Check for updates

Linguistic Radicalisation of Right-Wing and Salafi Jihadist Groups in Social Media: a Corpus-Driven Lexicometric Analysis

Pia Müller[1] · Stefan Harrendorf[1] · Antonia Mischler[2]

Abstract

Social media groups, for example on Facebook, WhatsApp or Telegram, allow for direct exchange, communication and interaction, as well as networking of different individuals worldwide. Such groups are also used to spread propaganda and thus allow for self-radicalisation or mutual radicalisation of their members. The article reports selected results from a research project analysing online communication processes of extremist groups. Based on data from group discussions in social media, corpus linguistic analyses were carried out, examining quantitative relationships between individual lexical elements and occurring regularities. To this end, four different corpora were built. These consist of data collected in right-wing and Salafi jihadist groups of a low or medium radicalisation level on Facebook and VKontakte via fake profiles, and of group communication in forums, messenger apps and social networks of highly radicalised persons, which were extracted from files of (e.g. terrorism) cases prosecuted in Germany. Quantitative linguistic analyses of social media data continue to be challenging due to the heterogeneity of the data as well as orthographic and grammatical errors. Nevertheless, it was possible to identify phenomenon specific sociolects that point to different levels of linguistic radicalisation. Based on the results of the analyses, the article discusses the prospects, problems and pitfalls of lexicometric analyses of online communication, especially as a tool for understanding radicalisation processes.

Keywords Right-wing extremism · Salafi jihadism · Radicalisation · Online communication · Social media · Corpus linguistics · Lexicometric analyses

✉ Stefan Harrendorf
   stefan.harrendorf@uni-greifswald.de

[1] University of Greifswald, Domstr. 20, 17489 Greifswald, Germany

[2] Kriminologische Zentralstelle, Wiesbaden, Germany

Introduction

Extremist groups use language strategically to infiltrate social discourses, trying to normalise their ideologies for potential new followers, while allowing for continued radicalisation of the previous ones (e.g. Struck et al., 2020). Combined with the qualities and characteristics of online communication, this turns digital social spaces into stomping grounds and strategic control centres for extremists of all kinds.

In the age of social media, potentially everyone—including propagandists of ideologies of inequality—has the chance to participate in online discourses. Social media platforms are easily accessible, generate a wide reach and enable exchange across all borders. This makes them so interesting for political extremist groups who use platforms like Facebook, ВКонтакте (VKontakte) or Telegram for advertising their own cause, for recruiting new members, for networking and financing opportunities or even for strategic discourse participation and shifting the boundaries of what can be said (Guhl & Davey, 2020; Harrendorf et al., 2019; Winter et al., 2020). Krzyżanowski and Ledin (2017) therefore regard social media spaces and discourses as a link between a "civil" discourse for society as a whole and an "uncivil" niche discourse in which extremist groups have so far remained in "media isolation" (Freter & Zimpelmann, 2015).

The discourses taking place here are not only reflections of a social reality, but they are equally involved in its construction—they do not only depict reality, but also constitute reality itself. This has consequences for the perception of the actors and viewers involved in the discourse (Niehr & Böke, 2004). This is also illustrated by the results of qualitative content analyses that preceded this study and focussed on social media groups from the fields of right-wing extremism and Salafi jihadism. Followers in both areas communicatively construct a social reality in which the respective idealised in-group is (putatively) exposed to an existential threat by various out-groups and social processes. As a consequence, some radicalised groups even see themselves forced to take (counter-)action by vigilantism or terrorist attacks (Mischler et al., 2019). This article focuses on the linguistic level of extremist online communication in the fields of right-wing extremism and Salafi jihadism. We pose the question of how the language used in radicalised or sympathising social media groups differs from a wider social use of language. Thereby, we aim to identify the linguistic characteristics (keywords and key terms) of online communication in the two phenomenon areas.

First, we present our basic theoretical assumptions on the effects of language and its role in radicalisation processes, also considering the characteristics of online communication. We then explain the quantitative methods utilised to study language use patterns in general and the corpus linguistic procedure applied here in particular (which also contains qualitative elements of content analysis). The analysed data originate from social media groups and were collected in the context of a research project on online radicalisation, focusing on the areas of right-wing extremism and Salafi jihadism. In the process, we encountered several methodological challenges which are also being discussed. This is followed by a presentation of the results of our study: Our present corpora from both phenomenon areas were referenced against broader web-based text corpora. Analysis reveals keywords and terms that are characteristic of the language used. They represent elements of a separate sociolect and indicate different levels of linguistic radicalisation. Based on the results of the analyses, the article discusses the prospects, problems and pitfalls of corpus linguistic analyses of social media data, especially as a tool for understanding radicalisation processes.

The considerations and analyses presented here derive from the project "Qualitative und quantitative Analyse internetbasierter Propaganda" ("Qualitative and quantitative analysis of Internet-based propaganda"), a sub-project of the research network "Radikalisierung im digitalen Zeitalter" ("Radicalisation in the digital age"; RadigZ). The project was funded by the German Federal Ministry of Education and Research from 2017 to 2020. For details on the overall project structure and aims, see Schröder et al. (2020) and Kudlacek et al. (2017).

Theoretical Framework: Radicalisation, Ideologies and Online Communication

Radicalisation is understood here as a processual development towards extremism (Beelmann, 2019; Beelmann and Lehmann, 2020). We define the latter as follows: "Extremism [is] characterized by attitudes, value systems, and actions that are marked by a significant deviation from certain socio-political norms and actively strive to establish new norm systems" (Beelmann and Lehmann, 2020, p. 5*, translated by the authors*). We define extremism not in terms of existing systems but in terms of the basic values of universal human rights, the principle of democracy and the rule of law as reference point from which extremists deviate. Hereby, we do not define extremist action as crucial but rather the values on which attitudes and actions are based (Beelmann and Lehmann, 2020).

While social media on the one hand offer a favourable starting point for propagandists to spread their ideologies to wider circles, online radicalisation on the other is a result of the interaction of individuals with their specific interests, attitudes and characteristics with the situational conditions of the medium itself (Bock & Harrendorf, 2014; Mischler et al., 2019): Online communication under certain circumstances contributes to the building and perpetuation of social (group) identities (cf. Spears & Lea, 1994; Spears & Postmes, 2015) and further intensifies polarisation tendencies (see Lee, 2007).

The Favourable Conditions for Extremist Radicalisation in Social Media

Already a few years ago, Sageman (2008) has stressed the relevance of the Internet for radicalisation processes (in case of his study: towards Islamist terrorism), and while further studies have shown that online media will usually not cause radicalisation autonomously, it is more than plausible that online media discourses and discussions have a catalysing and facilitating effect on radicalisation processes (for a literature review, see Meleagrou-Hitchens & Kaderbhai, 2017).

This catalysing function also seems to be confirmed by recent right-wing terrorist attacks in Christchurch, Poway, El Paso, Halle (all in 2019) or Hanau (2020), which made the interplay between extremist communication and the conditions of online communication that promote radicalisation shockingly clear. A more rapid internationalisation of right-wing terrorism, driven forward by online communication, can be observed (Albrecht & Fielitz, 2020; Macklin, 2019), in which the perpetrators of the later terrorist attacks explicitly refer to those who came before them and especially to the terrorist attack of Anders Behring Breivik as a kind of "archetype" (Ayyadi, 2019; Köhler, 2019). Such attacks would not have been possible without online communication—as a platform for radicalisation or for sharing propaganda and live streams of the attacks and for the acquisition of weapons or construction plans for weapons (Harrendorf et al., 2020). For Salafi

jihadism, Sageman (2008) already stressed the relevance of the Internet for radicalisation processes, and evidence can also be found from terrorist attacks in, inter alia, Nizza (2016), Orlando (2016), Boston (2013) or at Frankfurt Airport (2011). As a further example, the so-called Islamic State (IS) in its heyday produced professionally made propaganda magazines (first Dabiq until 2016, then Rumiyah until 2017) and distributed them online (Fritzsche, 2019; Vergani & Bliuc, 2015).

The options for social connections online, most prominently on the social media channels of Instagram, Twitter, Facebook, Telegram, etc., are almost endless, yet the platforms foster the grouping of people by similarity of their ideas, beliefs or convictions. This is, firstly, due to effects of self-selection (Sageman, 2008), which are comparable to the situation offline (see Haslam & Reicher, 2007). Online, these effects are not only generated by the users themselves, but also enhanced by algorithms used by search engines and social media websites to suggest content to users which is similar to the content an individual has consumed before.[1] Secondly, this effect is amplified by confirmation bias, meaning that individuals tend to seek and interpret information confirming their already made preconceptions (Del Vicario et al., 2017; Zimmer et al., 2019), and homophily (Mahrt, 2019), which illustrates that individuals are more likely to surround themselves with other people who share the same or similar views as themselves. This can lead to the illusion of an overall societal acceptance of one's beliefs, reinforcing them and rendering them more or less self-evident for those active in such social groups (Sageman, 2008).[2] It is plausible that the relevance of the effect increases the more polarized, radical, odd or extraordinary the topical focus of a discussion is (Mahrt, 2019).

Furthermore, users can make use of active identity management, which allows for group- and situation-specific adaptations, reaching from complete anonymity over pseudonymity and adapted or modified identities to a more or less realistic self-disclosure (Döring, 2010, 2018). Since online groups also provide for and are organised around a salient social identity, communication in these environments fosters (further) depersonalisation (i.e. the shift of focus from personal identity to social identity) and thus an adaptation of the participants to social group norms and context norms, which, in case of radical or extremist groups, means radicalisation (Harrendorf et al., 2019, 2020). These effects, described in the Social Identity Model of Deindividuation Effects (SIDE) (Spears & Lea, 1994; Spears & Postmes, 2015), are in line with the Social Identity Approach (Tajfel & Turner, 1986; Turner et al., 1988) and, more specifically, the Self-Categorisation Theory (Turner et al., 1988). They have been confirmed in various experiments (Postmes & Spears, 1998; Postmes et al., 2001), and it is plausible to assume that they are also of importance for online radicalisation processes (just see Rieger et al., 2020). The Social Identity Approach also provides an explanation for the social-psychologically well-studied effect of group polarisation. It can be observed that groups whose views on a topic already tended in a certain direction prior to their joint discussion develop even more strongly towards the corresponding extreme

---

[1] The existence of this so-called filter bubble, a term coined by Pariser (2011), has, however, been called into question in the last years, e.g. by Bruns (2019). Although "Pariser's ideas make logical sense" (Chitra & Musco, 2019, p. 2), the results regarding its existence are mixed at best (just see Cardenal et al., 2019; Chitra & Musco, 2019; Moeller & Helberger, 2018; Zimmer et al., 2019).

[2] Such (algorithm-amplified) self-selection processes help building up echo chambers (Sunstein, 2001), in which the own opinions and ideas are shared by a large majority of people (doubting Bruns, 2019). Differently from filter bubbles, echo chambers are not mere results of technical processes and can also exist without any influence of pre-selecting algorithms, just due to user self-selection (Zimmer et al., 2019). Research results seem to confirm the existence of such rooms (Garimella et al., 2018; Mahrt, 2019).

afterwards (i.e. as a group average) (Myers, 2010). According to the Self-Categorisation Theory, this is due to the fact that individuals adapt to the opinion of a perceived prototypical group member and thereby differentiate themselves as much as possible from perceived out-groups (Turner et al., 1988). This illustrates how group processes also have a potential to promote radicalisation in digital social spaces. An approach that analyses the content of extremist communication must therefore examine the realities and identities constructed in such groups.

Language and Radicalisation: Extremist Online Discourses and Their Linguistic Features

According to discourse theoretical concepts, discourses not only depict assumed realities, but also have the effect of constituting meaning (Fariclough et al., 2011; Keller, 2011; Niehr & Böke, 2004). Herein lies both their socially constructive character and their potential for mutual radicalisation of group members. In order to be able to examine realities drawn in this way, it is necessary to find out what is claimed as "real" in the discourses (Keller, 2011, p. 72).

Both phenomena, right-wing extremism and Salafi jihadism, are essentially based on ideologies of inequality that allow for an exaggeratedly positive communicative and linguistic evaluation of the in-group and a harsh devaluation of out-groups (Harrendorf et al., 2019; Mischler et al., 2019). Ideologies are generally "[…] systems of shared beliefs, ideas, and symbols that help us make sense of the world around us" (Alvarez, 2008, p. 216). They facilitate the positioning of the self within a complex social world by offering orientation in the form of interpretation schemes that individuals can adopt (Hall, 1989), contain clear rules for action and support individuals in the categorisation of right and wrong, good and evil. By devaluing out-groups and attributing superior qualities to the in-group, ideologies of inequality offer their followers the opportunity to improve the social identity of their own group by means of "social creativity" (in the meaning of Social Identity Theory: Tajfel & Turner, 1986) and to present themselves as a superior collective in being and action. Such ideologies offer the ideal background for in-group-out-group categorisations, providing the necessary stereotypes (Staub, 2001). Moreover, as an instrument of social creativity, they enable the development of a positive social identity even when one's own living conditions are rather unfavourable (Staub, 2001).

Realities are constructed in corresponding social media groups, in which the in-group, stereotyped as an ideal, is exposed to a (assumed) massive threat from specific out-groups. Whereas in the field of Salafi jihadism, membership of the in-group is determined on the basis of an interpretation of religion (lived "correctly" in the sense of ideology) embodied in the "ummah", in the field of right-wing extremism, membership depends on ethnic and/or cultural background, the "Volksgemeinschaft" (national community) (see also Harrendorf et al., 2019). Both in-groups see themselves as victims of out-groups, which are construed as morally inferior opponents, against which they have to defend themselves in the form of revolting or vigilante terrorism (Bibbert et al., 2017; Mischler et al., 2019; Müller & Mischler, 2020).

Rafael and Ritzmann (2019, p. 13) divide the according extremist narratives into three "essential components of hate speech propaganda": Firstly, the narrative of victimisation, which secondly legitimises radical redemption, leading to, thirdly, identity-forming narratives that satisfy the individual need for significance—for example the narrative of the strong, warlike man who defends his in-group until death.

