

Vehemence and Victims: Using generic lexicons to investigate emotions in historical Dutch parliamentary debates on war victim legislation (1945-1990)

Milan van Lange & Ralf Futselaar

NIOD Institute for War, Holocaust, and Genocide Studies (Amsterdam, NL)

Twitter: @milanvanlange, @Kinkirarufu

This paper analyses digitized historical parliamentary debates on post-war welfare legislation for victims of the German occupation of the Netherlands. We confront a well-researched historical case with an additional, computational approach. This case also allows us to explore and evaluate the use of generic emotion lexicons in an innovative approach to historical research.¹

1. The historical case

In Dutch historiography a rough periodization exists for the ways in which Dutch politics and society has dealt with the lasting consequences of the German occupation. “Victims”, broadly understood, form the backbone of this periodization. Firstly, there is a common assumption that the immediate post-war years were a period of intense attention for (personal) war experiences and their consequences. The following period, roughly covering the 1950s, is often characterised as a period of silence; a forward-looking phase in which there was little attention for victims, let alone space for expressing war-related emotions. Although less pronounced, this silence continued into the 1960s. The “silent period” is often contrasted with the phase beginning in the early 1970s. This phase has been described as a period of ‘emancipation of emotions’, together with a renewed interest in personal experiences, individual suffering, and emotional consequences.² Some authors have generalized such shifts into an overwhelming ‘emotionalization of [Dutch] society’ in this period.³ In this paper, parliamentary debates are analysed to investigate if this periodization can be discerned (in language use) in political debates about the rights of and welfare provisions for victims of the German occupation in the Netherlands.

This paper has a twofold goal. First, we want to gain more insights in the role emotions played in discussing and shaping war-related welfare legislation. Secondly, we apply and evaluate emotion lexicons in a diachronic historical investigation. By confronting close reading with computational text analysis, the following historical questions will be addressed: how emotional was the parliamentary discussion on war victims, compared to other contemporary debates? How did this change over time? And how can this be explained? In answering these questions, attention will be paid to a stereotypical ‘silence of the 1950s’ and the ‘emancipation of emotions’ from the 1970s onwards. Studying emotions from an historical perspective raises not only substantive but also epistemological questions: how can we observe, interpret, and compare emotionality in the context of historical research?

2. Emotions in history

Investigating emotions in ethically charged discussions from the relative recent past gives rise to some fundamental issues. Closeness in time and the often gruesome suffering are bound to elicit empathies and antipathies among researchers. The fact that the debates in question have been highly politicised, and are to an extent ongoing, mean that few if any historians are neutral with regard to this topic. Whether this has hindered objectivity, is the first question we want to ask. Secondly, the emotions of others are difficult to investigate, and all the more so in retrospect. They often manifest themselves in

¹ This conference paper is based on a dissertation/book chapter in progress. Findings and conclusions are preliminary.

² Jolande Withuis, *Erkenning: van oorlogstrauma naar klaagcultuur* (Amsterdam: De Bezige Bij, 2002); Jolande Withuis, ‘Opkomst en neergang van PTSS’, *Directieve therapie* 30, no. 3 (1 October 2010): 149–59, <https://doi.org/10.1007/BF03096232>; Jolande Withuis and Annet Mooij, ‘From Totalitarianism to Trauma - A Paradigm Change in the Netherlands’, in *The Politics of War Trauma: The Aftermath of World War II in Eleven European Countries*, ed. Jolande Withuis and Annet Mooij (Amsterdam University Press, 2010), 193–216; Oosterhuis, Harry, ‘Mental Health, Citizenship, and the Memory of World War II in the Netherlands (1945–85)’, *History of Psychiatry* 25, no. 1 (2014): 20–34.

³ Henri Beunders, *Publieke tranen: de drijfveren van de emotiecultuur* (Amsterdam: Contact, 2002); Henri Beunders, *Hoeveel recht heeft de emotie?: Straffen in de slachtoffercultuur* (Amsterdam University Press, 2018).

volatile forms of communication – such as tone of voice, hand gestures, or facial expressions. Such signifiers are mostly lost for the period under scrutiny here. This leaves us with the fraction of all historically expressed emotions that have been recorded and preserved: the emotions that have been put into words. We therefore investigate the recordings of emotional language in historical texts as carriers of past emotional expression.

