UvA-DARE (Digital Academic Repository)

Clouded reality: News representations of culturally close and distant ethnic outgroups

Kroon, A.C.; Trilling, D.; van der Meer, T.G.L.A.; Jonkman, J.G.F.

DOI
10.1515/commun-2019-2069
Publication date
2020
Document Version
Final published version
Published in
Communications : The European Journal of Communication Research
License
Article 25fa Dutch Copyright Act

Link to publication

Citation for published version (APA):
Kroon, A. C., Trilling, D., van der Meer, T. G. L. A., & Jonkman, J. G. F. (2020). Clouded reality: News representations of culturally close and distant ethnic outgroups. *Communications : The European Journal of Communication Research*, *45*(s1), 744–764. https://doi.org/10.1515/commun-2019-2069

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:19 Jun 2021

Communications 2020; 45(s1): 744–764 **DE GRUYTER** MOUTON

Anne C. Kroon*, Damian Trilling, Toni G. L. A. van der Meer and Jeroen G. F. Jonkman

Clouded reality: News representations of culturally close and distant ethnic outgroups

https://doi.org/10.1515/commun-2019-2069

Abstract: The current study explores how the cultural distance of ethnic outgroups relative to the ethnic ingroup is related to stereotypical news representations. It does so by drawing on a sample of more than three million Dutch newspaper articles and uses advanced methods of automated content analysis, namely word embeddings. The results show that distant ethnic outgroup members (i. e., Moroccans) are associated with negative characteristics and issues, while this is not the case for close ethnic outgroup members (i. e., Belgians). The current study demonstrates the usefulness of word embeddings as a tool to study subtle aspects of ethnic bias in mass-mediated content.

Keywords: prejudice, cultural distance, word embeddings, automated content analysis

News media and racial prejudices

Racial prejudices is argued to be at the heart of the increasingly unfavorable public opinion climate regarding ethnic outgroups in European democracies (Gorodzeisky and Semyonov, 2016). Though the sources of racial prejudice in societies are multifaceted, mass media has been shown to critically contribute to the establishment and re-activation of biased perceptions of outgroup members (e. g., Mastro, 2009). Stereotypes, defined as the shared beliefs about group members' traits (Greenwald, 1995), can be formed through media exposure –

***Corresponding author: Anne C. Kroon**, Amsterdam School of Communication Research (ASCoR), University of Amsterdam, Amsterdam, The Netherlands, E-Mail: A.C.Kroon@uva.nl.
Damian Trilling, Amsterdam School of Communication Research (ASCoR), University of Amsterdam, Amsterdam, The Netherlands, E-Mail: d.c.trilling@uva.nl.
Toni G. L. A. van der Meer, Amsterdam School of Communication Research (ASCoR), University of Amsterdam, G.L.A., Amsterdam, The Netherlands, E-Mail: vanderMeer@uva.nl.
Jeroen G. F. Jonkman, Amsterdam School of Communication Research (ASCoR), University of Amsterdam, Amsterdam, The Netherlands, E-Mail: J.G.F.Jonkman@uva.nl.

even in the absence of interpersonal contact with outgroup members (Mutz and Goldman, 2016; Ramasubramanian, 2007). Particularly in ethnically segregated western societies, the selective presentation in news messages of specific characteristics, issues, and opinions associated with ethnic outgroup members can have consequences for the development of audiences' (non-)prejudiced beliefs and stereotypical associations.

Yet, not all ethnic outgroups are alike or portrayed in the same way by the media. Previous research on media portrayals of ethnicity and race documents large variation in the manner in which diverse ethnic outgroups are portrayed in various media environments (Mastro, 2009; Mutz and Goldman, 2016). For instance, in the US, media reports of Latino and Black Americans over-emphasize issues related to illegality and crime, and this tends to align with negative societal predispositions towards these groups (e. g., Mastro, 2009). In Europe, immigrants – especially those from Arabic and Muslim descent – tend to be represented as a threat to the national security and economic welfare of host countries (Ahmed and Matthes, 2017; Eberl et al., 2018). Albeit far less prominent, alternative framing of immigrants has also been documented, representing European immigrants as highly skilled and/or using positive sentiment (Bleich, Stonebraker, Nisar, and Abdelhamid, 2015; Blinder and Jeannet, 2018).

Cultural distance and stereotypical media portrayals

An important factor that could (partially) account for the variation in stereotypicality of media portrayals of outgroups is the cultural distance between the ethnic outgroup relative to the ethnic ingroup (Allport, Clark, and Pettigrew, 1954). Experimental research documents the intervening role of perceived cultural distance in social categorization processes, ultimately affecting the motivation to maintain central divides between those that are like "us" and those that are like "them" (Mahfud, Badea, Verkuyten, and Reynolds, 2018; Taijfel and Turner, 1979; Van Osh and Breukelmans, 2012). Tapping into social identity sentiments, groups that are regarded as more culturally distant are believed to have different norms, values, beliefs, and worldviews, and are often seen as a threat to the host land's identity, in turn prompting a range of unfavorable intergroup outcomes, such as prejudice and unfavorable attitudes (Van Osh and Breukelmans, 2012). Conversely, groups characterized by a closer cultural distance trigger feelings of similarity and shared social identities, with more favorable intergroup perceptions as a result (Ye, Zhang, Huawen Shen, and Goh, 2014).

Extrapolating these findings to racial prejudice in media coverage, it can be anticipated that negative stereotypical representations in the media pertain especially to more culturally distinct outgroups than those outgroups culturally closer to the national ingroup. Media studies typically suggest that dominant social values and ideologies of ethnic ingroups are placed in the foreground of the media agenda, resulting in the underrepresented and negatively-skewed portrayal of culturally distant minorities (Ahmed and Matthes, 2017; Eberl et al., 2018). Culturally close outgroup members, considered largely equal regarding social norms and ideologies of the majority group, are not likely to stand out as different nor threatening, and may consequently benefit from more favorable media evaluations.

Although explicit, blatant and offensive racial statements may circulate in (social) media environments (Matamoros-Fernández, 2017), subtly communicated and transmitted forms of bias in mainstream media content have been recognized as especially powerful (Mastro, Behm-Morawitz, and Kopacz, 2008; Weisbuch, Pauker, and Ambaday, 2009). In agreement with publicly supported egalitarian norms, news sources and journalists may refrain from blatantly and overtly associating (distant) outgroups with stereotypical attributes (see Gaertner and Dovidio, 2005).

