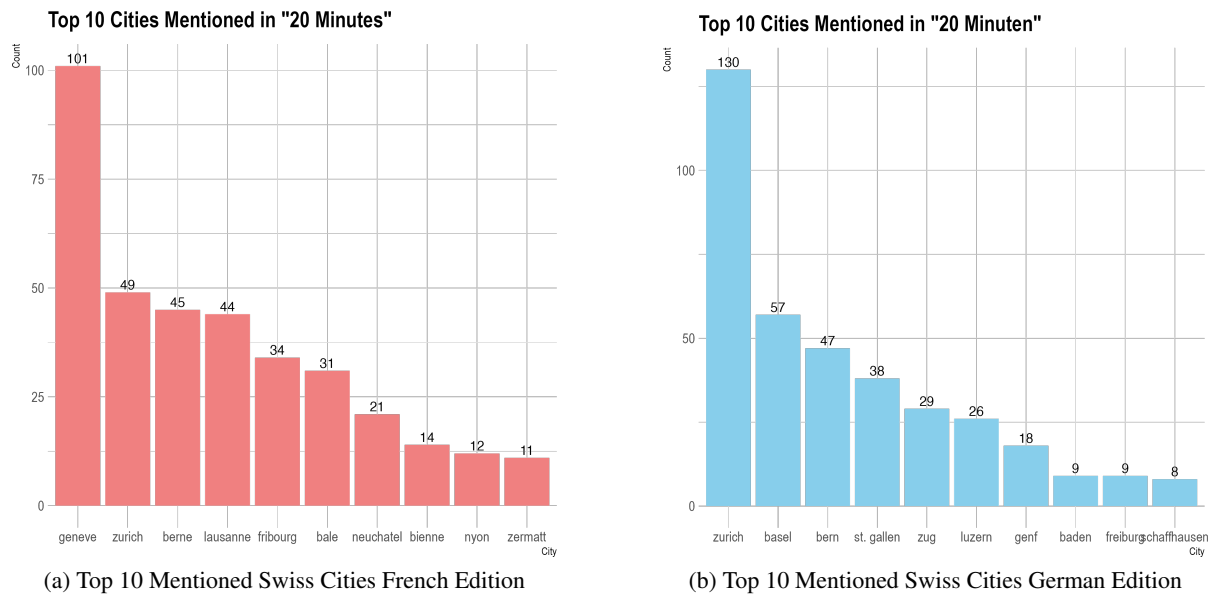


Figure 2: Top 10 Mentioned Swiss Cities in German and French Language Editions

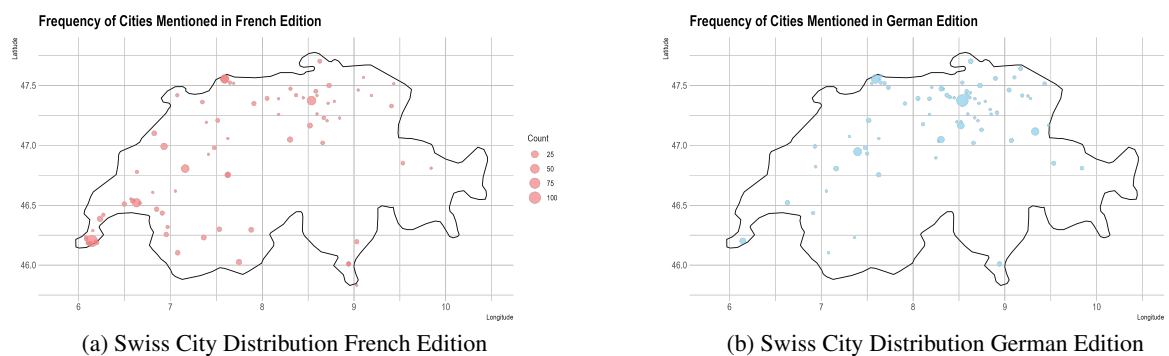


Figure 3: Swiss City Distribution in German and French Language Editions

4.2 Country Analysis, RQ2

The analysis of the ten most important countries mentioned in the German and French editions of "20 Minuten" revealed that the two editions had a broadly similar focus. In the German edition, Switzerland, Germany and the USA were mentioned most frequently. The French edition focused predominantly on Switzerland, the USA and France. This reflects the cultural and geopolitical influences on reporting

discussed in Chapter 2.1.

