

Tracking Attitudes Towards Immigration in Swedish Media

Jacobo Rouces, Lars Borin, Nina Tahmasebi

Språkbanken Text, University of Gothenburg, Sweden
{jacobou.rouces, lars.borin, nina.tahmasebi}@gu.se

Abstract. We use a gold standard under construction for sentiment analysis in Swedish to explore how attitudes towards immigration change across time and media. We track the evolution of attitude starting from the year 2000 for three different Swedish media: the national newspapers *Aftonbladet* and *Svenska Dagbladet*, representing different halves of the left–right political spectrum, and the online forum *Flashback*.

1 Introduction

The topic of immigration has been at political center stage for the last two decades, across Europe and many Western countries in general and certainly in Sweden. During the last decade, political parties whose defining issue is reducing or opposing immigration have evolved from extra-parliamentary movements to having significant representation in their national parliaments (Jungar and Jupskås, 2014; Rydgren and van der Meiden, 2018), often becoming the third or second largest parties in percentage of votes, and in some cases like Norway, they have entered coalition governments. There has been debate about to which extent this has influenced the position of parties across the rest of the political spectrum, as well as the general public discourse on the topic of immigration (Harmel and Svåsand, 1997; Hovden et al, 2018).

In this paper, we analyze the sentiment towards immigration in three Swedish media across different times from the year 2000. The media analyzed are the following:

- *Aftonbladet* (<http://www.aftonbladet.se/>), a daily newspaper with a daily circulation of 154,900 (2014). It describes itself as “independent social-democrat”.¹
- *Svenska Dagbladet* (SvD) (<http://www.svd.se/>), a daily newspaper with a daily circulation of 143,400 (2013). It describes itself as “independent conservative”.²
- The *Flashback Forum* (<https://www.flashback.org/>), a Swedish internet forum, with an Alexa ranking of 9,978, the 42nd in Sweden (2018).³

The paper is organized as follows. In Section 2 we describe the data annotated from the three different sources, in Section 3 we present and discuss the results of our analysis, and in Section 4 we conclude and look ahead.

¹ <https://www.svd.se/lo-saljer-aktier-i-aftonbladet>

² <https://www.svd.se/susanna-popova-tillbaka-pa-svd-ledare/i/senaste/om/ledare>

³ <https://www.alexa.com/siteinfo/flashback.org>

2 Annotated Data

We have obtained annotations of attitudes towards immigration as part of our ongoing effort to build a gold standard for sentiment analysis in Swedish (Rouces et al, 2018). We have employed 9 annotators to annotate 3,870 documents that we retrieved from Aftonbladet, SvD and Flashback, containing in total 1,574,226 words in 108,132 sentences. All annotators had a background including at least undergraduate training in (Swedish or general) linguistics. They were provided with an annotation manual and some example documents annotated by one of the authors. The documents to be annotated were selected randomly from among a larger set of documents retrieved as containing one among a list of 60 terms related to immigration, shown in Table 1. In the case of the two newspapers, the text types were limited to editorials and opinion pieces.

anhöriginvandring, antirasism, antirasist, antirasistisk, arbetskraftsinvandring, arbetstillstånd, assimilation, assimilering, asylpolitik, asylrätt, asylskäl, asylsökande, asylsökningsfråga, dublinförordning, etnicitet, etnisk, familjeanknytning, familjeåterförening, flykting, flyktingbarn, flyktingfamilj, flyktingläger, flyktingmottagande, flyktingmottagning, flyktingpolitik, flyktingström, flyktinsmuggla, flyktinsmuggling, flyktingrörelse, flyktingvåg, folkgrupp, främlingsfientlig, främlingsfientlighet, hederskultur, identitetshandling, immigration, immigrera, integration, integrationsprocess, invandrare, invandring, invandrings, invandringen, invandringens, invandra, invandrat, invandringskritisk, invandringskvot, invandringspolitik, invandringsvåg, islamisering, islamofobi, kvotflykting, massflykt, massinvandring, migrationsverket, multikultur, nyanlända, rasism, rasist

Table 1: Immigration-related terms used to retrieve documents from sources

As part of the annotation task, annotators were instructed to annotate each paragraph in each document with an attitudinal value towards the topic of immigration. The sentiment value could be one among 5 possible options: (‘very negative’, ‘negative’, ‘neutral’, ‘positive’, ‘very positive’). The following explanation was used in the annotation manual to define the meaning of positive and negative in the context of this topic: “We define a positive sentiment towards the topic of immigration as being in favour of more ‘open borders’ policies, negative sentiment as being in favour of ‘closed borders’ policies, and neutral sentiment as being in favour of keeping the immigration-related policies at the moment of the writing. ‘Very positive’ and ‘Very negative’ describe a stronger sentiment in this regard.”

