

Rachel McCullough, Daniel Drylie, Mindi Barta, Cass Dykeman, and Daniel Smith

CoDEC-M: The multi-lingual manosphere subcorpus of the Corpus of Digital Extremism and Conspiracies

Abstract: In 2023, the U.S. Surgeon General warned the public of the current “loneliness epidemic” and its potential consequences on physical and mental health. One possible consequence of this epidemic is the growth of a movement defined by loneliness and isolation: the incel (“involuntary celibate”) movement. This warning presents a worrying glimpse at the future, as the incel movement, along with other parts of the manosphere, is one that espouses violently misogynist rhetoric which is intrinsically linked to right-wing extremism. While linguistic studies have been conducted on the speech of incels and other constituent movements of the manosphere, few of these studies look at the language of these communities from a cross-cultural and cross-linguistic perspective. To address this gap, we have created CoDEC-M, a subcorpus of the Corpus of Digital Extremism and Conspiracies (CoDEC). CoDEC is an open-source, open-access corpus made up of several subcorpora documenting different online spaces where extremists and conspiracy theorists gather. CoDEC-M is our response to the growing interest in the manosphere and the gap in scientific knowledge on the language used in its non-English speaking communities.

In this paper, we use the text analysis software Sketch Engine to compare the top twenty keywords and bigrams in the English and Russian sections of CoDEC-M ranked by their keyness score. In doing so, we have uncovered evidence of language transfer between these two segments of the manosphere via direct borrowings from English into Russian and thematic overlap between keywords and bigrams that refer to gender, dating, and physical appearance. We have also uncovered and define a number of neologisms unique to each dataset and examine the real-world impact of the manosphere in English- and Russian-speaking

Rachel McCullough, Bolante.NET, e-mail: rechalmccullough@protonmail.com

Daniel Drylie, Old Dominion University, e-mail: ddrylie@odu.edu

Mindi Barta, Oregon State University, e-mail: bartam@oregonstate.edu

Cass Dykeman, Oregon State University, e-mail: cass.dykeman@oregonstate.edu

Daniel Smith, Bolante.NET, e-mail: danielvsmithpsyd@gmail.com

countries in support of our argument that non-Anglo portions of the manosphere warrant further analysis.

Keywords: corpus linguistics, multilingual corpora, language and gender, computer mediated discourse, Russian language, sociolinguistics

1 Introduction

As the 2020's unfold, we have witnessed misogynist domestic terrorism emerge as a serious threat to women. Of 32 ideologically-motivated acts of mass violence committed between 2016 and 2020, the perpetrators of four attacks (12.5%) were named by the U.S. Secret Service as being associated with the incel or “involuntary celibate” movement (NTAC 2023). In other words, more than one in ten perpetrators of ideologically-motivated mass violence in the United States are not only motivated by misogyny, but belong to one community in particular: the incel movement. Paired with this statistic is the adoption of language and rhetoric originating from incel communities into popular discourse on social media platforms like Reddit, YouTube and TikTok.¹ Alarming as these two facts are in conjunction, the present study does not intend to examine their relationship. Instead, we are interested in the spread of the incelosphere beyond the English-speaking parts of the web – specifically, how self-identified incels on the Russian-speaking internet communicate, what they have in common with their English-speaking counterparts, and some possible implications of any language transfer we find.

The influence of misogynist rhetoric on popular media of the 2020's has been palpable. According to a 2023 report, over 40% of surveyed men under 30 trust a prominent figure or movement associated with the manosphere (Barker et al. 2023). The list of trustworthy sources referenced includes controversial figures like Andrew Tate and Jordan Peterson, web forums like Reddit's r/TheRedPill, and hate groups like the Proud Boys. Alarming, the shared feature among all members of this list is a virulently anti-woman ideology and rhetoric that sometimes correlates with real world violence, as described above.

“Incel” is a compound-clipping of “involuntary celibate,” an initially gender-neutral phrase coined in the late 1990's by a woman who also shared her struggles to find a romantic partner on a website she created, *Alana's Involuntary Celi-*

¹ For more on the integration of incel rhetoric into mainstream social media apps, see Solea and Sugiura (2023).

bacy Project (Taylor 2018). Decades later, contemporary online spaces that use the “incel” label are typically male-dominated and known for extremely misogynistic attitudes, with some even restricting membership to exclude female incels (known as “femcels”; Incels 2017). Since shifting to its present male-as-norm state, the movement has brought like-minded men experiencing loneliness together to share grievances over sexual and romantic exclusion.

However, this sense of community has not proven beneficial to these men or society at large. Known members of the movement have committed acts of violence against women, like the 2018 Toronto van attack or the mass shooting in a Tallahassee yoga studio in the same year. While these incidents are frequently represented in news headlines, mass violence is not the only form of misogynist action taken by members of these communities. Two terrorists described in *Mass Attacks in Public Spaces: 2016–2020* participated in online misogynist communities and nonconsensually distributed sexual photographs of women to terrorize them before acting out violently (NTAC 2023). Men themselves are also suffering and at risk of harm due to these ideologies. The mental health of young men is at particular risk, as observed by Equimundo, with nearly half of all survey respondents aged 18–30 reporting thoughts of suicide in the previous two weeks (Barker et al. 2023). While concern for misogynist violence is significant, the prevalence of suicide notes posted by self-identified incels indicates another fatal consequence of this movement’s growth.²

While this cultural context comes from predominantly English-speaking communities, misogyny is a polyglot. Lonely adults from various language communities have shared similar concerns online long before English-speaking incels dominated the headlines, like the German-speaking *Absolute Beginner* community described by Sprenger (2014) or the French-speaking *Virginité-tardive* forum, which hosts threads dating back to 2007. In the last decade, though, self-identified incel spaces outside the online anglosphere have been appearing as the movement has gained notoriety (e.g., the Italian language *Forum Dei Brutti* and the Russian language /incel/ thread on 2ch.hk). These spaces are distinct from the previously mentioned *Absolute Beginners* and *Virginité-tardive* forums in that they identify specifically with the incel movement, borrowing the name of the predominantly English-speaking group for themselves. In light of this development, we ask: How much linguistic overlap is there between the speech of non-English speaking incel communities and the their English-speaking counterparts, and what is this overlap like?

2 For an examination of suicide notes from incel community members, see Daly and Laskovtsov (2021).

This study seeks to address how the speech patterns and ideologies of English-speaking incels have or have not permeated into the language of non-English speaking incel communities. As a case study, we will compare and contrast the language of speakers in both English- and Russian-speaking incel communities, and identify common rhetorical themes, shared beliefs, and lexical borrowings. While we cannot guarantee the national origin or place of residence of the users of either 2ch.hk or incels.is, we will consider these two populations as representative of incels as they exist on the anglophone web (on incels.is) and as they exist on the Russian-speaking web (commonly referred to as “Runet”; on 2ch.hk). Our reason for selecting Russian-speaking incels to compare to English-speaking incels is two-fold. First, the availability of data: the ongoing /incel/ thread on 2ch.hk provides a sizable amount of data in Russian from self-identified incels. Second, we are especially interested in Russian speakers in contrast with speakers of the languages of North America and Western Europe because of the popular notion that Russian culture exists separate from “Western” or other European cultures. Does this supposed cultural difference impact the speech and attitudes of lonely, disaffected young men? To what degree do we see evidence of language contact between incel populations that speak different languages?

