Title
Examining structural and semantic predictors of announced sarcasm on r/AskReddit

Permalink
https://escholarship.org/uc/item/34k6x5jk

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors
Hew, Joshua
Tarighat, Aida
Corley, Martin
et al.

Publication Date
2024

Examining structural and semantic predictors of announced sarcasm on r/AskReddit

Joshua Hew
University of Edinburgh
Department of Psychology
Edinburgh EH8 9JZ

Aida Tarighat
University of Edinburgh
UKRI CDT in NLP
Edinburgh EH8 9AB

Martin Corley
University of Edinburgh
Department of Psychology
Edinburgh EH8 9JZ

Zachary Horne
University of Edinburgh
Department of Psychology
Edinburgh EH8 9JZ

Abstract

People sometimes explicitly announce that they are being sarcastic. The announcement appears to be particularly common in text-based conversations where prosodic cues are more difficult to identify. In certain cases, the tone of a comment is sufficient to determine non-literal meaning. However, what happens in the absence of these features, or when context forces us to explicitly caveat our sarcasm? In this study, we examined Reddit comments from `r/AskReddit` for the features that are present in comments tagged with "/s", a convention on the platform for users to denote sarcasm. We found that a host of cues which mimic prosody, and other aspects of figures of speech, were inconsistent predictors of announced sarcasm. In contrast, when talking about sociomoral topics such as politics or race, users were more likely to tag their comments with "/s". This suggests that users are more likely to announce sarcasm in text-based conversations where misinterpretation would be socially detrimental.

Keywords: sarcasm; announced sarcasm; natural language processing; sociomoral; psycholinguistics; Reddit

Introduction

Non-literal language such as sarcasm and irony is commonplace in everyday interactions, where ironic language is estimated to constitute as much as 8% of interpersonal interactions (Gibbs, 2000). Sarcastic language is also present on social media platforms such as Twitter (Ling & Klinger, 2016) and Reddit (Khodak et al., 2017). Though their prevalence online averages at 2%, it can reach as high as 20% after widespread media coverage of scandals or disasters (Sykora et al., 2020). As the digitally native Generation Z is poised to come of age (Pew Research, 2020), cultural commentators have pointed to sarcasm as one of the defining characteristics of their humour (Mercado, 2019; Ehrlich, 2020). Together, these factors motivate research into a greater understanding of written sarcasm online.

Sarcasm is the use of remarks that clearly mean the opposite of what they say, often with humorous intent. Sarcasm is a negative form of irony, used to enhance a speaker's negativity towards a subject matter by deliberately uttering the opposite (Colston, 2017). Despite some publications using irony and sarcasm as interchangeable concepts (Mikhailova, 2020; Ptáček et al., 2014), their difference has been shown empirically. Ling & Klinger (2016) found that the sentiment values of sarcasm and irony differed, where sarcastic tweets had higher positive sentiments compared to their ironic and regular counterparts, in line with Colston's (2017) categorisation.

Humans are able to identify and process irony via contextual cues and paralinguistic markers (Deliens et al., 2018). To determine irony, the content of an utterance is compared to background context or world knowledge. For instance, one can compare the utterance that "it's a wonderful day today" to the reality that it is raining to infer irony. Tone and prosody are also markers of ironic content, for instance, by lengthening or stressing syllables (Burgers & van Mulken, 2017). A speaker can also contrast their prosody with the semantics of an utterance, for example by using a flat intonation for an enthusiastic sentence. Facial expressions such as a "blank face" have also been shown to signal irony (Attardo et al., 2003).

In research on sarcasm detection in written text, models have been trained on features that bear similarities to the features humans rely on (e.g., analogues to prosody). For instance, Farha et al. (2022) found that sentiment incongruity within a sentence was relatively easy to detect by sarcasm classifiers, where sentences were in the form of "I love (positive sentiment) failing (negative sentiment) exams". Other approaches use syntactic or typographical markers to imitate prosody. For instance, written fillers (um, hmm) were found to predict the perception of sarcasm in Twitter data (Tarighat et al., 2022). Similarly, the presence of exaggeration and tag questions, which correspond to exclamation points and question marks, (Wolf et al., 2022), along with letter repetition, excessive punctuation (Mikhailova, 2020), internet slang (HAHA, LOL: Yunitasari et al., 2019), and emojis (Thompson & Filik, 2016) are all possible cues of written sarcasm.