The manifestation of bias in news coverage, then, is more likely to surface in subtle and implicit forms. A promising way to investigate implicit forms of bias in media content is to investigate words that share semantic and syntactic similarity with different outgroups. In particular, one may tap into hidden forms of bias by investigating synonyms and words that often occur in the same linguistic context. Aversive forms of ethnic bias become apparent when stereotypic attributes, such as *criminal* or *offender*, are systematically used interchangeably or analogous to particular ethnic groups.

Current study focus and contribution

The investigation of systematic bias in the use of stereotypical attributes analogous to social categories has recently become possible due to the introduction of word embeddings in the field of Artificial Intelligence (AI) (e. g., Mikolov, Corrado, Chen, and Dean, 2013). The current research note explores these NLP-techniques to investigate hidden forms of bias in news coverage of culturally close and distant outgroups in the Dutch news media environment. We trained such embedding models on the total population of news articles that appeared in the five Dutch newspapers with the highest circulation rate (*de Volkskrant*, *NRC Handelsblad*, *Trouw*, *Algemeen Dagblad*, and *De Telegraaf*) for the period 2000–2015 (covering more than three million news items). Subsequently, we investigated which words have been ‘learned’ to resemble references to outgroup members. In order to investigate whether distant outgroups also appear more often in stereotypical news contexts, we complemented our analysis with a co-occurrence analysis with multidimensional scaling. Finally, we illustrated our findings using high-dimensional visualization techniques.

The current study makes several contributions to the literature. First, and methodologically, the current study shows how word embeddings can be used to explore and identify hidden ethnic bias in mediated content. This state-of-the-art algorithmic technique mapping relations between words allows for the large-scale identification of bias while maintaining high levels of accuracy (Bolukbasi, Chang, Zou, Saligrama, and Kalai, 2016a; Caliskan, Bryson, and Narayanan, 2017; Garg, Schiebinger, Jurafsky, and Zou, 2018). Second, and on a general theoretical level, the study explores the assumptions about the nature of bias in news coverage across two ethnic categories. In conclusion, the here-reported findings contribute to the formulation of expectations about the stereotypicality of news coverage, which may be generalizable beyond specific ethnic groups.

This research note will continue as follows. First, we will briefly outline possible differences between media outlets with regard to the portrayal of ethnic groups. We will then continue to the Method section, where we explain the general idea behind distributed word embeddings, illustrate their application, and argue why word embedding techniques offer a promising toolkit for researchers interested in media bias. We then discuss the results and reflect on our findings in the Discussion section.

Stereotypicality of news content across outlets

In addition to mapping the implicit nature of stereotypical news coverage, the current study aims to identify sources of variation in these portrayals, so to locate under which circumstances audiences are most likely to be exposed to biased representations of ethnic groups. Scholarship in the domain of media stereotyping has identified outlet type as an influential factor affecting the stereotypicality of news stories. The current study explores differences in the use of ethnic stereotypes in tabloid and quality newspapers. The use of stereotypes, as easily accessible heuristics and oversimplification of complex realities, aligns well with tabloids' reporting style of simplification and limited word count. In addition, as tabloid newspapers often have close proximity to right-wing populist parties, they may give voice to anti-foreigner sentiments. Empirical evidence supports the notion that ethnic minorities are represented in stereotypical terms in European tabloid newspapers (Arendt, 2010; Kroon, Kluknavská, Vliegenthart, and Boomgaarden, 2016; Van Dijk, 2000). As a consequence, we anticipate stronger bias regarding distant outgroups in popular newspapers compared to quality newspapers.

Method

Classical methods of automated content analysis that measure whether a given word co-occurs with another may give a clear indication of the news contexts in which specific minority groups appear (see, for example, Jacobs, Damstra, Boukes, Swert, and Boukes, 2018; Ruigrok and Atteveldt, 2007). Yet, co-occurrence analyses are less suitable for the detection of stereotypical synonyms and analogies due to the focus on the *direct* textual context and its inherently deductive nature.

Recently, advancements in natural language processing, and, in particular the use of word embeddings, have made the refined analysis of (hidden) bias in texts possible (Bolukbasi et al., 2016a; Mikolov, Corrado, et al., 2013). Word embeddings are a current state-of-the-art algorithm for capturing, understanding and analyzing aspects of word meaning. The premise of word embeddings is based on the principle of distributional similarity, which Firth (1957) tellingly summarized as "[y]ou shall know a word by the company it keeps" (p. 11). Taking collections of unlabeled sentences as input, word embeddings are trained to learn the meaning of words by analyzing the context in which these appear. The word-embedding algorithm thus transforms human language into meaningful numerical representations in the form of valued vectors. Each word is represented by a vector of – in our case – hundred dimensions. Just as one can easily compute the distance between two objects in a three-dimensional space, one can also compute the distance between two words in a hundred-dimensional space.

We analyze our data using the word2vec algorithm as implemented in the Gensim package for Python (Mikolov, Corrado et al., 2013; Mikolov, Yih, and Zweig, 2013). First, this tool builds a vocabulary using unlabeled sentences from the training dataset. Second, word2vec learns word representations – or, more precisely, embeddings – by training an unsupervised, shallow neural network model. In this training process, embeddings are determined based on the direct context[1] in which specific words occur in the different sentences in the training data. With sufficient instances of contextual similarity, the model will learn that

1 The size of this context, i. e., the number of surrounding words which the algorithm considers, depends on the fixed window size. The word embedding model presented in this paper was trained using a window size of 5.

words are associated. In fact, the main idea behind word embeddings is that words closer together in a vector space share semantic meaning. Hence, two synonyms would occupy similar positions in the vector space and have a distance close to zero, whereas very different words should have vector locations further apart. Most similar words can be synonyms, but also represent words that are used in comparable contextual and topical domains (Bolukbasi, Chang, Zou, Saligrama, and Kalai, 2016b). For example, if the word *cereal* occurs frequently in sentences with the word *breakfast*, the model will learn that these words share semantic meaning.