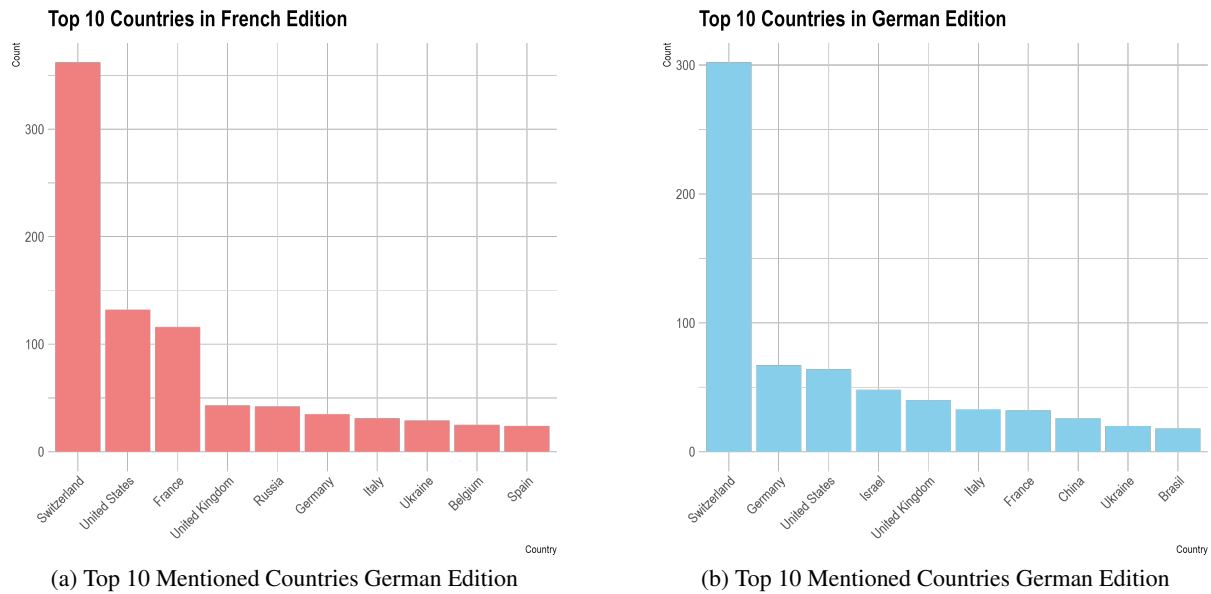


Figure 4: Top 10 Mentioned Countries in German and French Language Editions

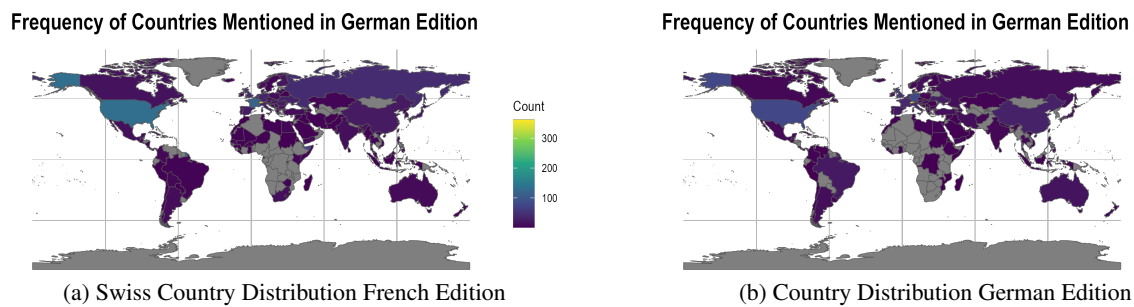


Figure 5: Distribution in German and French Language Editions

Germany and France were mentioned in both language versions. In addition, other countries such as the United Kingdom, Ukraine and Italy were frequently mentioned in both editions. In the German edition, Israel, China and Brasil were mentioned in the Top 10, while the French edition mentioned Belgium, Russia and Spain.

Building on these observations, a series of chi-square tests were performed to analyse the frequencies in more detail. These tests showed that the proportion of articles that mentioned only Switzerland was significantly ($p < 1.532e^{-05}$) higher in the German edition (32.9%) than in the French one (24.0%). In addition, the proportion of articles that did not mention any country was significantly higher ($p = 9.576e^{-06}$) in the German edition (27.4%) compared to the French edition (18.9%). This suggests that

the German edition may have a greater focus on domestic or non-geographically specific topics.

In the French edition, the proportion of articles in which only non-Swiss countries were mentioned was significantly ($p < 2.2e^{-16}$) higher (46.9%) than in the German edition (23.2%). This further indicates a broader international focus in the French edition. When looking at articles in which only Switzerland's neighbours are mentioned, the French edition (12.9%) has a higher proportion than the German edition (6.0%), with a p-value of $4.941e^{-07}$.

Articles mentioning countries that are neither Switzerland nor its neighbouring countries were also more frequent in the French edition (52.9%) than in the German edition (44.5%), with a p-value of 0.0002548.

