sciendo
DOI 10.2478/jazcas-2021-0060

INCOME, NATIONALITY AND SUBJECTIVITY IN MEDIA TEXT

IRENE ELMEROT
Department of Slavic & Baltic Studies, Finnish, Dutch & German, Stockholm University, Stockholm, Sweden

ELMEROT, Irene: Income, nationality and subjectivity in media text. Journal of Linguistics, 2021, Vol. 72, No 2, pp. 667 – 678.

Abstract: This article takes a bird's eye view of how positive or negative sentiments in the news press about countries and nationality nouns seem to reflect the country's general income groups. The study focuses on the four income groups classified by the World Bank and their co-occurrence with positively and negatively classified adjectives from the Subjectivity Lexicon for Czech. A search in the journalistic subcorpus of the SYN series, release 8 of the Czech National Corpus, results in a time line covering three decades. Previous research on subjectivity has either focused on other parts of the Subjectivity Lexicon or on fewer adjectives from other languages. In this article, it is argued that the income groups are treated in descending order, i.e., the higher the income, the more positive the sentiment. Even when the most influential groups in the top and bottom are removed, the result holds. Discourse concerning global war and peace, and the security of different nations, is also detected as a result.

Keywords: income groups, news press, sentiment, nationality, corpus linguistics, Czech language

1 INTRODUCTION

The focus of this article is the overall representation of different economic groups of countries and nationals in the Czech news press over three decades. The study concerns linguistic othering through the positive or negative value of adjectives co-occurring with nouns for these countries and nationals. In this study the "others" are considered to be part of a different income group from Czechs on an average, national level, whereas the same income group as that of the Czech Republic is considered to be the "ingroup". Discourses from the "Western" side of the Iron Curtain entered the Czech Republic at the end of 1989, after the transition from a Communist one-party to a multi-party state. The printed news media was then, and for many years to come, important in providing some kind of relation between the inhabitants, on the one hand, and what Fairclough [1] calls this new "outside", on the other. Many of the former Soviet-led countries are neighbours of the Czech Republic, and even today receive more news coverage due to their geographical location and cultural proximity ([2], [3]) than they do in news articles written in countries farther away. Most of these countries are now economically stronger than before 1989 and,

like the Czech Republic, are classified by the World Bank as High Income countries, with the exceptions of Bosnia and Herzegovina, Bulgaria, Kosovo, Montenegro, Romania, and the Russian Federation, which are all classified as Upper-middle Income countries [4]. In this article, discourse is seen as recurring, prominent claims about something (or someone), that makes a difference in the way the receivers of the discourse talk or write about these things or (groups of) persons [5].[1]

The purpose of this article is to examine whether theories proposed for other languages, such as those of Fairclough [1, p. 29] (see above) and Wodak [5, p. 10] on discursive constructions of national identities (see below), and Chovanec and Molek-Kozakowska [8] on the fluctuation of othering over time, also hold for the printed news press in the Czech Republic. The time span is 1990–2018, covering almost 30 years of democratic governance since the former Communist one-party state. The research aim is to conduct a corpus-based analysis of linguistic othering over time, using a search in which pre-defined adjectives from the Czech Subjectivity Lexicon [9] that co-occur with pre-defined nouns (see below under Material) are collected into a dataset extracted from the corpus. The dataset includes some variables from the corpus, such as publication year and title of the source, but also includes variables using World Bank income classifications.

1.1 Previous research and the contribution of this study

Veselovská describes studies that combine sentiment and discourse in extensive detail [10] and which essentially explain how language is not only a construction of human interaction but an influencer thereof. The sentiment analysis in the Czech Subjectivity Lexicon by Šindlerová et al. [11] is mainly concerned with evaluative verbs, which is one reason for this study to focus on adjectives. Paradis et al. [12] study 42 antonymic adjective pairs in English, focusing on their semantic profiles, whereas this study considers 773 adjectives co-occurring with the specified nouns.