Word embeddings have proved particularly useful in modeling diverse lexical tasks, such as information retrieval, as well as the identification and prediction of sentiment and topics (see for an example in the field of communication science: Rudkowsky et al., 2018). Yet, while word embeddings encode relevant semantic information, they also inherently reflect biases and stereotypes if those are present in the training dataset (Bolukbasi et al., 2016a). Of interest to the study's aim, hidden biases in texts can be accurately detected and analyzed using word embeddings, a claim that is supported by a series of recently published studies in the field of computational sciences (Bolukbasi et al., 2016a; Caliskan et al., 2017; Garg et al., 2018; Tulkens, Hilte, Lodewyckx, Verhoeven, and Daelemans, 2016). These studies show that words close to specific social categories in the vector space may reveal bias. For example, if the vector representation of the word *she* is close to representations of words such as *receptionist*, *nanny*, or *housekeeper*, while the word *he* is close to *financier*, *warrior*, or *broadcaster*, this could suggest the presence of gender stereotypes (Bolukbasi et al., 2016a).

Training and data

For the purpose of the current study, we trained embeddings on all news items that appeared in the five Dutch newspapers with the highest circulation rate: *de Volkskrant*, *NRC Handelsblad*, *Trouw*, *Algemeen Dagblad*, and *De Telegraaf* for the period January 2000 up to and including December 2015. This resulted in a final sample of 3,316,494 news articles. By training the embeddings on the total population of news articles over this 16-year period, we derived word meanings over a substantial period of time, herewith providing a critical test of our hypotheses of the nature of ethnic stereotypes. In addition, and acknowledging potential differences across newspaper types, we trained two separate sets of embeddings on news items from quality (*de Volkskrant*, *NRC Handelsblad*, *Trouw*, [n = 1,777,024 news items]) and popular (*Algemeen Dagblad*, and *De Telegraaf* [n = 1,539,470 news items]) newspapers.

Selection of culturally close and distant ethnic outgroups

The study considers close and distant ethnic outgroups in the case of the Netherlands. Although this country is traditionally seen as highly tolerant towards ethnic outgroups, its political climate towards ‘others’ has toughened significantly over the past two decades (Erisen and Kentmen-Cin, 2017; Vasta, 2007). As the residents from a neighboring country, we consider Belgians as culturally close ethnic outgroups. The Dutch share strong historical, cultural, social and lingual backgrounds with Belgium (Polek, Wöhrle, and van Oudenhoven, 2010). Differences in perceived happiness, life satisfaction, and subjective well-being are relatively small between the neighboring countries (Inglehart and Klingemann, 2000).

For our selection of culturally distant outgroups, the current study focuses on a major ethnic-minority group in the Netherlands: Moroccans. Traditionally this ethnic group does not share historical, cultural or linguistic ties with the Dutch. The vast majority of people from Moroccan descent tend to be Muslims, while the Netherlands is traditionally a Christian country (even though nowadays the share of practicing Christians has declined). Vastly documented negative stereotypes that pertain to Muslims in the Western world are therefore likely to apply to this ethnic group (González, Verkuyten, Weesie, and Poppe, 2008; Richardson, 2004; Savelkoul, Scheepers, Tolsma, and Hagendoorn, 2011; Strabac and Listhaug, 2008). Moreover, the Dutch report experiencing substantial cultural differences with people from Moroccan descent (Van Osh and Breukelmans, 2012).

Analysis word embeddings

We thus explore bias in Dutch newspapers by retrieving the hundred most proximate words to *Belgian(s)* and the hundred most proximate words to *Moroccan(s)*. A Python script was written to retrieve the hundred nearest neighbors to the selected close and distant outgroups from the embedding models. These words are considered most indicative for coverage about these groups, and therefore powerful in uncovering potential bias (see Agrawal and Awekar, 2018). An example of a neutral attribute close to an ethnic group would be a different nationality; when the word *Belgian* appears close to the word *Italian*, this indicates that both words are interchangeably in sentences such as: “*The [Belgian / Italian] governments are meeting...*” An example of a stereotypic attribute close to an ethnic group is *criminal*: When used interchangeable with references to *Moroccan(s)*, it suggests bias: “*The police have apprehended a [Moroccan / criminal]...*”

The authors thoroughly analyzed and categorized these words for the embeddings trained on tabloid newspapers, quality newspapers, and all newspapers simultaneously. For each word, the authors indicated whether or not it was negatively valenced. All the negatively valenced words were compiled into a list of 63 unique words. The words reflect dimensions of criminality, illegality, scam, prostitution, and violence in general. Please consult Appendix B for the complete list.

We visualized the results of our word embedding analysis using a machine-learning algorithm for data visualization, namely t-distributed stochastic neighbor embeddings (t-SNE). This popular method for visualizing high-dimensional data allows data points to be plotted on a two-dimensional map to identify patterns (Van der Maaten and Hinton, 2008). t-SNE uses a Gaussian distribution to create a probability distribution to map relationships between data points in a high-dimensional space. Due to its ability to preserve local structures, t-SNE will tend to plot points near each other that are close in the high-dimensional space.

Analysis of co-occurrence

Although the interchangeability of words to indicate targets (e. g., *Moroccans*) and negative attributes (e. g., *criminals*) strongly hints towards implicit bias in news coverage, one may wonder whether negative attributes are also present in news articles that mention close and distant outgroup members. A straightforward manner to assess whether these groups appear in stereotypical news contexts is the analysis of co-occurrences between targets groups and words that indicate stereotypical contexts. Therefore, and in a second step, co-occurrences were calculated between references to the target groups[2] and the compiled list of 63 negatively valenced words based on our embeddings analysis (as included in Appendix B). By investigating whether the target groups are mentioned in news articles that deal with criminality, trafficking and violence, we can draw conclusions about the stereotypicality of the direct news context in which these groups appear. We visualize the results of the co-occurrence analysis using multidimensional scaling, a technique that visualizes the similarity between data points.

2 The following search string was used to select the relevant articles: "Marokkaan OR Marokkanen OR Marokkaans OR Marokkaanse OR Belg OR Belgen OR Belgisch OR Belgische".