The underlying narratives involve dystopian ideas: in both areas the world views are based on apocalyptic scenarios, according to which the respective groups are threatened with destruction and extinction (Fielitz et al., 2018). Like a self-fulfilling prophecy, these views (seemingly) confirm the threat towards the in-group and justify the ideological mind-set. They are both cause and effect of living in an ideological environment that constantly confirms the ideology it produces (Fielitz et al., 2018). Through apocalyptic scenarios, they legitimise resistance as necessary and violence as self-defence.

The ultimate threat to the in-group, as imagined by the extreme right, is that the "own" people ("Volk") will ultimately be alienated and replaced by other groups of people. The underlying conspiracy narrative is the "great replacement" (Marcks & Pawelz, 2020; see, e.g. Meiering et al., 2020), which assumes that the "own" people would be gradually replaced by non-white and Muslim people, which would initially lead to a loss of identity and eventually to, as it is called in ideologised German, the "Volkstod" ("death of the people"). It is, however, assumed in right-wing ideology that the "great replacement" would not happen arbitrarily but would be controlled by Jewish people as part of an even greater plan to weaken Europe and North America (Davey & Ebner, 2019). Therefore, specially non-white, Muslim and Jewish people are marked as enemies and an ultimate threat to the existence of the in-group, giving the in-group legitimacy to use violence against them and to "defend" themselves.

A similar mechanism applies to Salafi jihadism (Harrendorf et al., 2019; Meiering et al., 2020; Mischler et al., 2019). Salafi jihadists use experiences of discrimination Muslims make in the diaspora and embed them in the context of a general oppression and threat to the in-group by "the West" and infidels (Mahood & Rane, 2017; Meiering et al., 2020). "The West" is often prominently embodied by the USA, which is reflected in anti-American narratives that overlap with anti-Semitic narratives due to the USA's relationship with Israel. In addition to the infidels in general and the USA, Jewish people and the Israeli state are also among the out-groups that are seen as threat to the in-group according to Salafi Jihadist ideology (Müller & Mischler, 2020).

Although both in-groups see their existence threatened, they construct themselves as superior in the case of resistance, which also goes hand in hand with constructs of martial masculinity (Meiering et al., 2020). When individuals are exposed to such agitation and hate speech and surround themselves with it, this has a decisive effect on their cognition and emotions: From a psychological perspective, Bilewicz and Soral (2020) conclude that social media use is significantly associated with being exposed to more hate speech or at least deviant language. This can encourage those who have already internalised prejudiced attitudes to express them publicly. Furthermore, emotions such as empathy towards other groups can gradually be replaced by intergroup contempt. This contempt is then both motivation and consequence of hate speech (Bilewicz & Soral, 2020). In the worst case, the radicalisation process, also catalysed by online communication, will eventually lead to extremist, violent actions the ideologies seem to prescribe.

Communicative radicalisation in niche discourses of extremist groups online can therefore be seen as one important factor in radicalisation processes, although it is neither a necessary nor a sufficient condition for violent extremism and hate speech.

An extremist discourse is necessarily also reflected in the linguistic specifics of the respective texts, representing a sociolect[3] group members in both areas share with each other. Assuming that both ideologies ignite on the conflict line between in-groups and

---

[3] A sociolect describes similarities of language use in groups. Members of one group express themselves more similarly than members of different groups; see Louwerse (2004). A sociolect is therefore group-specific, and a group can adopt it as identity-forming (Durrell, 2004). In our case, for example, scene-specific expressions also serve to pick up and strengthen the group identity.

out-groups, it is reasonable to assume that such conflict lines as well as the description of in-groups and out-groups are reflected in communication.

In their keyword analysis, Chelvachandran and Jahankhani (2019) focus on known extremist Salafi users on Twitter and compare their tweets with also highly radicalised propaganda magazines by the IS. With regard to the thematic orientation of the datasets, they conclude that Twitter communication refers more often to current events, while IS magazines focus on an ideologised construct of religion. The authors assume that the frequencies of the individual keywords provide further explanatory potential and suggest keyword in context analyses as further research avenues.

Cohen et al. (2014) examined traces left by "lone wolf terrorists" on the Internet prior to their acts in order to identify signals that could potentially be used to prevent such crimes. They identified several "linguistic markers for radical violence" (Cohen et al., 2014, p. 253). In addition to "leakage", an actual suggestion or announcement of a crime, they also examined markers for "fixation" and "identification". A fixation on a topic or a person is indicated when users write about them significantly more often, use certain according keywords more frequently or gather facts excessively. Identification, on the other hand, is indicated when users describe the in-group with positive adjectives while describing perceived out-groups with negatives ones. Furthermore, activists expressed more anger and grievance when something negative happened to the in-group and more joy in occasion of positive events (Cohen et al., 2014).

Examining extremist Salafi discourse patterns on Facebook, Buoko et al. (2021) show that posts tend to address stories of oppression. At the same time, users focused more on out-groups also in terms of nomination and predication, describing them as "shirk", "kuffar" (both Arabic for "non-believers"), "non-believers" (written in French, language of examination) or "immoral" (Buoko et al., 2021, p. 11).

This prominent "negative othering" of out-groups is also in line with the results by Baker and Vessey (2018), who compare English and French propaganda magazines of the so-called IS and al-Qaeda. The authors conducted a quantitative corpus linguistic approach combined with a qualitative discourse analysis similar to ours. Besides a focus on out-groups perceived as non-believers, they identify Allah as a central category. For the latter, the writers of the magazines claim to act on his behalf knowing what he wants (Baker & Vessey, 2018, p. 264).

Finally, Schwarz-Friesel (2019) carried out a corpus linguistic study on anti-Semitism on the Internet, based on sub-corpora extracted from commentaries on news webpages, Facebook, Twitter, YouTube, blogs, forums, etc. She identified an increase of anti-Semitism in social media between 2007 and 2018 and also found a semantic radicalisation of the discourse.

The results of these studies support the assumption that radicalisation is reflected in online communication and can be identified in extremist posts on a linguistic level—be it with regard to the topics or in- and out-group references. Corpus linguistic analyses can therefore be used to identify indicators of linguistic radicalisation in communication data of extremist social media groups. The central question that motivates this paper is: what linguistic differences can be identified between communication in (extreme) right-wing and Salafi jihadist groups and public discourse? And, with regard to our data, that is consisting of self-collected, openly accessible social media groups and clandestine conversations and group chats of convicted radicalised individuals: will it be possible to identify differences between these two data pools, which confirm our assumption of different levels of radicalisation?

Methodology

General Overview on the Study Design

As mentioned in the "Introduction", the results presented and discussed here stem from the project "Qualitative und quantitative Analyse internetbasierter Propaganda" ("Qualitative and quantitative analysis of Internet-based propaganda"), a sub-project of the research network "Radikalisierung im digitalen Zeitalter" ("Radicalisation in the digital age"; RadigZ). In this project, we focused on discourses and discussions in German-speaking social media groups on Facebook, VKontakte, WhatsApp, etc. and in online forums. The groups and forums we focused on all had an orientation towards right-wing extremism or Salafi jihadism and can be formally divided into three different categories, based on how easy it is to access them. From this, we cautiously derive different levels of radicalisation, as we assume that the more the level of radicalisation increases, the more clandestine the group members communicate with each other and the higher the barriers to access are:

> an entrance level, on which interested users typically have their initial contacts with radicalising group processes and extremist discourses and materials in easily accessible social media groups
>
> a medium level, which includes groups for persons who have already progressed in their radicalisation process towards extremism but are still publicly accessible or only require a formal membership request, which is easily granted
>
> and an upper level of highly radicalised, "closed" groups of users who are on the verge of transferring their extremist ideas into actions (e.g. planning terrorist attacks) or have already committed such acts, only accessible after some kind of prior acquaintance and trust between the administrator and the new member

The data collection for the levels 1 and 2 was carried out with the help of fake user profiles on the social media platforms Facebook and VKontake in the period August 2017 to February 2019. These user profiles were adapted in a way that they would fit to persons who might be interested in getting into contact with Salafi jihadist or, respectively, right-wing extremist ideology. Since the access to groups in both fields of extremism is not gender-neutral (just see Pearson, 2018; Rippl & Seipel, 1999), female and male profiles were used. For ethical, legal and methodological reasons, the profiles were not used to directly participate in extremist discourse, but only to request access to groups and befriend other, fitting user profiles. To build the fake networks, we started with a keyword search, liked relevant pages such as those of far-right parties or Salafi jihadist preachers, joined groups that support them and sent friend requests to people who were also in the groups we found or who liked the same content. In addition, the platforms' algorithms suggested similar groups or pages to us and showed us other users interested in the same content. We included groups in our data collection if the communication corresponded to the respective ideology, for example if the users made a corresponding differentiation into in- and out-groups or discussed questions about how to live out the ideology correctly.

To collect comparable data, we selected eleven key events which were socially controversial and of potential ideological relevance for both extremisms. However, differently than expected, from Salafi jihadist groups, no relevant communication could be found for the observation point Xenophobic protests in Freital (2015) and only one thread about New Year's Eve in Cologne (2015/2016). A list of key events and the number of discussions,

words and tokens[4] analysed for groups on levels 1 and 2 can be found in Table 1. The key event "Berlin wears Kippa" is also analysed by Schwarz-Friesel (2019) in a corpus linguistic study on anti-Semitism on the Internet. She concludes that more than 15 per cent of the comments on the solidarity action "Berlin wears a kippa", collected from Facebook pages of three German newspapers, have anti-Semitic references. Here, classic anti-Semitism makes up the majority and is conjointly articulated with Israel-related anti-Semitism.

Regarding these data, it is impossible to know if the participants in group discussions have a consistent right-wing extremist or Salafi jihadist world view. However, it is possible to analyse the communication regarding patterns of interpretation typical for extremist ideologies. The identified patterns can be assigned to a predominantly low to medium level of radicalisation. Although highly radicalised statements were also made in these contexts sometimes, and it cannot be ruled out that highly radicalised people also communicate in some of the groups studied, the analysed communication mostly was more moderate than in groups for which data were collected from case files. For level 3 groups, on the other hand, the method of data collection guarantees that most discussants in these highly radicalised groups share an actual extremist identity.

Data for level 3 were collected retrospectively, relying on case files of the prosecutorial offices of the federal states and of the Public Prosecutor General at the Federal Court of Justice on terrorism, extremist violence and hate speech. The data obtained in this way originates from the period 2008 to 2016, with one exception for right-wing extremism, where the communication dates back to 2005. On the one hand, it includes smaller groups on messenger services such as Telegram, in which conversations from two up to 15 people were conducted. On the other hand, the data include communication from relevant forums, in which the number of readers cannot be determined, similar to the open material. Table 2 shows the number of cases, discussions, words and tokens analysed for level 3 groups. As the data collection for this level is based on case files, group discussions have not been selected based on key events. In contrast to the corpora built from "open" material, the size of the respective corpora shows a greater imbalance. Especially case no. 2 in right-wing extremism and case no. 7 in Salafi Jihadism show a high word count. Hence, the results will be particularly shaped by the communication in these two cases.

This study builds on previous qualitative content and discourse analyses carried out in the project RadigZ (see above). They focused on the identification of ideologised patterns of interpretation (or ideologised narratives) and different argumentation and communication strategies, also in comparison between the two extremisms (cf. Harrendorf et al., 2019; Mischler et al., 2019), which shows that both groups are united in their enmity against an enlightened and democratic society that values human rights. Regardless of their differences, in their core, they are based on group-focused enmity, aiming to degrade certain out-groups while intending to improve the social identity of the in-group. An additional focus of qualitative analysis was laid on memes as a special way to communicate extremist messages (cf. Harrendorf et al., 2020; Müller & Mischler, 2020). The memes vary on a spectrum from pop cultural references that appeal to newcomers to drastic depictions that are likely to entertain mainly already radicalised individuals.

This contribution applies quantitative lexicometric analysis to study the lexical level of the communication data, asking how the use of language in both phenomena differs from a broader social use of language and whether characteristic keywords and key terms can be identified at this level.

---

[4] Tokens are the smallest units a corpus consists of (Sketch Engine, 2020b). They represent word forms, punctuation, digits and abbreviations. Therefore, a corpus contains more tokens than words.

Table 1 Number of key events, analysed discussions, words and tokens (open material)

Social/political key event		Right-wing extremism			Salafi jihadism		
		Discussions	Words	Tokens	Discussions	Words	Tokens
Federal parliamentary election (2017)	1	19	19,045	258,84	23	107,675	140,225
Solidarity movement "Berlin wears kippa" (2018)	2	23	32,561	44,252	6	17,208	22,415
Terrorist vehicle-ramming attack in Berlin (2016)	3	29	34,128	46,382	25	17,284	22,514
Xenophobic protests in Freital (2015)	4	21	16,881	22,942	0	0	0
Trump and Jerusalem (2018)	5	18	14,269	19,393	15	14,603	19,022
Murder case "Susanna" (2018)	6	12	15,505	21,073	8	13,092	17,054
Vehicle-ramming attack in Münster (2018)	7	18	19,360	26,311	9	8,620	11,229
9/11	8	3	15,380	20,903	5	5,458	7,110
HoGeSa (Hooligans against Salafists) (2015)	9	9	15,175	20,624	3	2,174	2,833
Mass sexual assault on New Year's Eve in Cologne (2015/16)	10	14	12,381	16,827	1	4,161	5,420
Violent riots in Chemnitz (2018)	11	22	22,557	30,656	24	44,222	57,603
Total		**188**	**217,242**	**295,247**	**119**	**234,497**	**305,425**

Table 2 Number of cases, analysed discussions, words and tokens (closed material)

Case number	Right-wing extremism			Salafi jihadism		
	Discussions	Words	Tokens	Discussions	Words	Tokens
1	4	26,089	35,651	15	8,336	15,086
2	74	343,978	470,037	3	1,539	2,786
3	25	81,564	111,456	1	2,747	4,972
4	23	33,014	45,113	7	36,010	65,162
5				5	19,377	35,065
6				76	66,977	121,200
7				45	168,815	305,480
8				3	18,419	33,331
9				4	8,425	15,247
10				16	52,835	95,609
11				7	11,298	20,445
12				44	29,558	53,488
13				2	5,353	9,687
14				2	4,331	7,838
Total	**126**	**484,645**	**662,257**	**230**	**434,020**	**785,396**

Quantitative Corpus Linguistics and Lexicometric Analyses

Quantitative corpus linguistic methods allow the analysis of comprehensive text corpora regarding language usage patterns. The method focuses on the language used in large amounts of text, instead of individual texts, to make statistically noticeable structures visible (Baker et al., 2008; Bubenhofer, 2009; Kutter, 2018). Corpus linguistic researchers assume that these very structures can be used to identify the constitutions of meaning and significance within a given corpus (Dzudzek et al., 2009; Weber, 2015). According to Bubenhofer (2009), language usage patterns also provide information about the nature, functioning and limits of a discourse.