A comparison of parallel subcultures from two different digital linguistic traditions will demonstrate how the internet age has allowed this specific variety of misogyny to spread to other languages. It will also provide a greater understanding of the factors that contribute to the formation of these communities. While we are only comparing two communities for the present study, the model we propose for comparing keywords and phrases across two subcorpora in different target languages has potential for use in the comparison of other non-English speaking sections of the manosphere (e.g., the *Forum Dei Brutti* or other under-researched communities).

Three research questions were designed for the present study. These were:

1. When compared to mainstream online discourse, what conversational subjects and vocabulary are associated with the speech of the English incelosphere?
2. When compared to mainstream online discourse, what conversational subjects and vocabulary are associated with the speech of the Russian incelosphere?
3. How do the conversational subjects and vocabulary of the English- and Russian-speaking incels compare with one another?

The formation of research questions was guided by existing literature on the English- and Russian-speaking portions of the manosphere. To address these questions, we will compare the top 20 keywords and phrases – specifically bigrams – in data from English- and Russian-speaking incel forums to general web discourse in the

target language. We will also compare and contrast how the resultant datasets overlap at the word level and their semantic or thematic overlap.

2 Literature review

To establish the greater context of this study, we will first give a brief explanation of the greater “manosphere” as it exists in the 2020’s online anglosphere and its relationship to incels in specific. Second, we will briefly examine the current and past state of gender roles in English- and Russian-speaking communities offline. While we specifically base our data collection and analysis on a target language rather than nationality due to the international nature of line communication, we will limit this examination of gender offline to North America and Russia for the sake of brevity, due to their outsized representation in popular culture and online.

2.1 “The Manosphere” and red pill ideology

The manosphere is a name for a group of overlapping communities centered around men’s perspectives and issues. These range from fitness and dating advice, to men’s rights, to antifeminist and incel communities. Incels are one of the most visible communities in the manosphere, producing misogynist content that has drawn massive mainstream media attention in the late 2010s and early 2020s – including coverage from the *BBC*, the *New York Times*, and *Al-Jazeera* (Griffin 2021; Chokshi 2018; Cunha 2020). Crucial to the understanding of the ideologies of the manosphere is red pill ideology.

In the context of the manosphere, red pill ideology refers to an awakening among men to the supposed deceit and cultural brainwashing of feminism (further described in Ging 2019). The red pill ideology of the manosphere has spread rapidly across the web from its origins in an eponymous subreddit (r/TheRedPill), a space Debbie Ging describes as “dedicated to antifeminism and the defense of rape culture” (2019: 645). Ging also notes a spread of both red pill ideology and its language into other communities that constitute the manosphere, specifically naming “red pill terminology” as a unifying feature between otherwise disparate male-centric communities that have begun to coalesce in some ways due to a shared sense of identity (645).

2.2 Incels and their speech

Members of the incel movement define themselves by their lack of sexual success with women despite a desire for it. While many manosphere communities espouse red pill ideology, self-identified incels are primarily concerned with the *black pill*, a red pill-derived ideology founded in hopelessness and pessimism. While redpilled members of the manosphere believe that they can partake in the sexual marketplace by manipulating it and working within it, blackpilled incels feel defeated at the hands of supposed systemic misandry (described in Fernquist et al. 2020; Preston et al. 2021).

Not all incels are blackpilled: some of them attempt to “ascend” from incel-dom – that is, enter a romantic or sexual relationship. Blackpilled community members, however, will attempt to dissuade this behavior by arguing that one’s incel status is genetically predetermined (Quiroz 2022). In fact, to ascend would seem to invalidate their very identity, which they have based on a supposed inability to find companionship. This loss of incel status would almost certainly result in being ousted from the community that had long served as a comfort for these vulnerable young men, as non-incels are explicitly banned from having an account on incels.is (see “Rules and FAQ”). Accordingly, maintaining and promoting the worldview of hopelessness and perpetual victimhood associated with blackpilled incels described by Fernquist et al. (2020) and Preston et al. (2021) can promote continued group membership.

The way incels speak is an important indicator of group membership and in-group status. Incels have developed an insular community of practice with its own subculture and what some refer to as a “cryptotect” enmeshed in their misogynist ideology, which can obfuscate their conversations from newcomers (Gothard 2021: 2–6). This often opaque language both furthers a sense of in-group identity among incels and marks non-group members as outsiders if they fail to conform. Self-identified incels are present in numerous parts of the world, according to self-reported survey data from incels.is.³ Not all incel communities speak English exclusively, though English-speaking incel communities appear to be the most widely studied, with most of the literature cited in this very paper referring to studies of English-speaking incel communities (e.g., Baele et al. 2023; Daly and Laskovtsov 2021; and Ging 2019, among others). There exist some exceptions – see Voroshilova and Pesterev’s (2021) study of Russian-speaking incels and Fernquist et

3 SergeantIncel (2020)’s survey indicated that 42.8% of users were European and 38% were American, with most of the remainder logging on from Central America, South America, and Asia.

al.'s (2020) section on Swedecels, for example – but few of these directly compare the speech of multiple incel communities with different primary languages.

Previous studies on incels identified in-group jargon, including terms shared with other parts of the manosphere – particularly the greater red pill community – and those originating from the incel community (see Gothard 2021; Moonshot 2020). These studies typically focus solely on anglophone communities of the manosphere and their rhetoric. This study, however, focuses on incels specifically and compares the words and phrases most frequently used among Russian-speaking incels to those used by English-speaking incels. Similar cross-linguistic studies of English- and Russian-speaking incels like Voroshilova and Pesterev's (2021) are rare and have not, to the authors' knowledge, covered lexical distribution and difference.

2.3 Misogyny and gender in North America

Part and parcel of the rise of conservatism in North America that began in the 2010's is the normalization of misogyny in mainstream media and politics. Described by DiBranco as a “gateway drug” for the recruitment of disaffected White men into racist communities,” the misogyny exemplified by incel communities has ramifications beyond these spaces (2017: 15). This is especially visible in the United States, where the 2016 U.S. presidential election witnessed red pill communities grow vocal in their support of then-candidate Donald Trump, a figure community members held up as an “alpha male” who, when elected, would combat the feminist agenda in their stead (Dignam and Rohlinger 2019: 603). Dignam and Rohlinger further describe the mobilization of Reddit's red pill communities in support of Trump during the 2016 election, claiming that Trump “caused hypermasculinity, blatant misogyny, and violent tough talk to resurge in popularity on the national stage” (2019: 603). While not directly inciting violent misogynist acts, the continued platforming of such rhetoric by successful politicians and those that follow in their footsteps may serve to legitimize underlying misogyny.

Contemporaneously, several high profile acts of misogynistic terrorism in North America have taken place, including the 2014 Isla Vista shooting and the 2018 Toronto van attack, among others (further described in O'Donnell 2021). As a result of these attacks, the threat of misogynist terrorism has been acknowledged by the U.S. Secret Service and covered extensively in American media via profiles of the manosphere, especially incels (NTAC 2023; Bosman et al. 2019; Townsend 2022). The perpetrators of these incidents have been revered by incels on public web forums and even, in the case of Elliot Rodger, acknowledged in the manifestos of those who go on to commit similar acts of mass violence (Moonshot 2020: 11).

In light of both the political context and recent incidents of misogynist terrorism, it should not be surprising that the speech of English-speaking incels is markedly misogynistic. Pelzer et al. (2021: 211–212) found that the language on incels is was 20% more toxic than control data.⁴ More specifically, attacks on women for their gender identity, race, ability, and sexuality (Czerwinsky 2023) are characteristic of discourse in incel communities.