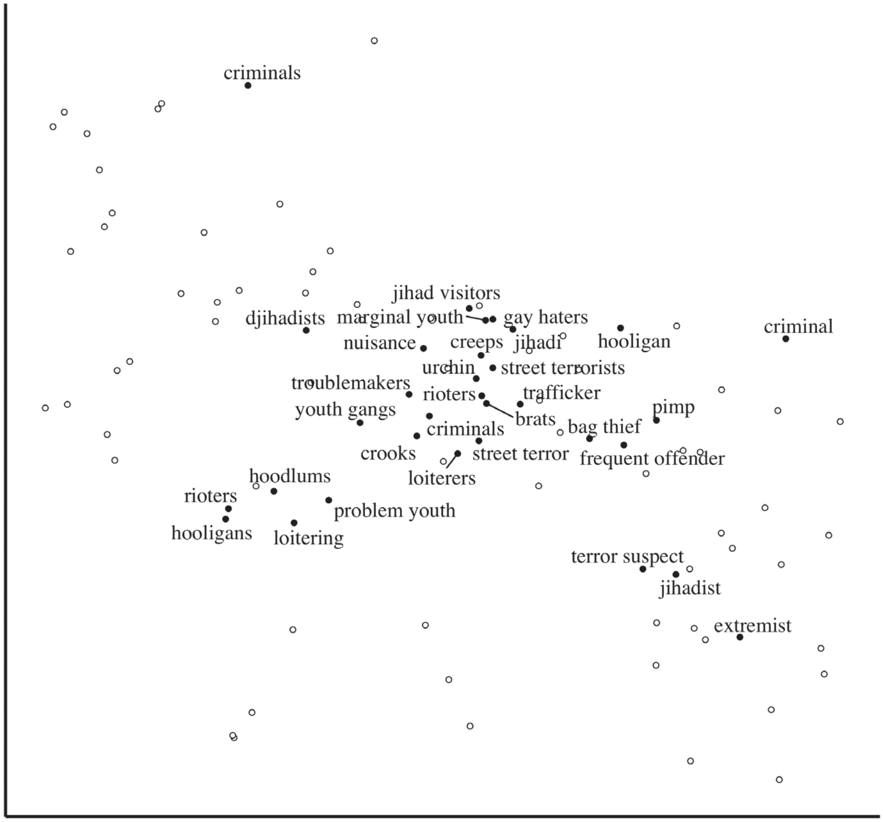
Results

Word embeddings. We discuss the results of the embeddings trained on the five largest Dutch newspapers. Both Belgian(s) (N = 45,931) and Moroccan(s) (N = 28,066) figured often in the news, warranting the quality of the embeddings specific to these groups (Schnabel, Labutov, and Mimno, 2015). Tables 1 and 2 display the findings. As can be seen, Belgians are mostly associated in the news with other European nationalities, such as the *French*, *German*, *Italians*, and *Spaniards*. In addition, references to Belgians were used analogous to sports-related terms in news stories; in particular cyclists, evidenced by proximate words such as *cross cyclists* (veldrijders), *top favorites* (topfavorieten), and *teammate* (teammaat). Notably, none of the hundred most proximate words carry a clear negative connotation.

Turning to our inspection of the hundred most proximate words to Moroccan(s), the data reveal a strikingly different picture. Moroccans often appear in the vicinity of other non-European ethnic minorities, such as *Turks*, *Antilleans*, and *Surinamese*. In addition, we find that around thirty percent of the hundred most proximate words are clearly negatively valenced: Dutch newspapers associate Moroccans with words such as *brats* (rotjochies), *criminals* (crimineeltjes), *loitering* (hangjongeren), *crooks* (boefjes), and *rioters* (relschoppers). Please consult Appendix A for a complete list of the hundred most proximate words for both Belgians and Moroccans.

Figure 1 summarizes these findings. As can be seen, Moroccans are often associated with negatively valenced.

The difference in representations of close and distant outgroups across newspaper types are discussed next. The embeddings trained on popular and quality newspapers, respectively, were used to retrieve the hundred most proximate words for both Belgian(s) and Moroccans. The percentages of negatively valenced words among the hundred most proximate words for both outgroups across newspaper types are presented in Table 2. As can be seen, Belgians are neither associated with negative words in popular nor quality newspapers. Moroccans, on the contrary, were used interchangeably with negative terms most often in popular newspapers (40 %) compared to quality newspapers (13 %). Popular newspapers describe Moroccans more often in relation to low status, criminality, and hostility, illustrated by words such as *troublemakers*, *drug addicts*, *youth gangs*, and *illegality*. These negative associations result in unwarranted and strongly negative representations of Moroccans in popular newspapers. The findings indicate that mainly popular newspapers are responsible for spreading negative associations regarding this distant outgroup category.



Note. Empty data points represent neutral words, filled data points represent negatively valenced words. For reasons of readability, only negatively-valenced words are labeled.

Figure 1: Two-dimensional vector representation of the hundred most proximate words to Moroccans.

Co-occurrences. In a next step, and in order to investigate the stereotypicality of the direct *news context* in which outgroups appear, co-occurrences were calculated between the target groups and the stereotypical attributes identified in our analysis using word embeddings. The results of the co-occurrence analysis are presented in Table 3. In line with the results of the analysis using word embeddings, we find that Moroccans appear more often in stereotypical news contexts than Belgians. More specifically, the results show that of the 3362 news articles that mention Moroccan(s), 19.2 % also mention at least one of the stereotypical attributes. In contrast, only 5 % of the 11081 news articles that mention Belgians contain stereotypical attributes. Figure 2 displays the results of the multidimen-

Table 1: Top 20 most proximate words to Belgian(s) and Moroccan(s).

Input	Word	Cosine similarity	Input	Word	Cosine similarity
	French	0.75		Turk	0.81
	German	0.74		Antilleans	0.77
	Italians	0.74		Turks	0.73
	Spaniards	0.74		Surinamese	0.72
	Norwegians	0.73		Surinamer	0.72
	Swede	0.71		Muslim	0.69
	Frenchman	0.71		**Brats**	**0.68**
	Czechs	0.70		**Jihadist**	**0.67**
	Austrian	0.70		**Criminals**	**0.66**
Belgian +	Germans	0.69	Moroccan +	Algerian	0.66
Belgians	Flames	0.69	Moroccans	Moroccan	0.65
	Danes	0.68		**Loitering**	**0.65**
	Swiss	0.68		Foreigners	0.65
	Veteran	0.68		Immigrant	0.65
	Luxembourger	0.68		Muslims	0.65
	Greek	0.68		Antillean	0.65
	Dutchman	0.67		Boys	0.65
	Czech	0.67		**Crooks**	**0.64**
	Portuguese	0.67		Moluccas	0.64
	Aussies	0.67		**Rioters**	**0.64**

Note. Top 20 most proximate words to Belgian(s) and Moroccan(s) according to embeddings trained on *de Volkskrant*, *NRC Handelsblad*, *Trouw*, *Algemeen Dagblad*, and *De Telegraaf*, for the period 2000–2015. Words in bold carry a negative connotation.

