

Anthony Verardi
 LING 2235 Fall 2019
 Language, Gender, and Society
 Dr. Scott Kiesling

TheRedPill: Examining a Breeding Ground for Toxic Gender Ideologies

I. Introduction

Since the proliferation of various options for digital communication electronically, from chat forums (Baudinette, 2017) to in-game team-speak (Collister, 2017) to text messaging and more (Gershon, 2010; Webster, 2018), there has been no shortage of interest from the scholarly world in investigating the virtual delivery of speech. This interest, however, has generally only led to investigating digital landscapes that could be considered safe or low-stakes,¹ as in any of the above. Linguists have yet to undertake scholarly analyses of language used in dangerous and toxic “male supremacist” or racist spaces, including online chat forums, blogs, and manifestos left by the perpetrators of large-scale violence against gendered, racial, and/or ethnic groups.

The current study delves into the written speech of a community called “TheRedPill,” (hereafter “TRP”) hosted on the website Reddit. TRP’s purpose is to “discuss **men's identity**, sexual strategy, and options in the context of our current global culture **for the benefit of men**,” as stated in their community details. Formed in 2012, TRP is one of many “subreddits,” communities devoted to content around one overarching theme, that is populated by members of “incel” and Men’s Rights Activism (MRA) groups. Incel is a contraction of “involuntarily celibate” coined in 1997 (Taylor, 2018) by a lonely woman struggling to connect with men on a romantic level, and in fact originally included “anybody of any gender who was lonely, had never had sex or who hadn't had a relationship in a long time.” The term was later hijacked by men who felt they had been wronged by women who considered them undatable. Men’s Rights Activism, relatedly, is a pseudo-intellectual movement that holds Feminism in

¹ This is not meant to suggest that posting in an online dating forum for Queer people is not safe or low-stakes, but rather that *the speech* occurring in those spaces is safe and low-stakes compared to the use of language observed in incel and MRA spaces

high contempt, blaming the emasculation of men as a central element of their woes. While acknowledging that a rare few Men's Rights Activists voiced legitimate grievances, the Southern Poverty Law Center includes them under the blanket hate ideology of male supremacists.²

The TRP community can be considered a hotbed for an ideology of "toxic masculinity," a concept most famously attributed to Connell and Messerschmidt (2005). Language is a crucial element of this subreddit, as evidenced by the fact that they maintain a community glossary.³ Because the TRP community has been quarantined by Reddit (which means accessing it requires consenting to view "shocking or highly offensive content"), it is impossible to know how many members it currently has; a 2017 estimate placed it at around 200,000⁴ members and growing. Despite the quarantine, the community is still very much active to this day.

In spite of multiple instances of mass femicide attributed to the incel community (for example, the 2014 Isla Vista killings and both the 2018 Tallahassee shooting and Toronto van attack), there remains a gap in scholarly literature on the particular speech habits of this small-but-vocal demographic, or any like it. The relationship between language and ideology has been explored in detail by Eckert (2008), who posits that "variation constitutes an indexical system that embeds ideology in language and that is in turn part and parcel of the construction of ideology." Previous linguistic work on chat forums (Baudinette, 2017) has examined how language is used by speakers of Japanese to establish both their own (in this case, gay) identity and their ideal partner's. Baudinette's conclusions reinforce the notion that although this type of virtual speech is distanced from its speaker by both anonymity (Collister, 2016) and temporality (Ochs, 1992), it is nonetheless a way in which people can both reify their own self-conceptualizations and affirm how they conceptualize others, either as individuals or as broad categories.

² <https://www.splcenter.org/fighting-hate/extremist-files/ideology/male-supremacy>

³ https://www.reddit.com/r/TheRedPill/comments/2zckqu/updated_glossary_of_terms_and_acronyms/

⁴ <https://www.newstatesman.com/science-tech/internet/2017/02/reddit-the-red-pill-interview-how-misogyny-spreads-online>

To this end, I have approached the current analysis with the following research question in mind: how is gendered language used by members of the TRP community, particularly regarding what types of ideologies this language is indexical of? In other words, what are the hallmarks of the gendered language employed in these forums, and what might they tell us about how members of this speech community construct identities for themselves and others? Given this community's status as a breeding ground for ideologies of toxic masculinity, it stands to reason that their speech would evidence a negative bias towards women.

