Social Meme-ing: Measuring Linguistic Variation in Memes

Naitian Zhou,[1] **David Jurgens**[2] and **David Bamman**[1]
[1]University of California, Berkeley
[2]University of Michigan
{naitian,dbamman}@berkeley.edu, jurgens@umich.edu

Abstract

Much work in the space of NLP has used computational methods to explore sociolinguistic variation in text. In this paper, we argue that memes, as multimodal forms of language comprised of visual templates and text, also exhibit meaningful social variation. We construct a computational pipeline to cluster individual instances of memes into templates and semantic variables, taking advantage of their multimodal structure in doing so. We apply this method to a large collection of meme images from Reddit and make available the resulting SEMANTICMEMES dataset of 3.8M images clustered by their semantic function. We use these clusters to analyze linguistic variation in memes, discovering not only that socially meaningful variation in meme usage exists between subreddits, but that patterns of meme innovation and acculturation within these communities align with previous findings on written language.

1 Introduction

One objective in variationist sociolinguistics is to study how social factors contribute to differences in the way people use language. Work in natural language processing has followed this tradition, offering large-scale analyses of how language use is conditioned on geography, (Eisenstein et al., 2010; Hovy and Purschke, 2018; Demszky et al., 2021), community (Del Tredici and Fernández, 2017; Zhu and Jurgens, 2021b; Lucy and Bamman, 2021) and time (Hamilton et al., 2016). This work is important not only because language variation often exposes shortcomings in NLP tools, which are primarily developed for standard language varieties (Blodgett et al., 2016), but also because variation often embeds **social meaning**. We make inferences about people's social class, regionality, gender, and much more based on the way they talk (Campbell-Kibler, 2009; Zhang, 2005), and we strategically use language to actively construct and perform identities (Labov, 1963; Bucholtz and Hall, 2005).

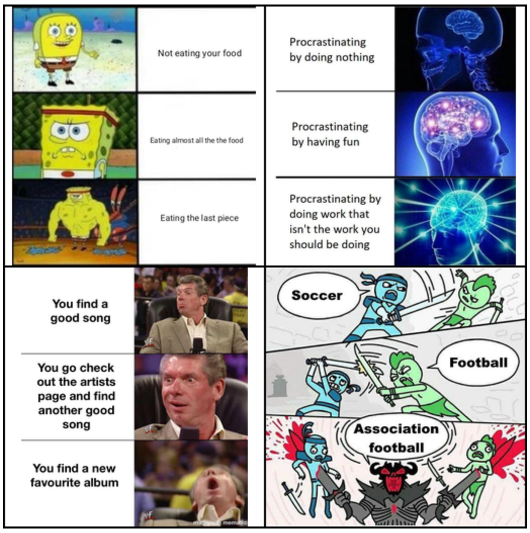

Not eating your food
Eating almost all the the food
Eating the last piece
Procrastinating by doing nothing
Procrastinating by having fun
Procrastinating by doing work that isn't the work you should be doing
You find a good song
You go check out the artists page and find another good song
You find a new favourite album
Soccer
Football
Association football

Figure 1: Meme templates can be visually diverse, but often provide the same semantic function; in this case, all four templates show a scalar increase.

Most of this work has focused on lexical or morphosyntactic variation in written texts. However, language exists beyond text or speech. In face-to-face interaction, multimodality in language has been construed as features like co-speech gesture, facial expression or body movement (Perniss, 2018). In online communication, previous work has extended the term to include the interplay between images and text (Kress and Leeuwen, 2001; Zhang et al., 2021; Hessel et al., 2023). Understanding text in isolation is insufficient to understanding how we communicate online.

In the space of multimodal online language, memes are interesting for their compositionality. They consist of a base image (the **template**) as well as superimposed text (which we refer to as the **fill**). For example, the "Drake" template depicted in figure 2 serves the semantic function of expressing a preference relation between the fills. This same Drake template can be used to express preference relations between a range of fills; at the same time, multiple different templates can share the same or

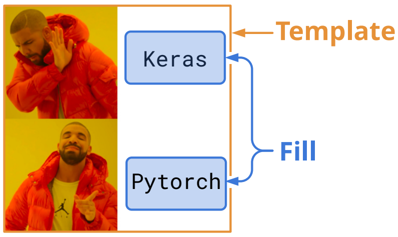


Figure 2: Memes are multimodal constructions where the base image template and additional text fills both have semantic value.

similar semantic function, as illustrated in fig. 1 for the function of "scalar increase." We refer to this set of functionally equivalent templates as a **semantic cluster**.