Despite work on the cues identified above, written sarcasm remains difficult to identify for both humans and machines (Farha et al., 2022); anecdotal experience makes it clear that conveying and understanding sarcasm over text can be difficult. Even friends with whom we communicate frequently can misinterpret sarcastic messages as literal, or literal messages as sarcastic (Gibbs, 2000). These observations may explain why in text-based conversations there is increased reliance on messages *announcing* when one is being sarcastic. In social media conversations, we also see this phenomenon: many internet forums, such as Reddit, have adopted a convention of indicating a message as sarcastic by typing "/s" at the end.

Online platforms are characterised by pseudonymity and anonymity (Proferes et al., 2021), where common ground between users can be difficult to establish. Proper interpretation of sarcasm requires an amount of shared experience and knowledge between participants in a conversation (Kreuz, 1996). In light of this, Bamman & Smith (2015) have suggested the "#sarcasm" hashtag acts as a marker to clarify

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society*.

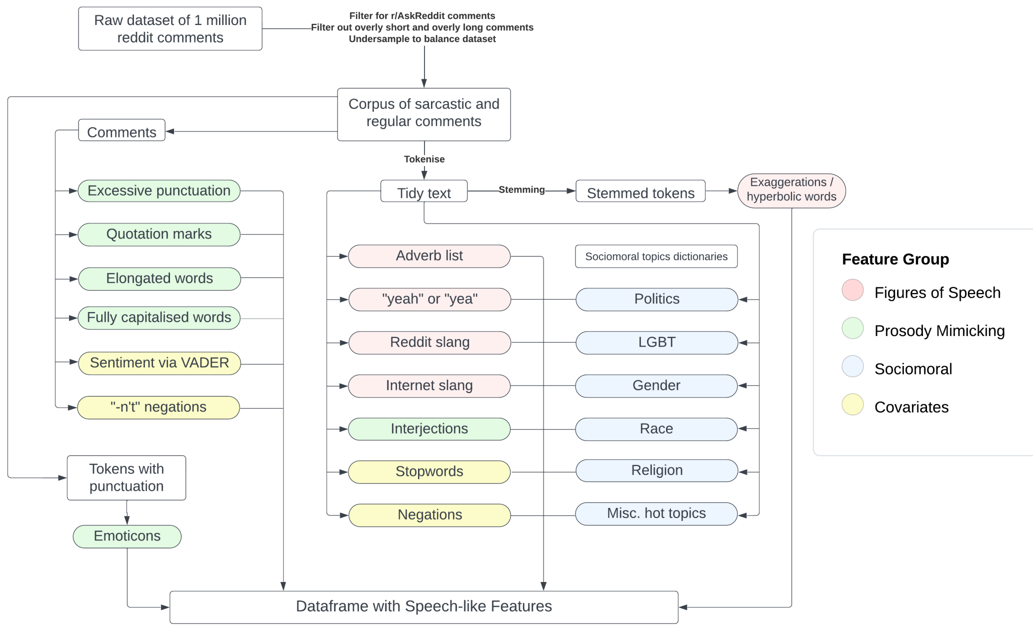

Figure 1: Data preprocessing flowchart.

communicative intent in the absence of common ground on Twitter. In fact, Joshi et al. (2016) found that political topics such as gun control were more prevalent in tweets denoted with "#sarcasm" compared to neutral topics on food or music. Building on this idea, we predict that announced sarcasm should occur more frequently when people are discussing sociomoral topics – broadly, topics concerning political, social, or moral issues – because of the risk that literal interpretations of what they say will generate social condemnation.