The four income groups of the world's economies, as defined by the World Bank, are rarely used in linguistic studies, but more often in other fields, such as United Nation reports [13] or medical articles ([14], [15]). Language issues relating to these groups or countries have, however, been studied ([16], [17]). Theoretically, this article follows Fairclough's note that there is a need for condensation when researching relations in politics and economics, simply because those subjects are complex in themselves [1, p. 18]. Another of the ideas behind this study is to examine a larger whole (i.e., a bird's eye view over three decades) to see which details emerge, since they may show us the prominent part(s) of the rhetoric [ibid., p. 18] used about these groups over time. It also aims to connect to the ideas in Wodak et

[1] One example is how a male-dominated discourse has meant that many languages have male-only connotations of certain professions [6]; another is how a European-dominated discourse made people believe that all Africans were lazy [7].

al. [5, p. 10] regarding connotations of national identities. The co-occurring adjectives and nouns in this study show how the countries (and their citizens) in these income groups are, and have been, reflected in Czech printed news media over three decades. The binary classification reflected in (positive) peaks and (negative) troughs for the four groups is based on the recently mentioned Subjectivity Lexicon.

2 RESEARCH QUESTIONS AND MATERIAL

The research question of the study is:

In what way does the binary sentiment of the adjectives correlate with the income group?

The hypothesis for this study is that the higher the income group, the more positive their overall level. To demonstrate this, the binary classification of adjectives is represented in the form of graphs.

2.1 Material

Two main sources have been used to create the dataset: the adjectives in the Subjectivity Lexicon, and a journalistic subcorpus of the SYN series, release 8, containing 4,499,370,372 tokens (running words, excluding punctuation) from the Czech National Corpus (see [18]). These adjectives[2] in the Lexicon are classified into one of three categories (positive, negative or both) depending on the subjective sentiment associated with them. For this analysis, one particular adjective, *bohatý* 'rich', has been excluded, as this could otherwise automatically drive differences between the income groups (since the antonym *chudý* 'poor' is not included in the lexicon). The income groups are taken from the World Bank classification of economies of June 2019 ([4], [19]). The countries are categorized into four groups: High, Upper-middle, Lower-middle or Low Income countries. Further, only those nationalities that have been registered as staying for more than 90 days in the Czech Republic since 1994 [20], and the Czechs themselves, were included in the dataset for this study. Such an official list – of nationalities that have had a long-term presence in the country – points to the nationalities that may be present in any news text corpus from the country in focus.

The subcorpus used contains a majority of nationwide daily newspapers, regional editions from Bohemia and Moravia, and several news- and lifestyle-related magazines. The total number of titles is almost 200 from the very end of 1989 to the end of 2018, although there is reasonable coverage mainly from 1991 onwards. For more details and specific titles, see [18]. Because the SYN-series corpora are updated on a yearly basis, this method is repeatable.

[2] The material for this study is accessible at: https://www.su.se/english/profiles/irel5167-1.364672.

For this article, the material is a subset of 4,408,853 data points (observations) extracted from the subcorpus described above. It includes all the countries and nationalities that co-occur with the adjectives also described above. Certain variables were chosen to accompany the linguistic variables Noun and Adjective. Table 1 shows examples of variables and values from the Low Income group, containing co-occurrences of *vážný* + *Severní Korea* ('serious/significant + North Korea'), *důležitý* + *Afghánistán* ('important + Afghanistan'), *milý* + *Ugaňan* ('nice + Ugandan man') and *zabitý* + *Syřanka* ('killed + Syrian woman').

sent	**neg-adj**	**noun**	**adj**	**fq**	**adjfq**	**title**	**pub-year**	**form**	**wb-income**
POS/ NEG	A	severní korea	vážný	1	0.25	Hospodářské noviny	2006	COU	WB-LI
POS/ NEG	A	afghánistán	důležitý	1	1	Týden	2012	COU	WB-LI
POS	A	uganďan	milý	1	1	Deníky	2007	MASK	WB-LI
NEG	A	syřanka	zabitý	1	1	Právo	2016	FEM	WB-LI