2.4 Misogyny and gender in Russia

While singular incidents of mass violence against women such as those described in the previous section are not found in Russian headlines, systematic discrimination against women via legislation have made international news. For example, a piece of legislation was signed into law in 2017 that decriminalized domestic violence, so long as it is the accused's first offense (HRW 2017). This type of legislation is aligned with the “traditional family values” agenda pushed by Vladimir Putin, Russia's longtime authoritarian leader with a political persona built on an exaggerated performance of masculinity (described further in Novitskaya 2017).

Russia's misogyny problem does not end with the institutionalization of conservative “family values”. Emboldened by this type of legislation and popular media coverage of Putin's domestic policies, the Runet has been slowly developing an “atmosphere of permissiveness for homophobic voices and misogynistic rhetoric” (Lokot 2019: 217). Like North America, Russia is also host to social movements espousing regressive gender norms. Russian misogynist movements have been buoyed by their acceptance in online spaces, reminiscent of their English-speaking counterparts emboldened by Gamergate.

Consider the Male State (*Мужское государство*), an anti-feminist movement whose page on the Russian social network VKontakte had over 150,000 members before being banned in 2020 due to reported calls for violence (Meduza 2020). According to Gaufman (2022), the organization has since been declared an extremist organization and consequently banned from operating in Russia. Members of the Male State have conducted harassment campaigns against and released the personal information of prominent women and feminists (Gaufman 2022). The group's acts of misogyny are not limited to online harassment campaigns and doxxing: in one documented instance, a female blogger was physically attacked by a

⁴ For the purpose of the cited study, “toxicity” was determined by a BERT model which measured toxic language by targeting insults, fantasies of violence, and bigoted language, among other types of negativity directed toward outsiders, other forum users, and the speakers themselves.

follower of the Male State in 2020 for posting sexual content online (Bellingcat 2021). While some media outlets have described the “supposed fall” of the Male State, the community continues to operate on the privacy-oriented messaging app Telegram after its ban from VKontakte – albeit with a smaller audience.

3 Method

3.1 Design

In this study, we use a set of three corpora to target both word and multi-word unit keyness, specifically of bigrams.⁵ All of the processes described in this section (e.g., preprocessing, generating keyness scores) were performed in Sketch Engine, a corpus management system. Word and multi-word keyness factor into the collection of keywords and key phrases, respectively. For the purpose of this study, keywords and phrases are those that appear in the target corpus more frequently than they would in mainstream discourse on the web, represented by a reference corpus. Keywords and phrases can then be used to begin to understand the “contents, style and discourse of a corpus” (further detailed in Moreno-Ortiz 2024). While keywords can certainly provide information about the prevalent topics and the type of discourse found in CoDEC-M, analysis of concordance lines and collocates would provide further nuance to our understanding of these texts from the incelsphere and are a topic of interest for future study.

With respect to the corpora used, this study uses a single study corpus consisting of two parts – the manosphere subcorpus of the Corpus of Digital Extremism and Conspiracies, henceforth CoDEC-M – in conjunction with two reference corpora. To create the first iteration of CoDEC-M, we collected and compared web data consisting of around 2 million tokens in English and around 3.6 million tokens in Russian.⁶ The size of these subsections was determined by the space available on our Sketch Engine account, which was limited to a total of 7 million tokens.

The English section of CoDEC-M was collected using the Selenium Python package to scrape the entirety of incels.is. This consisted of far more tokens than our

⁵ While the full results include multi-word units with more than two words, these have been excluded from our analysis as we chose to target bigrams specifically.

⁶ The Russian data was collected from 2ch.hk on 25 July 2023, and the English data was collected from incels.is on 23 February 2023. This data is available for download on this project’s Github page: <https://github.com/ddryl001/codec> (last accessed 14 February 2025).

Sketch Engine account would store (46,000 texts), so we randomly selected 5% of these texts to create an extract for analysis. To avoid biasing the dataset toward any period of time or subsection of the website, the 2,300 texts files were selected at random. While the size of these subsections of CoDEC-M are somewhat disproportionate, this will not affect our analysis. Because each subsection of the corpus is relatively large (at least 2 million words) and the algorithm used to determine keyness score is a ratio of relative frequencies, this size discrepancy is negligible.

The Russian section of CoDEC-M was collected using Beautiful Soup to scrape manually retrieved links for the 472 most recent threads from 2ch.hk's ongoing /incel/ thread on the /sex/ board. Recent threads were chosen because they are the most accessible to users, as 2ch.hk does not have a formal archive of all of its threads – the most recent instance of the /incel/ thread simply links back to the previous iteration, and so on. The scraping process for the Russian data could not be automated using Selenium due to a bevy of broken links and nonstandardized archival practices typical of imageboards. Because of this, 2ch.hk links were collected manually and limited such that the resultant body of texts would not take up over 4 million of the 7 million tokens available on the Sketch Engine account.

The Russian subcorpus of CoDEC-M uses ruTenTen17 as a reference corpus, while the English subcorpus uses enTenTen21.⁷ The data contained in ruTenTen17 in particular is not as new as we would have preferred for the study of such a nascent community as Russian-speaking incels, with the most recent iteration of ruTenTen scraped in 2017. Additionally, the TenTen corpus family is not restricted to a single genre of texts from the Web, while our target corpora are all from discussion forums, which is suboptimal. This genre problem is also present in enTenTen21. However, ruTenTen17 is the most recent, widely available large corpus of Russian language data from the Web. These factors, along with the availability of the TenTen corpus family on Sketch Engine, led to our ultimate decision to use this corpus family as reference corpora. When analyzing the Russian data, we consulted several resources developed by native speakers to provide English translations for our results.

Once compiled, the target corpora were preprocessed and run through the Wordlist and Keywords tools in Sketch Engine to determine the top keywords and bigrams via relative frequency (RF) and keyness. Preprocessing of both English and

⁷ These corpora are part of the TenTen Corpus Family, a multi-language family of corpora made up of web data hosted by Sketch Engine.

Russian data included lemmatization⁸ of both the English and Russian data⁹ and stop word removal, with custom stop words added for the Russian dataset to remove formatting text from 2ch.hk – these can be found in full on this project’s page at <https://osf.io/ytrsx> and include words related to date, time, and post authors that are included with every post on 2ch.hk.

3.2 Data analysis

Only the top 20 keywords and phrases in the English and Russian language datasets are presented in this paper, along with a keyness score and – for the Russian items – English translation. While the remainder of the 1,000 keywords and phrases provided by Sketch Engine warrant study, we only present the top 20 items from each category in this paper for the sake of parsimony. The full results can be found on this project’s OSF page.¹⁰ Also found on this project’s OSF page are brief explanations of the words and phrases from Tables 1–4 that cannot be found in mainstream dictionaries. These are more words more typically used by internet users and young people (including members of incel communities) that may be unknown to outsiders.

To avoid keyness results being dominated by the appearance of a word in just a few texts, Average Reduced Frequency (ARF) has been factored into these results and has been reported in our full results tables. Using ARF to factor in dispersion, we followed the example set by Venuti and Fruttaldo (2019) and excluded from Tables 1–4 any keywords or multi-word terms with an ARF score less than 2.00. The full results for this study also contain each keyword and phrase’s absolute and relative frequency in both the target and reference corpora.