In this work, we follow the variationist sociolinguistics tradition by treating templates as *variants* and semantic clusters as *variables*, observing how social factors might contribute to the distribution among these variants. To conduct this analysis, we develop a method for identifying semantic clusters by exploiting the visual structure of meme templates and the linguistic structure of meme fills. We use this to create the SEMANTICMEMES dataset of 3.8M Reddit memes[1] grouped into semantic clusters and validated with a human evaluation. Finally, we use these semantic clusters to perform a series of case studies demonstrating their use in studying linguistic variation, linguistic innovation, and linguistic acculturation. We find that:

1. socially meaningful variation in template choice exists between subreddits;
2. subreddits that first introduce a new template continue to use it more than others; and
3. users who stay in a subreddit for longer tend to use templates distinctive to that subreddit.

These findings illustrate the ways in which memes function as multimodal acts of communication, and how methods from computational sociolinguistics can shed light on meaningful variation within them.

2 Methods

To study variation in meme use, we need to identify the meme variables that organize a collection of meme **instances**—the individual memes that are created and posted online by specific people at specific moments in time. We create a pipeline that visually clusters meme **instances** into **templates** (i.e., the same memes that differ by variation in fills) by exploiting the visual similarity between them; and linguistically clustering meme **templates** into **semantic clusters** by exploiting the similarity among the fills used in different templates. Fig. 3 provides an overview of the process, which involves first clustering instances into templates (§2.1), and then clustering templates into variables (§2.2).

2.1 Visually clustering instances

Our process starts with a set of meme instances, which we wish to group based on visual similarity; this process serves to group memes into their base templates as well as filter out many non-meme images. This is difficult due to the massive number of images as well as the amount of variation in zoom, crop, borders and other visual details. We lay out the steps of the process here, but provide further details and example images in Appendix A.

We first preprocess images to remove any solid color framing elements to isolate the base image, then follow Zannettou et al. (2018) and Morina and Bernstein (2022) in extracting templatized memes by running a perceptual hashing algorithm.

We then compute the pairwise Hamming distance between hashes that occur more than 10 times, discard any pairs where the distance was greater than a cut-off $d_{\text{max}} = 10$. We use the Leiden clustering algorithm to perform clustering (Traag et al., 2019).[2] The Leiden algorithm iteratively finds well-connected subgraphs; we construct a graph where image hashes were vertices and the edge weight was $e_{ij} = d_{\text{max}} - d_{ij} + 1$ for vertices i and j, where d_{ij} was the Hamming distance between them.

The clustering algorithm splits aggressively—instances with similar base images may be split across multiple templates due to variations in the zoom, crop, and borders. We find the next step, which clusters based on the fill text, serves as a remedy by placing many of these duplicate templates into the same semantic cluster. Appendix A contains examples of template clusters.

2.2 Linguistically clustering templates

Given a set of meme templates, we want to identify clusters of those templates that have a similar

[1] We make data and code under available under the MIT license at https://github.com/naitian/semantic-memes

[2] We found that using DBSCAN, as was done in prior work, resulted in many images being put into a single noisy cluster.

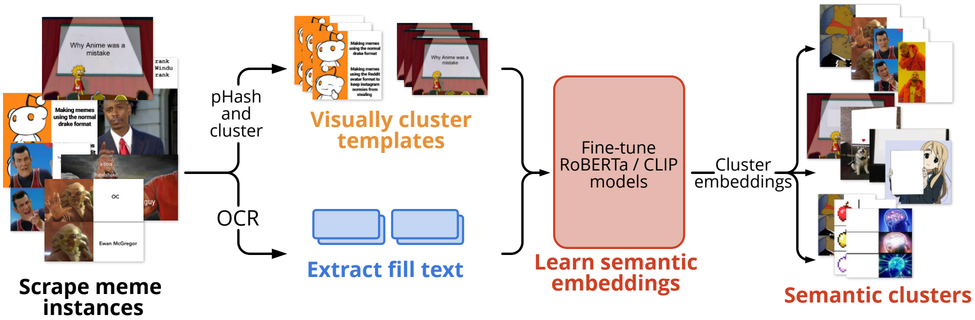


Figure 3: We group visually identical meme instances into templates, and extract the fills using OCR. This data is used to learn semantic embedding representations of templates, which we use to generate semantic clusters.