3. Generic emotion lexicons

In behavioural sciences it has become widely-accepted that the words we use can be considered reflections of people's mental or psychological states.⁴ Based on this idea, the expression of emotions in written or transcribed historical text can form an empirical baseline for making something seemingly intangible (emotions), identifiable and observable. To do so, this paper relies on the use of a generic word list (NRC EmoLex) representing different categories of basic emotions (joy, anger, sadness, etc.). This lexicon is generic in the sense that it is not based on a single (type of) dataset or developed for a specific application.⁵ The use of such lexicons to identify, analyse, and evaluate manifestations of emotions, can be considered as a rather crude method. On the level of a single particular sentence, it may not always be reliable in determining whether a certain emotion is manifest. But, when a lexicon is applied to larger corpora, they offer an opportunity for comparison between texts of different origin or from a different period. In this paper we reflect on the value of a generic lexicon in determining whether a certain part of a corpus has more (or fewer) emotional manifestations, compared to other fragments. We consider it an advantage that the NRC EmoLex is created outside the particular historical research context. It offers an additional perspective to personal judgments and interpretations of the sources by historians. In addition, identification of emotional language is transparent, traceable, and replicable – as the emotion lexicons are open source. In this paper, word lists for anger and sadness are used, as these emotions are expected to be important markers connected to (war) victimhood.⁶

4. Computational workflow

This investigation relies mainly on the analysis of the digitised, enriched, and machine-readable version of the Dutch parliamentary proceedings (*Handelingen der Staten-Generaal*).⁷ Different topics discussed in the debates are manually annotated based on the original descriptions of the documents. To confront this text collection with the emotion lexicons, this investigation relies on RStudio and the R-programming language.⁸ The text corpus is first loaded into the R-environment. Not all language's varieties and complexities are necessary (or even desirable) for this analysis. Therefore, cleaning and unification of the texts is performed using the Quanteda R-package: all characters are reduced to lower case and interpunction and frequent but meaningless stopwords (the, a, etc.) are removed.⁹ In addition, all words are lemmatised. In this process all occurrences of variations in spelling and syntactical form

⁴ James W Pennebaker, *The Secret Life of Pronouns: What Our Words Say about Us* (New York: Bloomsbury Press, 2013); Mark Dechesne and Bryn Bandt-Law, 'Terror in Time: Extending Culturomics to Address Basic Terror Management Mechanisms', *Cognition and Emotion* 33, no. 3 (3 April 2019): 492–511, <https://doi.org/10.1080/02699931.2018.1460322>.

⁵ Saif Mohammad, 'From Once Upon a Time to Happily Ever After: Tracking Emotions in Novels and Fairy Tales', *Proceedings of the 5th ACL-HLT Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, 24 June 2011, 105–14; Saif M. Mohammad and Peter D. Turney, 'Crowdsourcing a Word-Emotion Association Lexicon', *Computational Intelligence* 29, no. 3 (2013): 436–465; Peter Boot, Hanna Zijlstra, and Rinie Geenen, 'The Dutch Translation of the Linguistic Inquiry and Word Count (LIWC) 2007 Dictionary', *Dutch Journal of Applied Linguistics* 6, no. 1 (19 October 2017): 65–76, <https://doi.org/10.1075/dujal.6.1.04boo>.

⁶ Thomas Brudholm and Johannes Lang, 'Introduction: Emotions and Mass Atrocity', in *Emotions and Mass Atrocity: Philosophical and Theoretical Explorations*, ed. Thomas Brudholm and Johannes Lang (Cambridge and New York: Cambridge University Press, 2019), 14.

⁷ M. Marx et al., 'Thematic collection: PoliticalMashup and Dutch Parliamentary Proceedings 1814-2013' (Data Archiving and Networked Services (DANS), 2012), <https://doi.org/10.17026/dans-zg8-9x2v>.

⁸ R Core Team, *R: A Language and Environment for Statistical Computing* (Vienna: R Foundation for Statistical Computing, 2019), <https://www.R-project.org/>; RStudio Team, *RStudio: Integrated Development Environment for R* (Boston, MA, 2018), <http://www.rstudio.com>.