Corpus linguistic analyses are particularly suitable for empirically guided procedures. We decided for a corpus-driven approach in which the corpora are investigated inductively (Baker et al., 2008; Bubenhofer, 2009; Dzudzek et al., 2009; Weber, 2015). This means that the analysis is based on the available data, which determine the procedure to a certain

Table 3 Word counts of focus and reference corpora

Focus corpora	Word count	Tokens
rex_low-med	217,248	295,247
rex_high	484,648	662,257
sj_low-med	234,502	305,425
sj_high	434,028	785,396
Reference corpora	Word count	Tokens
Time Stamped German JSI Corpus 2014–2020 (05/10/2020)	5,863,062,257	7,011,918,484
German deTenTen 2013	16,526,335,416	19,808,173,163
German Web 2013 sample	193,838,751	231,458,549

extent. Bubenhofer (2009) even argues that in the case of large corpora, it is inevitable to approach the corpus without any fixed assumptions at first. This is the most likely way to discover linguistic structures with effects that are not obvious in discourse and would not have been expected beforehand. Dzudzek et al. (2009) also plead for such an approach to explore the differences and similarities of linguistic reference structures.

For the present study, an inductive corpus-driven approach makes sense for further reasons. Firstly, there is only few lexicometric research on extremist language in social media available so far (see the studies cited under the "Language and Radicalisation: Extremist Online Discourses and Their Linguistic Features" section). A theory-based analysis is therefore hardly possible, and, accordingly, this is an exploratory study. Rather, the aim of this work is to contribute to empirically based theory formation. But nonetheless, as shown above, this study is informed by a broader current state of research. Therefore, we expect to find specific in- and out-group-termini in our data. Secondly, our data show some hurdles that also call for an inductive approach: Grammar and spelling errors, data noise and difficult data quality made this study challenging (see also corpus formation in the "Method and Challenges of Keyword and Key Term Analysis" section).

Although a primarily inductive approach was chosen, deductive elements can certainly be used in the analysis process. If promising structures or keywords can be identified in the data, these can then form the basis for more in-depth qualitative investigations of the corpus combining them to wider themes and making sense of identified linguistic patterns (Baker et al., 2008).

In a first step, we identify keywords and key terms that are over-represented in the focus corpus compared to a reference corpus. "Keyness is defined as the statistically higher frequency of particular words or clusters in the corpus under analysis in comparison with another corpus" (Baker et al., 2008, p. 278; see also Culpeper & Demmen, 2015). Doing so, we are able to identify characteristics in our focus corpora that give us information about discussed topics and in- and out-group constructions that shape the extremist world view. For this, we combine elements of analyses of frequencies and co-occurrences, relying on the software Sketch Engine (see "Method and Challenges of Keyword and Key Term Analysis").

Second, we take a closer look at the identified keywords and terms in their contexts by analysing their concordances. These analyses focus on character strings before and after keywords or word sequences identified as relevant (Baker et al., 2008). They are therefore suitable for interpreting the context of the respective elements. In our study, they enable us to examine the embedding of specific terms in the text more closely, e.g. if keywords are used in ideologised contexts or if commentators distance themselves from them.

Method and Challenges of Keyword and Key Term Analysis

For our analyses, we used the corpus manager and text analysis software Sketch Engine and carried out keyword and key term analyses. While keywords represent single words, key terms consist of multi-word units, both occurring disproportionately often in the focus corpus, compared to a reference corpus. Keywords mainly consist of adjectives and nouns, as other word types are usually distributed similarly frequently between different corpora (Sketch Engine, 2020a). Key terms are adapted to the typical format of terms and expressions in the language under investigation on the basis of their syntactical phrase structure. In German, for example this means that the identified multi-word units often consist of adjectives and nouns or nouns and nouns (Sketch Engine, 2020a). For our analysis, we set the key term length to two words.

The keywords and terms are then ranked based on a score: The frequency per million (fpm) is calculated for each word. These values of a word from both corpora are then related to each other by dividing the fpm of the focus corpus by the fpm of the reference corpus (Kilgarriff, 2012). In this way, the keyness score is obtained, ranking the characteristic words of the focus corpus.

Based on the datasets described under the "General Overview on the Study Design" section, two corpora were created for each of the two phenomenon areas: a corpus based on the data at a formally low to medium level of radicalisation (levels 1 and 2; "rex_lowmed" for right-wing extremism and "sj_lowmed" for Salafi jihadism) and a corpus based on the data at a formally high level of radicalisation (level 3; "rex_high" and "sj_high").

As mentioned above, the data collected posed challenges in terms of both form and content. Formal challenges stem from the data source. The initial data from which the "lowmed" corpora were collated was collected in the form of PDF screenshots made with FireShot Pro, which were afterwards subject to OCR, using Adobe Acrobat Pro DC. The texts from the PDF were finally extracted as plain text (TXT) files and compiled to corpora which were then analysed using the Sketch Engine software.

While text recognition quality for these PDFs stemming from the own data collection files was, in general, good, there were still problems of "data noise", since OCR also recognises text from sidebars, banners, advertising, etc. and also often tries to read images as letters or signs. However, this could quite easily be identified later on and be removed. Bigger problems occurred with data extracted from case files: The chats, forums and social media group communications stored in these files sometimes had a very bad graphical quality (e.g. tiny, blurred or pixelated printouts or copies) and were thus not suitable for OCR at all or, although they still could be used, had a comparatively high error rate (wrongly identified letters and words); this led to a lower data quality of the "high" corpora, especially regarding "sj_high" (which is purely coincidental, since the quality of copies in case files depends on the respective investigators and prosecutors in charge). In terms of content, many user comments showed a high rate of typos, orthographical and grammatical errors or other forms of untypical use of grammar and language. The latter made more sophisticated linguistic analyses (which would also focus on syntactical phrase structure, etc.) impossible, but did not hinder the analyses presented here. Misspelled words remain as a big and not satisfyingly solvable problem. In many cases, such misspelled words cannot be identified as the words they meant to be by automatic word counting procedures. All result lists will show a systematically lower number of total word counts than manual counting would have identified. This issue is more relevant for the "high" corpora, but also needs to be taken into account for "lowmed".

Another problem resulted from Arabic words in the Salafi jihadist corpora. If these were written in Arabic letters, OCR did not recognise them. While most of the Arabic words in the corpora were used in a transliterated, Latin alphabet version, these words could not be understood by Sketch Engine, which often led to the software attributing to them the wrong part of speech, e.g. "verb" instead of "noun". Key term analysis therefore turned out to be impossible for the Salafi jihadist corpora. Insofar, we had to restrict ourselves to keyword analysis.

To work out the characteristics of a specific corpus, it needs to be compared with a reference corpus (see above). In our case, an ideal reference corpus would be one that contains social media communication over a similar period of time from different (non-extremist) groups. This would serve to map a broader social discourse and thus offer the possibility of investigating in which aspects our (extremist) discourses differ from it. In this way, characteristics of extremist language could most directly be explored. There are, however, numerous hurdles regarding a freely available social media corpus, especially concerning non-standardised and time-consuming data collection procedures (Mayr & Weller, 2017)

as well as copyright regulations. It was therefore not possible to compare our data to a "matching" social media reference corpus. Instead, we used the Time Stamped German JSI Corpus 2014–2020 as our preferred reference corpus. This corpus, developed at the Jožef Stefan Institute (JSI) at the University of Ljubljana, contains German RSS news feeds and is updated daily and fed into Sketch Engine (Bušta et al., 2017). Therefore, it should also contain data on the key events we have studied. It can be seen as a "public" discourse on the individual topics and many other issues of social and political relevance and thus reflects the broad societal debate. The JSI corpus was used for our keyword analyses, whereas, due to restrictions of the JSI dataset, key terms could only be explored by referencing them to the German Web 2013 sample, as a part of the German deTenTen 2013 corpus, which consists of crawled pages of German domains, Wikipedia and several forums from 2013. Table 3 gives an overview over the word counts of the focus and reference corpora used.

We chose a simple math value of +1 to emphasise terms that occur in the focus corpora and are rather rare in the reference corpora. With a higher value, more widespread terms would have been emphasised in the analysis. With a value below 1, terms that appeared even less frequently in the reference corpus would have been emphasised.

Results: Keywords and Terms of Right-Wing Extremist and Salafi Jihadist Social Media Discourses

In the following, our corpora are compared to the reference corpora to explore the characteristics of the respective extremist discourses. We focus on individual keywords or terms that seem particularly relevant with regard to their background or context. A complete list including translations and explanations can be found in (Tables 4, 5, 6, 7, 8 and 9). Starting with right-wing extremism, we first present results for keywords and terms on a formal low to medium level and then on a higher level of radicalisation. For this article, we have translated all keywords, terms and contextual texts into English, yet we additionally provide the original German and Arabic versions. In the word clouds, we use the English translations only for German words, while we kept the Arabic terms in their original language here, to make clear that they were also not German in the original texts. Translations from Arabic to both German and English were made relying on the expertise of Arabic-speaking research assistants.

Keywords and Terms in Right-Wing Extremist Corpora

Comparing the Right-Wing Extremist Corpus for Radicalisation Levels 1 and 2 (rex_lowmed) with the German JSI and deTenTen Corpora

Table 4 lists the first 50 keywords resulting from a referencing with the German JSI 2014–2020 corpus, sorted by their keyness score; Fig. 1 summarises these results more qualitatively using a word cloud in which the font size indicates the keyness of the words (i.e.: the larger the font, the higher the keyness). A few words reflect the thematic priorities in the data collection process, for example "kippa", "Freital" or "Breitscheidplatz". Others represent in- and out-group references that are not necessarily linked to specific key events and therefore seem to give insights into elements of a broader sociolect.

"Zionist" is listed as a disproportionately frequent word. It is mainly used in the key events of "Berlin wears kippa" and "Trump and Jerusalem", which, inter alia, refers to the opening of the US Embassy in Jerusalem, suggesting at least a thematic relevance. Further analysis

Fig. 1 Keywords in the right-wing extremist corpus for radicalisation levels 1 and 2 (rex_lowmed)

Fig. 2 Key terms in the right-wing extremist corpus for radicalisation levels 1 and 2 (rex_lowmed)

Table 4 Characteristic keywords in corpus "rex_lowmed" (minimum frequency: 30) in comparison with the reference corpus German JSI 2014–2020

Term	Translation or explanation	Score	Frequency	Reference frequency
01) Zionist	Zionist	140.060	59	3,023
02) Kippa	Kippah	123.800	76	7,574
03) ahuuu	Used as interjection	122.860	36	4
04) Freital	A German town that became known because of massive anti-refugee protests, right-wing extremist riots and a right-wing terrorist group	85.717	61	9,907
05) Moslem	Muslim	85.450	72	12,997
06) Gutmensch	Do-gooder	78.160	43	6,103
07) Antifa	Antifa	61.600	43	9,618
08) BRD	Federal Republic of Germany	52.510	51	16,081
09) Scheiß	Shit	49.600	35	9,823
10) widerlich	Disgusting	45.450	36	11,876
11) nix	Nothing	40.440	90	45,714
12) Salafisten	Salafists	40.140	44	19,070
13) scheißen	To shit	38.790	34	13,893
14) Breitscheidplatz	the site of the terrorist attack by Anis Amri in Berlin, Dec. 2016	38.780	32	12,683
15) Arsch	Arse	38.500	37	15,890
16) Jude	Jew	37.190	304	186,063
17) Idiot	Idiot	36.030	37	17,454
18) Araber	Arab	29.870	35	20,915
19) Hooligan	Hooligan	29.590	38	23,573
20) Pack	Riffraff	28.940	36	22,624
21) dumm	Stupid	26.580	101	82,945
22) Islam	Islam	22.470	129	128,793
23) hetzen	To agitate	21.800	31	26,911
24) Antisemitismus	Anti-Semitism	21.290	74	75,361
25) Demo	Protest	20.710	85	90,201
26) Israeli	Israeli	20.660	40	39,044

Table 4 (continued)

Term	Translation or explanation	Score	Frequency	Reference frequency
27) Palästinenser	Palestinian	20.450	66	69,524
28) Dreck	Dirt	20.010	35	34,650
29) blöd	Dumb	18.830	43	47,293
30) lächerlich	Ridiculous	17.840	36	41,027
31) lügen	To lie	16.960	37	44,930
32) UNO	UNO	16.560	36	44,760
33) Nazi	Nazi	16.090	72	99,024
34) Volk	The people	13.910	187	310,795
35) vergewaltigen	To rape	13.550	38	59,730
36) Jerusalem	Jerusalem	13.140	45	74,367
37) fressen	To devour	13.030	49	82,307
38) Chemnitz	A German city and site of (extreme) right-wing protests and attacks in Summer 2018	12.850	78	136,818
39) Propaganda	Propaganda	12.840	31	50,542
40) Mörder	Murderer	12.810	38	63,576
41) Merkel	Angela Merkel	10.790	361	783,131
42) Lüge	Lie	10.730	34	68,441
43) Israel	Israel	10.210	152	344,803
44) jüdisch	Jewish	8.920	95	245,131
45) Muslim	Muslim	8.430	60	161,862
46) Religion	Religion	8.270	55	150,818
47) DDR	German Democratic Republic	8.210	53	146,161
48) Lkw	Truck	7.880	77	224,446
49) Deutsche	Germans	7.690	222	674,777
50) Terrorist	Terrorist	7.570	45	134,174

reveals that "Zionists" are also discussed in the debate on the "Federal parliamentary election (2017)" and the "Terrorist vehicle-ramming attack in Berlin (2016)"—both contexts in which no thematic connection is obvious. Anti-Zionism propagates an anti-Israeli concept of the enemy and denies the state of Israel its right to exist and the right of national self-determination of Jews within this territory. The concept of anti-Zionism cannot be clearly distinguished from that of anti-Semitism. It is often mixed with "latent anti-Semitism" and could rather be defined as "Israel-related anti-Semitism" (Ranan, 2018, p. 38). Schwarz-Friesel and Reinharz even conclude that all anti-Zionist narratives under the guise of criticism of Israel rather serve to openly express anti-Semitic statements. Even if the focus is on the state of Israel, it ultimately serves as a symbol of Jewish life (Schwarz-Friesel & Reinharz, 2017).