semantic function—i.e., that are used to assert a similar relation among the text in the fills (such as a comparison function exemplified by the Drake meme in fig. 2). These semantic clusters are the linguistic variables of analysis: discrete sets of variants which share a semantic function but vary in the social meanings they index.

We apply the key intuition that people will use certain templates to make certain classes of statements (comparison, declaration, surprise); as with any other language, fills that are "grammatical" for one template may be nonsensical in another. Templates that share similar sets of fills, then, may perform a similar function over them.

To cluster templates using this principle, we extract the fill text from meme instances belonging to a template (§2.2.1), learn semantic representations for templates based on the distribution of text fills (§2.2.2), and cluster those representations (§2.2.3).

2.2.1 Extracting fill text

We extract text (along with the bounding boxes containing it) from meme instances using EasyOCR.[3] We use the order of the bounding boxes as a rough signal for the position and ordering of the text, but do not incorporate the bounding coordinates directly into the models described below.

Some meme templates contain text in the base image. To prevent these from trivializing the semantic embedding task, we remove bounding boxes with text that was identical in over 90% of the memes in a given template cluster.

2.2.2 Learning semantic embeddings

We examine four methods for learning semantic embeddings of memes, each described in more detail below: a RoBERTa (Liu et al., 2019) classifier fine-tuned to predict the template given the fill text; a CLIP model (Radford et al., 2021) fined-tuned on (fill text, image) pairs; the vector difference between fine-tuned and pretrained CLIP embeddings (CLIP-diff); and concatenating CLIP-diff and RoBERTa embeddings (Concat).

Text-only RoBERTa. In the text-only model, we fine-tune a RoBERTa model on a sequence classification task to predict a distribution over templates given the fill text as input. We separate text in different bounding boxes in a meme with a separator token when passing it into the model to impose a rough, linear notion of space.

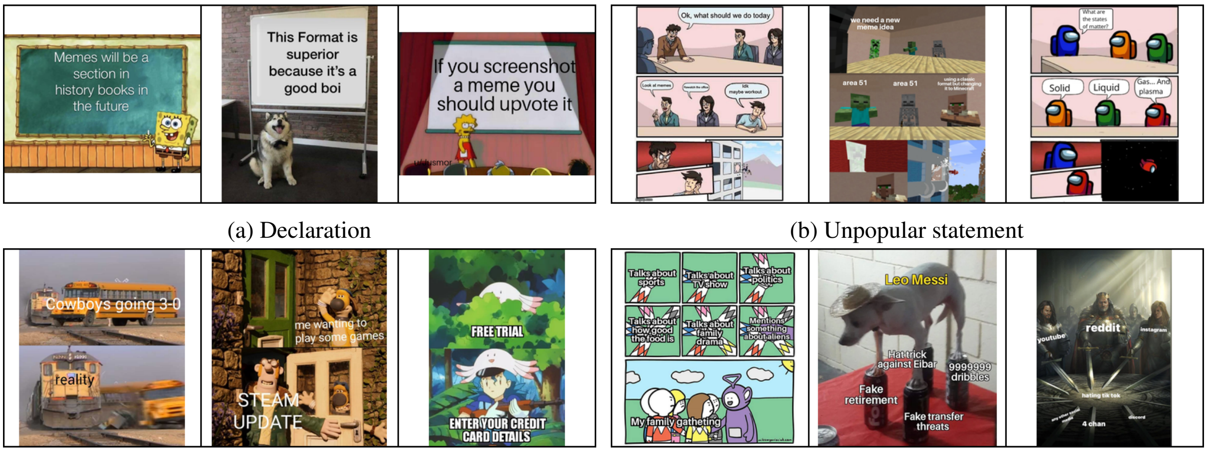
After fine-tuning, we take the weights of the final classification layer $W \in \mathbb{R}^{768\times N}$ as the embeddings, where N is the number of templates. Intuitively, RoBERTa is an encoder model that projects the fill text into a latent semantic space. The final classification layer can be thought of as a projection from that latent space into the discrete space of templates. Therefore, the transposition of the weight matrix can be viewed as a mapping from templates into the latent semantic space.