Research on humour about race is compatible with this hypothesis; speakers often caveat their utterances with phrases like "it was only a joke" to indicate that they do not seriously hold problematic views (Sue & Golash-Boza, 2013). There is also evidence that sarcasm can be used to disguise aggressive comments in hate speech contexts, suggesting that non-literal devices change how other users perceive the comment (Frenda et al., 2018). This finding could motivate why an announcement of sarcasm might be in the user's best interest to clarify communicative intent in situations involving polarising topics. Therefore, we propose that the use of "/s" works analogously to "#sarcasm", where users feel the need to announce their sarcasm when talking about sociomoral topics to prevent misinterpretation in their online exchanges.

Here, we focus on investigating the predictors of announced sarcasm, by modelling the structural and semantic features that predict the tagging of a comment with an explicit "/s" marker. We use a large-scale dataset from Reddit where messages have been tagged as sarcastic or not, in particular, the subreddit `r/AskReddit`. Our candidate set of predictors is derived from well-established psycholinguistic literature on the predictors of sarcasm; however, our investigation is exploratory because even if structural and prosodic features are predictors of *unstated* sarcasm, this may not be the case for announced sarcasm. Announced sarcasm may be predicted by other aspects of semantic content, sentiment, topic, and context. Indeed, we hypothesised that when talking about these topics the choice to *overtly mark* sarcasm with "/s" would be higher. Our findings would clearly contribute to our understanding of the psychology of sarcasm both off- and online.

Data and Code

We use a Reddit corpus containing both sarcastic and regular comments (Khodak et al., 2017). The corpus is self-annotated because users themselves denote their sarcasm by attaching "/s" at the end of their comments. The authors selectively included comments from users who have used the "/s" notation in the past to ensure these users were aware of this convention on Reddit. Of the one million comments in the dataset, we used around 65,600 comments from `r/AskReddit` for our analysis. We chose this subreddit as it included a sizeable amount of comments, and did not have discourse focused on any one topic. Because we hypothesised an important predictor of announced sarcasm was whether the topic of the comment was sociomoral or not, variability in the topics represented was essential to the investigation.

Sarcastic comments have been linked to hurtful language with the intent of aggression by users (Frenda et al., 2022), and have been found to be similar to abusive or hateful comments, so differentiating between the two is difficult without additional context or world knowledge (Nobata et al., 2016). Thus, we chose not to exclude sexually explicit (NSFW) or violent content as these could be more prevalent in sociomoral topics on Reddit. Comments that were either very short (fewer than 10 words) or very long (over 100 words) were not included in this analysis. We sampled a balanced number of sarcastic and non-sarcastic comments to allow for an easier interpretation of model classification performance. This dataset, along with all the associated dictionaries and code used for analysis can be found at `https://osf.io/s2xyn/`.

Methods

Data Preprocessing

Based on previous computational and psycholinguistic research, we extracted a set of 20 features from sarcastic and non-sarcastic comments. The full feature extraction process for this model can be found in Figure 1. Example sarcastic comments drawn from the dataset that contain these features can be found in Table 1. We grouped the extracted features into the following categories: *figures of speech* represented by the use of (for instance) "yeah" and internet slang; *prosody-mimicking features* represented by the use of emoticons and elongated words; presence of *sociomoral topics* (e.g., politics or other polarising issues); and finally a set of covariates that were not of theoretical interest but were previously established to be predictors of sarcasm in online text. Features that would be expected to have more than one occurrence in a comment, such as stop words, were coded as continuous variables. Other features were coded as binary variables. For instance, quotation marks were coded as a yes/no binary variable because few comments had two or more pairs of quotation marks. Most of the features we extracted used dictionaries, adapted from several sources: negated and exaggerated words were retrieved from ChatGPT (OpenAI, 2023); Reddit slang was adapted from `r/TheoryOfReddit` glossary (i_post_gibberish, 2016); internet slang was retrieved from Penugonda (2021); emoticons (Sakib, 2021) from Kaggle; interjections from Holen's (n.d.) website; and a list of adverbs was included as a possible set of predictors after a preliminary descriptive analysis (using term frequency-inverse document frequency) suggested these words were more common in sarcastic than non-sarcastic comments. We also observed frequent occurrences of the word "yeah" in sarcastic comments from this preliminary investigation; so, we included this as a predictor in our analyses to more formally assess its contribution as a predictor of sarcasm.