II. Method

The present analysis was conducted by using Python to scrape the top 1,000 posts of all time from TRP, which were then saved into an Excel file for analysis as well as processed using the Natural Language Toolkit (NLTK). The sample size limit was set at 1,000 posts due to a limitation enforced by Reddit's Application Programming Interface (API) that only allows for up to 1,000 posts to be pulled at a time. "Top posts" have been defined here in terms of the number of "upvotes," or "likes," that a post has received, dating from the forum's creation in 2012 until the day the data were scraped. The following items were collected for each post:

1. The title of the post
2. The number of upvotes the post had received over time
3. The number of comments the post had received over time
4. The body text of the post itself

Because Reddit does not attach information about the poster's identity to their posts, no demographic data is available for the participants behind these data points. It should, however, be safe to generalize that they are uniformly masculine-identifying persons and largely, if not also uniformly, heterosexual (or at the least, attracted to women). Otherwise, no assumptions should be made about their ages, races,

ethnicities, nationalities, socioeconomic background, level of education, geographic locations, or other demographic factors.

After gathering the textual data items and saving them as separate variables in Python, the titles and body paragraphs were then tokenized into individual words and punctuation marks using NLTK's built-in word tokenize function. The titles amounted to a raw total of 14,412 word tokens from 3,228 unique word types, while the text of the body paragraphs included 1,123,967 individual word tokens from 35,981 unique word types.

From the raw text, the data were then further cleaned by factoring out 179 common "stopwords" (see Appendix A for the full set) that are highly frequent in natural language corpora. Furthermore, non-alphanumeric symbols such as punctuation were also purged from the data for analysis using NLTK, since their inclusion can skew token counts. This was done in order to provide a better idea of how often a number of target words (masculine and feminine address terms) occurred in relation to the non-stopword, non-alphanumeric text. It is important to note here that this culled set of word tokens was only used for examining the frequencies of specific words, not the textual analysis; the actual Excel data remained untouched by this process. After removing stopwords and non-alphanumeric characters, there were 6,966 title words (48.33% of the original) and 437,889 body words (38.96% of the original), totaling 2,891 and 24,824 unique token types, respectively.

III. Results

Based on gendered address terms uncovered during an earlier project, the following lists of masculine and feminine words were initially considered for analysis. Their resulting frequencies in the trimmed corpus are presented accordingly in Table 1.

After tabulating these words that had been pre-selected for analysis, it seemed advisable to let the data speak for itself. To that end, the top 50 most common words from titles and body texts were pulled along with their frequencies, as presented in Tables 2 and 3 below. This has the dual benefits of

both bringing to the surface any interesting trends that may otherwise have been neglected and helping to mitigate my own confirmation bias (Kiesling, 2019) from seeping into the conclusions of the present analysis. In examining the word tokens based on their frequencies in the data alone and setting an arbitrary limit that can easily be adjusted, an element of intentional curation is removed from the process.

Table 1. Gendered address terms and frequencies within the cleaned corpus

TRP Masculine Word Frequencies			TRP Feminine Word Frequencies		
Word Token	Titles (6,966)	Bodies (437,889)	Word Token	Titles (6,966)	Bodies (437,889)
Man	48	1691	Woman	45	1781
Dude	2	136	Bitch	6	226
Guy	36	1236	Girl	33	1611
Bro	1	50	Lady	0	35
Dick	6	233	Slut	5	83
Boy	6	102	Cunt	1	20
Sir	1	15	Whore	3	80
Brother	1	46	Chick	2	90
Douche	0	23	Ma'am	0	0
Gentleman	0	9	Female	20	366
Male	34	638	Honey	0	23
Asshole	2	127	Babe	0	13
Daddy	0	33	Miss	1	54
			Gal	0	12
			Girlfriend	15	215