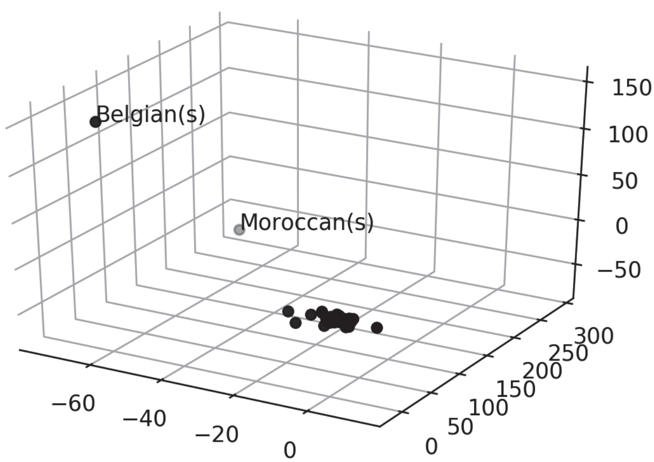
Table 2: Percentage of negative associations with close and distant outgroups across newspaper types.

	All newspapers	Tabloid newspapers	Quality newspapers
Belgian(s)	0 %	0 %	0 %
Moroccan(s)	33 %	40 %	13 %

sional scaling representation of the co-occurrence matrices. As can be seen, references to Moroccans appear closer to negative attributes than references to Belgians.

Table 3: Results of the co-occurrence analysis.

	N news articles about target groups	*N* mentions of the target group	*N* articles that mention both target group and stereotypical words	*N* co-occurrences between target group and stereotypical attributes within news articles	% of news articles in which target group appears in stereotypical context
Moroccan(s)	3362	5530	645	899	19.2 %
Belgian(s)	11081	15960	558	724	5.0 %



Note. Unlabed datapoints refer to negative attributes (full list is included in Appendix B).

Figure 2: Multidimensional scaling representation of the co-occurrences between Moroccan(s), Belgian(s), and negative attributes.

Discussion

This study has investigated representations of culturally close and distant ethnic outgroups in Dutch newspapers. It has done so by drawing on more than three million Dutch newspaper stories and by using word embeddings, an advanced algorithm for capturing, understanding and analyzing aspects of word meaning.

Based on the premises of frameworks of social and group identities (Taijfel and Turner, 1979), it was hypothesized that culturally distant outgroups are more negatively represented by newspapers than culturally close outgroups. The results confirm this expectation: The data show that Dutch newspapers associate Belgians, here considered a close ethnic outgroup, with neutral or sport-related terms. Particularly, Belgians were frequently associated with words such as *cyclists*, or *team mate*. On the contrary, the results reveal that Moroccans, considered a distant ethnic outgroup, are depicted in relation to negative issues and characteristics: A considerable number of words frequently used in close proximity to this group carry a clear negative connotation. More specifically, references to this group are replaceable with terms such as *criminals*, *crooks*, *jihadists*, and *rioters*. Especially popular newspapers prominently use references to Moroccans interchangeably with negative and unwarranted attributes. This finding confirms previous evidence for the stereotypical nature of news stories about ethnic outgroups in these types of newspapers (Arendt, 2010; Kroon et al., 2016; Van Dijk, 2000). Co-occurrence analyses confirm that Moroccans, as compared to Belgians, appear in the news in close proximity to stereotypic attributes.

It is important to note that such news representations are not inconsequential. The linkage between targets (i. e., social groups) and attributes (i. e., issues, characteristics) in media messages, also referred to as mediated associations (Arendt and Karadas, 2017) can cause audience members to see these concepts as related. Van Atteveldt (2008) reasons that "[e]ven if the two are not related or are even explicitly dissociated, this tells us something about the worldview of the source of the messages containing both concepts, and it can cause the receiver of those messages to relate the two concepts" (p. 65). Experimental evidence in the media stereotyping domain supports this argument: Studies consistently find that exposure to stereotypical mediated associations establishes and re-activates cognitive linkages between target groups and attributes, herewith increasing the availability and accessibility of stereotype-congruent associations in the memory (Arendt, 2013; Cho, Gil de Zuniga, Shah, and McLeod, 2006; Verhaeghen, Aikman, and Gulick, 2011).

This study has demonstrated the usefulness of word embeddings in studying the representations of minorities in Dutch news coverage. Following in the footsteps of innovative studies in the field of AI (e. g., Bolukbasi et al., 2016a; Caliskan et al., 2017; Garg et al., 2018), this study has explored the benefits of word embeddings to detect subtle bias in large bodies of news media data. Yet, some limitations should be acknowledged. First, as word embeddings are the outcome of a rather advanced and complex training process, its results might not always be intuitively and straightforwardly interpreted. Second, as we only considered a single close and distant outgroup, the here-reported findings cannot simply be transferred to other nationalities, countries, or newspapers. Future studies should test differences in media representations between culturally similar and distinct groups among a large sample of ethnicities. In addition, due to the static nature of our analysis, it remains unclear to what extent bias in news coverage has become more or less pronounced with time.

The current study has demonstrated that Moroccans, being a culturally distant ethnic minority group in the Netherlands, are unwarrantedly associated in news stories with problems, criminality, and hostility. Changing the media representation of this and other culturally distant minority groups could decrease harmful stereotypes about these groups in society.

References

Agrawal, S., & Awekar, A. (2018). Deep learning for detecting cyberbullying across multiple social media platforms. *Lecture notes in Computer Science (Including subseries lecture notes in Artificial Intelligence and lecture notes in Bioinformatics)*, *10772 LNCS* (Table 2), 141–153. doi:10.1007/978-3-319-76941-7_11

Ahmed, S., & Matthes, J. (2017). Media representation of Muslims and Islam from 2000 to 2015: A meta-analysis. *International Communication Gazette*, *79*(3), 219–244. doi:10.1177/1748048516656305

Allport, G. W., Clark, K., & Pettigrew, T. (1954). The nature of prejudice. Cambridge, MA: Addison-Wesley Pub. Co.