Tab. 1. Example of variables and values from the dataset for the Low Income group

The explanatory variables [23, p. 7] include the following:

- **sent** for whether they have positive or negative sentiment – or both,
- **negadj** for whether the adjectives in the corpus source text are negated (N) or not (A),
- **fq** for the number of occurrences in one and the same issue of a newspaper or magazine,
- **adjfq** for the distance-adjusted frequency of each adjective-noun co-occurrence, see details under Method,
- **title** for the name of the newspaper or magazine where the co-occurrence is found,
- **pubyear** for the year of publication,
- **form** for country noun (COU) or nationality in masculine (MASK) or feminine (FEM) form,
- **wbincome** for World Bank income classification: High (HI), Upper-middle (UMI), Lower-middle (LMI) or Low (LI) Income.

The adjusted frequency is explained in more detail under Method below. The reason for adding a variable for the adjectival negation (A for non-negated and N for negated), is that one type of negation in the Czech language is expressed by a prefix: *moderní* 'modern, trendy' is negated as *nemoderní* 'old-fashioned, outdated'. The meaning changes [21], and the sentiment is interpreted here as the opposite. In the corpus, the lemma is the same for both word forms, which is why a distinction had to be made.

3 METHOD

3.1 Calculating Sentiment Values by category

Before creating Figure 1 below, a variable called Sentiment Value, "sent-value", was added, created by

- representing co-occurrences with a positive sentiment by the number 1, a negative sentiment by −1, and those classified as both by 0,
- multiplying this by the adjusted frequency, and then
- calculating the mean of the resulting number per year per group.

The adjusted frequency is calculated as follows:

$$\text{adjFq} = \text{SUM}\ (\text{fq} \times 1/\text{distance})$$

This is a sum of the proximity (following definition 2.11 in [22, p. 361]) of the adjective and the noun, weighted by the inverse value of their distance. Hence, co-occurrences that are close get a higher number, and those farther apart get a number that is low enough to represent their assumed (non-)prominence in the original news text. This is done to measure the potential influence of the adjective on the node noun, based on findings about proximity in Czech [22] (in particular chapters 2 and 5). In Table 1 above, the co-occurrence of *vážný* 'serious/significant' with *Severní Korea* 'North Korea' has an adjusted frequency of 0.25, which means that the noun (or bigram) and adjective are 4 words apart, and are thus given lower relevance when calculating the assumed sentiment for these nouns. The co-occurrences of *důležitý* 'important' with *Afghánistán* 'Afghanistan' and *milý* 'nice' with *Ugandan* 'Ugandan man' have an adjusted frequency of 1, which means they are adjacent to each other and thus assumed modifiers. In short: the lower the adjusted sentiment frequency, the lower the impact on the results.

Low values are included since they may nevertheless form part of the general discourse concerning the sentiment expressed in texts about the specific nouns, but they are given less weight. The adjusted frequency is utilized, rather than searching for pre- or post-modifying adjectives, since there may well be more than one word between the adjective and its modified noun in Czech, and adjectives in the near vicinity may also be highly relevant for this analysis. The Czech SYN-series corpora are not (yet) syntactically annotated, which means that this is the best approximation available.

In order to obtain a normalized number, the mean is finally calculated per year. This is needed since the number of data points per year differs (see "Composition of the corpus SYN version 8" graph in [18]). The smaller the sample, the more clearly visible the changes in sentiment value. A few outliers [23, p. 9], sometimes even

a single one, may change the total in a more visible way when the number of data points is low. For comparatively small samples, such as the Low Income group, which only makes up 2.5 percent of the whole dataset for this study, singular events may thus cause extreme peaks and troughs. Smoothed lines are, therefore, chosen for the graphs to make it easier for the human eye to interpret the differences over a long time span.