The keyword in context analyses (KWIC) shows that users use the word in an anti-Semitic manner throughout. They use narratives of "leading Zionists" ("führende Zionisten") and characterise them as wanting to "sow war and discord among peoples" ("Kriege und Zwietracht unter den Völkern sähen").

The word "Jew" ("Jude") also shows a clearly negative usage. In the key event "Berlin wears kippa", the word appears most frequently for thematic reasons alone. In some cases, there are clear anti-Semitic references, but in some cases, in the context of this key event, solidarity is expressed, too, the latter, however, only with regard to victims of "Islamist violence" (because in that case, the event referred to an incident in which two men wearing kippas were attacked by a Muslim migrant and was meant to show solidarity with the victims and with Jews in Germany in general). For the key events that do not have a thematic reference to Judaism, such as 9/11 or "Terrorist vehicle-ramming attack in Berlin (2016)", the word is used almost exclusively in anti-Semitic contexts. The word "anti-Semitism" ("Antisemitismus") itself is also used disproportionately often and is almost exclusively linked to the key event "Berlin wears kippa". This is not, however, a response to anti-Semitic statements, but once again an expression of alleged solidarity with victims of anti-Semitism by perpetrators defined almost exclusively as Muslim, serving to reproduce anti-Muslim racism.

It is therefore not surprising that the word "Muslim" ("Moslem") is also considered a keyword, which represents out-group classification. "Islam" appears in nine of eleven key events—often without any obvious thematic reference to these key events. "Islam" is used to construct a threatening scenario, which is partly compared to National Socialism. In particular, the immigration of Muslim refugees is constructed as threatening, as if they would bring a civil war. Anti-Muslim racism is frequently present in this context. Anti-Muslim racist arguments could also be frequently identified in the communication regarding the key event "Hooligans against Salafists". The interjection "ahuuu" is closely linked to these racist arguments on that topic. With "ahuuu", battle cries were imitated at demonstrations and online. This is inspired by the movie "300", in which self-sacrificing Spartans fight against a Persian superiority. The film is popular in the right-wing scene, especially in the parts which refer to the Identitarian Movement (Identitäre Bewegung) (Vieregge, 2014). In rare cases, however, it is possible to identify counterarguments. But this does not seem to be characteristic for the use of the word in the groups under study.

Other out-groups (according to right-wing ideology) also turn out to be linguistically characteristic for the discourse: for example "Antifa", "Salafists" ("Salafisten") and "Arabs" ("Araber"). The word "Do-gooder" ("Gutmensch") is also significant, as it is most clearly derived from a right-wing vocabulary and refers to those people who, contrary to a (extremely) right-wing world view, stand up for a pluralistic society. There are numerous explanations about the term as a political slogan of the right in Germany (Auer, 2002; Hanisch & Jäger, 2011; Niehr & Reissen-Kosch, 2019). Right-wing extremists use these out-group descriptions as pejorative terms that can be modified almost inexhaustibly (see Scharloth, 2021).

Table 5 Characteristic key terms in corpus "rex_lowmed" (minimum (frequency: 10) in comparison with the reference corpus German web sample 2013

Term	Translation or explanation	Score	Frequency	Reference frequency
01) Frau Merkel	Ms Merkel	49.800	38	371
02) deutsches Opfer	German victim	31.360	10	26
03) Zentralrat der Juden	Central Council of Jews in Germany	31.190	15	153
04) eigenes Volk	Own people	25.210	13	182
05) tote Hose	Referring to "Die Toten Hosen", a famous German punk band	23.090	14	254
06) deutsche Frau	German woman	21.110	11	188
07) deutsches Volk	German people	16.890	20	711
08) normaler Mensch	Normal human being	14.270	10	334
09) Angela Merkel	Angela Merkel	10.500	21	1,359
10) eigene Meinung	Own opinion	8.890	10	677
11) jüdische Gemeinde	Jewish community	7.020	11	1,030
12) junger Mann	Young man	2.530	10	2,963
13) anderes Land	Other country	2.100	11	3,987
14) ganze Welt	Whole world	1.700	12	5,435
15) großer Teil	Big part	1.380	12	6,778
16) erste Linie	Referring to "in erster Linie", meaning "primarily"	1.090	11	7,885

Also, the word "riffraff" ("Pack") seems to be characteristic for the vocabulary in the communication processes. It has a clear negative meaning (similar to the English translation) in German language and, according to Duden (2022a), describes a group of people who are despised and rejected as "degenerate". The KWIC analyses show that "riffraff" is used in combination with any out-groups, be they refugees, politicians, democratic or left-wing people.

Contrary to the assumption in the user guide of Sketch Engine that corpora are rarely distinguished by their verbs, one verb on the list stands out: "to rape" ("vergewaltigen"). When viewed in context, it becomes clear that the narrative of German women being raped (in the corpus: exclusively) by Muslim men/refugees is used across seven of eleven key events.

The analysis of key terms (Table 5; also see Fig. 2) consisting of multi-word units results in a list of 16 terms, which is headed by the expression "Ms Merkel" ("Frau Merkel"). The name of the German chancellor "Ms Merkel" (whole name "Angela Merkel" is also listed) is generally used in a pejorative manner, as the KWIC analyses make clear. As previous qualitative analyses have also shown, in the groups analysed, she is blamed for all perceived grievances (Kopke, 2017; see also Küpper et al., 2016; Mischler et al., 2019). She is also massively devalued as a person. To cite two examples from right-wing oriented groups of levels 1 and 2: "Stupid bitch, this Ms Merkel", "The political MELTDOWN would be that in Germany!!!! a person invited by Ms Merkel and not entitled to asylum bestially raped a JEW!!!!".[5]

Another out-group name turns out to be "young man" ("junger Mann"). At first glance, this is a neutral term, but the KWIC analyses show that in the vast majority of cases, it is used to designate young male asylum seekers who would, according to the ideology shared in the groups, pose a threat, particularly to German women. The latter term (*"German woman" = "deutsche Frau"*) is thus also found in the list of corpus characteristics.

In contrast to the keywords, there are stronger references to the in-group in the list of key terms, with terms such as "German victim" ("deutsches Opfer"), "own people" ("eigenes Volk"), "German woman" ("deutsche Frau") and "German people" ("deutsches Volk") being used disproportionately often. It is noticeable here that the first three terms are all attributed to the in-group as victims, as KWIC analyses show. "German people", on the other hand, is also used in a mobilising way: "WE, the German people!!!!", "We will live, since the German people will defend itself".[6]

Comparing the Right-Wing Extremist Corpus for Radicalisation Level 3 (rex_high) with the German JSI and deTenTen Corpora

The list of keywords from the corpus with a formally high level of radicalisation (Table 6, Fig. 3) is clearly marked by vocabulary that is related to the National Socialist regime, its crimes and the Holocaust in the narrower and broader sense. The thematic orientation points in the direction of neo-Nazism as part of the extreme right-wing scene, whose followers identify with the ideology of National Socialism. The word "gas chamber" ("Gaskammer") is the most obvious, but also "Höß",[7] "concentration camp"

---

[5] These quotes are translated analogously, and capitalization and punctuation kept close to the original; all misspellings and grammatical errors in the German version are original, as well as capitalization and punctuation: "Dummes Weib, diese Frau Merkel", "Der politische SUPERGAU ist noch, dass in Deutschland!!!!, ein von Frau Merkel eingeladener, nicht asylberechtigter eine JÜDIN!!!! bestialisch vergewaltigt".

[6] See comment in Footnote 6. German version: "WIR, das Deutsche Volk !!!!", "Wir werden leben, denn das deutsche Volk wird sich zur Wehr setzen".

[7] Rudolf Höß (1901–1947) was, among other macrocriminal tasks for the Nazi regime, the commander of the extermination camp in Auschwitz from 1940 to 1943.

Table 6 Characteristic keywords in corpus "rex_high" (minimum frequency: 30) in comparison with the reference corpus German JSI 2014–2020

Term	Translation or explanation	Score	Frequency	Reference frequency
01) Gaskammer	Gas chamber	155.980	187	5,689
02) BRD	Federal Republic of Germany	126.270	276	16,081
03) Rennicke	Frank Rennicke, right-wing extremist singer-songwriter	118.020	80	223
04) Antifa	Antifa	71.460	112	9,618
06) Höß	Rudolf Höß, commander of the Extermination camp Auschwitz	55.460	71	6,625
05) Wiesel	Elie Wiesel, Holocaust survivor, publicist and lecturer	63.079	62	3,480
07) Pastörs	Udo Pastörs, a cadre of the right-wing extremist party NPD	54.740	40	847
08) Volksgenosse	Compatriot with same ethnical background	52.690	37	552
09) Auschwitz	Auschwitz	51.530	231	40,322
10) NPD	A right-wing extremist party	51.140	232	40,891
11) vergasen	To gas	48.000	44	2,822
12) Krematorium	Crematorium	47.890	65	7,456
13) Moslem	Muslim	47.240	89	12,997
14) Holocaust	Holocaust	47.060	219	42,134
15) Zionist	Zionist	47.030	44	3,023
16) Revisionist	Revisionist	45.060	31	425
17) Stauffenberg	Claus von Stauffenberg, partook in the assassination attempt on Hitler on 20 July 1944	42.080	47	4,947
18) Jude	Jew	41.230	756	186,063
19) Nationale	Nationalists	36.650	40	4,703
20) zionistisch	Zionistic	36.020	46	6,659
21) Neger	n-word	33.830	42	6,298
22) Gutmensch	Do-gooder	29.510	36	6,103
23) Buchenwald	Buchenwald	28.820	46	10,064
24) KZ	Short for: concentration camp	28.627	83	23,778
25) Hitler	Hitler	28.544	259	88,722
26) Himmler	Himmler	28.290	34	5,922

Table 6 (continued)

Term	Translation or explanation	Score	Frequency	Reference frequency
27) Henker	Executioner	27.840	32	5,375
28) Zyklon	Referring to Zyklon B, the gas used to kill in extermination camps	27.680	34	6,208
29) Nationalist	Nationalist	25.940	82	26,555
30) Nazi	Nazi	24.580	247	99,024
31) Vernichtungslager	Extermination camp	24.360	38	9,730
32) Scheiße	Shit	23.470	48	14,841
33) Burka	Burqa	21.810	36	10,717
34) Führer	A reference to Adolf Hitler	21.620	129	56,126
35) Massenmord	Mass atrocity	21.560	36	10,917
36) StGB	Criminal law	21.070	30	8,344
37) Adolf	Adolf Hitler	21.040	124	55,363
38) Alliierte	Allies	19.950	52	20,800
39) Konzentrationslager	Concentration camp	19.850	69	29,943
40) Arsch	Arse	19.630	42	15,890
41) Tagebuch	Diary	19.320	77	35,318
42) scheißen	To shit	18.990	37	13,893
43) Churchill	Churchill	18.700	31	10,845
44) Lüge	Lie	18.650	133	68,441
45) Stalin	Josef Stalin	18.240	44	18,790
46) Vernichtung	Extinction	18.120	53	24,193
47) Patriot	Patriot	18.080	38	15,525
48) Gruß	Greeting	17.940	92	47,362
49) Kamerad	Comrade	17.460	172	97,083
50) Wehrmacht	Wehrmacht	16.820	40	18,461

("Konzentrationslager") or "KZ" (abbreviation for "concentration camp"), "extermination camp" ("Vernichtungslager"), "Auschwitz", "to gas" ("vergasen"), "crematorium" ("Krematorium") and "Dachau"[8] are characteristic for the corpus.

A glance at the KWIC analyses of the word "gas chamber" reveals that those who communicate doubt their existence and functioning or even completely negate it, which is a common position of right-wing extremists, especially of those identifying with neo-Nazism (Benz, 2016; Weitzman, 2006). In the discussions, eyewitness statements are cited and reassessed, some by survivors (e.g. Elie Wiesel) and some by perpetrators. Users who take a stand against Holocaust denial cite, for example quotations from Höß, the commander of the Auschwitz extermination camp. These, in turn, are often negated by deniers, who claim that the statements had been forced under torture.

Other users conduct calculations about the size of gas chambers and the number of people murdered in them or assume that gas chambers were only built afterwards by German prisoners of war ("Sachsenhausen, another concentration camp in which after the war a gas chamber has been built by German prisoners; similar to Dachau").[9] Such discussions seem to correspond with Cohen et al.'s (2014) observation of a fixation on one set of issues as a marker for radicalised communication.

In a few cases, however, the general line is also called into question, and it is stated that it is now necessary to agree whether or not gas chambers in particular or the Holocaust in general existed. Similarities can be seen with the use of "concentration camp". While their existence is not widely questioned, their function is.

"Hitler" is not only used in this context but more general with regard to different aspects of the National Socialist regime. In many cases, veneration for him becomes apparent: "Of course, Hitler is not dead. The Führer will live in the hearts of the Germans forever", "Give Hitler in this country with its system built upon lies another hundred years. Then he will also be a folk hero".[10] His person is mentioned with positive connotations throughout the communications, clearly showing that these are discussions among severely radicalised right-wing extremists.