Based on prior research showing empirical differences in positive sentiment score between sarcastic and regular comments (Ling & Klinger, 2016), we also included sentiment as a variable to be controlled for in the model. Rather than using a simpler AFINN sentiment dictionary (Nielsen, 2011), we used the valence-aware dictionary and sentiment reasoner (VADER: Hutto & Gilbert, 2014) as a lexicon and rule-based model specifically designed for text on social media. The model better accounts for linguistic features that modify sentiment that would have otherwise been ignored by dictionary-based analysis methods. For instance, it better handles negations such as "no" or "never" and intensifiers such as "really" or "very." It also accounts for slang, emoticons, and punctuation in its calculation of sentiment.

Table 1: Example sarcastic comments from the dataset that contain speech-like or sociomoral (topic) features.

Feature	**Example Comment**
Excessive punctuation	since you love to give things away and love your gf **...** i suppose you love giving away your gf**....**
Elongated words	Excuse me I took half a semester of psych 101 I'm **tooootaly** an expert on this subject
Interjections	**Wow** that doesn't sound like it would be unbiased or anything
"yeah"/"yea"	**Yea** it was the teenagers making memes that got trump elected
Adverbs	**Clearly** the necklace was so bad that it convinced the girlfriend to cheat
Politics	But you need to spread "democracy" and "freedom" for those other countries to be more like yours...
Race	How could a white person get shot when he's hiding behind all that privilege
Non-sociomoral	Yeah, because advertisers aren't known for being intrusive and unrelenting or anything.

Along with sentiment, we also used a dictionary-based method to annotate the topic discussed in each comment and parent comment pair. Six dictionaries of the following topics: politics, gender, LGBTQ, race, religion, and other miscellaneous hot topics, were adapted from Priniski & Horne (2018) and extended with input from ChatGPT (OpenAI, 2023). We considered a topic to be present if any word from each of the dictionaries appeared in either the target comment or the parent comment. For example, a political topic is present in a comment if the word "Trump" appears in the parent comment. Topics were coded independently so a comment could be labelled as concerning more than one sociomoral topic.

Analytic Strategy

Although the features we examined were inspired by the existing psychological and computational social science literatures, our analyses are exploratory. For this reason, we take a machine learning approach in which we simultaneously model all of the features of interest, along with covariates that prior authors have claimed predict sarcasm in online text data (Ling & Klinger, 2016), in a Bayesian logistic regression model. Because of the exploratory nature of our investigation, we use a prior distribution that strongly regularises the model's estimates (here, a horseshoe prior). A horseshoe prior is a shrinkage prior similar to the LASSO (Carvalho et al., 2009); however, it is symmetric, infinitely spiked at zero, and has fat, Cauchy-like tails. The horseshoe prior has been

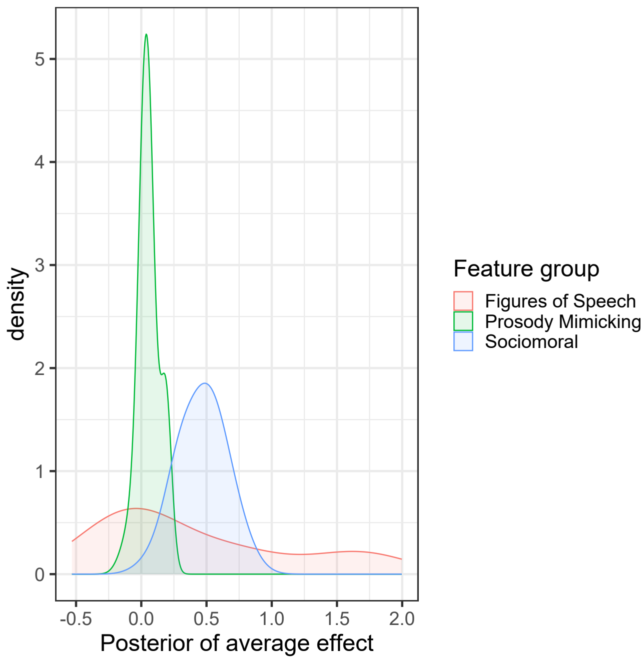

Figure 2: Model estimated average effects of structural and semantic features predicting sarcasm on `r/AskReddit` based on 4000 draws from the posterior distribution.