4 ANALYSES

World bank category	*Number of data points*
High Income	3,001,847
Upper-middle Income	990,576
Lower-middle Income	306,052
Low Income	110,378

Tab. 2. Number of data points in the dataset that are classified into each income group

Countries and people from countries that are classified as having high income represent such a large proportion of the dataset, 68 percent, that they form an ingroup by themselves, which is seen in Table 2. It is also apparent that the number of data points follows other income levels in the set.

4.1 Adjectives over the whole period

Aggregated over the whole period, there are certain patterns in the type of adjectives that are attributed to different income groups. This is shown in Table 3, where the most frequent adjectives are shown per income group.

High Income	*%*	*Upper-middle Income*	*%*	*Lower-middle Income*	*%*	*Low Income*	*%*
velký (‘large, big’)	35.82	velký	34.76	velký	36.57	velký	29.08
dobrý (‘good’)	17.85	poslední	16.86	poslední	17.08	poslední	16.58
poslední (‘last, final’)	15.85	**dobrý**	15.23	**dobrý**	12.88	teroristický	9.26
silný (‘strong, forceful’)	5.59	silný	6.01	**mírový** (‘peaceful’)	6.13	**dobrý**	8.72
rád (‘happy, delighted’)	5.00	**možný**	5.37	silný	5.28	**mírový**	7.76
základní (‘fundamental, basic’)	4.30	důležitý	4.96	**možný**	5.01	válečný (‘war, martial’)	7.75

High Income	*%*	*Upper-middle Income*	*%*	*Lower-middle Income*	*%*	*Low Income*	*%*
možný (‘possible, feasible’)	4.20	špatný	4.39	důležitý	4.50	bezpečnostní (‘safety, security’ [adj.])	5.90
špatný (‘bad, wrong’)	3.88	**lidský** (‘human, humane’)	4.27	teroristický	4.26	**možný**	5.80
důležitý (‘important’)	3.85	**rád**	4.19	špatný	4.24	špatný	4.60
nízký (‘low, reduced’)	3.66	těžký (‘hard, severe’)	3.91	**rád**	4.05	**lidský**	4.56

Tab. 3. The ten most frequent adjectives 1990–2018 per income group. Percentage of these ten. **Bold** = positively classified, underlined = negatively classified, normal style = classified as both. Translations into English are shown the first time the adjective occurs in the table

A few words are needed about the most frequent adjectives to better understand the content of the corpus in general. The adjective that occurs most frequently with the highest income groups is *velký* ‘large’. This multifunctional adjective [24, p. 122] co-occurs with the highest income group 395,141 times in the dataset, which is slightly more than one third of the total of these ten, and 13 percent of the total number of adjectives for this group.[3] In the opposite corner of Table 3, the tenth most frequent adjective for the lowest income group, *lidský* ‘human’ or ‘humane’, co-occurs with that group 1,640 times, which is only about 4.5 percent of these ten, and 1.49 percent of the total for that income group. There is thus a difference in the number of adjectives between the different groups, which explains why the income groups must be studied separately.

Some highly frequent adjectives are present in all groups, but the less-frequent are of more interest here: *silný* ‘strong, forceful’ appears less frequently in the Low-middle Income group, and not at all in the Low Income group; *rád* ‘happy, delighted’ shows a similar trend; *mírový* ‘peaceful’ and *teroristický* show an opposite trend, only being present in the two lower income groups, and increasingly so. *Bezpečnostní* ‘safe, secure’ is only present in the lowest income group, which will be discussed under Conclusions.

We also notice that the three highest income groups all have two adjectives that can be both positive and negative, and that they also have four of each sentiment classification. The lowest income group actually has four negative and five wholly positive adjectives in the top ten, but the fact that the negatively classified *poslední*

[3] *Velký* and *poslední* ‘last, final’ (adjectives that are in the top 3 for all groups) are indeed also among the most frequent adjectives in Czech news language [25, p. 11 and 13].