8 A note on lemmatization and CoDEC-M: while it was performed, Sketch Engine’s lemmatization function seems to not do well with lexical innovation and slang. This is true for both the English and Russian data; in the full English keyword data, for example, both “incel” and “incels” can be found, and in the full Russian keyword data, we see both *инцел* (“incel”) and *инцелы* (“incels”).

9 A note on lemmatization of Russian data: while no white paper is available for Sketch Engine’s Russian lemmatizer, correspondence with the developers revealed that it holds space for certain inflected forms (e.g., gender lemmas and lemmas of degree), which can be seen in Table 4.

10 The data presented in the spreadsheet of this study’s full results includes manual annotation performed during analysis, primarily consisting of color-coding performed during cleaning and translation notes for the Russian data. A key is provided in the full results. This annotation was performed for at least the top 20 items of each category, but additional translation notes are provided for the remaining top 100 keywords and phrases in Russian.

The difference in magnitude between the study and reference corpora was addressed using Kilgarriff’s simple maths, a ratio of relative frequencies with a smoothing function (Kilgarriff 2009). This smoothing factor offers a balance between identifying common and rare words with respect to keyness (Kilgarriff 2012). The smoothing factor selected for this study was 1, the default setting in Sketch Engine, because we were interested in terms that are both frequently used and characteristic of the community.

Keyword overlap was measured via the Jaccard Similarity Index. This commonly-used index indicates the number of shared elements $|A \cap B|$ divided by the unique number of elements $|A \cup B|$ (da F. Costa 2021). This results in a score between 0 and 1, where a higher score indicates a higher degree of overlap.

4 Results

4.1 The English data

Table 1 contains the top 20 keywords in order of keyness from the English section of CoDEC-M, followed by their keyness score. No duplicate word stems were present, and no words were omitted from the results presented here. For the results in their entirety, please visit this project’s OSF page at <https://osf.io/ytrsx>.

Table 1: Top 20 English keywords by Simple Maths score (RQ1).

Word	Simple Maths	Word	Simple Maths
foid	252.037	cuck	46.448
incel	249.685	tbh	45.559
normie	112.083	blackpill	39.550
chad	99.381	ugly	39.148
fuck	75.115	cope	34.882
nigger	58.452	mog	33.901
jfl	50.840	iq	32.583
whore	49.448	curry	32.367
shit	47.448	ascend	32.212
subhuman	46.714	fakecel	30.640

Table 2 contains the top 20 bigrams from the English section of CoDEC-M, ordered by keyness score. While they are present in the complete results, five phrases were removed due to being part of boilerplate text; 12 were removed because they were usernames; three were removed due to suspicion of spamming; and one was removed because it was not a bigram.

Table 2: Top 20 English bigrams by Simple Maths score (RQ1).

Word	Simple Maths	Word	Simple Maths
white woman	13.103	black pill	6.357
white man	9.097	average look ¹²	5.963
good look ¹¹	8.636	white girl	5.941
white foid	8.206	white knight	5.784
ugly man	7.153	i dont	5.730
incel forum	7.099	black man	5.623
white guy	7.057	virtue signal	5.551
video game	6.736	asian woman	5.288
average height	6.660	short man	5.209
foid worship	6.609	social circle	5.154

4.2 The Russian data

Table 3 contains the top 20 keywords in the Russian section of CoDEC-M in order of keyness, followed by their English translations and their keyness score. Translations were initially obtained via transliteration and/or machine translation and were further refined by consulting Russian dictionaries and language learning

¹¹ Because lemmatization was performed as a preprocessing step, the bigrams “good-looking” and “average-looking” are rendered “good look” and “average look”, respectively.

¹² See previous footnote.

resources,¹³ Russian speakers, and the *Incels Wiki*'s page on the Russian incelosphere.¹⁴

Table 3: Top 20 Russian keywords by Simple Maths score (RQ2).

Word	Translation	Simple Maths	Word	Translation	Simple Maths
тян	<i>chan</i> ; a young woman	935.261	скуфидрон	an unappealing man, often bald or overweight	138.582
инцел	incel*	830.620	опухший	swollen, puffy ¹⁵	130.147
чед	Chad*	390.290	ирл	irl*	129.181
ебало	(<i>mat</i>) mouth	233.049	лвл	<i>lit.</i> lvl* ("level"); age	127.649
спок	chill, calm down (v.)	231.235	двачую	<i>lit.</i> twice; seconded (as in agreement) ¹⁶	123.876
нахуй (cf. хуй)	(<i>mat</i>) fuck it	188.297	нормис	normies*	114.207
пиздец (cf. пизда)	(<i>mat</i>) fuckload; or, clusterfuck	164.862	блять (cf. блядь)	vulgar exclamation	108.317
ебать	(<i>mat</i>) to fuck	157.79	кунов (cf. кун)	<i>kun</i> 's (belonging to a young man)	106.935
бетабакс	betabucks*	144.278	максилла	maxilla	105.402
всратый	<i>lit.</i> shitted-in; ugly	138.732	подкатывать	hit on; pick up	103.714

¹³ The dictionaries consulted for the results in Table 3 were Barron's *Dictionary of Russian Slang and Colloquial Expressions* and Terminy.info's *Словарь молодежного сленга*, a community-sourced online dictionary that was able to account for more recent lexical innovations.

¹⁴ While the *Incels Wiki* is not a source reviewed or edited by professionals, it is the only community-developed resource known to the authors that describes language used by Russian-speaking incels. Terms are only available in transliteration rather than Cyrillic, but those that overlapped with our results were identified via transliteration.

¹⁵ Literally means 'swollen'. This same root can be used figuratively to indicate boredom or tiredness in one sense, and eccentricity in another; see Šljachov and Adler (2006: 185). However, concordance lines from CoDEC-M indicate that this word is primarily used to refer to a swollen face or a person with such a face.

¹⁶ Online slang originating on 2ch but now used across Russian language web forums. Used to agree to a prior message, comparable to "^^this" as used on English-language forums (*Двачую* n.d.).

Eighteen words were not included in Table 3 due to repeated word stems.¹⁷ Some of these words were inflected (e.g., *инцел*, ‘incel’; and *инцелы*, ‘incels’), but others were derived from words whose roots are already in the top 20 keywords.¹⁸ One spelling variant (*чэд*, cf. *чед*, or ‘Chad’) and one abbreviation via clipping (*скуф*, cf. *скуфидрон*, or ‘skuf’, cf. ‘skufidron’) were also removed. Asterisks in the “Translation” column indicate that the Russian words are direct borrowings from English transliterated into Cyrillic. These consist of the transliterations of *incel*, *Chad*, *beta-bucks*, *normies*, *irl*, and *lvl*. We also note in the “Word” column which words are derived from *mat* (Russian profanity).

Table 4 contains the top 20 bigrams in the Russian section of CoDEC-M ordered by keyness, followed by their English translations. These translations were also obtained via machine translation and further refined by consulting the previously-named resources. One phrase was not included in this table because it was identified as spam, and two were not included because they are not bigrams.

Table 4: Top 20 Russian Bigrams by Simple Maths score (RQ2).