Table 2: Population-level effects by feature group.

	Estimate	SE	Lower CI	Upper CI
Intercept	0.07	0.05	−0.03	0.17
Figures of Speech				
Adverbs	0.71	0.07	0.57	0.85
Hyperbolic words	−0.01	0.02	−0.04	0.02
Internet Slang	0.02	0.04	−0.06	0.10
Reddit Slang	-0.27	0.07	−0.40	−0.13
"yeah"/"yea"	1.67	0.10	1.48	1.86
Prosody Mimicking				
Elongated words	0.05	0.02	0.01	0.09
Emoticons	−0.58[a]	0.20	−0.98	−0.19
Excessive punctuation	0.01	0.02	−0.03	0.05
Fully capitalised	0.07	0.02	0.03	0.12
Interjections	0.19	0.02	0.14	0.23
Quotation marks	−0.02	0.07	−0.18	0.12
Sociomoral Topics				
Gender	0.29	0.04	0.21	0.38
LGBT	0.26	0.14	−0.01	0.54
Misc. hot topics	0.45	0.09	0.27	0.63
Politics	0.52	0.06	0.39	0.64
Race	0.71	0.09	0.53	0.88
Religion	0.55	0.09	0.37	0.73
Covariates				
Negations	0.21	0.02	0.16	0.26
Sentiment	0.08	0.04	0.00	0.17
Stop words	−0.66	0.05	−0.76	−0.57

[a] Due to the rarity of emoticons, this coefficient is based on very few observations relative to the overall sample size.

shown to be well suited to accommodate cases where a large number of regression coefficients are zero, but a minority of coefficients are potentially large (Carvalho et al., 2009). We specified a horseshoe where the student-t prior of the local shrinkage parameters had three degrees of freedom to improve sampling performance in Stan. We also fit a model in which this distribution was specified with one degree of freedom, which produced nearly identical results yet led to divergent transitions suggesting the sampler was misbehaving.

The default specifications of the horsehoe prior may underregularise non-zero regression coefficients; so, we followed the recommendations of Piironen & Vehtari (2016) and Piironen & Vehtari (2017) by specifying both the ratio of zero to non-zero parameters we crudely expected (parameter ratio = .1), and by controlling the amount of regularisation exerted by 'slab' in the distribution (e.g., the behaviour of the tails of the distribution; slab scale = .5). We also conducted a series of prior sensitivity analyses to ensure our conclusions were not dependent on unknown idiosyncrasies caused by our prior distribution decisions. For instance, we made the parameter ratio more skeptical of non-zero regression coefficients or altered the strength of the regularisation exerted in the slab of the model. The details of our models are on the Open Science Framework `https://osf.io/s2xyn/`.

Results

First, we replicated prior work suggesting that positive sentiment, stop words, and negations were all credible predictors of text-based sarcasm (Table 2), although sarcastic comments here were only slightly more positive than non-sarcastic comments. Second, we observed that although some structural features were clear predictors of announced sarcasm, many were not. For instance, use of adverbs, "yeah", and interjections all credibly predicted sarcasm; however, other prosody-like aspects of text (e.g., exaggeration, excessive punctuation) did not (Table 2; Figure 2). In contrast, and consistent with our prediction that announced sarcasm may be more common when the topic is serious and thus the consequences of misperception are more severe, every sociomoral topic was a credible predictor of sarcasm (Table 2; Figure 2). A model comparison approach confirmed these results: adding sociomorality to a model containing all the structural predictors and covariates substantially improved out-of-sample predictive performance, Δ expected log pointwise predictive density = 114, SE = 15.

These results confirm that in situations where it may be more problematic if someone misperceives a sarcastic comment as literal (such as when talking about polarising issues), people are more likely to announce they are being sarcastic.