'last, final' and *teroristický* have such a large share of the total makes them prominent in the discourse and creates a downward trend. What Table 3 shows is thus a prominence of words relating to war and peace for the lowest income group. This gives us a hint about the discourse surrounding the so-called war on terror that emerged at the beginning of the present millennium. It seems that there are some adjectives here that clearly have an "inscribed" attitudinal value, i.e., a stable negative value over time and context [26, p. 2], and that may create a clearly negative picture of the negative persona some nationalities are given in the Czech press during these years. Now let us scrutinize the trends for the four income groups.

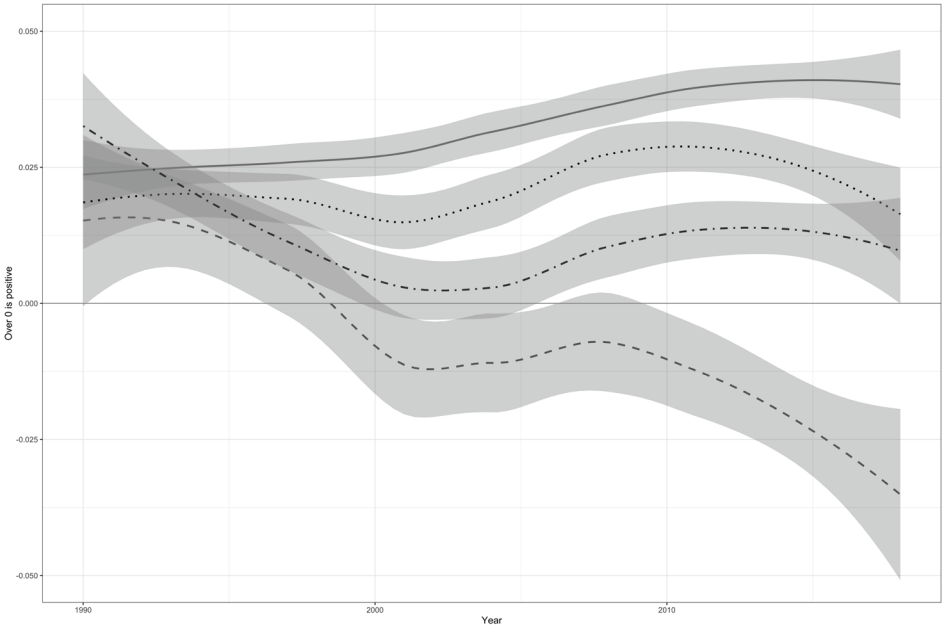

Fig. 1. Sentiments trends for the four income groups: High (solid —), Upper Middle (dotted …), Lower Middle (dot-dashed .-.-.) and Low (dashed - -) Income. The smooth, grey line equals the standard error with 95 % confidence interval. 0, which marks a balance between positive and negative adjectives, is shown on the horizontal line

The lower the income, the more negative the overall sentiment in Figure 1. As the corpus grows larger and more reliable in reflecting a general trend (from around 2004), this result becomes clearer, except possibly for the Upper-middle Income group, which closes in on the Lower-middle Income group towards the end of the period. The economic crisis in 2008 does not seem to have affected the three top income groups, whereas the trend for the lowest income group slumps from that year onwards. For this reason, the group with the lowest income is examined in closer detail.