Multimodal CLIP. In learning the embeddings, the text-based RoBERTa model does not have direct access to the image features in the templates. We experiment with using both the image and text data by fine-tuning a CLIP model.

We fine-tune CLIP using a contrastive loss between the embedding of a meme instance and its fill text. To prevent the model from cheating by reading the text in the image, we sample a meme instance with different text but the same template. This fine-tuning step modifies the image embedding to align with fill text, which implicitly describes the semantic function of the meme, instead of with the pretraining dataset of image captions, which explicitly describe the contents of the image.

CLIP generates embeddings of meme instances.

[3] https://github.com/JaidedAI/EasyOCR



(a) Declaration

(b) Unpopular statement

(c) Surprise narrative

(d) Similarity / parts of a whole

Figure 4: Examples from semantic clusters generated from RoBERTa embeddings; visually diverse clusters emerge even for complex semantic functions like a surprise narrative.

To generate template embeddings, we randomly sample up to ten instances of a template as input for the image embedding module. We then compute the average image embedding of those instances. We don't embed the fill text for this step, since fill text greatly varies between meme instances that use the same template, but the image templates should be more or less visually identical.

CLIP-diff. It is possible that the fine-tuned model contains a notion of meme semantics that is in tension with the pretraining task of image captioning. To isolate the meme-specific knowledge learned by fine-tuning, we calculate the difference between the fine-tuned CLIP embedding of an image and the embedding from the base CLIP model.

CLIP-diff + RoBERTa. Finally, we concatenate the CLIP-diff and RoBERTa embeddings to incorporate both the visual features from CLIP as well as the semantics learned by the RoBERTa model.

2.2.3 Semantic clusters

To group templates into meme variables, we use Leiden clustering on the template representations from the embedding models. Some prior work has shown that embeddings from later layers of language models can contain "rogue dimensions" that dominate the dot product and lower representational quality. Following recommendations to mitigate this behavior, we first standardized the template embeddings before calculating the cosine similarity (Timkey and van Schijndel, 2021).

We construct an adjacency matrix from the top 50 nearest neighbors for each template embedding, weighting edges as a function of the ranked cosine similarity. We provide details about this process in Appendix B. We generate **semantic clusters** by running the Leiden algorithm on this graph.

3 SemanticMemes Dataset

We used the pipeline described above to generate semantic clusters from a dataset of 27.9M images collected from Reddit (Baumgartner et al., 2020). We scraped image posts from the top 1000 most active subreddits with "meme" in name (e.g. `r/HistoryMemes`). Temporally, the dataset spans the decade between 2011 and mid-2021.

We fine-tuned both the RoBERTa and CLIP models for three epochs on memes whose template appeared at least 100 times in the dataset. We used an 80/10/10 split of train, dev and test data, ensuring there was no leakage of fill text between splits.

Using the pipeline with the RoBERTa model results in 784 semantic clusters spanning 6,384 templates and over 3.8M meme instances. Figure 4 shows some templates that appear in the same semantic cluster. The dataset includes posts to 655 subreddits by 908,917 users. We include examples and descriptive statistics for clusters generated with each of the embedding models in the appendix (Figures 17–21).

A qualitative examination of the clusters showed that the largest, most commonly occurring clusters were highly interpretable with clear semantics. For some of the less common clusters, it was more difficult to assign a clear semantic meaning, but even these clusters often had a coherent quality or affect. A potential line of qualitative future work would be to better understand and identify the unifying features of these clusters.

4 Evaluation

We evaluate the coherence and visual diversity of semantic clusters derived from each model using human judgment. We design an evaluation task in which annotators are presented with a pair of templates, and randomly vary if the templates are drawn from the same or different semantic clusters.