Bigram	Translation	Simple Maths	Bigram	Translation	Simple Maths
тёмная триада	dark triad	33.913	линия волос	hairline	15.370
сын шлюхи	son of a whore	27.697	средний рост	average height	14.514
невольное воздержание	involuntary abstinence	27.064	будка чеда	strong jaw ¹⁹	13.533
пониже рост	shorter height	24.627	глазе жертвы	prey eyes ²⁰	13.376

¹⁷ The words excluded from the results presented in Table 3 are as follows: *тянки*, *тянок*, *тянка*, *тянками*, and *тянку* (all derived from *тян*, ‘chap’); *инцелы*, *инцелов*, *инцела*, and *инцелом* (all derived from *инцел*, ‘incel’); *чеда*, *чеды*, and *чедов* (all derived from *чед*, ‘Chad’); *похуй*, *хуйня*, *нихуя*, and *хуй* (all derived from *хуй*; because *нахуй* is the form with the highest keyness score, it was included in the table); *ебанный* (derived from *ебать*); and *скуфыня* (derived from *скуфидрон*).

¹⁸ Words inflected with derivational affixes were not included in Table 3 because of our team’s lack of a native Russian speaker – without a native speaker, we were unlikely to achieve the level of nuance in translation that these inflected words would warrant.

¹⁹ Literally translated as ‘chad’s booth’ or ‘chad’s box’; refers to a strong, square jawline.

²⁰ Literally translated as ‘victim’s eye’.

Bigram	Translation	Simple Maths	Bigram	Translation	Simple Maths
осознанное воздержание	conscious abstinence	24.191	процентом жира	fat percentage	13.002
желание секса	desire for sex	23.411	зона глаз	eye area	12.818
фаза знакомства	dating phase	19.149	черта лица	facial feature	12.788
размер хуя	dick size	17.754	ментальный чед	mental Chad	12.536
наибольший хуй	biggest dick	16.465	основная теория	basic theory	12.374
линия роста	growth line ²¹	16.331	большинство тян	most girls ²²	12.109

4.3 Overlap

With respect to RQ3, the number of shared elements in Tables 1 and 3 ($|A \cap B| = 4$) divided by the unique number of elements ($|A \cup B| = 36$) produced a Jaccard Similarity Index score of 0.111, which is rather low. However, when we expand our view to the datasets as a whole, we see more overlap. Of the Russian keywords in Table 3, 40% of them ($n=8$) are found in the full set of English keywords, and in the top English keywords in Table 1, 70% of them ($n=14$) are found in the full set of Russian keywords.²³ Overlap between bigrams was not possible to calculate in a meaningful way, and was therefore not calculated.

²¹ Part of the phrase *линия роста волос*, ‘hairline’.

²² Literally translated as ‘most chans’.

²³ Because translating all 1,000 Russian keywords would be too laborious for this pilot study, we elected to seek out the top 20 English words in translation in the greater Russian dataset – because the full Russian dataset is not translated, we are unable to provide a Jaccard Similarity Index score for the full keyword dataset.

5 Discussion

At the outset of this study, we sought to uncover the top 20 keywords and phrases used by English- and Russian-speaking incels and compare these words and phrases. In the following sections, we will look in greater detail at the trends in both the English and Russian language data, as well as compare and contrast the data from Tables 1–4.

It must be noted that the primary limitation of this study has been our team's lack of a native Russian speaker – while a number of Russian speakers and Russian language data sources have been consulted over the course of this project, our analysis and any future analysis of Russian-speaking incels would greatly benefit from the knowledge of a native speaker.

5.1 Trends in the English incelosphere

Looking to the English keywords and phrases shown in Tables 1 and 2, we observed the following: 1) English-speaking incels are responsible for a sizeable amount of lexical innovation; 2) a number of words and phrases indicate that a sense of community and in-group identity has taken shape among English-speaking incels; 3) use of profanity – including slurs – is the norm; and 4) there is a thematic focus on social categories, particularly race and gender.

Other than redefining *incel* to be male-as-default, English-speaking incels were early adopters of lexical innovations like *foid* (woman; pejorative); *mog* (to dominate or outclass); *curry* (South Asian person; pejorative); *ascend* (have sex or enter a romantic relationship); and the prolific use of the *-cel* suffix. In Table 1, we see *fakecel*, a non-incel who claims incel group membership, but the full dataset also includes terms like *volcel*, *ricecel*, and *mentalcel*, among others.

This unique lexicon, while cryptic to outsiders, has been extensively documented by community members on the *Incels Wiki* and is, like the presence of the incels.is forum itself, indicative of a strong group identity. The designation of in-group members (*incel*), outsiders (both *normie* and *Chad*), and wannabes (*fakecel*) helps establish this identity. Additionally, the phrase *incel forum* is found in Table 2, representative of not only the existence of a (digital) space for the community to gather, but also meta-commentary about the existence of such a space.

While the presence of profanity is not as extensive as it is in the Russian data, it is still notable due to the presence of pejoratives: the top 20 English keywords include an anti-Black slur (the 'n-word'), as well as *fuck* and *whore*. The first of these words is especially jarring, as the use of this word is both highly salient and highly

taboo in English-speaking communities.²⁴ Other language that could be considered profane or pejorative is also present in Table 1, including *cuck*, the newly-coined pejoratives *curry* and *foid*, and *jfl* (“just fucking laugh” or “just fucking lol”).

The slurs found in the top 20 English keywords are also indicative of the final theme to be discussed with respect to the English data: a focus on society and social categories. Race and gender – as separate issues and in conjunction – are all heavily featured in the results. In addition to the extensive use of the n-word and the pejorative *curry* seen in Table 1, we also see a focus on race within the key phrases in Table 2. For example, seven of the phrases in Table 2 refer to a specific race of women or men (e.g., *black man*, *asian woman*). Gender is also a topic of concern: nine of the key phrases explicitly refer to people by their gender (e.g., *short man*, *ugly man*). Additionally, we observe discussion among English-speaking incels on how women and men treat each other (e.g., *cuck*, *foid worship*, *white knight*) and perceive one another, especially physically (e.g. *ugly*, *good look*, *average look*).

5.2 Trends in the Russian incelosphere

Regarding the Russian keywords and phrases identified in Tables 3 and 4, several trends are apparent. In this discussion, we will focus on the following: an intense focus on physical appearance – often the physical appearance of men – and the prevalence of three types of lexical items: profanity, online jargon, and borrowings from other languages.

The most apparent theme, especially in Table 4, is the scrutiny of physical appearance, particularly that of men: men are concerned with their hairlines (*линия роста*, ‘growth line’; *линия волос*, ‘hairline’), penis size (*размер хуя*, ‘dick size’; *наибольший хуй*, ‘biggest dick’), and height (*пониже рост*, ‘shorter height’). More generally, we also see discussion of body fat (*процентом жира*, ‘fat percentage’), eye shape (*глазе жертвы*, ‘prey eyes’; *зона глаз*, ‘eye area’), and other facial features (*будка чеда*, ‘strong jaw’; *черта лица*, ‘facial feature’). Additionally, some of the language used is typical of an academic register, like the phrase *основная теория* (‘basic theory’) or scientific names for physical features (*максилла*, ‘maxilla’ or ‘jaw’). Additionally, we see a reference to the dark triad (*тёмная триада*), which refers to a psychological theory of three negative personality traits some-

²⁴ While we cannot guess the race of those using this word, use of this slur with a “hard R,” as seen in Table 1, is highly salient and almost always pejorative, unlike the truncated variant that has been reclaimed by in-group members.

times conflated with sexual and financial success. While noteworthy, academic keywords and phrases are outweighed by profanity and slang associated with youth culture.