4.2 Removal of an influencer

The Low Income group shows a clear trough around the year 2001. The time line recovers somewhat in the following years, but then deepens again around the year 2008. Let us, therefore, examine the noun + adjective co-occurrences that cause the dips during the period 2000–2009. This is shown by the adjusted senti-values in Table 4, visualized per year for the top and bottom ten of this group. It then becomes clear that one country and nationality alone creates the main reason for both the peaks and troughs in the Low Income group: Afghanistan. The nouns *Afghánistán 'Afghanistan'*, *Afghánec* 'Afghan man' and *Afghánka* 'Afghan woman' make up 34,115 of the 110,378 observations, i.e., 31 percent, for the entire period 1990–2018. It turns out that by far the biggest reason for the 2001 trough is the co-occurrence of the noun *Afghánistán* and the adjective *teroristický* 'terrorist [adj.]'. During 2001 and 2002, these two words in co-occurrence create an enormous trough, as seen by the sentiment numbers in Table 4. The co-occurrence returns in 2003, 2004, 2006 and 2007, although not as strongly.[4]

Most positive, Low income				*Most negative, Low Income*			
Year	*Noun*	*Adj*	*Total Adj Sent-v*	*Total Adj Sent-v*	*Noun*	*Adj*	*Year*
2001	afghánistán	mírový	34.4	-76.8	afghánistán	teroristický	2001
2001	afghánistán	možný	27.9	-41.0	afghánistán	poslední	2001
2002	afghánistán	mírový	24.7	-28.6	afghánistán	zvláštní	2001
2008	kongo	hluboký	20.1	-24.1	afghánistán	válečný	2001
2001	afghánistán	bezpečnostní	19.4	-19.8	afghánec	obyčejný	2001
2008	afghánistán	bezpečnostní	15.7	-13.8	afghánistán	tajný	2001
2002	afghánistán	dobrý	14.2	-12.8	afghánistán	těžký	2001
2002	afghánistán	bezpečnostní	14.1	-44.1	afghánistán	teroristický	2002
2003	afghánistán	mírový	13.9	-19.9	afghánistán	poslední	2002
2001	afghánistán	významný	12.4	-13.9	afghánistán	teroristický	2003

Tab. 4. The most positive and the most negative co-occurrences for the Low Income group 2001–2009

The negatively classified *adjectives are* countered by the positively classified *mírový* 'peace-, peaceful', and *bezpečnostní* 'secure, safe(ty-)', but the positive adjectives in Table 4 do not reach the same numbers. When added, as done above in Figure 1, the positivity of 34.4 (for *mírový*) plus 19.4 (for *bezpečnostní*) for the year

[4] This still holds even after removing the multifunctional adjectives [24, p. 122] for the negatively classified *poslední* 'last' and *obyčejný* 'ordinary, average', and for the positively classified *dobrý* 'good' and *možný* 'possible, believable'.

2001, and 24.7 (*mírový*) and 14.1 (*bezpečnostní*) for the year 2002 nevertheless does not make up for the negativity caused by *teroristický* in the same years.

This result might give reason to analyse the ingroup – Czechia and its people in a Czech context – as well. The outcome of that shows that in Figure 1, *Česko* 'Czechia', *Čech* 'Czech man' and *Češka* 'Czech woman' make up 547,332, i.e., 18 percent, of the total of 3,001,847 observations in the High Income group. The "obvious" ingroup of Czechs and their country is thus not as influential in their group as the outgroup noun of Afghanistan, and when removed, the High Income trend line nevertheless has a similar form.[5]

5 CONCLUSIONS

World Bank statistics rarely figure in linguistic articles, but their classification of income groups around the world has been used here alongside linguistic data to compare sentiment towards the countries and nationalities of these groups. Prominent parts of the overall rhetoric [1, p. 18] were indeed revealed when the data were scrutinized as a whole: the obvious ingroup of Czechs somewhat drove their income group's positive sentiment results, and an outgroup from a country far away from the ingroup drove their income group's negative results even more. Figure 1 thus shows that the higher the national income, the more positive the overall representation of the nationalities in these data. Even when the two most influential (most positive and most negative) noun groups are removed and the amount of source data swells around the year 2000, the sentiment trend lines remain separated.