In addition to clear references to historical National Socialism, the corpus is also characterised by vocabulary that reveals communication on current issues of the extreme right (current in the sense of the times when the data material was seized by criminal justice agencies). For example, the operative and former chairman of the extreme right-wing "National Democratic Party of Germany" (Nationaldemokratische Partei Deutschlands; NPD) Udo Pastörs plays a role, as does the right-wing extremist singer-songwriter Frank Rennicke.

Out-groups are called "Antifa", "Do-gooder" ("Gutmensch") and "Jew" ("Jude"), as in rex_lowmed. The German version of the n-word ("Neger") is also used. As the previous remarks already suggest, people of Jewish faith are massively devalued. The KWIC analyses illustrate the entrenchment of profound anti-Semitism. "Do-gooder" points to a right-wing sociolect, as it is mainly used pejoratively by right-wing populists and extremists (although some conservatives might still use it, too). Similarly, the German n-word is nowadays used almost exclusively in right-wing populist and extremist context with an offensive and defaming meaning (detailed Technau, 2018).

Regarding the in-group, words such as "Volksgenosse" (which refers to a compatriot in an ethnocentric meaning), "Nationals" ("Nationale", meaning people with a nationalistic orientation in the right-wing extremist sociolect), "Nationalist" (similar to the latter), "Patriot" and

---

[8] Referring to the concentration camp in Dachau near Munich.

[9] See comment in Footnote 6. German version: "Sachsenhausen, noch ein KZ, in dem nach dem Krieg von deutschen Gefangenen eine Gaskammer gebaut worden ist; wie auch in Dachau".

[10] See comment in Footnote 6. German version: "Natürlich ist Hitler nicht tot. Der Führer lebt in des Deutschen Herz für immer fort", "Gebt Hitler in diesem Land mit seinem Lügensystem noch hundert Jahre. Dann wird auch er ein Volksheld sein".

Fig. 3 Keywords in the right-wing extremist corpus for radicalisation level 3 (rex_high)

"comrade" ("Kamerad") appear to be characteristic. While "Volksgenosse" was already used in German before the 1930s in an "empathetically exaggerated" equivalent of the German "Landsmann" (Schmitz-Berning, 2007, p. 661), i.e. "compatriot" in ethno-nationalist groups, it was especially in widespread use during the Nazi era. The term is used for members of an in-group defined by their shared biological ancestry and nationality (Schmitz-Berning, 2007). It illustrates the depiction of a racistly constructed in-group in terms of ideology. The discussants use this term to describe themselves or other group members. "Comrade" is taken up in relation to other members of the in-group as well and implies militaristic tendencies.

The list of key terms (Table 7; Fig. 4) more or less provides a thematic overview over the discussions.

As the keyword analyses already imply, the Holocaust, Jewish personalities and Judaism represent a large part of the communication that took place. In context, it becomes clear that the terms or topics of discussion also give clues to expressed anti-Semitism of any kind. The diary of Anne Frank is widely regarded as a forgery and "marketing strategy" in the "rex_high" corpus: "Certainly it is tragic that Anne Frank died in Bergen-Belsen due to typhoid. But the whole cult about her is typical Jewish style".[11] Elie Wiesel, Holocaust survivor and publicist, is characterised as a "consternation jew" (a made-up word from the sociolect; "Betroffenheitsjude"), "agitator and biggest liar" ("Hetzer und größter Lügner"). Accounts of the Holocaust are doubted or negated. Furthermore, the corpus is characterised by the expression "Zionist Regime" ("zionistisches Regime") as a description for Israel. Alongside the latent anti-Semitism inherent in the concept of anti-Zionism, the anti-Semitic meaning is already evident from the vocabulary, since "regime" has negative connotations in German usage (Duden, 2022b). The term thus directly devalues the idea of an independent Jewish state as illegitimate.

With "German people" ("deutsches Volk"), an in-group description ranks among those terms that can be seen as particularly characteristic for the corpus. When viewed in context, it

---

[11] See comment in Footnote 6. German version: "Gewiss ist es tragisch, daß Anne Frank in Bergen-Belsen an Typhus verstarb. Aber der ganze Kult um sie ist typisch jüdische Mache".

Table 7 Characteristic key terms in corpus "tex_high" (minimum frequency: 10) in comparison with the reference corpus German web sample 2013

Term	Translation or explanation	Score	Frequency	Reference frequency
01) Anne Frank	Jewish child and Nazi victim, known for her diary	61.750	98	240
02) Thor Steinar	Rght-wing extremist fashion lable	41.800	33	50
03) Elie Wiesel	Holocaust survivor, publicist and lecturer	40.830	28	14
04) deutsches Volk	The German people	39.560	106	711
05) Adolf Hitler	Adolf Hitler	24.514	46	331
06) nationaler Widerstand	National resistance	23.780	18	43
07) Holger Apfel	Former leader of the right-wing extremist party NPD	21.950	16	34
08) jüdische Lehre	Jewish doctrine	21.600	14	6
09) Otto Frank	Anne Frank's father	21.490	14	7
10) deutscher Boden	German soil	19.630	23	190
11) europäischer Jude	European Jew	19.330	17	88
12) deutsches Reich	German Reich	18.340	49	715
13) islamische Weltliga	Islamic world league	17.610	11	0
14) zionistisches Regime	Zionist regime	17.030	11	8
15) Darstellung des Holocaust	Representation of the Holocaust	15.970	10	2
16) Blaue Narzisse	A right-wing extremist magazine and website	15.910	10	3
17) Udo Voigt	Former leader of the right-wing extremist party NPD	15.540	11	31
18) nationales Lager	Nationalist camp	15.510	10	9
19) jüdisches Opfer	Jewish victim	15.420	11	33
20) bedingungslose Kapitulation	Unconditional capitulation	14.480	10	26
21) Million Jude	Million Jews	13.890	10	37
22) deutsche Frau	German woman	12.220	14	188
23) deutsches Kind	German child	11.280	12	161
24) politischer Gegner	Political opponent	10.960	15	268
25) eigenes Volk	Own people	9.860	11	182
26) rote Armee	Red Army	9.810	14	291
27) drittes Reich	Third Reich	9.300	28	846

Table 7 (continued)

Term	Translation or explanation	Score	Frequency	Reference frequency
28) jüdischer Glaube	Jewish faith	9.260	10	171
29) deutscher Soldat	German soldier	8.610	16	445
30) jüdisches Volk	Jewish people	7.690	11	299
31) ganzes Volk	The whole people	6.610	10	332
32) weißes Haus	White House	6.220	12	480
33) deutsche Stadt	German city	5.970	18	862
34) deutsche Bevölkerung	German population	5.650	10	428
35) alter Mann	Old man	5.170	17	963
36) andere Meinung	Different opinion	4.980	11	587
37) deutsche Geschichte	German history	4.600	11	654
38) eigenes Land	Own country	3.470	11	943
39) damalige Zeit	At that time	3.450	10	850
40) behindertes Kind	Handicapped child	3.410	10	861
41) Zweiter Weltkrieg	World War II	3.130	33	3,528
42) heutiger Tag	Present day	2.470	13	1,703
43) deutsche Sprache	German language	2.420	18	2,464
44) junger Mann	Young man	2.150	19	2,963
45) Erster Weltkrieg	World War I	2.010	10	1,626
46) gute Idee	Good idea	1.930	15	2,612
47) anderes Land	Another country	1.710	20	3,987
48) ganze Welt	Whole world	1.640	26	5,435
49) heutige Zeit	Present time	1.540	13	2,877
50) junge Frau	Young woman	1.070	10	3,258

Fig. 4 Key terms in the right-wing extremist corpus for radicalisation level 3 (rex_high)

becomes clear that those who communicate, in accordance with right-wing extremist ideology, construct the "people" on the basis of racist aspects and at the same time see it as massively threatened: "vilest multi-cultural destruction of the ethnic German people", "In the whole world a campaign of horrific lies and hate propaganda is led against Germany and the German people".[12] The same applies to the term "own people" ("eigenes Volk"). As already in the corpus rex_lowmed, these terms and similar terms, such as "German woman" ("deutsche Frau") or "German child" ("deutsches Kind"), are almost exclusively attributed to the in-group, presenting it as threatened and depicting it as victimised. The term "German soldier" ("deutscher Soldat") is also partly used in contexts of victimisation after the Second World War; partly revered, as German soldiers who "bravely defended their homeland"[13] and partly seen as a potential to represent the interests of the in-group in the future: "What would happen if German police officers and German soldiers would join the side of the German people?".[14]

---

[12] See comment in Footnote 6. German version: "übelste multi-kulturelle Zerstörung des ethnisch deutschen Volkes"; "Gegen Deutschland und das deutsche Volk wird in der ganzen Welt ein Greuellügen und Haßpropagandafeldzug geführt".

[13] See comment in Footnote 6. German version: "tapfer ihre Heimat verteidigten".

[14] See comment in Footnote 6. German version: "Was würde denn passieren, wenn deutsche Polizisten und deutsche Soldaten sich auf die Seite des deutschen Volkes stellen würden?".

There is no doubt that the corpus "rex_high" reflects a higher level of radicalisation, not only on a formal but also on a linguistic level, as can be seen, for example, in out-group terms such as the N-word. Thematically, the group members deal here more intensively with aspects that shape ideology, pointing into a similar direction as the results of Chelvachandran and Jahankhani (2019) for Salafi jihadism. This is made clear by the thematic references to the Holocaust, which is sometimes denied entirely, sometimes downplayed, and a stronger focus on different aspects of the Nazi regime. All of this confirms that indeed the "rex_high" corpus collates highly radicalised right-wing extremist communication.

The keyword analyses complement our previous qualitative content analyses (Harrendorf et al., 2019; Mischler et al., 2019). The words and concepts characteristic for the corpora provide an insight into the extreme right-wing ideology: from the construction of an ethnically homogenous and threatened in-group to massive devaluations of out-groups typical for right-wing ideology.

Keywords in Salafi Jihadist Corpora

We already explained under the "Method and Challenges of Keyword and Key Term Analysis" section why lexicometric analyses with Salafi jihadist data turned out to be complex and problematic. As key term analysis in Sketch Engine depends on the correct identification of the part of speech, it had to be skipped here, and we can only present the results of keyword analysis.

Comparing the Salafi Jihadist Corpus for Radicalisation Levels 1 and 2 (sj_lowmed) with the German JSI Corpus

Table 8 and Fig. 5 list the first 50 keywords resulting from a referencing with the German JSI 2014–2020 corpus. Even at first glance it becomes obvious that these words are much less connected to the specific key events (see Table 1) but clearly reflect a German-Arabic religious discourse. Some of the keywords also show a clear connection to Salafi jihadist ideology. Especially for religious terms, it is necessary to analyse keywords in context. This is the only possibility to identify whether a word is used in line with its common religious meaning or with a Salafi jihadist notion.

The most typical word for the corpus is "Allah". The word also appears as "allahu", the latter, for example, referring to the well-known religious phrase "allahu akbar" ("God is great"). "Allah" is used in all communications for all key events. Typically, "Allah" is referred to for protection, like in "May Allah protect you!" ("Möge Allah euch beschützen!"), which is a common religious phrase and reflects a general religious discourse. The abbreviation "swt" after the name of Allah means "subhanahu wa taala", which translates as "glorious and exalted is He" and can also be found on the list of keywords; it is used for reverence.

The keyword "Allah" is more frequent regarding three key events: "Federal parliamentary election", "Trump and Jerusalem" and "violent riots in Chemnitz". Here, KWIC analysis shows a different context in which the name is used, which points towards Salafi jihadist ideology. In connection with the Federal parliamentary election 2017, it is discussed whether "true" Muslims are allowed to vote, since democracy is not the reign of Allah. Anti-democratic comments are very common in the analysed threads, for example: "this [i.e.: lawmaking] is ONLY THE RIGHT OF ALLAH AND NOT THAT OF HUMANS ALLAH SETS THE LAWS NOBODY ELSE YOU DROWN IN SHIRK AND YOU

Table 8 Characteristic keywords in corpus "sj_lowmed" (minimum frequency: 30) in comparison with the reference corpus German JSI 2014–2020

Term	Translation or explanation	Score	Frequency	Reference frequency
01) Allah	Allah	883.670	960	17,937
02) Kuffar	Unbeliever	281.278	92	522
03) Shirk	Idolatry; polytheism	250.460	77	73
04) Gelehrte	Scholar	209.380	183	13,087
05) Akhi	Brother	183.810	56	20
06) swt	Abbreviation of Subhanahu wa-ta'ala = glorious and exalted is He	180.450	59	529
07) kufr	Unbelief; paganism	165.665	51	98
08) Taghut	Idolatry	154.760	47	5
09) Fatwa	Legal opinion	141.870	56	2,085
10) Koran / Quran	Quran	121.963	150	21,281
11) Muslim / Moslem	Muslim	129.658	1,027	174,889
12) Sunnah	Traditional customs and practices	121.450	37	39
13) Sure	Sura	119.700	55	3,571
14) Ummah	Community	119.360	37	162
15) Zionist	Zionist	98.870	43	3,023
16) islamfeindlich	Anti-Islamic	92.080	41	3,264
17) Islam	Islam	91.100	542	128,793
18) Haram	Forbidden	79.600	100	21,773
19) Scharia	Sharia	79.150	52	8,104
20) Prophet	Prophet	78.840	106	23,785
21) Salam	Peace (used as greeting)	77.940	31	2,194
22) Heuchler	Hypocrite	72.970	30	2,506
23) Gesandte	Envoy	63.750	35	5,664
24) amin	Amen	61.024	53	12,959
25) Vers	Verse	56.920	71	21,605
26) Imam	Imam	54.770	60	18,145
27) Ungläubige	Unbelievers	53.870	34	7,558
28) muslimisch	Muslim (adj.)	37.860	198	112,496
29) Kopftuch	Headscarf	37.760	79	40,928
30) Sünde	Sin	35.580	43	20,789
31) Ausländer	Foreigner	30.680	174	122,595
32) Araber	Arab	29.700	36	20,915
33) hetzen	To incite	28.480	42	26,911
34) ungültig	Invalid	28.250	42	27,189
35) Religion	Religion	26.350	182	150,818
36) Palästina	Palestine	25.470	34	23,753
37) benehmen	Behaviour	24.390	39	29,782
38) islamisch	Islamic	24.140	254	233,334
39) Nazi	Nazi	23.600	112	102,235
40) wählen	To vote	23.290	967	940,427
41) Gesetzgebung	Legislation	19.930	30	27,714
42) PKK	The Kurdistan Worker's Party	18.810	36	37,043

Table 8 (continued)

Term	Translation or explanation	Score	Frequency	Reference frequency
43) übel	Evil	17.360	78	95,884
44) hassen	To hate	16.560	39	47,151
45) au	Used as interjection	15.960	39	49,191
46) dumm	Stupid	15.300	60	82,945
47) Jerusalem	Jerusalem	14.670	52	74,367
48) Moschee	Mosque	14.400	57	83,812
49) Geschwister	Siblings	13.630	54	83,874
50) vergewaltigen	To rape	13.100	38	59,730

Fig. 5 Keywords in the Salafi jihadist corpus for radicalisation levels 1 and 2 (sj_lowmed) (Arabic terms were not translated in the word clouds. When we identified different spellings for individual terms in the results, we summarised them in the most common spelling.)