The profanity in Table 3 may be understated due to the elimination of duplicate word stems, which are numerous due to the nature of Russian profanity (*mat*). There are four lexical roots from which the bulk of *mat* is derived. These are *хуй* ('dick'), *пизда* ('cunt'), *ебать* ('fuck'), and *блять* ('whore'). While we have provided simple translations for these four words in parentheses, the versatility with which they may be used far exceeds that of English profanity and cannot be overstated. *Mat*, known for its ubiquity among the working class and youth culture (Erofeyev 2003), appears to also be ubiquitous among users on 2ch.hk. Not only are each of the four pillars represented in Tables 1 and 2, but five of the omitted words were derived from one of these four words. Additionally, three of the bigrams in Table 4 (*сын шлюхи*, 'son of a whore'; *размер хуя*, 'dick size'; *наибольший хуй*, 'biggest dick') feature one of these words.

Lexical items from youth and internet culture are also prevalent in our results: several top keywords are more modern than the colloquial Russian covered in traditional dictionaries, including *ирл* ('irl; in real life') or *подкатывать* ('hit on'). These words may also come from subcultures based around hobbies closely associated with young people. Examples include *лвл* ('lvl; level'), used to refer to age – used by gamers – or the pair of borrowings *тян* ('chan; young woman') and *кунов* ('kun; young man'), initially used by anime enthusiasts (see *Словарь молодежного сленга*).

Additionally, words originating from online subcultures like incels themselves and 2ch.hk users are also attested in Table 3. Words like *скуфидрон* ('*skufidron*') and *всратый* ('shitted-in') are claimed by *Incels Wiki* users to be "current terminology" in the Russian incelosphere (2022). *Двачую* (lit. 'twice'; 'seconded') is an expression of agreement and a play on the imageboard's name (*Двач*), likely representative of users' affiliation with the site where the /incel/ thread is hosted.

Last among the words we see most frequently in Tables 3 and 4 are borrowings from other languages, primarily English. These words typically surfaced as transliterated jargon used by English-speaking incels (e.g., *betabucks*; *normies*), suggesting that Russian-speaking incels are consuming content created by English-speaking incels. These borrowings appear to be used in a similar context to their English counterparts, and some are so integrated in the speech of Russians-speaking incels that they are incorporated into metaphors, as in the bigram *будка чеда*, seen in Table 4. Literally translated, this phrase means 'Chad's booth' or 'Chad's square,' but concordance lines from CoDEC-M show it being used figuratively to refer to a strong or square jawline.

Also present in the data are one Latin borrowing (*максилла*, a transliteration of ‘maxilla,’ used primarily in the sciences in Russian and English) and two borrowings from Japanese. The Japanese borrowings differ from the English borrowings in that they derive not from lexical words, but from Japanese honorific suffixes for young women and men (*-chan* and *-kun*, respectively). As borrowings, these honorifics are instead used and inflected as nouns by Russian speakers to refer to their corresponding demographic groups, as indicated in Table 3.

5.3 Overlap

Despite the fact that the Jaccard Similarity Index score was quite low, this is not the only means by which we can consider linguistic overlap or even evidence of language contact. This low score is due in part to lexical items without a one-to-one translation, including expressions that are a single word in Russian but translated as a multi-word unit in English (e.g., *нахуй*, ‘fuck it’). To examine other dimensions of overlap between these two datasets, we shall first consider any overlap of the previously-discussed themes, followed by a discussion of loanwords.

Looking first at Tables 1 and 3, we see considerable overlap in the words with the three highest keyness scores: a word that refers to women has the highest keyness score (*foid* in English and *мян* in Russian), *incel* has the second highest score (*инцел* in Russian), and a term for an out-group member holds the third highest score (*normie* in English and *чед* in Russian, both populations often contrasted with incels). This thematic overlap indicates similar topics of conversation between English- and Russian-speaking incels, even though the translations may not be exact. The tendency of Russian-speaking incels to borrow words from the English-speaking incel community also indicates that similar ideologies are circulating among both populations (e.g., married men are used by women for their money – exemplified by *betabucks*). Both datasets also indicate similar attitudes toward profanity – i.e., it is the norm in both spaces – and place emphasis on physical appearance. This is particularly true of height, which 48.5% of incels.is (then incels.co) users perceived as a “factor significantly preventing [them] from finding a partner” (SergeantIncel 2020).

While both datasets focus on physical appearance (possibly in relation to either a desire for or lack of romantic success), the Russian data trended toward a focus on specific physical features, while the English data focused mostly on race. While one might instinctively attribute this to the legacy of whiteness and history of institutionalized racism in the United States, only 38% of incels.is members surveyed in 2020 reported that they were from the United States, so this is unlikely to be the

case.²⁵ It is beyond the scope of this study to posit *why* English-speaking incels discuss race more than Russian-speaking incels, but it may be due to the greater prevalence of American culture – including notions of race and whiteness – on the English-speaking internet.

One final similarity between these two datasets is the presence of words that were coined or popularized in the English-speaking manosphere: *foid*, *incel*, *normie*, *Chad*, *jfl*, *blackpill*, *mog*, *curry*, *ascend*, and *fakecel* are found in the English data, while transliterations of *incel*, *Chad*, *betabucks* (also spelled “betabux”) and *normies* are found in the Russian data. Each of these terms is documented in Moonshot’s “Incels: A Guide to Symbols and Terminology” as being characteristic of the speech of self-identified incels. It is evident that the English-speaking incels are prolific lexical innovators, with 55% (n=11) of the top 20 keywords represented in Moonshot’s *Guide*, but what of the Russian-speaking incels? While 30% of the words in Table 3 are borrowed from English-speaking incels, there are also Russian words in Tables 3 and 4 that could be characteristic of their speech. This particularly true of *бюдка чед* (‘strong jaw’), *опухший* (‘swollen’), *всратый* (‘ugly’), and *скуфидрон* (‘*skufidron*’), because they belong to the same semantic domain where much lexical innovation has occurred with English-speaking incels: physical appearance.

Many of the words Moonshot and other researchers associate with incel culture bleed into mainstream English language internet subcultures, especially those unique to imageboards. So, it is reasonable to suggest that some of the Russian keywords may be similarly characteristic of Russian-speaking incels.²⁶ Consider *лвл* (‘lvl’), *двачую* (‘seconded’), *тян* (‘chan’), and *кун* (‘kun’), all of which are associated with being a user of 2ch.hk (*Словарь молодежного сленга*; Uglova 2020). As a parallel, terms like *kek* and *mang* associated with 4chan users can be found in the full English results, and may have similarly developed an association with English-speaking incels through repeated use.

If these words are all considered part of the Russian incel cryptolect, the keywords in Table 3 would consist of 13 words (65% of keywords) unique to or characteristic of incels, including borrowings. This is a similar ratio of markedly incel-adjacent terms to the English keyword data, and suggests that Russian-speaking incels

²⁵ Incidentally, the majority of users (42.8%) were from Europe, per SergeantIncel, “Survey Results”.

²⁶ Additionally, while some uses of *mat* may also be specific to Russian-speaking incels, the complexity of Russian profanity and its diversity of meaning suggests that without an L1 or near L1 Russian speaker immersed in contemporary youth culture, we lack the cultural context to determine whether or not this is the case.

are similarly prolific with respect to lexical innovation, whether it is through borrowing from English, adaptation from other internet subcultures, or other means.