Table 2. Top 50 most common title word tokens and their frequencies from TRP

women	137	guide	34	new	24	feminist	19	story	16
Men	71	male	34	time	23	married	18	husband	16
Get	57	girl	33	years	22	guys	18	frame	16
Red	51	life	30	gets	22	stop	18	divorce	16
Pill	50	trp	28	rape	22	good	17	fucking	16
Man	48	shit	28	girls	21	sexual	17	wife	16
Woman	45	make	27	year	20	beta	17	3	15
Guy	36	reddit	25	female	20	way	17	girlfriend	15
Like	35	want	25	one	19	game	16	things	15
Sex	35	never	25	getting	19	fuck	16	date	15

Table 3. Top 50 most common body word tokens and their frequencies from TRP

women	4509	woman	1781	think	1319	something	1167	could	983
like	3455	even	1753	never	1273	back	1166	right	955
get	3104	life	1744	see	1265	work	1117	day	950
one	2549	man	1691	things	1236	feel	1112	also	947
men	2496	good	1683	guy	1236	take	1089	guys	944
time	2407	know	1631	sex	1235	fuck	1080	well	887
want	2183	girl	1611	shit	1230	really	1075	always	868
people	1940	go	1578	need	1211	first	1074	give	852
would	1858	way	1472	much	1206	girls	1064	around	852
make	1824	going	1384	say	1186	every	1014	many	846

Lastly, based on work in corpus linguistics by Baker (2015), a number of terms were selected from across all three tables for an examination of their discursive environments. They were then used to produce concordances from within the text in order to examine the wider speech contexts in which they were being used. For the titles, these terms included “women”, “men”, “male”, “female”, “bitch”, and “dick”. These tokens were chosen out of a desire to include two pairs of “neutral” terms (women, men; female, male) and one “negative” set (bitch, dick). Three concordances of 100 characters were selected by hand from among the candidates for each of these words, as presented below in Table 4. Items have been edited for presentability, but not content; in some cases, the program used to select concordances returned prior or following title text as well, which has been trimmed from either side when necessary. In some cases, text has been added to either side to lend a more complete sense of the message.

Table 4. Concordances in post titles

three ways to consciously manipulate	women	before they subconsciously manipulate you.
Australian trial finds POSITIVE bias towards	women	in employment, government stops trial...
the guide to accepting yourself (even when	women	don't)
if you weren't sure there's a war on	men	, check it out: “hey, white guys: new year’s resolutions” -MTV
Fat girls to	men	is equivalent to what beta males are to females.
Male birth control trial was cancelled for	men	. Feminists beta-shame and find a way to
being unsafe to		make it all about themselves
I was asked to train a new	female	employee to do the promotion i got passed
		on
domineering, manipulative	female	coworker was trying to edge me out by
		flirting and ingratiating with Boss

drake university expels male student despite	female	admitting she forced him into sex act
Manipulative, entitled, feminist DA threatens her	male	uber driver - gets fired
sues Google, alleging intolerance of white	male	conservatives
THE BEST OF RATIONAL	male	YEARS ONE/TWO/THREE (If you haven't read this yet, do it now)
List of 4 books you've never read that will mercilessly rip the little	bitch	right out of your soul
How to not be a	bitch	: 2019 edition
"There is nothing I can say to you that will magically make you not a	bitch	" -GayLubeOil
be careful where you stick your	dick	, gentlemen.
Accept that all women desire to have a	dick	inside of them, no matter how cute and innocent they seem to be.
assume women want your	dick	.

Here, the examples that have been presented seem to be largely illustrative of this group's particularly misogynist ideology. If an example that was flattering to a particular gendered group had been uncovered in the title data, I would have made certain to present it here; however, no such cases were discovered in the title search. Similarly, three concordances each were selected for the same words in post bodies, as displayed in Table 5.

Table 5. Concordances in post bodies

also, higher agreeableness. this leads to	women	generally having a harder time negotiating salaries
I wanted things to make sense, but I found myself alienated. Not only by	women	and romantic interests, but also by friends and family.
If there is one thing that turns	women	off more than anything, it's neediness.
Because women are more dependent on	men	for survival than visa versa, they have collectively developed
overt communication - women are not like	men	, they communicate their intentions covertly.
so women often use this rational phase to trick	men	into doing things for them. holding frame
he took the ring bsck thst night. as always,	female	requ
it's part of the natural order. innate	female	hypergamy knows no bounds. it is a biological
like cardi b & nicki minaj promote the dominate	female	bitchiness existed long before feminism and how being a whore is how to get ahead.
evolved to obey and be obsessed with the alpha	male	because they need his protection and resources
need emotional support, consult a close trusted	male	friend. true brotherhood is out there. don't