⁹ Kenneth Benoit et al., 'Quanteda: An R Package for the Quantitative Analysis of Textual Data', *Journal of Open Source Software* 3, no. 30 (2018): 774, <https://doi.org/10.21105/joss.00774>.

of words are reduced to their linguistic basic form.¹⁰ Every document is then subsequently divided into consecutive 250-word segments (chunks). This process makes a proportional comparison between document-emotion-scores possible.¹¹

Next, a consequential process in the pre-processing of texts is performed: a so-called bag of words is created (Document Term Matrix, or DTM). Word order and sentence structure are completely discarded.¹² This reductionist representation of text is basically a table of lists with (occurrences of) words per document.¹³ In this process a Term Frequency – Inverse Document Frequency (TF-IDF) weighting is assigned.¹⁴ Each word has now its own ‘weighting’ for every occurrence. Those weights of words from the NRC EmoLex are added up, and thus form the score of an individual debate (or 250-word chunk) on a particular emotion. These scores are plotted in boxplots and diachronic graphs. Basic statistical tests are performed to inform conclusions about significance of findings. Finally, we bring together quantitative results and a more traditional approach to historical research by a close(r) reading and interpretation of the emotion words, the original historical documents, and secondary literature.¹⁵

5. Results

5.1 Historical developments in emotion scores over time

First, all 250-word chunks of every parliamentary discussion in the *Handelingen* dataset from the 1945-1990 period are scored using the above-mentioned lexicons for anger and sadness. The mean of all 250-word chunks belonging to a single debate is plotted as a red bar on the next page. Each bar represents a single discussion in parliament. Not only results for all war victim-related debates are plotted; also average scores of all other, non-relevant parliamentary debates are taken into account (in grey dots, see Graph 1 and 2).

¹⁰ M. Reynaert et al., ‘PICCL: Philosophical Integrator of Computational and Corpus Libraries’, [Http://www.Nederlab.Nl/Cms/Wp-Content/Uploads/2015/10/Reynaert_PICCL-Philosophical-Integrator-of-Computational-and-Corpus-Libraries.Pdf](http://www.Nederlab.Nl/Cms/Wp-Content/Uploads/2015/10/Reynaert_PICCL-Philosophical-Integrator-of-Computational-and-Corpus-Libraries.Pdf), 2015, <https://repository.ubn.ru.nl/handle/2066/150918>.

¹¹ Matthew L Jockers, *Text Analysis with R for Students of Literature* (Springer, 2016), 137–40.

¹² Justin Grimmer and Brandon Stewart, ‘Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts’, *Political Analysis*, 2013, 6, <https://doi.org/10.1093/pan/mps028>.

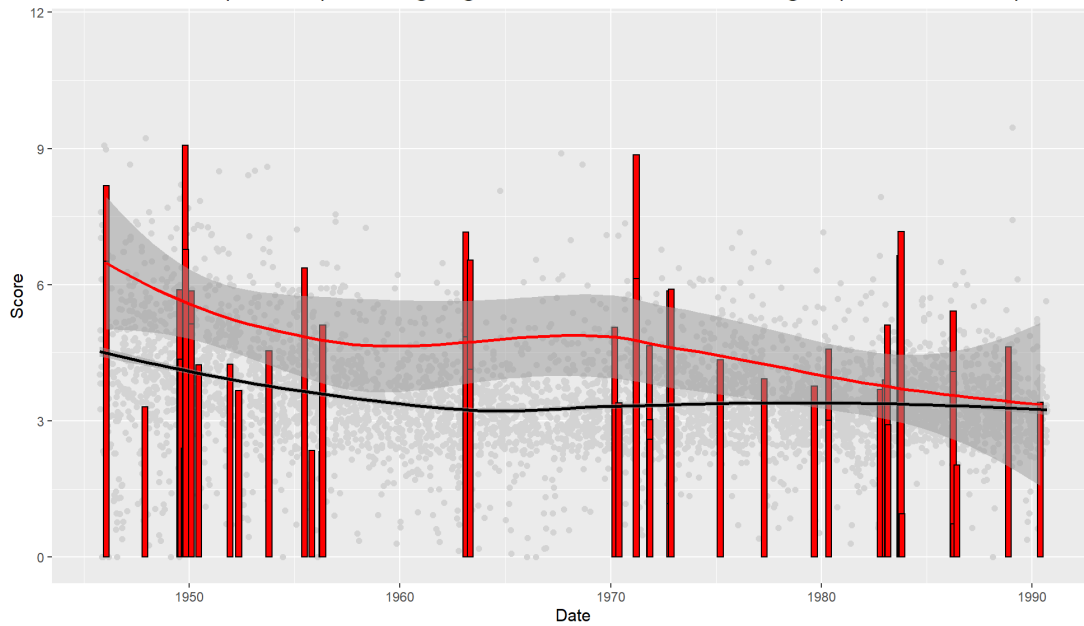
¹³ In practice the original order of words within the historical documents does not inform the analysis performed. Although there are examples of sentences in which a change in the specific word order also fundamentally changes its meaning, this is rare. A bag of words is in most of the computational text analyses sufficient enough to capture the meaning of text. See also: Grimmer and Stuart, ‘Text as Data’, 6.