6 Conclusion

While there is not a significant overlap of the top keywords used by English- and Russian-speaking incels at first glance, there is significant thematic overlap between the keywords and phrases collected. These include extensive use of profanity, language associated with mainstream online youth culture, and a fixation on physical appearance – particularly that of men. Additionally, the data from both communities contains lexical items that have become integrated into their respective communities that are comparatively absent from mainstream discourse, suggesting that both incel communities are lexically productive. Generally, the discussion of English-speaking incels seems to skew more toward discussing women, race, and community membership, while the Russian data suggests a more intense focus on physical appearance. Finally, the Russian data also includes a significant number of borrowings from English-speaking incels, indicative of language transfer between English- and Russian-speaking incel communities.

Our cross-linguistic comparison of incel communities illustrates some degree of contact between the English- and Russian-speaking incel communities. This language transfer, exhibited by the numerous borrowings in the Russian data, suggests a shared, toxic ideology between these two groups. The ideologies of both incel communities center around obsession with physical appearance and discontent with society, as demonstrated by the keywords and bigrams. One can also observe overlap between the themes discussed by the English-speaking incel community and rhetoric from the manifestos of prominent violent criminals affiliated with the subculture, particularly with respect to race and women (further discussed in O'Donnell 2021). As events like the 2018 Toronto van attack have shown, this discontent has been directed at out-group members such that it has a death toll.

While large-scale acts of misogynist terrorism like the mass shootings seen in North America are not yet attested in predominantly Russian-speaking areas, we have established that these communities are not bereft of misogynist violence: consider the targeted harassment campaigns of the Russian Male State movement described in Section 2.4. It is not within the scope of this study to gauge the capacity for violence of a group based on its speech, but the noted similarities are certainly worthy of notice as these two groups continue to operate in parallel. Mass media does not seem to reflect the same level of concern about ideologically-motivated violence from incel communities in larger Russian-speaking society compared to

the current concern about incels in English-speaking communities – North America, for example. Contrary to this, the current study illustrates that both groups of incels have comparable grievances, hold comparable views about society, and deserve comparable amounts of attention.

More in-depth study of the speech of Russian-speaking incels and their lexis would prove useful to the study of international incel movements; while we have provided a glossary of some salient terms that appeared in the Russian data on this project's OSF page, the full Russian dataset could be further harnessed to put together a wordlist of key lexical items used by Russian-speaking incels. Lastly, we would like to emphasize the importance of further cross-linguistic studies of non-English speaking incels with other non-English speaking communities and with English-speaking incels. Such studies could shed light on patterns across the greater incel subculture: how similar is the rhetoric, what level of lexical overlap can we find, and what themes emerge?

References

- Baele, Stephane, Lewys Brace & Debbie Ging. 2023. A Diachronic Cross-Platforms Analysis of Violent Extremist Language in the Incel Online Ecosystem. *Terrorism and Political Violence* 1–24. <https://doi.org/10.1080/09546553.2022.2161373>.
- Barker, Gary, Hayes Caroline, Heilman Brian & Reichert Michael. 2023. *The State of American Men: From crisis and confusion to hope*. Washington, DC: Equimundo.
- Bellingcat. 2021. Meet the Male State: Russia's Nastiest Online Hate Group. *Bellingcat*. <https://www.bellingcat.com/news/uk-and-europe/2021/10/20/meet-the-male-state-russias-nastiest-online-hate-group/> (last accessed 28 March 2024).
- Bosman, Julie, Kate Taylor & Tim Arango. 2019. A Common Trait Among Mass Killers: Hatred Toward Women. *The New York Times*. <https://www.nytimes.com/2019/08/10/us/mass-shootings-misogyny-dayton.html> (last accessed 28 March 2024).
- Chokshi, Niraj. 2018. What Is an Incel? A Term Used by the Toronto Van Attack Suspect, Explained. *New York Times*. <https://www.nytimes.com/2018/04/24/world/canada/incele-reddit-meaning-rebellion.html> (last accessed 28 March 2024).
- Costa, Luciano da F. 2021. Further Generalizations of the Jaccard Index. *arXiv*. <https://arxiv.org/abs/2110.09619> (last accessed 28 March 2024).
- Cunha, Darlena. 2020. Red pills and dog whistles: It is more than 'just the internet'. *Al Jazeera*. <https://www.aljazeera.com/opinions/2020/9/6/red-pills-and-dog-whistles-it-is-more-than-just-the-internet> (last accessed 28 March 2024).
- Czerwinsky, Alysa. 2023. Misogynist incels gone mainstream: A critical review of the current directions in incel-focused research. *Crime, Media, Culture: An International Journal*. <https://doi.org/10.1177/17416590231196125>.
- Daly, Sarah E. & Albina Laskovtsov. 2021. "Goodbye, My Friendcels": An Analysis of Incel Suicide Posts. *Journal of Qualitative Criminal Justice & Criminology*. <https://doi.org/10.21428/88de04a1.b7b8b295>.

- DiBranco, Alex. 2017. Mobilizing Misogyny. *Political Research Associates*. <https://politicalresearch.org/2017/03/08/mobilizing-misogyny> (last accessed 28 March 2024).
- Dignam, Pierce Alexander & Deana A. Rohlinger. 2019. Misogynistic Men Online: How the Red Pill Helped Elect Trump. *Signs: Journal of Women in Culture and Society* 44(3). 589–612. <https://doi.org/10.1086/701155>.
- Erofeyev, Victor. 2003. Dirty Words. *The New Yorker*. <https://www.newyorker.com/magazine/2003/09/15/dirty-words-2> (last accessed 28 March 2024).
- Fernquist, Johan, Björn Pelzer, Katie Cohen, Lisa Kaati, & Nazar Akrami. 2020. Hope, cope & rope: Incels i digitala miljöer [Hope, cope & rope. Incels in digital environments]. Memo 7040. Totalförsvarets forskningsinstitut (FOI).
- Gaufman, Elizaveta. 2022. It's a male, male world: rise and fall of "the Male State", a far-right misogynistic organization in Russia. *C-REX - Center for Research on Extremism*. <https://www.sv.uio.no/c-rex/english/news-and-events/right-now/2022/it%E2%80%99s-a-male-male-word-rise-and-fall-of-%E2%80%9Cthe-male-s.html> (last accessed 28 March 2024).
- Ging, Debbie. 2019. Alphas, Betas, and Incels: Theorizing the Masculinities of the Manosphere. *Men and Masculinities* 22(4). 638–657. <https://doi.org/10.1177/1097184X17706401>.
- Gothard, Kelly Caroline. 2021. *The incel lexicon: Deciphering the emergent cryptotext of a global misogynistic community*. Burlington, VT: University of Vermont MS Thesis. <https://scholarworks.uvm.edu/graddis/1465/> (last accessed 28 March 2024).
- Griffin, Jonathan. 2021. Incels: Inside a dark world of online hate. *BBC*. <https://www.bbc.com/news/blogs-trending-44053828> (last accessed 28 March 2024).
- Human Rights Watch. 2017. Russia: Bill to Decriminalize Domestic Violence. *Human Rights Watch*. <https://www.hrw.org/news/2017/01/23/russia-bill-decriminalize-domestic-violence> (last accessed 28 March 2024).
- Incels. 2017. Rules and FAQ. Forum post. *Incels.is*. archive.ph/PcMw8 (last accessed 14 March 2024).
- Incels Wiki. 2022. "Russian incelosphere". <https://archive.ph/bsqg3> (last modified 27 April 2023).
- Kilgariff, Adam. 2009. Simple Maths for Keywords. In *Proceedings of Corpus Linguistics Conference CL2009*. University of Liverpool, UK. <https://api.semanticscholar.org/CorpusID:124884692> (last accessed 28 March 2024).
- Kilgariff, Adam. 2012. Getting to Know Your Corpus. In Petr Sojka, Aleš Horák, Ivan Kopeček & Karel Pala (eds.), *Text, Speech and Dialogue*. Lecture Notes in Computer Science 7499, 3–15. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-32790-2_1.
- Lokot, Tetyana. 2019. Affective Resistance Against Online Misogyny and Homophobia on the RuNet. In Debbie Ging & Eugenia Siapera (eds.), *Gender Hate Online*, 213–232. London: Palgrave Macmillan. https://doi.org/10.1007/978-3-319-96226-9_11.
- Meduza. 2020. Паблик «Мужское государство» заблокировали за «призывы к насильственным действиям» [The "Male State" page has been taken down due to "calls for violence"]. *Meduza*. <https://meduza.io/news/2020/07/01/pablik-muzhskoe-gosudarstvo-zablokirovali-za-prizyvyy-k-nasilstvennym-deystviyam> (last accessed 28 March 2024).
- Moonshot. 2020. *Incels: A Guide to Symbols and Terminology*. <https://moonshotteam.com/resource/incels-a-guide-to-symbols-and-terminology/> (last accessed 28 March 2024).
- Moreno-Ortiz, A. 2024. Keywords. In *Making Sense of Large Social Media Corpora: Keywords, Topics, Sentiment, and Hashtags in the Coronavirus Twitter Corpus*, 59–102. Cham, Switzerland: Palgrave Macmillan. <https://doi.org/10.1007/978-3-031-52719-7>.
- National Threat Assessment Center. 2023. *Mass Attacks in Public Spaces: 2016–2020*. U.S. Secret Service. <https://www.secretservice.gov/sites/default/files/reports/2023-01/usss-ntac-maps-2016-2020.pdf> (last accessed 8 March 2024).