Qualitative Analysis

After establishing a link between sociomoral topics and the use of announced sarcasm, we wanted to explore and confirm the ways in which sarcasm was employed within these discussions. Furthermore, we wanted to investigate whether sarcasm use changed over time, as social media is characterised by rapid changes of internet slangs and microtrends. To do this, we randomly sampled a subset of political comments from the data (20 from each year, except 2009 and 2010 where we only had 4 and 7 comments respectively) to qualitatively look for possible chronological shifts in the sarcastic content (done independently by two of the authors). We found that political content is heavily influenced by the ongoing news at the time of posting, with US politics dominating the discourse with contributions from both American and international Redittors. For example, comments from the years 2009 to 2015 had discussions revolving Barack Obama's presidency, healthcare, and gun control. In 2016, users mainly discussed the 2016 US elections and Donald Trump's presidential campaign. We could see examples of all the categories of sarcasm goals listed in Kreuz et al. (1991). Most frequently, sarcasm was used to be funny or witty, to mock (make fun of certain political figures or groups), and to criticise or reprimand. For the latter, users commonly used sarcasm to point out an obvious scenario where the parent comment's logic was flawed. For example, in a discussion on reducing the government's constitutional power, a replying user employed sarcasm to argue that large corporations use their power and wealth to dominate the market and abuse their power. Finally, the common sarcastic linguistic features were present in many of them; for instance, they started with *yeah*/*yes*/*oh yeah*, and used emphasis (e.g., italics or caps lock) as well as non-verbal cues (e.g., "... its own people *gasp* the *horror*").

Discussion

This paper demonstrates a substantial predictor of announced sarcasm, in the form of semantic topic, that previously has not been the focus of research in either machine learning or psycholinguistics. By noting that announcing one's sarcasm was common in online dialogue, we were able to generate further predictions about written sarcasm based on drawing a distinction between announced and unstated sarcasm. Indeed, we observed that users were more likely to explicitly say that they were being sarcastic when discussing these topics, likely to establish their communicative intent and prevent misinterpretation, or loss of social standing.

Even though the affect of prosody-mimicking features on the probability of a comment being marked with "/s" is less pronounced compared to sociomoral topics (although some effects were large and credible), this might not be the case for unannounced sarcasm where users are nonetheless sarcastic yet do not mark their utterances explicitly. It is possible that "/s" convention circumvents the need for Reddit users to further mark out their comment as sarcastic via these speech-like markers, where "/s" is sufficient to do so. Further research could compare the incidences of these structural features between announced sarcasm and unannounced sarcasm.

Limitations and Future Work

Several limitations need to be acknowledged about this investigation. First, our dataset contained comments that were collected between 2009 and 2016. Social media changes rapidly, with microtrends lasting on the order of weeks rather than years. While our qualitative analysis revealed that users discuss topics or events that were prominent during a specific year (such as the 2016 US election), the way people converse online will undoubtedly have changed since 2016.

The norm of using "/s" itself is also peculiar, even if it is now accepted convention to use it when conveying sarcasm. Our study shows that prosodic features and sociomoral topics are associated with using "/s", but how users decide to use it and how they interpret it may still vary. For instance, users may use "/s" itself ironically for humorous purposes, or they may use it to mask their insensitive views. Therefore, the social function of this internet convention may not translate directly into everyday speech. Other complications also arise when using Reddit data. For instance, Khodak et al. (2017) note the possibility of false negatives, where users leave sarcastic comments without denoting them with "/s."

Our investigation focused on one general subreddit; however, other subreddits will differ in their topics, methods of discussion, and moderation styles. All of these facts shape the context and, in turn, the discussion occurring on these subreddits. One possibility then is that the use of sarcasm may differ depending on the subreddit, as well as how well moderated the subreddit is. In a poorly moderated subreddit where insensitive comments could be allowed to survive without consequences, users might feel less of a need to caveat their sarcastic utterances with "/s." Investigating the generalisability of our findings across subreddits is a question for future research. Additionally, a more comprehensive study of several subreddits would provide us with enough data to pursue alternative computational strategies (e.g., fine-tuning language models for sarcasm detection).