¹⁴ This statistical weighting is basically still counting words, but its implication is a bit more sophisticated: a TF-IDF score measures how distinctive or characteristic words from an emotion category (from the lexicons) are for a specific document – relative to all other documents in the collection. TF-IDF has as additional advantage that it takes into account a word’s commonness or rarity: relative rare words weigh heavier than very frequent terms. See also: Stephen Robertson, ‘Understanding Inverse Document Frequency: On Theoretical Arguments for IDF’, *Journal of Documentation* 60 (2004): 503.

¹⁵ Eventually this will be done more detailed in the book chapter resulting from (discussing) this conference paper.

Anger in Dutch parliamentary debates (1945-1990)

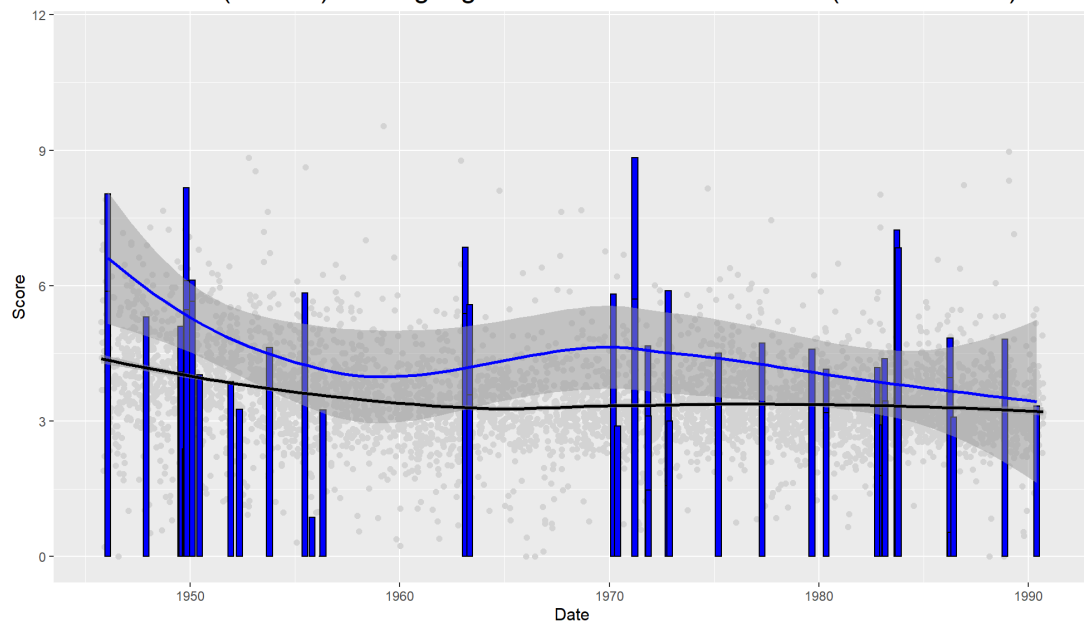
Mean scores (TF-IDF) on language use associated with anger (NRC EmoLex)



Graph 1: Scores of each individual war victim debate and the moving average (loess) on word use associated with anger (in red). All other parliamentary debates in grey (daily mean) and black (moving average).

Sadness in Dutch parliamentary debates (1945-1990)

Mean scores (TF-IDF) on language use associated with sadness (NRC EmoLex)