This article argues further that when looking beyond the binary classification of positive and negative sentiment, a certain discourse emerges. As income decreases from High through the two Middle groups to the Low Income group, the adjectives *dobrý* 'good' and *silný* 'strong', 'forceful' decrease; *mírový* 'peaceful' and *teroristický* increase, and *bezpečnostní* 'safety', 'security' [adj.] make an appearance. Despite several of these being classified as positive in the Subjectivity Lexicon, they together reflect a prominent discourse about the lowest income group being involved in war and terrorism. This is a continuation of the "new security" discourse [5, p. 66] combined with the discourse surrounding the "war on terror" [1, p. 32] of the 2000s. This might even be a case of "canonical pairings" [12, p. 155]. The connection between peacefulness, safety and terrorism could form the basis of a future study on stereotypical conventions (such as "the poor are more likely to face negative than supportive behavior", [27, p. 243]). Such study could also compare the findings of Cvrček and Fidler [28, p. 17] where the noun "migrant" is closely associated with certain regional identities and the adjective "violent".

[5] This figure can be found at: https://www.su.se/english/profiles/irel5167-1.364672 together with other data from this study.

That an obvious ingroup, in this case Czechs in Czech newspapers, forms a large part of the peak (more positive sentiment) in such a combined volume of text could be expected. Without them, the High Income trend nevertheless remains clearly positive. That a single co-occurrence (of the noun Afghanistan and the adjective terrorist) creates such a large part of the trough in the trend line for the Low Income group constitutes a reason to examine the details, but also to plan future studies to explore how certain low-income nationalities have been treated in Czech media, or perhaps even Slavic media in general. It seems to be a case of resonance (as in [29]), i.e., repeating of linguistic formulae whether or not the next author agrees with the connotation given to the formula(e) by the previous author. Another next step could be to dig deeper into the co-occurrences of some of the nouns or adjectives, without the restriction of the Subjectivity Lexicon. This resonance could be explored even further with current-day corpora of both written and spoken text.

ACKNOWLEDGEMENTS

The author would in particular like to thank Václav Cvrček at the department of the Czech National Corpus for his continuous assistance, including the extraction of the dataset. Grateful thanks also go to the anonymous reviewers of this article, who kindly made clear what was clouded.

References

[1] Fairclough, N. (2006). Language and globalization. Routledge: Abingdon.

[2] Fowler, R. (1991). Language in the news: discourse and ideology in the Press. Routledge: London.

[3] Bednarek, M. (2019). The Language and News Values of 'Most Highly Shared' News. In F. Martin and T. Dwyer (eds.), Sharing News Online – Commendary Cultures and Social Media News Ecologies, pages 157–188, Palgrave Macmillan: Cham.

[4] World Bank Group. (2020). World Bank Country and Lending Groups. Datahelpdesk. Worldbank.Org.

[5] Wodak, R., de Cillia, R., Reisigl, M., and Liebhart, K. (2009). The Discursive Construction of National Identity. 2nd ed. Edinburgh University Press: Edinburgh.

[6] Lindqvist, A., Renström, E. A., and Gustafsson Sendén, M. (2019). Reducing a Male Bias in Language? Establishing the Efficiency of Three Different Gender-Fair Language Strategies. Sex Roles, 81(1), pages 109–117.

[7] Rönnbäck, K. (2014). The Idle and the Industrious – European Ideas about the African Work Ethic in Precolonial West Africa. History in Africa, 41, pages 117–145.

[8] Molek-Kozakowska, K., and Chovanec, J. (2017). Media representations of the "other" Europeans. Common themes and points of divergence. In J. Chovanec and K. Molek-Kozakowska (eds.), Representing the Other in European Media Discourses, pages 1–22, John Benjamins Publishing Company: Amsterdam/Philadelphia.

[9] Veselovská, K. (2013). Czech subjectivity lexicon: A lexical resource for Czech polarity classification. In K. Gajdošová A. and Žáková (eds.), Natural Language Processing, Corpus Linguistics, E-Learning, pages 279–284, RAM-Verlag: Bratislava/Lüdenscheid.