They are asked to evaluate whether the two templates are 1) semantically similar and 2) visually similar. We define semantic similarity as being able to reasonably substitute the text from one template into the other with minor changes. We define visual similarity to include sharing a similar art style or source (e.g., two different templates featuring Spongebob). We include example pairs in the appendix; one strong source of visual similarity (cf. Appendix Fig. 19) are sets of templates that are largely identical in their form but that exhibit slight variation in size, crop, and margins.

We collect judgments for the top ten semantic clusters from each model most commonly represented in our dataset as well as a random selection of ten clusters from each model. For each cluster, we sample 10 pairs, and the same human evaluators provided judgment across all the models. We find strong interannotator agreement (Krippendorff's $\alpha = 0.75$, calculated across all models). We provide more details on this process, including the agreement scores for each individual model, in Section C.1 of the appendix.

From the human judgments, we calculate p_s (the probability that a pair of templates are semantically similar if they appear in the same cluster) and p_v (the equivalent measurement for visual similarity) for each model. To measure variation, it is more important each semantic cluster is semantically coherent and visually diverse, but less important that all relevant templates are surfaced within the same cluster. Following this reasoning, we focus on evaluating the semantic precision p_s.

Our goal in this work is to explore meaningful semantic variation across visually *diverse* memes, since memes that are visually similar (e.g., slight variations on the same template) have trivially similar semantics. Accordingly, we design a measure of *visually adjusted precision* based on Cohen's κ:

$$p_{\text{adj}} = \frac{p_s - p_v}{1 - p_v},$$

Model	Precision	Visual-adjusted
RoBERTa	**0.78**	**0.44**
CLIP	0.65	-0.09
CLIP-diff	0.69	0.18
Concat.	0.70	0.30

Table 1: Comparison of cluster quality for different embedding models. CLIP-based models yield clusters that are biased towards visual features.

Intuitively, this metric represents the extent to which the semantic clusters agree with annotator judgments of semantic similarity while controlling for correlations with visual similarity. A negative score means the model clusters based on visual similarity instead of semantic coherence.We calculate metrics on the model judgments over the set of all annotated pairs across models. This not only allows us to evaluate on a larger set of annotations, but also helps highlight differences between models.

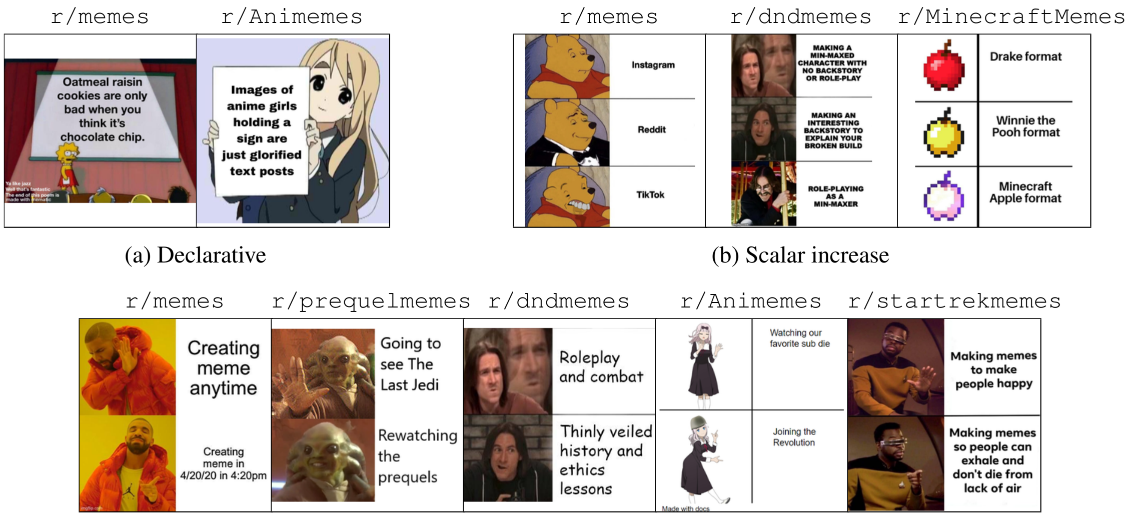
Table 1 presents the results of this evaluation. Introducing any visual features results in some clusters based on visual similarity instead of semantics; accordingly, RoBERTa clusters have the highest visually-adjusted precision (significant w.r.t to CLIP and CLIP-diff for a 95% bootstrap CI). We present more results in Table 4 in the appendix.