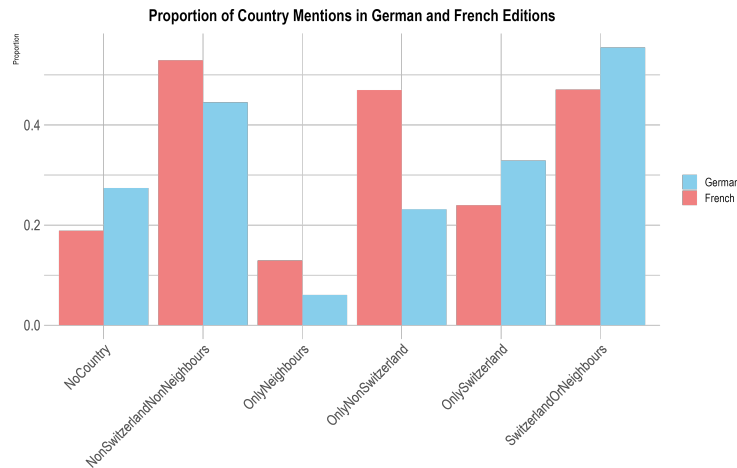


Figure 6: Proportion of Country Mentions in German and French Editions

4.3 Sentiment Variations Across Headlines, RQ3

To compare the prevalence of negative sentiment between the German and French editions, a two proportion z-test was performed. The test revealed no significant difference ($p = 0.6791$) in the proportion of negative headlines (58.4% in the German edition and 57.4% in the French edition). There was also no significant difference ($p = 0.6791$) in the comparison of positive headlines between the two editions (41.6% and 42.6% respectively). These results suggest that the editorial attitude towards sentiment is consistent across the different language editions.

4.3.1 Sentiment Distribution Within Editions

Conversely, the intra-edition sentiment comparison was significant. The chi-squared tests for both the German ($p = 5.255e^{-07}$) and French ($p = 1.655e^{-06}$) showed a statistically significant difference in the distribution of positive and negative headlines. This suggests that, within each language edition, there is a clear editorial bias towards a negative headline sentiment. These findings are consistent with the perspective that media headlines tend to reflect a negative bias (see Chapter 2.4).

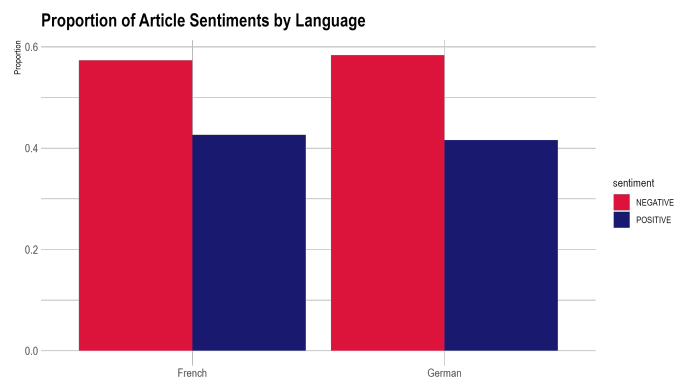


Figure 7: Proportion of Positive and Negative Article Sentiments by Language Edition

5 Discussion

5.1 Interpretation of Regional Focus in Media Coverage

The disproportionate focus on the German-speaking regions in the German edition of "20 Minuten", which is reflected in the minimal coverage of Romandie and Ticino, aligns with previous research (see Chapter 2.2) and further raises the question of the role of the media in promoting national cohesion. This editorial focus not only leaves out important Swiss population groups, but also raises the question of a fair representation of the different regional identities. This emphasises the need for a more balanced media presence that reflects the multilingual nature of Swiss society and supports an inclusive national narrative.

5.2 Patterns of International Coverage

The country analysis of "20 Minuten" editions reveals a similar international focus in both German and French editions, with Switzerland prominently featured, followed by other major Western nations. The German edition's emphasis on Germany and the French edition's on France is in line with established news value theory, according to which geographical proximity influences news selection (Galtung & Ruge, 1965). In this context, it would be interesting to further analyse what influence the overspilling broadcasters have on the perception of relevance of the respective population groups.

The German edition shows a higher frequency of exclusively mentioning Switzerland, whereas the French edition leans more towards international news. However, while these differences are statistically significant, it's important to note the potential for the Chi-squared test to overstate significance in large samples. Therefore, these findings should be interpreted with caution and explored further in more comprehensive studies for a deeper understanding of the editorial choices and their cultural implications.

5.3 Sentiment Analysis and Its Implications

The lack of significant differences in sentiment between the German and French editions indicates a consistent editorial policy that transcends language barriers. This could be due to the fact that both papers follow a similar editorial line and can be categorised as tabloid media. More research would need to be done to find out if different languages have an impact on news sentiment.

The greater tendency towards negative headlines within the editions is inline with Galtung and Ruge (1965) and confirms the claims of Shoemaker and Reese (2014) regarding the impact of negative news on audience engagement.