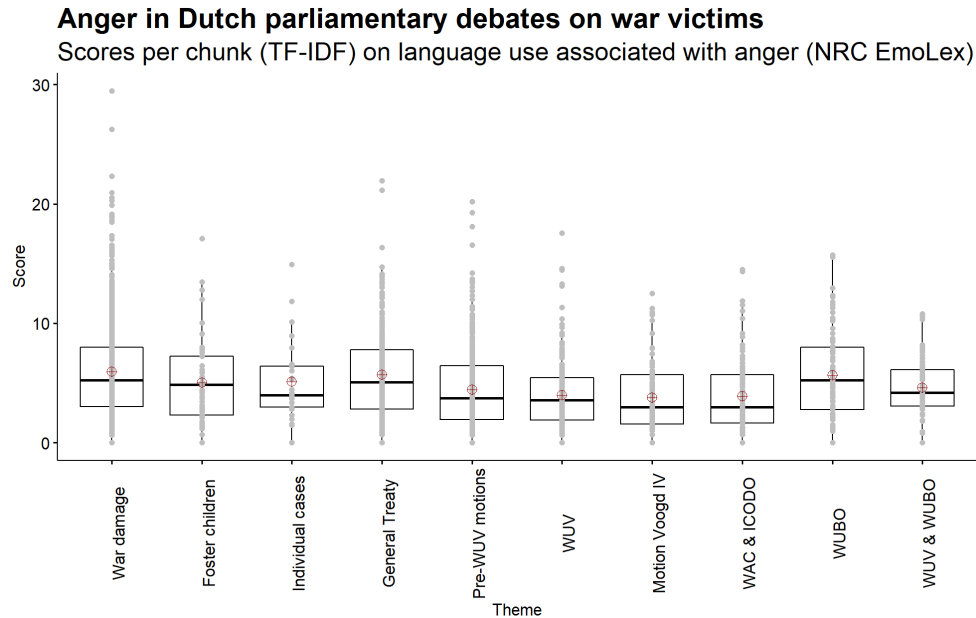


Graph 2: Scores of each individual war victim debate and the moving average (loess) on word use associated with sadness (in blue). All other parliamentary debates in grey (daily mean) and black (moving average).

Two observations are important here. First, there is no linear increase in emotional language use observed over time. Given the stereotypical received wisdom of a period of silence (1950s) followed by a period of sky-rocketing emotionality (1970s), the quantity of use of language associated with anger and sadness seems to be relatively stable over time. This goes both for the ‘average’ parliamentary debate between 1945 and 1990, and the parliamentary discussion of war victim legislation. Secondly, the individual debates on war victim legislation score mostly above the moving average of non-war victim related parliamentary discussions 1945-1990.

5.2 Variation between different thematic clusters of war victim debates

Next, the emotion scores of chunks of individual debates are merged into ten different thematic categories. The results per thematic category of relevant debates are shown in Boxplots 1 (anger) and 2 (sadness). The themes are chronologically sorted in the figures below, based on the first occurrence of the theme on the parliamentary agenda.¹⁶



Boxplot 1: Scores (per 250-word chunk) on anger word use in parliamentary debates on war victims

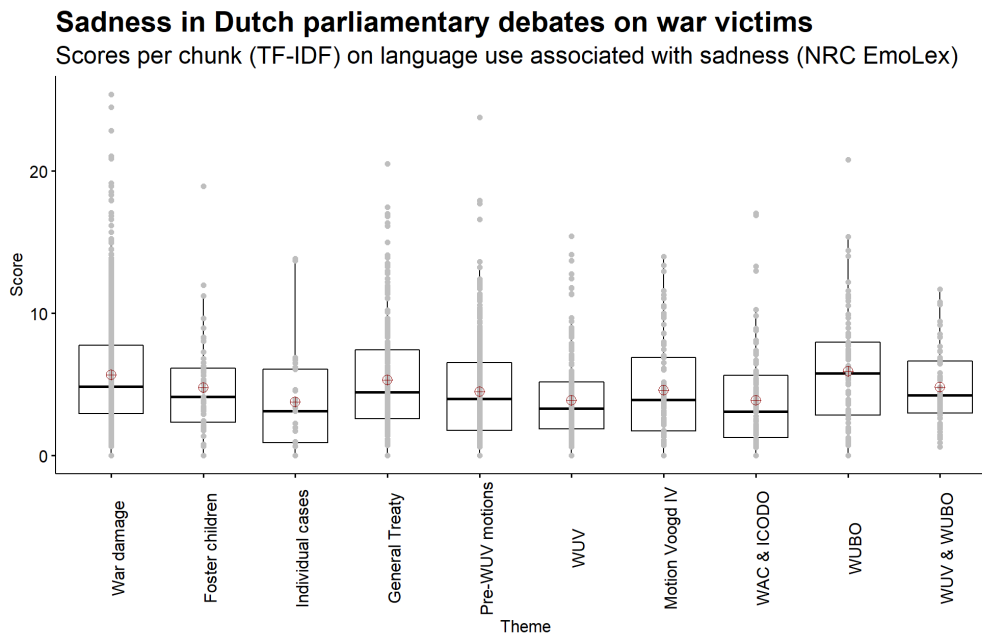
As can be seen in Boxplot 1, the distributions/boxes of the anger scores of the different thematic clusters of debates are relatively similar. There is no single thematic cluster in the parliamentary discussion of war victim legislation between 1945 and 1990 that really stands out. This also goes for the emotion sadness (see Boxplot 2). This rather moderate variation between the scores of the different clusters is also supported by the statistical evaluation of the scores. Statistical evaluation of the results is performed with a Pairwise Wilcoxon Rank Sum Test that is integral part of the R stats package.¹⁷ This test calculates a pairwise comparison between group levels. A confidence interval of 95% is used.¹⁸ Pairwise comparison of the anger scores of the clusters shows in the first place that statistical significant differences are only reported in a minority of the cases (12 out of 45). More meaningful, however, is that we found for all comparisons a small magnitude of effect. This also goes for sadness scores (significant in 13 out of 45 cases, magnitude of effect in all cases is here also small).¹⁹