Arendt, F. (2010). Cultivation effects of a newspaper on reality estimates and explicit and implicit attitudes. *Journal of Media Psychology*, *22*(4), 147–159. doi:10.1027/1864-1105/a000020

Arendt, F. (2013). Dose-dependent media priming effects of stereotypic newspaper articles on implicit and explicit stereotypes. *Journal of Communication*, *63*, 830–851. doi:10.1111/jcom.12056

Arendt, F., & Karadas, N. (2017). Content analysis of mediated associations: An automated text-analytic approach. *Communication Methods and Measures*, *11*(2), 105–120. doi:10.1080/19312458.2016.1276894

Bleich, E., Stonebraker, H., Nisar, H., & Abdelhamid, R. (2015). Media portrayals of minorities: Muslims in British newspaper headlines, 2001–2012. *Journal of Ethnic and Migration Studies*, *41*(6), 942–962. doi:10.1080/1369183X.2014.1002200

Blinder, S., & Jeannet, A. M. (2018). The “illegal” and the skilled: Effects of media portrayals on perceptions of immigrants in Britain. *Journal of Ethnic and Migration Studies*, *44*(9), 1444–1462. doi:10.1080/1369183X.2017.1412253

Bolukbasi, T., Chang, K.-W., Zou, J., Saligrama, V., & Kalai, A. (2016a). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. *NIPS*, 1–9. Retrieved March 20, 2018 from http://arxiv.org/abs/1607.06520.

Bolukbasi, T., Chang, K.-W., Zou, J., Saligrama, V., & Kalai, A. (2016b). Quantifying and reducing stereotypes in word embeddings. doi:10.1016/S0017-9310(97)00074-4

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora necessarily contain human biases. *Science*, *6334*(356), 183–186.

Cho, J., Gil de Zuniga, H., Shah, D. V., & McLeod, D. M. (2006). Cue convergence: Associative effects on social intolerance. *Communication Research*, *33*(3), 136–154.

Eberl, J., Meltzer, C. E., Heidenreich, T., Theorin, N., Lind, F., Berganza, R., …, Schemer, C. (2018). The European media discourse on immigration and its effects: A literature review. *Annals of the International Communication Association*, *842*(3), 207–223. doi:10.1080/23808985.2018.1497452

Erisen, C., & Kentmen-Cin, C. (2017). Tolerance and perceived threat toward Muslim immigrants in Germany and the Netherlands. *European Union Politics*, *18*(1), 73–97. doi:10.1177/1465116516675979

Firth, J. R. (1957). *Papers in linguistics, 1934–1951*. Oxford University Press.

Garg, N., Schiebinger, L., Jurafsky, D., & Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *PNAS*, 1–33. Retrieved April 20, 2019 from http://arxiv.org/abs/1711.08412.

Gaertner, S. D., & Dovidio, J. F. (2005). Understanding and addressing contemporary racism: From aversive racism to the common Ingroup Identity Model. *Journal of Social Issues*, (3), 615–639.

González, K. V., Verkuyten, M., Weesie, J., & Poppe, E. (2008). Prejudice towards Muslims in the Netherlands: Testing integrated threat theory. *British Journal of Social Psychology*, *47*(4), 667–685. doi:10.1348/014466608X284443

Gorodzeisky, A., & Semyonov, M. (2016). Not only competitive threat but also racial prejudice: Sources of anti-immigrant attitudes in European societies. *International Journal of Public Opinion Research*, *28*(3), 331–354. doi:10.1093/ijpor/edv024

Greenwald, A. G. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*(1), 4–27.

Inglehart, R., & Klingemann, H. D. (2000). Genes, culture, democracy, and happiness. In E. Diener & E. M. Suh (Eds.), *Culture and subjective well-being* (pp. 165–183). Cambridge, MA: The MIT Press.

Jacobs, L., Damstra, A., Boukes, M., Swert, K. De, & Boukes, M. (2018). Back to reality: The complex relationship between patterns in immigration news coverage and real-world developments in Dutch and Flemish newspapers (1999 – 2015). *Mass Communication and Society*, *21*(4), 473–497. doi:10.1080/15205436.2018.1442479

Kroon, A. C., Kluknavská, A., Vliegenthart, R., & Boomgaarden, H. G. (2016). Victims or perpetrators? Explaining media framing of Roma across Europe. *European Journal of Communication*, *31*(4). doi:10.1177/0267323116647235

Mahfud, Y., Badea, C., Verkuyten, M., & Reynolds, K. (2018). Multiculturalism and attitudes toward immigrants: The impact of perceived cultural distance. *Journal of Cross-Cultural Psychology*, *49*(6), 945–958. doi:10.1177/0022022117730828

Mastro, D. (2009). Effects of racial and ethnic stereotyping. In J. Bryant & M. B. Oliver (Eds.), *Media effects: Advances in theory and research* (pp. 325–341). New York, NY: Routledge.

Mastro, D. E., Behm-Morawitz, E., & Kopacz, M. A. (2008). Exposure to television portrayals of Latinos: The implications of aversive racism and social identity theory. *Human Communication Research*, *34*(1), 1–27. doi:10.1111/j.1468-2958.2007.00311.x

Matamoros-Fernández, A. (2017). Platformed racism: the mediation and circulation of an Australian race-based controversy on Twitter, Facebook and YouTube. *Information Communication and Society*, *20*(6), 930–946. doi:10.1080/1369118X.2017.1293130

Mikolov, T., Corrado, G., Chen, K., & Dean, J. (2013). Efficient estimation of word representations in vector space. *ArXiv*, 1–12.

Mikolov, T., Yih, W., & Zweig, G. (2013). Linguistic regularities in continuous space word representations. *Proceedings of NAACL-HLT*, 746–751.