DON'T EVEN CARE!!"[15] To vote in a democratic system is constructed as a sin, as a betrayal of Allah. Yet, some discussants also oppose, especially with a view to the increasing support for the right-wing populist party "Alternative for Germany" ("Alternative für Deutschland"; AfD), which is anti-Muslim racist and reminds some of the discussants of the Nazi party, the NSDAP. They feel that in such a state of emergency, it is religiously acceptable to vote to weaken the AfD. Also in this context, Allah is called upon for everybody regardless their religion: "May ALLAH protect us and the non-Muslims".[16]

Allah is also called upon in other instances as an acting entity to destroy certain outgroups, for example regarding the opening of the US embassy in Jerusalem: "May Allah destroy these Zionists".[17] Similar to right-wing extremism, anti-Zionist statements in Salafi jihadism also stand for an at least latent anti-Semitism (Ranan, 2018). Anti-Zionist narratives are used with reference to Palestinians and their homeland. In this context, Palestine is sometimes also identified as Arabic, Muslim territory which should, from the in-group perspective, not be ruled by non-Muslims (Ranan, 2018). The word "Jew", that just missed the 50 most characteristic keywords, also appears in this context as an out-group term, and overt anti-Semitic views come to the fore. At other points, the in-group compares their perceived discrimination with the anti-Semitism experienced by Jews, including a historical perspective that relativizes the Holocaust by denying its singularity: "Yesterday the Jews, today the Muslims".[18]

Another relevant keyword is "shirk", which refers to idolatry or polytheism (Ebrem et al., 2014) and was also identified in the abovementioned study by Buoko et al. (2021). Shirk is not a sin that can be forgiven by Allah but an offence that excludes a person from Muslim community forever. It is quite rare in everyday use of the Arabic language and points towards a religious, also Salafi jihadist context. This becomes, for example, clear for the quotation cited above, which was made in connection with the Federal parliamentary election and in which voting was seen as "shirk". Although nowadays a few Salafi parties exist in some democracies, much similar to right-wing extremist parties, they only use the system to further their own goals and wish to eventually overthrow it after winning the elections (Karagiannis, 2018). From an ideological standpoint, Salafists reject democracy (ibid.). As the citation shows, democratic elections are seen as a kind of idol to which people pray instead of praying to Allah. In our corpus, the word "shirk" is only used with reference to this key event.

The in-group is often called "ummah". Literally, this simply means "community", and from a religious perspective, it includes all Muslims (Dantschke, 2009). Yet, from a Salafi jihadist perspective, the "ummah" only consists of persons believing in the "true" Islam in a way that conforms with the ideology (Harrendorf et al., 2019; Mischler et al., 2019).

People outside the "ummah" are seen as unbelievers. These out-groups are named differently in the corpus, sometimes by the Arabic word for "unbeliever" in plural ("kuffar") or singular ("kafir"), sometimes in German as "Ungläubige". A similar use has been shown by Buoko et al. (2021). In the analysed ideological context, "kuffar" has a stigmatising meaning and refers to all who practice the "wrong" religion or are Muslim,

---

[15] See comment in Footnote 6. Translation: "dies ist NUR ALLAHS RECHT UND KEINES MENSCHEN ALLAH MACHT DIE GESETZE NIEMAND SONST IHR ERTRINKT SCHON IN SHIRK UND EUCH JUCKT ES NICHTMAL !!".

[16] See comment in Footnote 6. German version: "Möge ALLAH uns und auch die Nichtmuslime beschützen".

[17] See comment in Footnote 6. German version: "Möge Allah diese Zionisten vernichten".

[18] See comment in Footnote 6. German version: "Gestern die Juden, heute die Muslime".

but do not believe in the "true" Islam in the meaning of Salafi jihadism. Followers of Salafi jihadist ideology usually do not hold contact to "unbelievers" (Farschid, 2014) or those who commit "kufr" ("unbelief"; "paganism"). These terms can also mainly be identified in connection with the Federal parliamentary elections. In the same context, we also find "taghut" (idolatry; the worship of anything except Allah). The argumentation is similar: "Who chooses a legislator besides Allah commits Kufr",[19] or, according to other comments, "taghut".

"Sharia" (German transliteration: "Scharia") literally means "path to the watering place" but in religious context refers to a law given by God. Yet, as for any other law, there are different interpretations, some of which are opposed to or conflict with human rights (Ebrem et al., 2014). In our data, comments favour the sharia instead of democratic laws. Inter alia, there is some discussion about the question whether democracy might be used for a Salafist party to come to power and introduce the sharia (also see above); some even reject this. The word can also be found frequently in the discussion of the murder case "Susanna", where commentators call out for sharia law to deal with the case instead of German criminal law, which is seen to be too mild. As in the area of right-wing extremism, the verb "to rape" appears here, serving and discussing anti-Muslim narratives. In one case, one user in particular makes anti-Muslim racist comments; in other cases, the discussions take place in a group in which individuals from different faiths communicate.

Comparing the Salafi Jihadist Corpus for Radicalisation Level 3 (sj_high) with the German JSI Corpus

Similar to the open material, the name "Allah" is, once again, central (Table 9; Fig. 6). Apart from expressions like "inshalla" ("if God wills") or "mashalla" ("what God has willed"), which are very frequent, "Allah" is mainly called upon for protection, a result that was also found for "sj_lowmed". Many threads involve persons who have left Germany to join the IS and their contact persons. Their communication shows that they see their actions blessed or favoured by Allah. God also appears as an acting entity: "Allah will crush you", "Allah takes care of us".[20]

In these cases, Allah is seen as transcendence. Allah acts himself, through them. Their own fighting strength is given by the greatness and power of Allah (Ebrem et al., 2014). The transcendence of Allah is often associated with the exclamation "inshallah", an expression of the belief that nothing happens without God's will or God's approval and that God's will is above all. In other, non-extremist contexts, "inshallah" can also be translated as "hopefully", "perhaps" and "let's see", but this is not the typical meaning in our corpus.

"Kuffar" and "kafir" (the unbelievers, see above) are even more strongly rejected than in "sj_lowmed", culminating in homicidal ideation or even real homicide: "When Kuffar come … and I am sure that the explosion will get them, too, this is something wonderful for Allah. Because it is necessary to protect oneself and attack the enemy".[21]

"Jihad" is a keyword here, too. In general religious contexts, the word refers to the efforts on the path to God (Ebrem et al., 2014). In Salafi jihadism (and also in our corpus),

---

[19] See comment in Footnote 6. German version: "Wer einen Gesetzgeber neben Allah wählt, begeht Kufr".

[20] See comment in Footnote 6. German version: "allah macht dich platt"; "allah versorgt uns".

[21] See comment in Footnote 6. German version: "Wenn kuffar kommen … und ich sicher bin das ich sie mit in die luft fliegen lasse Ist es etwas wundervolles für allah. Denn man muss sich schützen und den feind angreifen".

Table 9 Characteristic keywords in corpus "sj_high" (minimum frequency: 30) in comparison with the reference corpus German JSI 2014–2020

Term	Translation or explanation	Score	Frequency	Reference frequency
01) Inshallah	If God wills	934.070	791	556
02) alaikum	Part of "as-salam alaikum", greeting; different spellings	690.862	586	571
03) Allah	Allah	669.811	1.871	17,937
04) salam	Part of "as-salam alaikum", greeting; different spellings	549.622	567	2,211
05) Kuffar	Unbeliever	408.579	344	522
06) Akhi	Brother	336.170	264	20
07) Alhamdulillah	Praise be to God	291.848	230	48
08) swt	Abbreviation of Subhanahu wa-ta'ala = glorious and exalted is He	269.560	227	529
09) Subhanallah	Glory to God	199.398	156	8
10) wallah	I swear by God	183.330	145	87
11) dawla	State (used for the IS)	119.729	94	56
12) Quran	Quran	119.090	95	168
13) Mashallah	What God has willed	115.956	91	55
14) barakatuh	Peace be upon you	88.854	69	0
15) Jihad	Struggle/effort, also Holy War	88.190	107	3,874
16) Baraka	Blessing	77.740	64	425
17) Mujahidin	Mujahideen	75.840	59	26
18) Ummah	Community	74.390	59	162
19) mom	Abbreviation of "moment"	70.710	86	3,920
20) Fitna	Temptation	70.300	55	72
21) omg	Abbreviation of "oh my god"	60.900	61	2,032
22) Dawa	Invitation	60.150	51	670
23) Hadith	Story (record of the traditions or sayings of Prophet Mohammed)	59.840	48	265
24) amin	Amen	58.220	130	12,959
25) fik	As part of "baraka allahu fik": "May God bless you!"	58.080	45	26
26) rahmatullahi	Mercy of God	57.020	44	0

Table 9 (continued)

Term	Translation or explanation	Score	Frequency	Reference frequency
27) rahmatu	God's grace is with you	57.010	44	1
28) Taghut	Idolatry	55.710	43	5
29) Schia	Shiite	54.830	44	278
30) bro	Abbreviation of "brother"	52.870	54	2,224
31) sach	Things	49.440	49	1,965
32) Prophet	Prophet	48.680	168	23,785
33) akbar	Greater	47.710	73	6,751
34) barakallahu	Part of "barakallahu feek/feeki!": "May god bless you!"	45.550	35	2
35) Brüder	Brothers	44.060	34	36
36) Schwestern	Sisters	43.830	34	73
37) bocken	"It's the business!"	43.360	45	2,400
38) Sunnah	Sunnah	42.780	33	39
39) Sheikh	Sheik	41.380	52	4,348
40) Qaida	Al Qaida	40.490	43	2,625
41) Mohammed	Mohammed	39.640	279	55,640
42) Bismillah	In the name of God	38.810	30	70
43) Imam	Imam	38.780	109	18,145
44) kheir	Good	38.690	30	92
45) Sham	The Levant	37.860	35	1,418
46) Amis	Abbreviation of "Americans")	37.640	34	1,231
47) Spaß	Fun	33.730	42	4,284
48) Menschen	Humans	32.480	30	1,441
49) boah	Used as interjection	30.720	33	2,787
50) Heuchler	Hypocrite	30.700	32	2,506

Fig. 6 Keywords in the Salafi jihadist corpus for radicalisation level 3 (sj_high) (Arabic terms were not translated in the word clouds. When we identified different spellings for individual terms in the results, we summarised them in the most common spelling.)

however, it means a *fight* on the path to God in order to extend the Salafi jihadist territory and spread "true" Islam. The word "jihad" refers here to a Holy War (Ebrem et al., 2014), aiming to eliminate the out-groups, which are seen as enemies: "The jihad simply is the cleansing of the people from fisq and kufr" (with "fisq" meaning "sinner"), "Jihad is when you bomb the Israeli elite".[22]

The keyword "Mujahidin" ("fighter") is derived from the Arabic term "jihad"; the word refers to a person who makes an effort, who, for example, sacrifices him- or herself for the family and takes care of them (Ebrem et al., 2014). From a Salafi jihadist perspective, a "mujahid" is someone who engages in the Holy War. Persons described as "mujahidin" are respected individuals from the in-group who are seen as role models: "Yet, the kuffar and Al Qaida are not alike: Mujahideen fight on the path of Allah and do not commit wrongdoing".[23] The example also shows the close connection to the Salafi jihadist concept

---

[22] See comment in Footnote 6. German version: "Der Jihad ist einfach die Reinigung des Volkes von Fisq und Kufr", "Jihad ist wenn du Israels Elite zerbombst".

[23] See comment in Footnote 6. German version: "Dennoch sind die Kuffar und Al Qaida nicht gleich: Mujahidin kämpfen auf dem Wege Allahs und begehen kein Unrecht".

of "jihad"—mujahideen are seen as warriors in the Holy War but also as respected heroes and rebels, who, if they die on the battlefield, are honoured as martyrs (Ebrem et al., 2014).

The term "dawla" or "daula" ("state") is also characteristic for the corpus. In the German texts we analysed, the use of this Arabic term was connected to the development of the IS dominion in Syria and Iraq and is used as a synonym for it. The word indicates the use of a radicalised language. The IS implemented a very strict sharia-based legal system, which was welcomed by the discussants in our "sj_high" corpus: "Because nobody intervenes with robbers and felonies as rigorously as dawla".[24]

There are significant differences between the corpus "sj_high" with a formally higher radicalisation level and the corpus "sj_lowmed". This becomes obvious if one analyses not only the keywords, but also the contexts in which they are used. For some of the words, the meaning and interpretation differ strongly, depending on these contexts. Some of the words are typical for Muslim religious practice. An extremist interpretation of such terms only becomes visible in KWIC analyses, which confirm a use in line with the Salafi jihadist ideology. The level of linguistic radicalisation that can be identified from such analyses is higher in "sj_high" than in "sj_lowmed". This is in line with expectations, since the "sj_high" corpus per definition and due to our methodology contains communication of more radicalised individuals. Yet, it is important to note that the results of this study show that this higher level of radicalisation is also reflected linguistically.