¹⁶ For more information about the (annotated) thematic clusters of parliamentary debates, see also: *Data Repository 'Vehemence and Victims'*, version 1 (Amsterdam: NIOD, 2020), <https://github.com/MilanvanL/vehemence-victims>.

¹⁷ R Core Team, *R: A Language and Environment for Statistical Computing*.

¹⁸ This means a p-value equal to or smaller than 0.05 is interpreted as statistical significant.

¹⁹ See results of Wilcoxon test in: *Data Repository 'Vehemence and Victims'*.



Boxplot 2: Scores (per 250-word chunk) on sadness word use in parliamentary debates on war victims

6. Historical evaluation and conclusion

Abovementioned results indicate two interesting findings regarding the parliamentary discussion of war victims and related legislation 1945-1990. Firstly, there was no significant or meaningful difference in scores on word use associated with the negative emotions anger and sadness in various thematic clusters of parliamentary debates on war victim-related legislation between 1945-1990. A closer reading of the debates confirms that debates in the early 1950s were already often imbued with strong manifestations of anger and sadness. These expressions were, for example, present in mentions of suffering of victims – this was often used as an argument in advocating for an generous and just dealing with material war damage.²⁰ In addition, no significant linear increase in the use of anger and sadness words between 1945 and 1990 can be reported. This goes not only for the war-victim related parliamentary debates, but also for parliamentary debate in general. Both our computational analysis and close reading of the sources finds no support for the idea of an ‘emotionalization’ of parliamentary discussion, that should have taken off from the 1970s onwards. When we take a closer look at the words associated with anger and sadness that are behind the scores, an interesting historical development can be observed. Although our analysis indicate that the semi-quantitative statement of ‘emotionalization’ does not hold, we did observe a development in the debates that has been described in literature as ‘emancipation of emotions’ regarding war victims. From 1970 onwards more words used to explicitly mention emotions (of others) are used in parliament, compared to earlier debates. Examples are the emergence of words like ‘bitter’, ‘fear’, or ‘grief’ in the top 20 sadness words in the debates.²¹

To conclude, the parliamentary debates on legislation related to victims of the German occupation of the Netherlands display a stable pattern in tone of discussion regarding the negative emotions anger and sadness. Emotion mining in this investigation is in the first place helpful as indicator for outliers (or lack thereof), and in describing diachronic trends on a very rough level. Close reading of the debates remained necessary to gain more fine-grained insights in what was actually going on. Although we did not find support for the idea of an ‘emotionalization’, our analysis indicates the historical development of an ‘emancipation of emotions’. Parliamentarians did not talked more and more emotionally over time, but did explicitly discuss emotions (of others) more often from 1970 onwards.

²⁰ ‘Handelingen Tweede Kamer 1949-1950’, n.d., 305.

²¹ Top 20 highest-scoring anger and sadness words per thematic cluster can be found in the online repository on Github. See: *Data Repository ‘Vehemence and Victims’*.