Mutz, D. C., & Goldman, S. K. (2016). Mass media. In J. F. Dovidio, M. Hewstone, P. Glick, & V. M. Esses (Eds.), *The SAGE handbook of prejudice, stereotyping and discrimination* (pp. 241–258). London, England: Sage Publications.
Polek, E., Wöhrle, J., & van Oudenhoven, J. P. (2010). The role of attachment styles, perceived discrimination, and cultural distance in adjustment of German and Eastern European immigrants in the Netherlands. *Cross-Cultural Research*, *44*(1), 60–88. doi:10.1177/1069397109352779
Ramasubramanian, S. (2007). Media-based strategies to reduce racial stereotypes activated by news stories. *Journalism & Mass Communication Quarterly*, *84*(2), 249–264. doi:10.1177/107769900708400204
Richardson, J. E. (2004). *(Mis) representing Islam: The racism and rhetoric of British broadsheet newspapers*. Amsterdam: John Benjamins Publishing.
Rudkowsky, E., Haselmayer, M., Wastian, M., Jenny, M., Emrich, Š., Sedlmair, M., …, Sedlmair, M. (2018). More than bags of words: Sentiment analysis with word embeddings. *Communication Methods and Measures*, *12*(2–3), 140–157. doi:10.1080/19312458.2018.1455817
Ruigrok, N., & Atteveldt, W. Van. (2007). Global angling with a local angle: How U.S., British, and Dutch newspapers frame global and local terrorist attacks. *Press/Politics*, *12*(1), 68–90. doi:10.1177/1081180X06297436
Savelkoul, M., Scheepers, P., Tolsma, J., & Hagendoorn, L. (2011). Anti-Muslim attitudes in the Netherlands: Tests of contradictory hypotheses derived from ethnic competition theory and intergroup contact theory. *European Sociological Review*, *27*(6), 741–758. doi:10.1093/esr/jcq035
Schnabel, T., Labutov, I., & Mimno, D. (2015). Evaluation methods for unsupervised word embeddings. *Proceedings of the 2015 conference on empirical methods in natural language processing*, 298–307.
Strabac, Z., & Listhaug, O. (2008). Anti-Muslim prejudice in Europe: A multilevel analysis of survey data from 30 countries. *Social Science Research*, *37*(1), 268–286. doi:10.1016/j.ssresearch.2007.02.004
Taijfel, H., & Turner, J. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA: Brooks-Cole.
Tulkens, S., Hilte, L., Lodewyckx, E., Verhoeven, B., & Daelemans, W. (2016). A dictionary-based approach to racism detection in Dutch social media.
Van Atteveldt, W. (2008). *Semantic network analysis: Techniques for extracting, representing and querying media content*. Charleston SC: BookSurge Publisher.
Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, *9*(2210), 2579–2605. doi:10.1007/s10479-011-0841-3
Van Dijk, T. A. (2000). New(s) racism: A discourse analytical approach. In S. Cottle (Ed.), *Ethnic minorities and the media* (pp. 33–49). Keynes, UK: Open University Press.
Van Osh, Y. M. J., & Breukelmans, S. M. (2012). Perceived intergroup differences as an organizing priniciple of intercultural attitudes and acculturation attitudes. *Journal of Cross-Cultural Psychology*, *43*(43), 801–821. doi:10.1177/0022022111407688
Vasta, E. (2007). From ethnic minorities to ethnic majority policy: Multiculturalism and the shift to assimilationism in the Netherlands. *Ethnic and Racial Studies*, *30*(5), 713–740. doi:10.1080/01419870701491770

Verhaeghen, P., Aikman, S. N., & van Gulick, A. E. (2011). Prime and prejudice: Co-occurrence in the culture as a source of automatic stereotype priming. *British Journal of Social Psychology*, *50*, 501–518. doi:10.1348/014466610X524254

Weisbuch, M., Pauker, K., & Ambaday, N. (2009). The subtle transmission of race bias via televised nonverbal behavior. *Science*, *5960*(326), 1711–1714. doi:10.1126/science.1178358

Ye, B. H., Zhang, H. Q., Huawen Shen, J., & Goh, C. (2014). Does social identity affect residents' attitude toward tourism development?: An evidence from the relaxation of the individual visit scheme. *International Journal of Contemporary Hospitality Management*, *26*(6), 907–929. doi:10.1108/IJCHM-01-2013-0041

Appendix A

Top 100 most proximate words to Belgian(s) and Moroccan(s) in Dutch newspapers

Table A1: Top 100 most proximate words to Belgian(s) and Moroccan(s).

Belgian + Belgians			Moroccan + Moroccans		
Dutch	**English translation**	**Cosine similarity**	**Dutch**	**English translation**	**Cosine similarity**
Fransen	French	0.746	Turk	Turk	0.810
Duitser	German	0.740	Antillianen	Antilleans	0.770
Italianen	Italians	0.738	Turken	Turkish	0.725
Spanjaarden	Spaniards	0.738	Surinamers	Surinamese	0.722
Noren	Norwegians	0.731	Surinamer	Surinamer	0.717
Zweed	Swede	0.710	Moslim	Muslim	0.695
Fransman	Frenchman	0.709	Rotjochies	**Brats**	0.678
Tsjechen	Czechs	0.700	Jihadist	**Jihadist**	0.669
Oostenrijkers	Austrian	0.696	Crimineeltjes	**Criminals**	0.663
Duitsers	Germans	0.692	Algerijn	Algerian	0.662
Vlamingen	Flames	0.685	Marokkaanse	Moroccan	0.654
Denen	Danes	0.684	Hangjongeren	**Loitering**	0.654
Zwitser	Swiss	0.681	Buitenlanders	Foreigners	0.650
Routinier	Veteran	0.678	Allochtoon	Immigrant	0.650
Luxemburger	Luxembourger	0.677	Moslims	Muslims	0.647
Griek	Greek	0.675	Antilliaan	Antillean	0.646
Nederlander	Dutchman	0.674	Jongens	Boys	0.645
Tsjech	Czech	0.673	Boefjes	**Crooks**	0.644
Portugezen	Portuguese	0.669	Molukkers	Moluccas	0.644
Aussies	Aussies	0.668	Reljongeren	**Rioters**	0.641
Vlaming	Fleming	0.666	Buurtvaders	Neighborhood fathers	0.641
Oostenrijker	Austrian	0.662	Tasjesdief	**Bag thief**	0.640
Slowaak	Slovak	0.661	Arabier	Arab	0.639
Schlecks	Schlecks	0.658	Hindoestanen	Hindus	0.638
Raborenner	Rabo rider	0.656	Imam	Imam	0.637
Engelsen	English	0.653	Roemenen	Romanians	0.632
Limburgers	Residents of Limburg	0.649	Marokkaans	Moroccan	0.631
Haantjes	Machos	0.649	Jood	Jew	0.631
Achterhoekers	Residents of the Achterhoek	0.648	Straatterroristen	**Street terrorists**	0.630
Deen	Dane	0.644	Allochtonen	Allochtonen	0.625
Finnen	Fins	0.643	Egyptenaar	Egyptian	0.621

Table A1 (continued)