Discussion and Conclusions

The lexicometric analyses of corpora built from online communication of groups that to a different degree conform with Salafi jihadist or right-wing extremist ideology show linguistic characteristics of the texts not only compared to the general discourse on the German web, but also in comparison between the corpora with the same ideological background, yet with different levels of radicalisation. Some of these differences are also due to the fact that the "lowmed" corpora were built around certain key events, while the "high" corpora were not. However, the corpora reflect the dominating ideological constructs of ingroups and out-groups.

Even in the corpora reflecting a formally lower level of linguistic radicalisation, it was possible to identify keywords which are not only characteristic for the corpus, but also indicate a radicalised sociolect in line with the respective ideologies. An example is the term "do-gooder" ("Gutmensch") in right-wing extremist or the term "shirk" ("idolatry") in Salafi-jihadist communication. KWIC analyses confirm that the characteristic keywords are mostly embedded into an ideologised discourse.

The corpora on a formally higher level of radicalisation show even stronger influences of extremist ideology. Very radical political positions are communicated in these groups. Regarding the thematic landscape, in "rex_high" groups, for example the Holocaust is a central topic and often it is relativised or even denied. The "sj_high" corpus strongly focuses on topics that circulate around the "jihad" in the extremist meaning of a "Holy War". As said before, it is not surprising that the individuals communicating in the "high" groups show a higher level of radicalisation, since this was intended by our methodological approach (see above). Yet, comparison between groups from radicalisation levels 1 and 2 on the one hand and radicalisation level 3 on the other proves that

---

[24] See comment in Footnote 6. German version: "Denn keiner greift bei Räubern und verbrechen so hart ein wie dawla".

we can also trace this higher radicalisation level linguistically with lexicometric corpus analysis. For right-wing extremism, words like "Volksgenosse" (a compatriot with the same ethnical background), "Patriot" or "comrade" ("Kamerad") as well as the use of the German n-word can be seen as typical for an even more radicalised sociolect. The same applies to words like "dawla" (as a synonym for the IS) or "mujahidin" (for the warriors of the jihad) in the Salafi jihadist corpus. This result is supported by KWIC analyses which show a very strong devaluation of out-groups.

The specifics of the corpora allow to draw conclusions about the ideologised construction of in- and out-groups. In line with Niehr and Böke (2004), we assume that concepts of reality and ideological fragments identified from these group communications do not only reflect the respective linguistic community, but also contribute to the construction of a collective, ideologised reality. In line with assumptions of the Social Identity Approach and the SIDE model, such divergent constructions of the world in which we live can also be spread, adopted and fortified via processes of social identification, e.g. by new group members. An ideologised definition of in-group and out-group identities is clearly identifiable in the material. Also, it can be shown that the elements of the sociolects reflected in keywords and terms differ between groups with a formally higher and a formally lower level of radicalisation. Criticism of the respective ideologies is seldom found in the material.

This does, however, not mean that counter speech is absent from the material. Since the method used here focuses on linguistic differences between the extremist corpora and non-extremist reference corpora, it cannot be expected that it is sensitive to counter speech (which would not bear the "typical" linguistical specifics of the extremist corpora). Yet, in our qualitative analyses, we clearly identified some instances of counter speech. We also carried out a quantitative semantic network analysis, which sheds further light on different form of group interaction, inter alia by counter speech. The results of this analysis will be published in a further article that is yet to be submitted.

Both the specifics of online communication and data quality issues, for Salafi jihadism also the use of terms stemming from another language (Arabic), made the study methodologically challenging and restricted the possibilities for analysis. Lexicometric analyses are, however, excellent to structure great amounts of data and provide valuable insights into the topics discussed and, in our case, also the construction of in-groups and out-groups. They are useful as a first step of analysis that can be followed by a more in-depth approach based on qualitative content and discourse analysis or can complete, like in this study, qualitative analyses that have been carried out beforehand (see Harrendorf et al, 2019, 2020; Mischler et al., 2019; Müller & Mischler, 2020). For social media data, however, it should not be used as a standalone method, because it can only analyse plain texts, while we find there heterogeneous forms of communication, relying not only on texts, but combining it with images, videos or animated GIFs. This variety can, as far as we can see, only be addressed by relying on multiple methods, including both qualitative and quantitative approaches.

Declarations

Competing Interests The authors declare no competing interests.

References

Albrecht, S., & Fielitz, M. (2020). *Rechtsterrorismus im digitalen Zeitalter*. Schriftenreihe des Instituts für Demokratie und Zivilgesellschaft. Schwerpunkt: Rechtsterrorismus. Retrieved January 19, 2022, from https://www.idz-jena.de/fileadmin/user_upload/PDFS_WsD6/15_Albrecht_Fielitz.pdf

Alvarez, A. (2008). Destructive beliefs: Genocide and the role of ideology. In A. Smeulers & R. Haveman (Eds.), *Maastricht series in human rights: Vol. 7. Supranational Criminology: Towards a Criminology of International Crimes* (pp. 213–232). Intersentia.

Auer, K. (2002). „Political correctness" - Ideologischer Code, Feindbild und Stigmawort der Rechten. *Österreichische Zeitschrift Für Politikwissenschaft, 31*(3), 291–303.

Ayyadi, K. (2019). *Die „Gamification" des Terrors - Wenn Hass zu einem Spiel verkommt.* Amadeu Antonio Stiftung. Retrieved January 19, 2022, from https://www.belltower.news/antisemitische-tat-in-halle-die-gamification-des-terrors-wenn-hass-zu-einem-spiel-verkommt-91927/

Beelmann, A. (2019). Grundlagen eines entwicklungsorientierten Modells der Radikalisierung. In E. Marks (Ed.), Prävention & Demokratieförderung: Gutachten und Stellungnahmen zum 24. Deutschen Präventonstag (pp. 181-209). Forum Verlag.

Beelmann, A., & Lehmann, L. (2020). Radikalisierung im digitalen Zeitalter: Handlungsempfehlungen an Politik, Praxis und Gesellschaft - Kurzfassung. Hannover: Kriminologisches Forschungsinstitut Niedersachsen e.V. (KFN).

Baker, P., Gabrielatos, C., KhosraviNik, M., Krzyżanowski, M., McEnery, T., & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse and Society, 19*(3), 273–306.

Baker, P., & Vessey, R. (2018). A corpus-driven comparison of English and French Islamist extremist texts. *International Journal of Corpus Linguistics, 23*(3), 255–278.

Benz, W. (2016). Die Funktion von Holocaustleugnung und Geschichtsrevisionismus für die rechte Bewegung. In S. Braun, A. Geisler, & M. Gerster (Eds.), *Strategien der extremen Rechten: Hintergründe - Analysen - Antworten* (2nd ed., pp. 221–228). Springer VS.

Bibbert, M., Mischler, A., Geng, B., & Harrendorf, S. (2017). Vorüberlegungen zur Analyse von Radikalisierungsverläufen im Internet. *Neue Kriminalpolitik, 29*(4), 388–397. https://doi.org/10.5771/0934-9200-2017-4-388

Bilewicz, M., & Soral, W. (2020). Hate speech epidemic. The dynamic effects of derogatory language on intergroup relations and political radicalization. *Political Psychology, 41*, 3–33. https://doi.org/10.1111/pops.12670

Bock, S., & Harrendorf, S. (2014). Strafbarkeit und Strafwürdigkeit tatvorbereitender computervermittelter Kommunikaion. *ZStW - Zeitschrift Für Die Gesamte Strafrechtswissenschaft, 126*, 337–381.

Bruns, A. (2019). Filter bubble. *Internet Policy Review*, *8*(4). https://doi.org/10.14763/2019.4.1426

Bubenhofer, N. (2009). Korpuslinguistik als Methode der Diskurs- und Kulturanalyse. *Walter De Gruyter*. https://doi.org/10.1515/9783110215854

Buoko, C., Naderer, B., Rieger, D., van Ostaeyen, P., & Voué, P. (2021). Discourse patterns used by extremist Salafists on Facebook: Identifying potential triggers to cognitive biases in radicalized content. *Critical Discourse Studies*, 1–22.

Bušta, J., Herman, O., Jakubíček, M., Krek, S., & Novak, B. (2017). *JSI Newsfeed Corpus*. The 9th International Corpus Linguistics Conference, University of Birmingham. Retrieved January 19, 2022, from

https://www.birmingham.ac.uk/Documents/college-artslaw/corpus/conference-archives/2017/general/paper382.pdf

Cardenal, A. S., Aguilar-Paredes, C., Galais, C., & Pérez-Montoro, M. (2019). Digital technologies and selective exposure: How choice and filter bubbles shape news media exposure. *The International Journal of Press/politics, 24*(4), 465–486. https://doi.org/10.1177/1940161219862988

Chelvachandran, N., & Jahankhani, H. (2019). A study on keyword analytics as a precursor to machine learning to evaluate radicalisation on social media. In Institute of Electrical and Electronics Engineers (Ed.), *12th International Conference on Global Security, Safety and Sustainability (ICGS3).* IEEE.

Chitra, U., & Musco, C. (2019). *Understanding filter bubbles and polarization in social networks.* Retrieved January 19, 2022, from http://arxiv.org/pdf/1906.08772v1

Cohen, K., Johansson, F., Kaati, L., & Mork, J. C. (2014). Detecting lingustic markers for radical violence in social media. *Terrorism and Political Violence, 26*(1), 246–256.

Culpeper, J., & Demmen, J. (2015). Keywords. In D. Biber & R. Reppen (Eds.), *Cambridge Handbooks in Language and Linguistics. The Cambridge Handbook of English corpus linguistics* (1st ed., pp. 90–105). Cambridge Univ. Press.

Dantschke, C. (2009). Zwischen Feindbild und Partner: Die extreme Rechte und der Islamismus. In S. Braun, A. Geisler, & M. Gerster (Eds.), *Strategien der extremen Rechten: Hintergründe - Analysen - Antworten* (1st ed., pp. 440–460). Springer VS.

Davey, J., & Ebner, J. (2019). *'The great replacement': The violent consequences of mainstreamed extremism.* London. Retrieved January 19, 2022, from https://www.isdglobal.org/wp-content/uploads/2019/07/The-Great-Replacement-The-Violent-Consequences-of-Mainstreamed-Extremism-by-ISD.pdf.

Del Vicario, M., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2017). Modeling Confirmation Bias and Polarization. *Scientific Reports, 7,* 40391. https://doi.org/10.1038/srep40391

Döring, N. (2010). Sozialkontakte online. In W. Schweiger & K. Beck (Eds.), *Handbuch Online-Kommunikation* (pp. 159–183). VS Verl. für Sozialwiss.

Döring, N. (2018). Sozialkontakte online. In W. Schweiger & K. Beck (Eds.), *Handbuch online-kommunikation* (pp. 1–28). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-18017-1_8-1

Duden. (o.J.a). *Pack.* Retrieved January 19, 2022a, from https://www.duden.de/rechtschreibung/Pack_Poebel_Gesindel

Duden. (o.J.b). *Regime.* Retrieved January 19, 2022b, from https://www.duden.de/rechtschreibung/Regime

Durrell, M. (2004). Sociolect. In U. Ammon, N. Dittmar, K. J. Mattheier, & P. Trudgill (Eds.), *Soziolinguistics. Soziolonguistik: An international handbook on the science of language and society. Ein internationales Handbuch zur Wissenschaft von Sprache und Gesellschaft* (2nd ed., pp. 200–206). De Gruyter.

Dzudzek, I., Glasze, G., Mattissek, A., & Schirmel, H. (2009). Verfahren der lexikometrischen Analyse von Textkorpora. In G. Glasze & A. Mattissek (Eds.), *Handbuch Diskurs und Raum: Theorien und Methoden für die Humangeographie sowie die sozial- und kulturwissenschaftliche Raumforschung* (pp. 233–260). transcript Verlag.

Ebrem, I. E., Karakaya, E., & Krieger, U. (2014). *Jugendszenen zwischen Islam und Islamismus.* Retrieved January 19, 2022, from https://www.lpb-bw.de/fileadmin/lpb_hauptportal/pdf/publikationen/islam_glossar.pdf

Fariclough, N., Mulderrig, J., & Wodak, R. (2011). Critical discourse analysis. In T. A. van Dijk (Ed.), *Discourse Studies* (pp. 357–379). SAGE.

Farschid, O. (2014). Salafismus als politische Ideologie. In S. Behnham & H. Fouad (Eds.), *Salafismus. Auf der Suche nach dem wahren Islam* (pp. 160–192). Herder Verlag.

Fielitz, M., Ebner, J., Guhl, J., & Quent, M. (2018). *Hassliebe: Islamfeindlichkeit, Islamismus und die Spirale gesellschaftlicher Polarisierung.* Jena. Retrieved January 19, 2022, from https://www.idz-jena.de/fileadmin/user_upload/IDZ_Islamismus_Rechtsextremismus.pdf

Freter, W., & Zimpelmann, H. (2015). Internet und Rechtsextremismus. In S. Beck, B.-D. Meier, & C. Momsen (Eds.), *Cybercrime und Cyberinvestigations: Neue Herausforderungen der Digitalisierung für Strafrecht, Strafprozessrecht und Kriminologie* (1st ed., pp. 119–124). Nomos. https://doi.org/10.5771/9783845266121-119

Fritzsche, M. (2019). Der sogenannte „Islamische Staat" und seine Hochglanzmagazine. Ein Gegenstand für die linguistische Forschung. *IDS Sprachreport*, *35*(1), 14–22. Retrieved January 19, 2022, from https://ids-pub.bsz-bw.de/frontdoor/index/index/year/2019/docId/8551

Garimella, K., Morales, G. D. F., Gionis, A., & Mathioudakis, M. (2018). *Political discourse on social media: echo chambers, gatekeepers, and the price of bipartisanship.* Retrieved January 19, 2022, from http://arxiv.org/pdf/1801.01665v2

Guhl, J., & Davey, J. (2020). *A safe space to hate: White supremacist mobilisation on telegram.* London. Institute for Strategic Dialogue. Retrieved January 19, 2022, from https://www.isdglobal.org/wp-content/uploads/2020/06/A-Safe-Space-to-Hate2.pdf

Hall, S. (1989). *Ideologie*. Ausgewählte Schriften. Argument Verlag.