Semantic clusters provide a strong separation between content (the semantic cluster) and style (the choice of template within a semantic cluster). In other words, the choice of semantic cluster is *what* a user is trying to say, and the choice of a template within that cluster is *how* they are saying it. In the remainder of the paper, we use the clusters generated from the RoBERTa embeddings, which have the highest visual-adjusted precision, for our case studies on linguistic variation and change.

5 Linguistic variation

The sociolinguistic study of variation centers around the linguistic variable, which captures different ways of saying the same thing. The specific choice a speaker make varies systematically based on information such as the speaker's identity, their relationship to interlocutors, sociopragmatic context, among many other factors (Tagliamonte, 2006). Through variation, language conveys *social meaning* (Nguyen et al., 2021).

There is a rich body of work that aims to analyze linguistic variation computationally. Often, the focus is on lexical variation (Bamman et al., 2014;


r/memes
r/Animemes
Oatmeal raisin cookies are only bad when you think it's chocolate chip.
Images of anime girls holding a sign are just glorified text posts
r/memes
r/dndmemes
r/MinecraftMemes
Instagram
Reddit
TikTok
MAKING A MIN-MAXED CHARACTER WITH NO BACKSTORY OR ROLE-PLAY
MAKING AN INTERESTING BACKSTORY TO EXPLAIN YOUR BROKEN BUILD
ROLE-PLAYING AS A MIN-MAXER
Drake format
Winnie the Pooh format
Minecraft Apple format
r/memes
r/prequelmemes
r/dndmemes
r/Animemes
r/startrekmemes
Creating meme anytime
Creating meme in 4/20/20 in 4:20pm
Going to see The Last Jedi
Rewatching the prequels
Roleplay and combat
Thinly veiled history and ethics lessons
Watching our favorite sub die
Joining the Revolution
Making memes to make people happy
Making memes so people can exhale and don't die from lack of air

(a) Declarative

(b) Scalar increase

(c) Comparison

Figure 5: Subreddits exhibit variation in the preferred templates within a semantic cluster. All are statistically significantly overrepresented in their respective subreddits, $p < 0.05$.

Zhang et al., 2017; Zhu and Jurgens, 2021a); semantic variation in online communities (Lucy and Bamman, 2021; Del Tredici and Fernández, 2018); or orthographic variation in online text (Eisenstein, 2015; Stewart et al., 2017). In our view of memes as language, we ask the same kind of question:

RQ1: Does the template choice within a semantic cluster vary systematically between communities?

Methods. The semantic clusters form our variable context, and set of templates within any given semantic cluster form a discrete set of choices with the same semantic value. We use the weighted log odds-ratio to compute the extent to which a template is specific to a given subreddit compared to all other subreddits, relative to the other templates in a semantic cluster (Monroe et al., 2017; Jurafsky et al., 2014). We find the templates that have a statistically significant association with a subreddit (z-score > 1.96); the semantic clusters these templates belong to are *in variation*: a community prefers one variant over the others in this cluster.

Results. We find 94 out of 784 semantic clusters exhibit statistically significant variation, spanning 391 different templates. Figure 5 shows how functionally similar memes take different forms in different communities.

Speakers use language to construct their social identities (Bucholtz and Hall, 2005). We find that, not only do subreddits prefer certain variants of a template over others, but they choose templates that index into a localized cultural knowledge, making cultural allusions to characters or celebrities.

For example, the orange Drake template (fig. 5c, left) is used frequently in general purpose meme subreddits like r/memes, but alternatives are used in other subreddits. One version that is specific to r/dndmemes (which discusses the role playing game Dungeons and Dragons) replaces Drake with Matthew Mercer, a voice actor who stars in a popular Dungeons and Dragons web series (fig. 5c, middle).

Linguistic variants usually become associated with identities through a gradual process in which the association slowly permeates public awareness (Eckert, 2008). In general, a phonological variable does not inherently index any given identity. However, the multimodality of memes permits greater expressiveness—a meme in r/Animemes might use the anime art style, indexing into the aesthetic of that community explicitly.