Belgian + Belgians			Moroccan + Moroccans		
Dutch	**English translation**	**Cosine similarity**	**Dutch**	**English translation**	**Cosine similarity**
Catalanen	Catalans	0.642	Zigeuners	Gypsies	0.620
Roemenen	Romanians	0.642	Jihadgangers	**Jihad visitors**	0.618
Noor	Norwegian	0.639	Afghaan	Afghan	0.618
Renners	Cyclists	0.637	Pedofiel	**Pedophile**	0.617
Ajacieden	Ajacids	0.633	Marokkaantjes	Moroccans	0.617
Sprinters	Sprinters	0.633	Koerd	Kurd	0.615
Italiaan	Italian	0.632	Arabieren	Arabs	0.614
Friezen	Friezes	0.632	Homohaters	**Gay haters**	0.613
Hoste	Hoste	0.631	Joegoslaven	Yugoslavs	0.612
Roemeen	Romanian	0.631	Ghanezen	Ghanaians	0.612
Noorderlingen	Northerners	0.630	Moslimjongeren	Muslim youth	0.609
Brabander	Resident of Brabant	0.626	Lastpakken	**Troublemakers**	0.608
Sloveen	Slovene	0.626	Berbers	Berbers	0.608
Hongaar	Hungarian	0.625	Islamieten	Islamists	0.607
Routiniers	Veterans	0.624	Jongen	Youth	0.607
Stybar	Stybar	0.624	Autochtonen	Natives	0.605
Luxemburgers	Luxembourgers	0.623	Voetbal-supporters	Football supporters	0.604
Brazilianen	Brazilians	0.620	Imams	Imams	0.604
Raborenners	Cyclist with team Rabobank	0.619	Hooligans	**Hooligans**	0.603
Topsprinters	Top sprinters	0.617	Skinheads	Skinheads	0.600
Spanjaard	Spaniard	0.617	Asielzoeker	Asylum seeker	0.600
Scandinaviers	Scandinavians	0.616	Raddraaiers	**Hoodlums**	0.598
Cancellara	Cancellara	0.615	Meisjes	Girls	0.595
Debutant	Debutant	0.615	Extremist	**Extremist**	0.594
Slovenen	Slovenes	0.614	Medelanders	Fellow citizens	0.594
Kopmannen	Leaders	0.614	Buitenlander	Outlander	0.593
Argentijnen	Argentinians	0.613	Immigranten	Immigrants	0.593
Kazak	Kazak	0.612	Probleem-jongeren	**Problem youth**	0.593
Brabanders	Residents of Brabant	0.611	Autochtone	Autochthonous	0.592
Zabel	Zabel	0.610	Joden	Jews	0.590
Profs	Pros	0.610	Rijksgenoten	Nationals	0.590
Limburger	Limburger	0.609	Iranier	Iranian	0.589
IJslander	Icelander	0.607	Rotjongens	**Creeps**	0.589
Renner	Cyclist	0.607	Hangjeugd	**Loiterers**	0.588
Veldrijder	Cyclocrosser	0.607	Afkomst	Descent	0.588

Table A1 (continued)

Belgian + Belgians			Moroccan + Moroccans		
Dutch	**English translation**	**Cosine similarity**	**Dutch**	**English translation**	**Cosine similarity**
Topfavorieten	Top favorites	0.607	Relschoppers	**Rioters**	0.588
Ploegmaat	Team mate	0.605	Irakees	Iraqi	0.587
Kopman	Leader	0.604	Straatterreur	**Street terror**	0.585
Kazach	Kazakh	0.604	Jongeren	Youth	0.584
Zwitsers	Swiss	0.603	Jongetjes	Little boys	0.582
Bask	Basque	0.602	Djihadist	**Jihadi**	0.581
Tijdritspecialist	Time trial specialist	0.602	Crimineel	**Criminal**	0.581
Feyenoorders	Feyenoord	0.601	Veelpleger	**Frequent offender**	0.580
Invallers	Substitutes	0.601	Bulgaren	Bulgaria	0.578
Ieren	Irish	0.600	Algerijnen	Algerians	0.578
Klassements-renners	GC riders	0.600	Pooier	**Pimp**	0.576
Jeugd-internationals	Youth internationals	0.600	Molukker	Moluccan	0.575
Walen	Walloons	0.600	Bekeerling	Convert	0.575
Impe	Impe	0.599	Somaliers	Somalians	0.574
Toprenners	Top riders	0.598	Somalier	Somalian	0.573
Veldrijders	Cross cyclists	0.598	Nederlanders	Dutch	0.572
Amsterdammers	People from Amsterdam	0.597	Voetbalsupporter	Football sup-porter	0.572
Slowaken	Slovaks	0.596	Terreurverdachte	**Terror suspect**	0.572
Ritwinnaar	Stage winner	0.596	Jihadisten	**Jihadists**	0.571
Pool	Pool	0.596	Hooligan	**Hooligan**	0.569
Vedetten	Vedets	0.595	Vrouwenhande-laar	**Trafficker**	0.569
Tukkers	Tukkers	0.595	Migrant	Migrants	0.567
Verheyen	Verheyen	0.595	Criminelen	**Criminals**	0.567
Jonkies	Young ones	0.595	Immigrant	Immigrant	0.566
Tilburgers	People from Tilburg	0.594	Gastarbeider	Guest worker	0.566
Topschutter	Top scorer	0.594	Straatschoffies	**Urchin**	0.566
Kittel	Kittel	0.593	Jochies	Boys	0.566
Australier	Australian	0.593	Griek	Greek	0.566
Spurter	Sprinter	0.592	Kickbokser	Kickboxer	0.565
Hongaren	Hungary	0.592	Djihadisten	**Djihadists**	0.565
Boonen	Boonen	0.591	Metin	Metin	0.565
Devolder	Devolder	0.591	Randgroep-jongeren	**Marginal youth**	0.563
Voigt	Voigt	0.591	Onruststokers	**Nuisance**	0.562
Jalabert	Jalabert	0.591	Jeugdbendes	**Youth gangs**	0.562

Appendix B

List of negative words derived from the word embeddings

terreurverdachte, jihadgangers, racisten, oorlogsmisdadiger, crimineeltjes, homohaters, extremist, prostituee, hangjongere, randgroepjongeren, criminelen, delinquenten, politieman, misdadiger, misdadigers, bedelaars, pedofielen, reljongeren, moordenaars, straatterroristen, terrorist, relschoppers, probleemjongeren, hangjongeren, extremisten, rotjochies, djihadisten, boefjes, jeugdbendes, geweldplegers, jihadstrijders, jihadisten, drugsdealer, tasjesdief, politieagente, terreurcel, crimineel, oplichter, pooier, hooligan, jihadist, onruststokers, pedofiel, hangjeugd, gedetineerde, zwervers, djihadist, strijder, straatschoffies, drugshandelaren, oorlogsmisdadigers, rotjongens, prostituees, straatterreur, politieagent, raddraaiers, maffiosi, bende, hooligans, lastpakken, vrouwenhandelaar, gangsters, veelpleger