1. make sure your study partner is an organized	male	who wants to get an a in class . absolutely avoid
sexually ungenerous, lazy, or just an entitled	bitch	that expects you to do all kinds of crap for her
ng around with your dick, you were her little	bitch	. You spent all day doing shit and keeping her
make her crave your attention. oh, and this	bitch	is completely lying about there not being anyw
or etc because they won't help you get your	dick	wet. they are human too. they have the freedom
attached to your dick. don't go waving your	dick	around unless it's explicitly asked for.
2. Do not be a dick for the sake of being a	dick	. Teasing girls should be light-hearted and fun

Since there were more body concordances to choose from for each word type, the task of selecting tokens for examination took on an element of randomness: items were selected by scrolling through 1,000 potential candidates, stopping at random points, and scanning the examples briefly. It became necessary to sample from this many lines because of how the program returns concordances, namely in linear order starting from the top post and going to the bottom. This meant that selecting from only the first ten or so concordances would have limited the search to one document, which I did not consider a methodologically valid sample for this purpose. Here, some examples were found that seemed to contain less toxic thought, and as such were included in the sampling. Generally, one concordance per word was pulled from near the top, one from the middle, and one from the bottom of the 1,000 concordance items.

IV. Discussion

Glancing only at the word tokens on their own, it would seem as though my initial hypothesis that members of TRP would write with a clear negative bias towards women is demonstrably false. In fact, after collapsing singular and plural forms into single categories, terms such as “male” and “female,” which invoke a biological sense, are the third most prominent tokens after “woman”, “man”, “guy” and “girl.” The latter four tokens do not carry any overt hostility in either form or function, unlike the profane terms “dick” and “bitch”. The frequencies of “dick” and “bitch” relative to any of the prior six

terms are notably smaller, as are the frequencies of “asshole”, “douche”, “slut”, “cunt”, and “whore”.

For the top 1,000 posts in a known MRA forum, there is decidedly little use of profanity or vulgarity across titles and body texts. *All* of the negative word tokens noted in the present analysis occur with less frequency than any of the more generic, neutral ones. The top tokens from the titles and bodies presented in Tables 2 and 3, in which neither “bitch” nor “dick” are found, corroborate this observation.

This finding may lead one to conclude that while the members of this speech community *are* perhaps committed to the fallacy of innately biological and insurmountable differences between masculine and feminine persons (Cameron, 2007),⁵ given their high usage of gendered language in titles, their speech is not more negative than that of the population at large. This is also a situation in which domain specificity plays a role, given that the subreddit is specifically focused on “men’s identity [and] sexual strategy.” One of the original goals of the present analysis was to compare the gendered address terms used by members of TRP against a sample taken on a college campus, but it quickly became clear that comparing such a large set of data alongside a smaller set was impractical. However, in examining the usage of these words in context, an understanding of the ideological weight behind their speech begins to coalesce even without a comparative sample.

Eckert’s concept of the *indexical field* (2008) is particularly useful here in contemplating the various (and at times contradictory) indexicalities of the selected word tokens above. As Eckert defines it, the indexical field is a “constellation of *ideologically related meanings* [emphasis added], any one of which can be activated in the situated use of the variable,” the variable here being a given word token and the context in which it occurs; in a sense, a noun and the adjectives or other modifying phrases that appear to collocate with it. From the concordances outlined in Tables 4 and 5, the following labels can be used to begin forming a rough sketch of an indexical field for the TRP data: “Manipulative,” “Victim,”

⁵ A friend with whom I was discussing my data quipped that it was “like [I had] fed a text generator random 19th-century naturalist texts and subbed certain words for “women” and “betas.”