6 Linguistic innovation

Equally as important as the study of synchronic linguistic variation is the study of diachronic linguistic change. Language change has been heavily studied in natural language processing (Rosenfeld and Erk, 2018; Martinc et al., 2020; Zhu and Jurgens, 2021b). We focus on understanding the innovation of meme templates within a semantic cluster.

RQ2: Do new meme templates co-exist with preexisting templates in the semantic cluster, or does

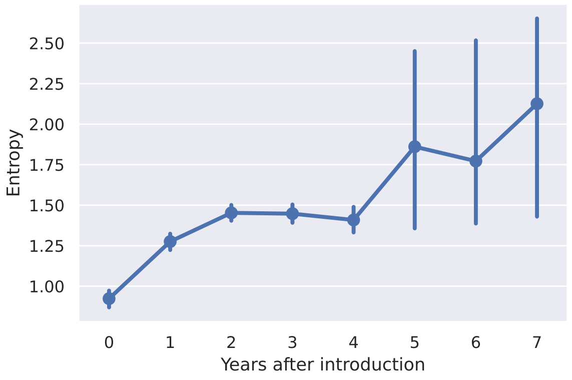


Figure 6: On average, semantic clusters diversify over time. Very old semantic clusters are rarer, leading to larger confidence intervals in later years.

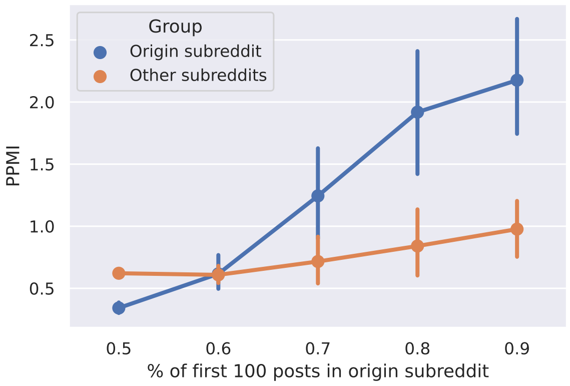


Figure 7: Communities that lead the introduction of a new template continue to use it more than others.

the most popular template monopolize the cluster?

When multiple templates that fulfill the same function appear, we expect there to be competition. Prior work has observed this competition between lexical choices, with two outcomes: new words replace old ones that serve the same function, but if similar words have discourse-relevant differences in meaning, they can coexist (Karjus et al., 2020).

Methods We measure the entropy of semantic clusters over time. If meme templates ultimately co-exist, we would expect entropy to increase; if a subset of templates dominate, we would expect the entropy to converge to a lower value.

For each semantic cluster, we group posts by the age of the semantic cluster in years at the time of posting. We define the "birth" of the cluster as when the first instance of a template in that cluster was posted. Within each year, we calculate the entropy of template distribution within each cluster.

It is possible that some clusters have low entropy early on due to data sparsity. To account for this, we filter to semantic clusters that have existed at least 5 years with at least 30 posts in all years, and resample with replacement within each year such that every year has the same number of posts. Ultimately, we conduct our analysis over 146 semantic clusters that span over 950K posts.

Results Entropy steadily increases in the years following a semantic cluster's initial introduction (Figure 6). This suggests that no one meme template grows to become the de facto template for all users; there is steady variation. This is supported by our findings in Section 5 that there are socially meaningful differences between variants.

RQ3: Do new templates diffuse widely or occupy a niche?

Language change is often socially motivated; a community can opt to use a particular variant to distinguish themselves from others (Trudgill, 1986; Giles and Powesland, 1975). Thus, we might expect meme templates to be most specific to the subreddits in which they were first introduced.

Methods We measure the extent to which template variants are ultimately specific to the subreddits that originated them.

We filter our dataset to templates which occur at least 200 times. For each template, we identify a set of "seed posts," which we define as the first 100 posts using the template. We then filter to templates with a subreddit that comprises the majority of the seed posts, which we call the "origin subreddit."