6 Limitations

One limitation of this study is the scope of the data collection. While the analysis includes a significant number of articles from the "20 Minuten" portals, the selection of articles was limited to a specific time period (10 November to 24 November 2023). This may not fully capture the broader trends and nuances

in coverage that could emerge over a longer period or at different times of the year.

Another limitation concerns the methodology used for sentiment and location analysis. Sentiment analysis relies on automated NLP tools that do not always correctly interpret the nuances and context of language, which can lead to misclassifications. Location analysis, on the other hand, was dependent on the accuracy and scope of the `spacy` NLP library and the PositionStack API. While a sample of articles was checked for accuracy of sentiment and location data, outliers and anomalies could still be present, potentially skewing the results.

Finally, I am a novice in media analysis using natural language processing and data analytics. Therefore, I may have overlooked or misunderstood certain aspects of the methodology or interpretation of the results. The results of this study should therefore be viewed as a learning process.

References

- BAKOM. (2018, August). SRG-Konzession. *Bundesamt für Kommunikation BAKOM*. <https://www.bakom.admin.ch/bakom/de/home/elektronische-medien/informationen-ueber-radio-und-fernsehveranstalter/srg-ssr/konzessionierung-und-technik-srg-ssr.html>
- Beier, A., Fiechtner, S., & Trebbe, J. (2020). Ordinary Citizens in Swiss Public Television News: Representing a National Public Sphere? *Journalism Studies*, 21(8), 1055–1075. <https://doi.org/10.1080/1461670X.2020.1771193>
- Brüggemann, M., & Königslöw, K. K.-v. (2013). Explaining cosmopolitan coverage. *European Journal of Communication*, 28(4), 361–378. <https://doi.org/10.1177/0267323113484607>
- Forschungszentrum Öffentlichkeit und Gesellschaft. (2019). *Jahrbuch Qualität der Medien 2019 (Jubiläumsausgabe)* (U. of Zurich, Ed.). Schwabe.
- Galtung, J., & Ruge, M. H. (1965). The Structure of Foreign News: The Presentation of the Congo, Cuba and Cyprus Crises in Four Norwegian Newspapers. *Journal of Peace Research*, 2(1), 64–90. <https://doi.org/10.1177/002234336500200104>
- Gerhards, J. (1998). Öffentlichkeit [The Public Sphere]. In O. Jarren, U. Sarcinelli, & U. Saxer (Eds.), *Politische Kommunikation in der demokratischen Gesellschaft: ein Handbuch mit Lexikonteil [Political Communication in the Democratic Society]* (pp. 268–274). Westdeutscher Verlag.
- Imhof, K. (2011). *Die Krise der Öffentlichkeit: Kommunikation und Medien als Faktoren des sozialen Wandels*. Campus.
- Newman, N., Fletcher, R., Eddy, K., Robertson, C. T., & Nielsen, R. K. (2023). *Reuters Institute Digital News Report 2023*. Reuters Institute. Retrieved January 3, 2024, from https://policycommons.net/artifacts/4164711/digital_news_report_2023/4973510/
- Publicom AG. (2015, December 11). Programmanalyse SRG-Radios 2014. Retrieved January 24, 2024, from <https://www.publicom.ch/2015/12/11/programmanalyse-der-srg-radios-2014/>
- Puppis, M. (2009). Media Regulation in Small States. *International Communication Gazette*, 71(1-2), 7–17. <https://doi.org/10.1177/1748048508097927>

- Rozin, P., & Royzman, E. B. (2001). Negativity Bias, Negativity Dominance, and Contagion. *Personality and Social Psychology Review*, 5(4), 296–320. https://doi.org/10.1207/S15327957PSPR0504_2
- Shoemaker, P. J., & Reese, S. D. (2014). *Mediating the message in the 21st century : A media sociology perspective* ([3rd ed.]). Routledge.
- Signer, S., Puppis, M., & Piga, A. (2011). Minorities, integration and the media: Media regulation and media performance in multicultural and multilingual Switzerland. *International Communication Gazette*, 73(5), 419–439. <https://doi.org/10.1177/1748048511405818>
- Swiss Federal Department of Foreign Affairs. (2023, December 28). *Medien*. Retrieved January 4, 2024, from <https://www.eda.admin.ch/aboutswitzerland/de/home/wirtschaft/taetigkeitsgebiete/medien.html>
- Udris, L., Vogler, D., Weston, M., & Eisenegger, M. (2023). Sourcing practices of online news media in Switzerland during the war in Ukraine. *Frontiers in Political Science*, 5. <https://doi.org/10.3389/fpos.2023.1089994>
- WEMF AG. *MACH Total Audience 2023-2*. MACH AG. 2023.
- Wilke, J., Heimprecht, C., & Cohen, A. (2012). The geography of foreign news on television: A comparative study of 17 countries. *International Communication Gazette*, 74(4), 301–322. <https://doi.org/10.1177/1748048512439812>