It is also necessary to design separate sarcasm studies for different platforms, specifically, Reddit and Twitter, due to the distinct interaction dynamics and management styles of each platform. The organisation of discussions in subreddits with original posts and subsequent comments facilitates context analysis on Reddit. However, on Twitter where we encounter more spontaneous posting behaviour, contextual analysis may not be as robust.

Conclusion

We have proposed that the use of "/s" to announce sarcasm would be affected by the topic of discourse, because the cost of being misunderstood would be higher when a discussion concerned sociomoral issues. Our findings suggest that users of Reddit are more likely to announce sarcasm in text-based interactions, especially when conversing about polarising topics, implying that they recognise the potential for misinterpretation and have a desire to clarify non-literal meaning in such contexts.

References

Attardo, S., Eisterhold, J., Hay, J., & Poggi, I. (2003). Multimodal markers of irony and sarcasm. *Humor - International Journal of Humor Research*, *16*(2). doi: 10.1515/humr.2003.012

Bamman, D., & Smith, N. (2015). Contextualized sarcasm detection on twitter. In *proceedings of the international aaai conference on web and social media* (Vol. 9, pp. 574–577).

Burgers, C., & van Mulken, M. (2017). Humor markers. In *The routledge handbook of language and humor* (1st ed., p. 385-399). United Kingdom: Routledge.

Carvalho, C. M., Polson, N. G., & Scott, J. G. (2009). Handling sparsity via the horseshoe. In *Artificial intelligence and statistics* (pp. 73–80).

Colston, H. L. (2017). Irony and Sarcasm. In *The Routledge Handbook of Language and Humor*. Routledge. (Num Pages: 16)

Deliens, G., Antoniou, K., Clin, E., Ostashchenko, E., & Kissine, M. (2018). Context, facial expression and prosody in irony processing. *Journal of memory and language*, *99*, 35–48.

Ehrlich, H. (2020). *The isolation and absurdity of gen z humor – arts + culture.* Retrieved from `https://culture.affinitymagazine.us/the-isolation-absurdity-of-gen-z-humor/`

Farha, I. A., Wilson, S., Oprea, S., & Magdy, W. (2022). Sarcasm detection is way too easy! an empirical comparison of human and machine sarcasm detection. In *Findings of the association for computational linguistics: Emnlp 2022* (pp. 5284–5295).

Frenda, S., Cignarella, A. T., Basile, V., Bosco, C., Patti, V., & Rosso, P. (2022). The unbearable hurtfulness of sarcasm. *Expert Systems with Applications*, *193*, 116398.

Frenda, S., et al. (2018). The role of sarcasm in hate speech. a multilingual perspective. In *Proceedings of the doctoral symposium of the xxxivinternational conference of the spanish society for natural language processing (sepln 2018)* (pp. 13–17).

Gibbs, R. W. (2000). Irony in talk among friends. *Metaphor and symbol*, *15*(1-2), 5–27.

Holen, V. (n.d.). *Dictionary of interjections (aww, oh, ah, eek, oops).* Retrieved from `https://www.vidarholen.net/contents/interjections/`

Hutto, C., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the international aaai conference on web and social media* (Vol. 8, pp. 216–225).

i_post_gibberish. (2016). *The glossary.* Retrieved from `https://www.reddit.com/r/TheoryOfReddit/wiki/glossary/`

Joshi, A., Jain, P., Bhattacharyya, P., & Carman, M. (2016). Who would have thought of that!': A hierarchical topic model for extraction of sarcasm-prevalent topics and sarcasm detection. *arXiv preprint arXiv:1611.04326*.

Khodak, M., Saunshi, N., & Vodrahalli, K. (2017). A large self-annotated corpus for sarcasm. *arXiv preprint arXiv:1704.05579*.

Kreuz, R. J. (1996). The use of verbal irony: Cues and constraints. In J. S. Mio & A. N. Katz (Eds.), *Metaphor: Implications and applications* (p. 23-38).

Kreuz, R. J., Long, D. L., & Church, M. B. (1991). On being ironic: Pragmatic and mnemonic implications. *Metaphor and symbol*, *6*(3), 149–162.