- Novitskaya, Alexandra. 2017. Patriotism, sentiment, and male hysteria: Putin's masculinity politics and the persecution of non-heterosexual Russians. *NORMA* 12(3–4). 302–318. <https://doi.org/10.1080/18902138.2017.1312957>.
- O'Donnell, Kelly M. 2021. Incel Mass Murderers: Masculinity, Narrative, and Identity. *Ohio Communication Journal*. Ohio Communication Journal 59. 64–76. <https://stanford.idm.oclc.org/login?url=https://search.ebscohost.com/login.aspx?direct=true&site=eds-live&db=ufh&AN=150697570> (last accessed 28 March 2024).
- Pelzer, Björn, Lisa Kaati, Katie Cohen & Johan Fernquist. 2021. Toxic language in online incel communities. *SN Social Sciences* 1(8). 213. <https://doi.org/10.1007/s43545-021-00220-8>.
- Preston, Kayla, Michael Halpin & Finlay Maguire. 2021. The Black Pill: New Technology and the Male Supremacy of Involuntarily Celibate Men. *Men and Masculinities* 24(5), 719–909. <https://doi.org/10.1177/1097184X211017954>.
- Quiroz, Joselyne. 2022. “Life Wasn’t Supposed to be This Way”: *Involuntary Celibacy Among Young Men Online*. Chicago, IL: University of Chicago MA Thesis.
- SergeantIncel. 2020. Survey Results - March 2020. Forum post. *Incels.co*. archive.vn/Vde6H (last accessed 14 March 2024).
- Sketch Engine. 2018. Keywords and terms – lesson. *Sketch Engine*. <https://www.sketchengine.eu/quick-start-guide/keywords-and-terms-lesson/> (last accessed 28 March 2024).
- Šljachov, Vladimir I. & Eve Adler. 2006. *Dictionary of Russian slang and colloquial expressions*. 3rd edition. Hauppauge, NY: Barron's.
- Solea, Anda Iulia & Lisa Sugiura. 2023. Mainstreaming the Blackpill: Understanding the Incel Community on TikTok. *European Journal on Criminal Policy and Research* 29(3). 311–336. <https://doi.org/10.1007/s10610-023-09559-5>.
- Speckhard, Anne, Molly Ellenberg, Jesse Morton, & Alexander Ash. 2021. Involuntary Celibates’ Experiences of and Grievance over Sexual Exclusion and the Potential Threat of Violence Among Those Active in an Online Incel Forum. *Journal of Strategic Security* 14, 2. 89–121. <https://doi.org/10.5038/1944-0472.14.2.1910>.
- Sprengrer, Robin. 2014. *Männliche Absolute Beginner: Ein kommunikationswissenschaftlicher Ansatz zur Erklärung von Partnerlosigkeit* [Absolute Beginner Men: a Communication Science Approach to an Explanation for Partnerlessness]. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-05924-8>.
- Taylor, Jim. 2018. The woman who founded the “incel” movement. *BBC*. <https://www.bbc.com/news/world-us-canada-45284455> (last accessed 28 March 2024).
- Townsend, Mark. 2022. Experts fear rising global ‘incel’ culture could provoke terrorism. *The Observer*. <https://www.theguardian.com/society/2022/oct/30/global-incel-culture-terrorism-misogyny-violent-action-forums> (last accessed 28 March 2024).
- Uglova, Julia. 2020, May 30. Вы знаете больше языков, чем кажется. Интернет заставил вас выучить их [You know more languages than you think. The internet made you learn them]. *hi-tech* (30 July, 2024).
- Venuti, Marco & Antonio Fruttaldo. 2019. Contrasting News Values in Newspaper Articles and Social Media: A Discursive Approach to the US Ruling on Same-Sex Marriage. In Barbara Lewandowska-Tomaszczyk (ed.), *Contacts and Contrasts in Cultures and Languages*. Second Language Learning and Teaching, 147–161. Cham, Switzerland: Springer International Publishing. https://doi.org/10.1007/978-3-030-04981-2_11.
- Voroshilova, Anzhelika I. & Dmitriy O. Pesterev. 2021. Russian Incels Web Community: Thematic and Semantic Analysis. In *2021 Communication Strategies in Digital Society Seminar (ComSDS)*, 185–190. St. Petersburg, Russia: IEEE. <https://doi.org/10.1109/ComSDS52473.2021.9422872>.

Словарь молодежного сленга. Двачую (Dvachuyu). *Словарь молодежного сленга [Dictionary of Youth Slang]*. <http://www.terminy.info/jargon/dictionaries-of-teen-slang/dvachuyu> (last accessed 28 March 2024).

Словарь молодежного сленга. Кун (Kun). *Словарь молодежного сленга [Dictionary of Youth Slang]*. <http://www.terminy.info/jargon/dictionaries-of-teen-slang/kun> (last accessed 28 March 2024).

Словарь молодежного сленга. Лвл (Lvl). *Словарь молодежного сленга [Dictionary of Youth Slang]*. <http://www.terminy.info/jargon/dictionaries-of-teen-slang/lvl> (last accessed 28 March 2024).

Словарь молодежного сленга. Тян тянка (Tyan tyanka). *Словарь молодежного сленга [Dictionary of Youth Slang]*. <http://www.terminy.info/jargon/dictionaries-of-teen-slang/tyan-tyanka>. (last accessed 28 March 2024).