Hanisch, A., & Jäger, M. (2011). Das stigma „gutmensch". *DISS-Journal. Zeitschrift Des Duisburger Instituts Für Sprach- Und Sozialforschung*, *22*, 3.

Harrendorf, S., Mischler, A., & Müller, P. (2019). Same same, but different: Extremistische Ideologien online. Salafistischer Jihadismus und Rechtsextremismus in Social Media. In A. Petzsche, M. Heger, & G. Metzler (Eds.), *Schriften zum internationalen und europäischen Strafrecht, Terrorismusbekämpfung in Europa im Spannungsfeld zwischen Freiheit und Sicherheit: Historische Erfahrungen und aktuelle Herausforderungen* (1st ed., pp. 273–306). Nomos Verlagsgesellschaft mbH & Co. KG. https://doi.org/10.5771/9783845298085-273

Harrendorf, S., Müller, P., & Mischler, A. (2020). Das Zeitalter des digitalen Extremismus? Einige Befunde zu politisch extremer Kommunikation in Social Media. *Zeitschrift Für Internationale Strafrechtsdogmatik, 15*(9), 411–420.

Haslam, S. A., & Reicher, S. (2007). Beyond the banality of evil: Three dynamics of an interactionist social psychology of tyranny. *Personality & Social Psychology Bulletin, 33*(5), 615–622. https://doi.org/10.1177/0146167206298570

Karagiannis, E. (2018). The rise of electoral Salafism in Egypt and Tunisia: The use of democracy as a master frame. *The Journal of North African Studies, 24*(2), 207–225. https://doi.org/10.1080/13629387.2017.1417124

Keller, R. (2011). *Diskursforschung: Eine Einführung für SozialwissenschaftlerInnen*. VS Verlag für Sozialwissenschaften.

Kilgarriff, A. (2012). Getting to Know Your Corpus. In P. Sojka, A. Horák, I. Kopeček, & K. Pala (Eds.), *Text, Speech and Dialogue: 15th International Conference* (pp. 3–15). Springer.

Köhler, D. (2019). The Halle, Germany, Synagogue attack and the evolution of the far-right terror threat. *CTC Sentinel, 12*(11), 14–20.

Kopke, C. (2017). Verschwörungsmythen und Feindbilder in der AfD und in der neuen Protestbewegung von rechts. *NK Neue Kriminalpolitik, 29*(1), 49–61. https://doi.org/10.5771/0934-9200-2017-1-49

Krzyżanowski, M., & Ledin, P. (2017). Uncivility on the web. Populism in/and the Borderline Discourses of Exclusion. *Journal of Language and Politics, 16*(4), 566–581. https://doi.org/10.1075/JLP.17028.KRZ

Kudlacek, D., Jukschat, N., Beelmann, A., Bögelein, N., Geng, B., Glitsch, E., et al. (2017). Radikalisierung im digitalen Zeitalter: Risiken, Verläufe und Strategien der Prävention. *Forum Kriminalprävention, Issue, 3*, 23–32.

Küpper, B., Häusler, A., & Zick, A. (2016). Die Neue Rechte und die Verbreitung neurechter Einstellungen in der Bevölkerung. In R. Melzer (Ed.), *Gespaltene Mitte - feindselige Zustände: Rechtsextreme Einstellungen in Deutschland 2016* (pp. 143–166). J.H.W. Dietz.

Kutter, A. (2018). Corpus analysis. In R. Wodak & B. Forchtner (Eds.), *The Routledge Handbook of Language and Politics* (pp. 169–186). Routledge.

Lee, E.-J. (2007). Deindividuation effects on group polarization in computer-mediated communication: The role of group identification, public-self-awareness, and perceived argument quality. *Journal of Communication, 57*(2), 385–403. https://doi.org/10.1111/j.1460-2466.2007.00348.x

Louwerse, M. M. (2004). Semantic variation in ideolect and sociolect: Corpus linguistic evidence from literary texts. *Computers and the Humanities, 38*, 207–221.

Macklin, G. (2019). The El Paso Terrorist Attack: The chain reaction of global right-wing terror. *CTC Sentinel, 12*(11), 1–9.

Mahood, S., & Rane, H. (2017). Islamist narratives in ISIS recruitment propaganda. *The Journal of International Communication, 23*(1), 15–35.

Mahrt, M. (2019). *Beyond filter bubbles and echo chambers: The integrative potential of the Internet*. digitalcommunicationresearch.de. https://doi.org/10.17174/dcr.v5.0

Marcks, H., & Pawelz, J. (2020). From mythos to victimhood to fantasies of violence: How far-right narratives of imperilment work. *Terrorism and Political Violence*, 1–18.

Mayr, P., & Weller, K. (2017). Think before you collect: Setting up a data collection approach for social media studies. In L. Sloane & A. Quan-Haase (Eds.), *Sage Handbook of Social Media Research Methods* (pp. 107–124). SAGE.

Meiering, D., Dziri, A., & Foroutan, N. (2020). Connecting structures: Resistance, heroic masculinity and anti-feminism as bridging narratives within group radicalization. *International Journal of Conflict and Violence (IJCV), 14*(2), 1–19.

Meleagrou-Hitchens, A., & Kaderbhai, N. (2017). *Research perspectives on online radicalization: Literature review 2006–2016*. Retrieved January 19, 2022, from https://icsr.info/wp-content/uploads/2017/05/ICSR-Paper_Research-Perspectives-on-Online-Radicalisation-A-Literature-Review-2006-2016.pdf

Mischler, A., Müller, P., Geng, B., & Harrendorf, S. (2019). Neue Wege in den Terrorismus? Deutungsmuster extremistischer Ideologien in Social Media. *RW Rechtswissenschaft, 10*(4), 481–524. https://doi.org/10.5771/1868-8098-2019-4-481

Moeller, J., & Helberger, N. (2018). *Beyond the filter bubble: Concepts, myths, evidence and issues for future debates*. Amsterdam. University of Amsterdam. Retrieved January 19, 2022, from https://pure.uva.nl/ws/files/29285427/beyond_the_filter_bubble_concepts_myths_evidence_and_issues_for_future_debates_1_.pdf

Müller, P., & Mischler, A. (2020). Kleine Schnipsel, große Wirkung - die Relevanz und Verwendung von Karikaturen und Memes in Kontexten extremistischer Online-Kommunikation. In C. Grafl, M. Stempowksi, K. Beclin, & I. Haider (Eds.), *Neue Kriminologische Schriftenreihe: Vol. 118. „Sag, wie hast du's mit der Kriminologie?": Die Kriminologie im Gespräch mit ihren Nachbardisziplinen* (pp. 457–477). Retrieved January 19, 2022, from http://www.krimg.de/drupal/files/9783964100191.pdf

Myers, D. G. (2010). Group Polarization. In J. M. Levine & M. A. Hogg (Eds.), *Encyclopedia of Group Processes and Intergroup Relations* (pp. 361–365). SAGE Publications.

Niehr, T., & Böke, K. (2004). Diskursanalyse unter linguistischer Perspektive — am Beispiel des Migrationsdiskurses. In R. Keller, A. Hiersland, W. Schneider, & W. Viehöver (Eds.), *Handbuch Sozialwissenschaftliche Diskursanalyse: Band 2: Forschungspraxis* (2nd ed., pp. 325–351). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-99764-7_12

Niehr, T., & Reissen-Kosch, J. (2019). *Volkes Stimme? Zur Sprache des Rechtspopulismus*. Bundeszentrale für politische Bildung.

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Viking.

Pearson, E. (2018). Online as the new frontline: Affect, gender, and ISIS-take-down on social media. *Studies in Conflict & Terrorism, 41*(11), 850–874. https://doi.org/10.1080/1057610X.2017.1352280

Postmes, T., & Spears, R. (1998). Deindividuation and antinormative behavior: A meta-analysis. *Psychological Bulletin, 123*(3), 238–259. https://doi.org/10.1037/0033-2909.123.3.238

Postmes, T., Spears, R., Sakhel, K., & de Groot, D. (2001). Social influence in computer-mediated communication: The effects of anonymity on group behavior. *Personality & Social Psychology Bulletin, 27*(10), 1243–1254. https://doi.org/10.1177/01461672012710001

Rafael, S., & Ritzmann, A. (2019). Backround: The ABC of hate speech, extremism and the NetzDG. In J. Baldauf, J. Ebner, & J. Guhl (Eds.), *Hate Speech and Radicalisation Online. The OCCI Research Report* (pp. 10–17).

Ranan, D. (2018). *Muslimischer Antisemitismus. Eine Gefahr für den gesellschaftlichen Frieden?* J.H.W. Dietze.

Rieger, D., Frischlich, L., & Bente, G. (2020). Dealing with the dark side: The effects of right-wing extremist and Islamist extremist propaganda from a social identity perspective. *Media, War & Conflict, 13*(3), 280–299. https://doi.org/10.1177/1750635219829165

Rippl, S., & Seipel, C. (1999). Gender differences in right-wing extremism: Intergroup validity of a second-order construct. *Social Psychology Quarterly, 62*(4), 381. https://doi.org/10.2307/2695835

Sageman, M. (2008). *Leaderless Jihad: Terror networks in the twenty-first century*. University of Pennsylvania Press Inc.

Scharloth, J. (2021). *Hässliche Wörter. Hatespeech als Prinzip der neuen Rechten*. J. B. Metzler.

Schmitz-Berning, C. (2007). *Vokabular des Nationalsozialismus* (2., durchgesehene und überarbeitete Auflage). De Gruyter.

Schröder, C. P., Goede, L.-R., & Lehmann, L. (2020). *Perspektiven von Studierenden: Ergebnisse einer Befragung zu den Themen Politik, Religion und Gemeinschaft im Rahmen des Projektes „Radikalisierung im digitalen Zeitalter (RadigZ)"* (Forschungsbericht / Kriminologisches Forschungsinstitut Niedersachsen e.V. (KFN), Nr. 156). Hannover: Kriminologisches Forschungsinstitut Niedersachsen e.V. (KFN).

Schwarz-Friesel, M. (2019). *Judenhass im Internet: Antisemitismus als kulturelle Konstante und kollektives Gefühl.*

Schwarz-Friesel, M., & Reinharz, J. (2017). *Inside the anti-semitic mind: The language of jew-hatred in contemporary Germany*. Brandeis University Press.

Sketch Engine. (2020a). *Keywords and term extraction - Identifying typical words*. Retrieved January 19, 2022, from https://www.sketchengine.eu/guide/keywords-and-term-extraction/#toggle-id-1

Sketch Engine. (2020b). *Token*. Retrieved January 19, 2022, from https://www.sketchengine.eu/my_keywords/token/

Spears, R., & Lea, M. (1994). Panacea or Panopticon? The hidden power in computer-mediated communication. *Communication Research, 21*(4), 427–459. https://doi.org/10.1177/009365094021004001

Spears, R., & Postmes, T. (2015). Group identity, social influence, and collective action online. In S. S. Sundar (Ed.), *The Handbook of the Psychology of Communication Technology* (Vol. 90, pp. 23–46). Wiley. https://doi.org/10.1002/9781118426456.ch2

Staub, E. (2001). Individual and group identities in genocide and mass killing. In R. D. Ashmore, L. J. Jussim, & D. Wilder (Eds.), *Social Identity, Intergroup Conflict, and Conflict Reduction* (p. 159). Oxford University Press.

Struck, J., Müller, P., Mischler, A., & Wagner, D. (2020). Volksverhetzung und Volksvernetzung: Eine analytische Einordnung rechtsextremistischer Onlinekommunikation. *Krim OJ. Kriminologie – Das Online-Journal* (2), 310–377. https://doi.org/10.18716/OJS/KRIMOJ/2020.2.12

Sunstein, C. R. (2001). *Echo chambers: Bush v Gore, impeachment, and beyond*. Princeton University Press.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In W. G. Austin & S. Worchel (Eds.), *The Nelson-Hall series in psychology. Psychology of Intergroup Relations* (2nd ed., pp. 7–24). Nelson-Hall Publishers.

Technau, B. (2018). Beleidigungswörter. *De Gruyter*. https://doi.org/10.1515/9783110563870

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1988). *Rediscovering the social group: A self-categorization theory*. Blackwell.

Vergani, M., & Bliuc, A.-M. (2015). The evolution of ISIS' language: A quantitative analysis of the language of the first year of dabiq magazine. *Sicurezza, Terrorismo E Scieta* (2), 7–20.

Vieregge, E. (2014). Neonazis, Identitäre, Salafisten – Eine Nachbetrachtung der eskalierten Kölner Hooligan-Demonstration. Retrieved January 19, 2022, from https://www.endstation-rechts.de/news/neonazis-identitaere-salafisten-eine-nachbetrachtung-der-eskalierten-koelner-hooligan-demonstratio.html

Weber, F. (2015). Diskurs - Macht - Landschaft. In S. Kost & A. Schönwald (Eds.), *Landschaftswandel - Wandel von Machtstrukturen* (pp. 97–112). Springer VS.

Weitzman, M. (2006). Antisemitismus und Holocaust-Leugnung: Permanente Elemente des globalen Rechtsextremismus. In T. Greven & T. Grumke (Eds.), *Globalisierter Rechtsextremismus?* (pp. 52–69). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90125-1_4

Winter, C., Neumann, P., Meleagrou-Hitchens, A., Ranstorp, M., Vidino, L., & Fürst, J. (2020). Online extremism: Research trends in internet activism, radicalization, and counter-strategies. *International Journal of Conflict and Violence (IJCV), 14*(2), 1–20. https://doi.org/10.4119/ijcv-3809

Zimmer, F., Scheibe, K., Stock, M., & Stock, W. G. (2019). Fake news in social media: Bad algorithms or biased users? *Journal of Information Science Theory and Practice, 7*(2), 40–53.