Ling, J., & Klinger, R. (2016). An empirical, quantitative analysis of the differences between sarcasm and irony. In *The semantic web: Eswc 2016 satellite events, heraklion, crete, greece, may 29–june 2, 2016, revised selected papers 13* (pp. 203–216).

Mercado, W. (2019). *The shocking and absurdist humor of gen z.* Retrieved from `https://medium.com/@williamfmercado/the-shocking-and-absurdist-humor-of-gen-z-f3d6e363d42b`

Mikhailova, P. (2020). Graphic markers of irony and sarcasm in written texts. In *Speech and computer: 22nd international conference, specom 2020, st. petersburg, russia, october 7–9, 2020, proceedings 22* (pp. 346–356).

Nielsen, F. Å. (2011). A new anew: Evaluation of a word list for sentiment analysis in microblogs. *arXiv preprint arXiv:1103.2903*.

Nobata, C., Tetreault, J., Thomas, A., Mehdad, Y., & Chang, Y. (2016). Abusive language detection in online user content. In *Proceedings of the 25th international conference on world wide web* (p. 145–153). Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee. Retrieved from `https://doi.org/10.1145/2872427.2883062` doi: 10.1145/2872427.2883062

OpenAI. (2023, Dec). Retrieved from `https://chat.openai.com`

Penugonda, G. S. (2021). *Chat / internet slang — abbreviations — acronyms.* Retrieved from `https://www.kaggle.com/datasets/gowrishankarp/chat-slang-abbreviations-acronyms`

Pew Research. (2020). *On the cusp of adulthood and facing an uncertain future: What we know about gen z so far.* Retrieved from `https://www.pewresearch.org/social-trends/2020/05/14/on-the-cusp-of-adulthood-and-facing-an-uncertain-future-what-we-know-about-gen-z-so-far-2/`

Piironen, J., & Vehtari, A. (2016). On the hyperprior choice for the global shrinkage parameter in the horseshoe prior, arxiv. *arXiv preprint arXiv:1610.05559*, *10*.

Piironen, J., & Vehtari, A. (2017). Sparsity information and regularization in the horseshoe and other shrinkage priors. *Electronic Journal of Statistics*, *11*(2). doi: 10.1214/17-ejs1337si

Priniski, J., & Horne, Z. (2018). Attitude change on reddit's change my view. In *Cogsci.*

Proferes, N., Jones, N., Gilbert, S., Fiesler, C., & Zimmer, M. (2021). Studying reddit: A systematic overview of disciplines, approaches, methods, and ethics. *Social Media+ Society*, *7*(2), 20563051211019004.

Ptáček, T., Habernal, I., & Hong, J. (2014). Sarcasm detection on czech and english twitter. In *Proceedings of coling 2014, the 25th international conference on computational linguistics: Technical papers* (pp. 213–223).

Sakib, A. S. (2021). *Chat / internet slang — abbreviations — acronyms.* Retrieved from `https://www.kaggle.com/datasets/gowrishankarp/chat-slang-abbreviations-acronyms`

Sue, C. A., & Golash-Boza, T. (2013). 'it was only a joke': how racial humour fuels colour-blind ideologies in mexico and peru. *Ethnic and Racial Studies*, *36*(10), 1582–1598.

Sykora, M., Elayan, S., & Jackson, T. W. (2020). A qualitative analysis of sarcasm, irony and related# hashtags on twitter. *Big Data & Society*, *7*(2), 2053951720972735.

Tarighat, F. S., Magdy, W., & Corley, M. (2022). Understanding fillers may facilitate automatic sarcasm comprehension: A structural analysis of twitter data and a participant study. In *Proceedings of the 26th workshop on the semantics and pragmatics of dialogue.*

Thompson, D., & Filik, R. (2016). Sarcasm in written communication: Emoticons are efficient markers of intention. *Journal of Computer-Mediated Communication*, *21*(2), 105–120.

Wolf, N., Rohde, H., & Corley, M. (2022). A really great study it is, isn't it? In *Xprag 2022.*

Yunitasari, Y., Musdholifah, A., & Sari, A. K. (2019). Sarcasm detection for sentiment analysis in indonesian tweets. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, *13*(1), 53–62.