“Domineering,” “Submissive,” “Unwitting,” and “Rational.” Half of these labels (“Manipulative,” “Domineering,” and “Rational”) emerged directly from concordances of gendered terms in either the titles or the body texts, the others being instead inferred from them (“Victim,” “Submissive”, and “Unwitting”). Despite being posted by different authors, there appear to be uniform indexical meanings behind how words like “female” and “male” are used in context, where “female” would be positioned in the indexical field at the intersecting point of “manipulative,” “domineering,” and, perhaps counterintuitively, “unwitting.” Male, conversely, is positioned with “victim,” “submissive,” and “rational,” although it may also tend towards the “manipulative” edge as well (as seen in “three ways to consciously manipulate women before they subconsciously manipulate you”). Manipulation in this sense, though, is characterized as a preemptive strike against a predeterminedly hostile other, not an act of aggression.

Most crucially, however, is the realization that these seemingly “neutral” terms are loaded with indexical weight by their surroundings and should accordingly not be taken as innocuous based on a consideration of their form alone. As Eckert (2008) notes, the denotative value of a lexeme can “absorb connotations through association with aspects of the context in which it is used and, most certainly, stances.” Based on these data, it appears that for members of the TRP community, gendered address terms have undergone a radically substantial change from how non-incels/MRAs regard these terms, positioning men as both eternally victimized and rationally superior to all others and women as conniving, scheming harpies who are somehow nonetheless incompetent under scrutiny.

Astonishingly, there *were* posts within the top 1,000 that contained language like the following and seemingly sought to contradict the dominant narrative of misogyny within the community (albeit still in the pursuit of seducing women):

“Do not put down or insult women for not sleeping with you. They are not evil, prudes, bitches, cunts, sluts, whores, or etc because they won't help you get your dick wet. They are human too. They have the freedom of choice and numerous of men to choose from. It is just a different world for women.”

“Don’t go waving your dick around unless it’s explicitly asked for. Don’t be that dick. Dicks are like pens. Everyone’s got one, most people’ll loan you one if you ask for it and unless it’s super unique, nobody cares about it.”

The first of these examples was in fact pulled from the 7th most popular post overall, with 2,393 upvotes and 317 comments. Looking for cases of speech that go against my own preconceived ideas of how this community thinks and behaves is not done in the enterprise of defending or excusing their speech habits, but rather as an endeavor on my part as a researcher to present these findings with some sense of fairness and objectivity. It does no good to try and argue with people who purport to be “objective” and “rational” without making an attempt at doing the same yourself.

V. Conclusion

For the general public, instinct may be to accept a claim that because members of TRP (or any other prominent male supremacist) are using such outwardly neutral terminology, they are not saying anything harmful; a sociolinguistic analysis should rightly tear this assertion apart. While the language used in these communities may read at a glance as relatively tame, worthy only of a quarantine and not an outright ban, the total toxicity of this community’s speech patterns is greater than the sum of its word token parts. The current analysis may best be regarded as an exploratory first step into looking at how language and toxic masculinity interact, even given a number of limitations at hand.

First and foremost, the texts scraped for the current analysis are all presentational forms and not necessarily meant to be indicative of interaction between interlocutors. For this type of interactional content, the comments would have to be scraped and maintained in a corpus alongside the original posts, and as a future direction, I would be very much interested in traveling this route. Technical limitations owing to my own limited experience with coding have also had an impact on the robustness of the analysis at hand; while Baker (2015) outlines other ways to consider semantic meaning in corpora besides concordances (namely two methods of analyzing collocations), I was unable to accomplish this with the computational resources at my disposal. A collocation analysis would ostensibly be highly

elucidating for the topic at hand, given that so much of my analysis hinges on the idea that gendered nouns are pulling in additional meaning from their neighboring adjectives and other phrases.

All things considered, I feel the need to acknowledge that for me, looking at these types of difficult, unpleasant, and depressing communities is done in the service of social progress, not solely for academic merit in its own right. Regarding the gap in the literature on explicitly harmful, toxic, and violent masculinity and language, it is my hope that socially-minded linguists (not necessarily even sociolinguists by training) will step up to try and propose mechanisms for change by way of language policy. The rules governing speech on the internet sit at an intersecting point of language and the law, and this current work is only one small part of a much grander question about how we conduct ourselves in public discourse.