[10] Veselovská, K. (2017). Sentiment analysis in Czech. Ústav formální a aplikované lingvistiky (Institute of Formal and Applied Linguistics): Prague.
[11] Šindlerová, J., Veselovská, K., and Hajič, J. (2014). Tracing Sentiments: Syntactic and Semantic Features in a Subjectivity Lexicon. Proceedings of the XVI EURALEX International Congress: The User in Focus, pages 405–414.
[12] Paradis, C., Löhndorf, S., van de Weijer, J., and Willners, C. (2015). Semantic profiles of antonymic adjectives in discourse. Linguistics, 53(1), pages 153–191.
[13] UNICEF. (2020). Averting a lost COVID generation: A six-point plan to respond, recover and reimagine a post-pandemic world for every child. UNICEF Division of Communication: New York.
[14] Roberton, T., Carter, E. D., Chou, V. B., Stegmuller, A. R., Jackson, B. D., Tam, Y. et al. (2020). Early estimates of the indirect effects of the COVID-19 pandemic on maternal and child mortality in low-income and middle-income countries: a modelling study. The Lancet Global Health, 8(7), pages e901–e908.
[15] Tan-Torres Edejer, T., Hanssen, O., Mirelman, A., Verboom, P., Lolong, G., Watson, O. J. et al. (2020). Projected health-care resource needs for an effective response to COVID-19 in 73 low-income and middle-income countries: a modelling study. The Lancet Global Health, 8(11), pages e1372–e1379.
[16] Ziai, A. (2016). Development discourse and global history: from colonialism to the sustainable development goals. Routledge: London.
[17] Moretti, F., and Pestre, D. (2015). Bankspeak. New Left Review, (92), pages 75–99.
[18] Křen, M. (2019). Corpus SYN version 8 – Czech National Corpus Wiki.
[19] World Bank's Data Help Desk. (2020). How does the World Bank classify countries? World Bank Group.
[20] Czech statistical office. (2019). Cizinci v ČR podle státního občanství v letech 1994–2018 (Foreigners in the CR by citizenship in the years 1994–2018 (as of 31 December)).
[21] Kováříková, D., Chlumská, L., and Cvrček, V. (2012). What Belongs in a Dictionary? The Example of Negation in Czech. In R. Vatvedt Fjeld and J. M. Torjusen (eds.), Proceedings of the 15th EURALEX International Congress, pages 822–827, Department of Linguistics and Scandinavian Studies: Oslo.
[22] Cvrček, V. (2014). Kvantitativní analýza kontextu. Nakladatelství Lidové Noviny/Ústav českého národní korpusu: Praha.
[23] Brezina, V. (2018). Statistics in Corpus Linguistics: A Practical Guide. Cambridge University Press: Cambridge.
[24] Cvrček, V., Laubeová, Z., Lukeš, D., Poukarová, P., Řehořková, A., and Zasina, A. J. (2020). Registry v češtině (Registers in Czech). Nakladatelství Lidové Noviny/Ústav českého národní korpusu: Praha.
[25] Čermák, F., and Křen, M. (2011). A Frequency Dictionary of Czech. Routledge: Oxon.
[26] Don, A. (2016). "It is hard to mesh all this": Invoking attitude, persona and argument organisation. Functional Linguistics, (3), pages Article number 9.
[27] Lindqvist, A., Björklund, F., and Bäckström, M. (2017). The perception of the poor: Capturing stereotype content with different measures. Nordic Psychology, 69(4), pages 231–247.
[28] Cvrček, V., and Fidler, M. (2021). No Keyword is an Island: In search of covert associations (Preprint), pages 1–28.
[29] Du Bois, J. W. (2014). Towards a dialogic syntax. Cognitive Linguistics, 25(3), pages 359–410.