7. Bibliography

- Benoit, Kenneth, Kohei Watanabe, Haiyan Wang, Paul Nulty, Adam Obeng, Stefan Müller, and Akitaka Matsuo. 'Quanteda: An R Package for the Quantitative Analysis of Textual Data'. *Journal of Open Source Software* 3, no. 30 (2018): 774. <https://doi.org/10.21105/joss.00774>.
- Beunders, Henri. *Hoeveel recht heeft de emotie?: Straffen in de slachtoffercultuur*. Amsterdam: University Press, 2018.
- . *Publieke tranen: de drijfveren van de emotiecultuur*. Amsterdam: Contact, 2002.
- Boot, Peter, Hanna Zijlstra, and Rinie Geenen. 'The Dutch Translation of the Linguistic Inquiry and Word Count (LIWC) 2007 Dictionary'. *Dutch Journal of Applied Linguistics* 6, no. 1 (19 October 2017): 65–76. <https://doi.org/10.1075/dujal.6.1.04boo>.
- Brudholm, Thomas, and Johannes Lang. 'Introduction: Emotions and Mass Atrocity'. In *Emotions and Mass Atrocity: Philosophical and Theoretical Explorations*, edited by Thomas Brudholm and Johannes Lang, 1–20. Cambridge and New York: Cambridge University Press, 2019.
- Data Repository 'Vehemence and Victims' (version 1). Amsterdam: NIOD, 2020. <https://github.com/MilanvanL/vehemence-victims>.
- Dechesne, Mark, and Bryn Bandt-Law. 'Terror in Time: Extending Culturomics to Address Basic Terror Management Mechanisms'. *Cognition and Emotion* 33, no. 3 (3 April 2019): 492–511. <https://doi.org/10.1080/02699931.2018.1460322>.
- Grimmer, Justin, and Brandon Stewart. 'Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts'. *Political Analysis*, 2013, 1–31. <https://doi.org/10.1093/pan/mps028>.
- 'Handelingen Tweede Kamer 1949-1950', n.d.
- Jockers, Matthew L. *Text Analysis with R for Students of Literature*. Springer, 2016.
- Marx, M., J. Van Doornik, A. Nusselder, and L. Buitinck. 'Thematic collection: PoliticalMashup and Dutch Parliamentary Proceedings 1814-2013'. Data Archiving and Networked Services (DANS), 2012. <https://doi.org/10.17026/dans-zg8-9x2v>.
- Mohammad, Saif. 'From Once Upon a Time to Happily Ever After: Tracking Emotions in Novels and Fairy Tales'. *Proceedings of the 5th ACL-HLT Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, 24 June 2011, 105–14.
- Mohammad, Saif M., and Peter D. Turney. 'Crowdsourcing a Word-Emotion Association Lexicon'. *Computational Intelligence* 29, no. 3 (2013): 436–465.
- Oosterhuis, Harry. 'Mental Health, Citizenship, and the Memory of World War II in the Netherlands (1945–85)'. *History of Psychiatry* 25, no. 1 (2014): 20–34.
- Pennebaker, James W. *The Secret Life of Pronouns: What Our Words Say about Us*. New York: Bloomsbury Press, 2013.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, 2019. <https://www.R-project.org>.
- Reynaert, M., M. van Gompel, K. van der Sloot, and A. P. J. van den Bosch. 'PICCL: Philosophical Integrator of Computational and Corpus Libraries'. http://Www.Nederlab.Nl/Cms/Wp-Content/Uploads/2015/10/Reynaert_PICCL-Philosophical-Integrator-of-Computational-and-Corpus-Libraries.Pdf, 2015. <https://repository.ubn.ru.nl/handle/2066/150918>.
- Robertson, Stephen. 'Understanding Inverse Document Frequency: On Theoretical Arguments for IDF'. *Journal of Documentation* 60 (2004): 2004.
- RStudio Team. *RStudio: Integrated Development Environment for R*. Boston, MA, 2018. <http://www.rstudio.com>.
- Withuis, Jolande. *Erkenning: van oorlogstrauma naar klaagcultuur*. Amsterdam: De Bezige Bij, 2002.
- . 'Opkomst en neergang van PTSS'. *Directieve therapie* 30, no. 3 (1 October 2010): 149–59. <https://doi.org/10.1007/BF03096232>.
- Withuis, Jolande, and Annet Mooij. 'From Totalitarianism to Trauma - A Paradigm Change in the Netherlands'. In *The Politics of War Trauma: The Aftermath of World War II in Eleven European Countries*, edited by Jolande Withuis and Annet Mooij, 193–216. Amsterdam: University Press, 2010.