For this annotation task, the annotators were further instructed to treat titles as paragraphs. The annotators were also instructed to provide a attitudinal value to documents as a whole.

Table 2 shows the number of documents and paragraphs annotated for each source.

source	documents	paragraphs	sentences	words
Aftonbladet	1,184	18,979	42,408	617,442
SvD	852	19,360	44,884	677,385
Flashback	1,834	12,021	20,840	279,399
Total	3,870	50,360	108,132	1,574,226

Table 2: Number of documents and paragraphs in the annotated corpus

3 Data Analysis

A total of 40 documents were annotated by all annotators, so inter-annotator agreements could be calculated. Krippendorff's alpha using an interval metric was 0.34 for document-level annotations and 0.44 for paragraph-level annotations.

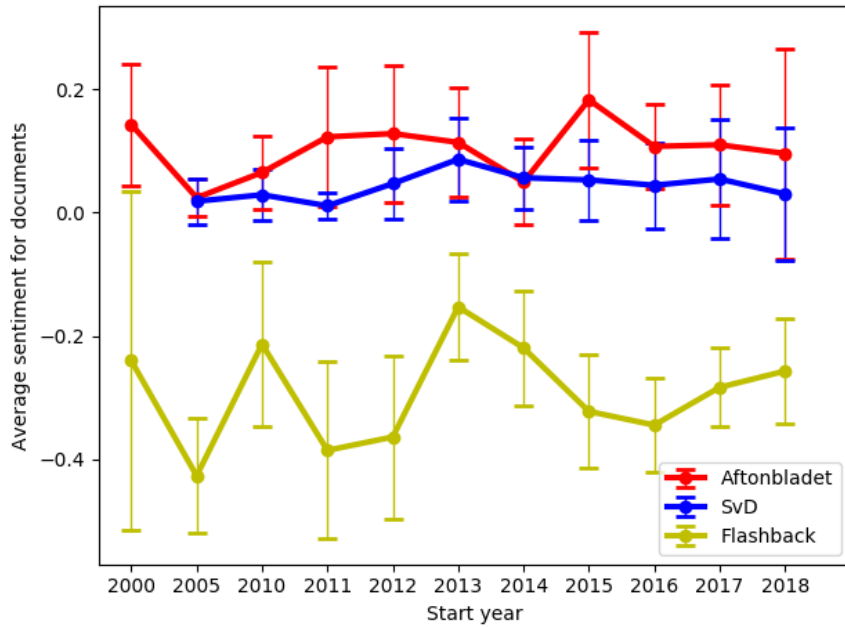
Figure 1a shows the average sentiment obtained for each source and each time period. This average is obtained using the following quantification of the sentiment values: ('very negative': -2, 'negative': -1, 'neutral': 0, 'positive': +1, 'very positive': +2). For the period 2000–2010 the number of data points was lower, so it was divided into two 5-year periods to achieve sufficient statistical significance. (i.e. otherwise, the confidence intervals in Figures 1a and 2a before 2010 would have been so wide that they would have overlapped with each other horizontally and vertically).

Figure 1b shows the distributions for the same sources and periods. Darker shades correspond to more frequent values. The neutral (0) category is omitted because for any time period most of the elements are annotated as neutral (it includes cases where the topic is not mentioned) and its inclusion would require a range for the shades that would not allow us to see the differences among other categories. However, for each time period the neutral value is equal to the difference between 1 and the sum of the other 4 categories.