VI. References

- Baker, P. (2015). Two Hundred Years of the American Man. In T. M. Milani (Ed.), *Language and Masculinities : Performances, Intersections, Dislocations*. New York, NY: Routledge.
- Baudinette, T. (2017). Constructing identities on a Japanese gay dating site: Hunkiness, cuteness and the desire for heteronormative masculinity. *Journal of Language and Sexuality*, 6(2), 232-261.
doi:<https://doi.org/10.1075/jls.6.2.02bau>
- Cameron, D. (2007). *The myth of Mars and Venus*. Oxford and New York: Oxford University Press.
- Collister, L. (2016). "At least I'm not Chinese, gay, or female": Marginalized voices in World of Warcraft. In L. Squires (Ed.), *English in Computer-Mediated Communication: Variation, Representation, and Change*. Boston, MA: De Gruyter Mouton.
- Connell, R. W., & Messerschmidt, J. W. (2005). Hegemonic Masculinity: Rethinking the Concept. 19(6), 829-859. doi:[10.1177/0891243205278639](https://doi.org/10.1177/0891243205278639)
- Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, 12(4), 453-476.
doi:[10.1111/j.1467-9841.2008.00374.x](https://doi.org/10.1111/j.1467-9841.2008.00374.x)

- Gershon, I. (2010). Breaking Up Is Hard To Do: Media Switching and Media Ideologies. *Journal of Linguistic Anthropology*, 20(2), 389-405. Retrieved from www.jstor.org/stable/43104270
- Kiesling, S. F. (2019). *Language, gender, and sexuality: an introduction*. Abingdon, Oxon and New York, NY: Routledge.
- Ochs, E. (1992). Indexing gender. In A. Duranti & C. Goodwin (Eds.), *Rethinking context: Language as an interactive phenomenon*: Cambridge University Press.
- Taylor, J. (2018). The woman who founded the 'incel' movement. *BBC Radio 5 Live*. Retrieved from <https://www.bbc.com/news/world-us-canada-45284455>
- Webster, L. (2018). "I wanna be a toy": Self-sexualisation in gender-variant Twitter users' biographies. *Journal of Language and Sexuality*, 7(2), 205-236. doi:<https://doi.org/10.1075/jls.17016.web>

VII. Appendix

A. Stopwords

{ 'out', 'yourselves', 'where', 'ain', 'both', 'we', 'against', 'shan't', 'from', 'that'll', 'needn't', 'him', 'are', 'too', 'didn't', 'ourselves', 'between', 'that', 'most', 'now', 'in', 'be', 'who', 'because', 'for', 'there', 'wasn't', 'were', 'while', 'should', 'my', 'you'd', 'don't', 'have', 'hers', 'am', 's', 'me', 'just', 'doing', 'into', 'do', 'down', 'you're', 'yours', 'on', 'm', 'these', 'y', 'hasn', 'did', 'nor', 'is', 'having', 'he', 'not', 'and', 'needn', 'she', 'our', 'some', 'them', 'until', 't', 'up', 'then', 'haven', 'myself', 'very', 'doesn't', 'yourself', 'than', 'didn', 'above', 'will', 'being', 'further', 'such', 'what', 'had', 'aren't', 'own', 'has', 'again', 'more', 'mightn't', 'at', 'through', 'but', 'her', 'don', 'once', 'it', 'with', 'all', 'o', 'does', 've', 'll', 'over', 'couldn't', 'isn't', 'theirs', 'its', 'only', 'she's', 'doesn', 'they', 'below', 'aren', 'themselves', 'you've', 'you'll', 'hasn't', 'ours', 'haven't', 'wasn', 'which', 'or', 'whom', 'no', 'been', 'his', 'hadn', 'it's', 'hadn't', 'by', 'himself', 'won't', 'i', 'couldn', 'any', 'you', 'under', 'why', 'shouldn', 'isn', 'was', 'the', 'off', 'of', 'wouldn', 'as', 'an', 'shouldn't', 'before', 'if', 'mustn', 'their', 'each', 'so', 'those', 'after', 're', 'wouldn't', 'ma', 'herself', 'd', 'few', 'here', 'same', 'during', 'your', 'about', 'how', 'when', 'can', 'other', 'weren', 'this', 'a', 'mustn't', 'should've', 'shan', 'won', 'mightn', 'weren't', 'itself', 'to' }