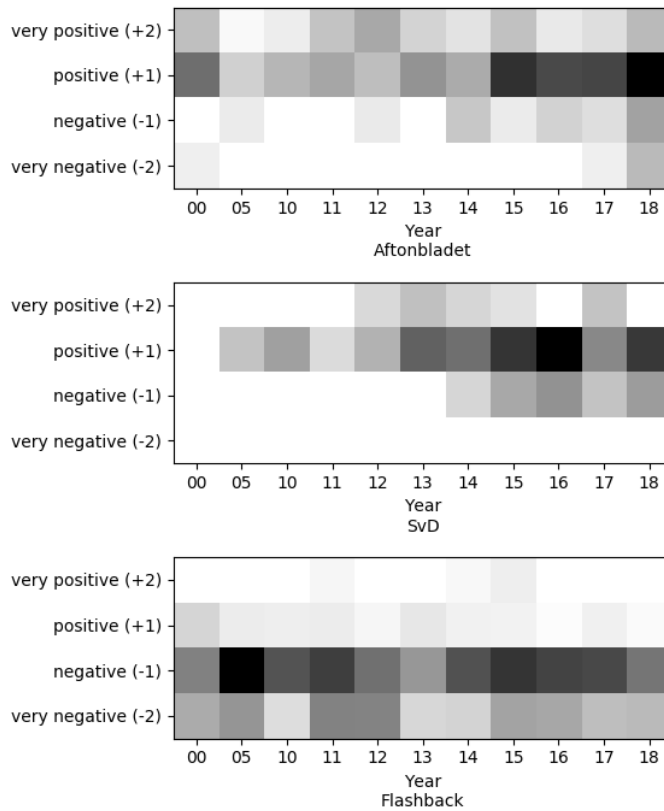
Figures 2a and 2b show the same statistics calculated over the sentiment associated to paragraphs independently (titles are included as paragraphs). The results are similar, which is an expected result since the emotion towards a topic expressed in a document should be positively correlated with the average emotion in its paragraphs. However, because most documents contain more than one paragraph, the paragraph-level sentiment values are statistically more significant (the confidence intervals are narrower), and we consider them our primary source for extracting conclusions.

It is clear from the data that the sources can be ordered by their general sentiment towards immigration – from more positive to less positive – as Aftonbladet > SvD > Flashback, with Aftonbladet and SvD being relatively close and Flashback taking a more distinctive negative attitude.

However, Aftonbladet and SvD held indistinguishable positions in the period 2005–2010 (There is no data for SvD for the first period 2000–2005). After that period Aftonbladet took a slightly more positive stance. In the year 2013, there is a shift in Flashback towards less negative stance. However, in the year 2014, a shift towards a more negative stance is

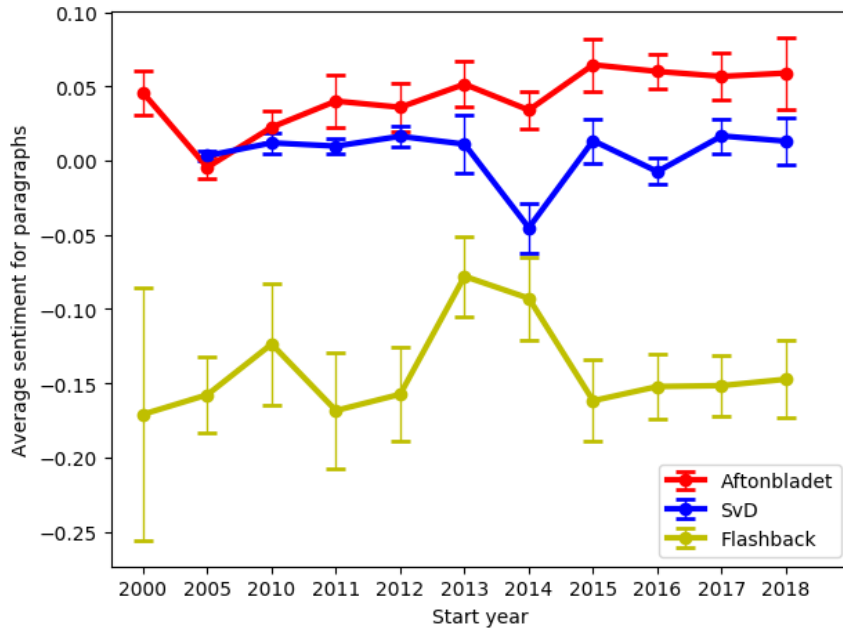


(a) Average document sentiment towards immigration for our three sources. 95% confidence intervals are shown for each time period.

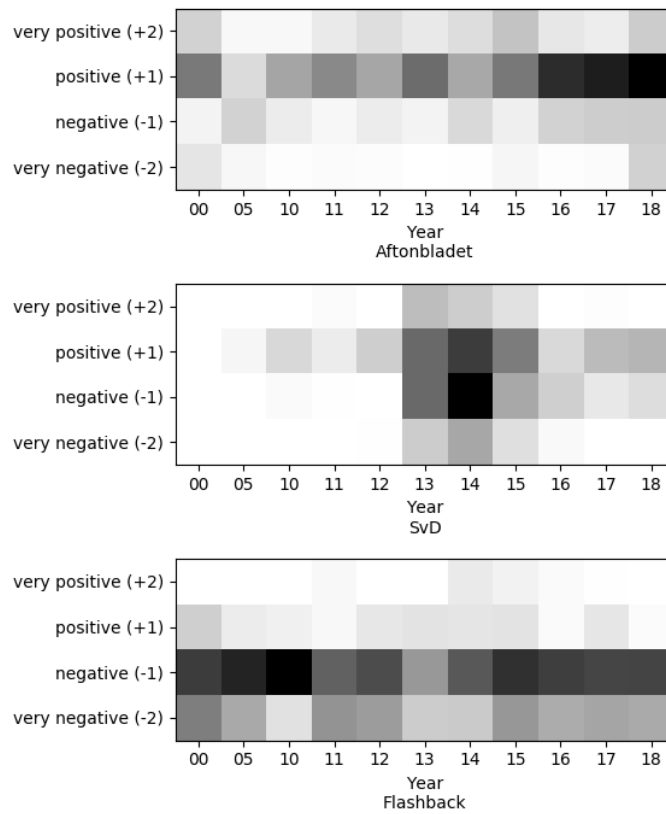


(b) Distribution of document sentiment towards immigration for our three sources.

Fig. 1: Document sentiment towards immigration.



(a) Average paragraph sentiment towards immigration for our three sources. 95% confidence intervals are shown for each time period.



(b) Distribution of paragraph sentiment towards immigration for our three sources.

Fig. 2: Paragraph sentiment towards immigration.

noticeable in all three sources. By the following year, the stance of both newspapers becomes more positive until it reaches the previous levels, while the negative shift becomes more pronounced in Flashback and stabilizes at 2011–2012 levels.

4 Concluding and Looking Ahead

In this paper, we have described some early findings from a preliminary quantitative analysis of a corpus consisting of newspaper and online forum texts about immigration and immigration policies in Sweden, manually annotated for sentiment on multiple levels. In this small study, we have looked only at document- and paragraph-level sentiment annotations, and not at all at the fine-grained phrase- and sentence-level aspect-based sentiment annotation that we also asked our annotators to add to the texts, and which make up the bulk of the annotations.

We are not political scientists, contemporary historians or sociologists, but it seems to us that the results of our study indicate that this kind of sentiment annotation could provide a valuable addition to the methodological toolbox of those disciplines, allowing large-scale analysis of opinion trends over time and opinion differences among media.

Our ultimate purpose in creating this corpus is to provide training and evaluation data – a gold standard – to be used in the development of automatic aspect-based sentiment analysis systems for Swedish text (Schouten and Frasincar, 2016). Since our preliminary results described above are in broad agreement with what we would have expected both for the time period and for the three different text sources, they bode well for the potential usefulness of this gold standard for Swedish sentiment analysis.

Acknowledgements

This work has been supported by a framework grant (*Towards a knowledge-based cultur-omics*; contract 2012-5738) as well as funding to Swedish CLARIN (*Swe-Clarín*; contract 2013-2003), both awarded by the Swedish Research Council, and by infrastructure funding granted to Språkbanken by the Swedish Research Council and the University of Gothenburg.

References

- Harmel R, Svåsand L (1997) The influence of new parties on old parties' platforms: The cases of the progress parties and conservative parties of Denmark and Norway. *Party Politics* 3(3):315–340.
- Hovden JF, Mjælde H, Gripsrud J (2018) The Syrian refugee crisis in Scandinavian newspapers. *Communications* 43(3):325–356.
- Jungar AC, Jupskås AR (2014) Populist radical right parties in the Nordic region: A new and distinct party family? *Scandinavian Political Studies* 37(3):215–238.

- Rouces J, Tahmasebi N, Borin L, Eide SR (2018) SenSALDO: Creating a sentiment lexicon for Swedish. In: Proceedings of LREC 2018, ELRA, Miyazaki, pp 4192–4198.
- Rydgren J, van der Meiden S (2018) The radical right and the end of Swedish exceptionalism. *European Political Science* .
- Schouten K, Frasincar F (2016) Survey on aspect-level sentiment analysis. *IEEE Transactions on Knowledge & Data Engineering* 28(3):813–830.