We modify the method from (Zhang et al., 2017) to measure the specificity of a template-subreddit pair by using the positive pointwise mutual information (PPMI) between templates and the subreddits in which they are used, matching other work in NLP (Church and Hanks, 1990; Jurafsky and Martin, 2009). Formally, we calculate

$$\mathrm{PPMI}(t; s \mid c) = \max\left(\log \frac{P(t \mid s, c)}{P(t \mid c)}, 0\right),$$

where $P(t \mid s, c)$ is the probability of template t appearing in subreddit s and semantic cluster c, $P(t \mid c)$ is the probability of template t in that cluster globally, and templates are only compared against others within the same semantic cluster. We calculate the PPMI over non-seed posts to measure the specificity of a template after its introduction.

Results For each template, we compare the PPMI for origin subreddits with the average PPMI of all other subreddits. Figure 7 shows a significant positive correlation between the proportion of seed posts that originated in the origin subreddit and the eventual specificity of template. These results support previous findings that lexical innovations succeed when filling in a social niche (Altmann et al., 2011; MacWhinney, 1989). We also find that large, generic subreddits (like `r/memes`, `r/dankmemes` and `r/meme`) have a significantly lower eventual PPMI than subreddits with fewer posts ($p < 0.001$, two-sample t-test), suggesting that templates originating in these large subreddits diffuse more widely.

7 Linguistic acculturation

Finally, we study how users alter their meme posting habits as they spend more time in a subreddit. Previous work on linguistic acculturation show that users adopt more community-specific language as they become enculturated within a community (Danescu-Niculescu-Mizil et al., 2013; Srivastava et al., 2018). We can ask a similar question here:

RQ4: Do veteran users in a subreddit use more community-specific templates?

Methods To answer this question, we measure the average specificity of a user's posts in successive months after they enter a community. We once again calculate the PPMI of templates as a measure of specificity; this time, we calculate the value over the full range of the dataset.

For each user in a subreddit, we bin their posts by 30-day windows starting with their first post in the subreddit (i.e., for each month after they joined), and compute the average PPMI of their posts for that time period. We filter the dataset to users with at least 10 lifetime posts and subreddits with at least 30 such users. To prevent extremely popular subreddits from unduly influencing the results, we sample up to 100 users from each subreddit to compute the average across all subreddits. This yields a total of 3,174 users in 130 subreddits.

Results We find that acculturated users use templates that are slightly more specific to the communities in which they post (Pearson's $r = 0.074, p < 0.001$), shown in Figure 8. This finding aligns with existing literature on linguistic acculturation as well as theories in new media that memes are cultural capital. The "correct" use of memes can demonstrate a user's assimilation into a shared language and identity (Nissenbaum and Shifman, 2017).

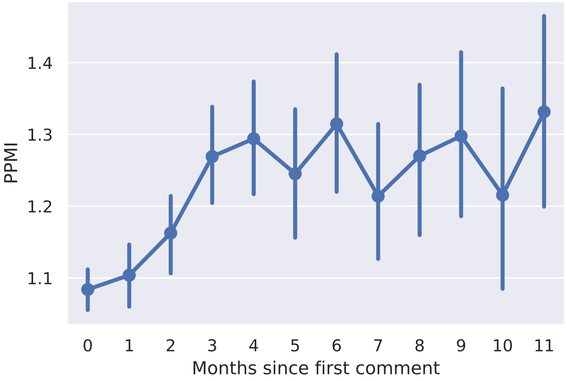


Figure 8: Veteran (acculturated) users employ more subreddit-specific meme templates.

8 Related work

Prior work on memes in NLP and social computing has largely focused on two tasks: meme understanding and modeling how memes originate and spread. Our work offers novel methods and perspectives at the intersection of these areas of research.

Meme understanding encompasses a number of discrete tasks, including classifying if memes convey harmful messages (Kiela et al., 2021; Qu et al., 2022), labeling emotion (Mishra et al., 2023), and detecting humor (Tanaka et al., 2022) or figurative speech (Liu et al., 2022) within them. While these can generally be framed as classification tasks, other work generates open-ended explanations of visual humor using large multimodal language models (Hwang and Shwartz, 2023; Hessel et al., 2023). Our work complements this existing body of research by inferring semantic variables in an unsupervised approach, leveraging the implicit structure within memes by modeling template semantics separately from the fills.

In modeling the internal structure of memes, our work draws on existing research examining the relationship between fills and templates to match semantic roles to entities within harmful memes (Sharma et al., 2023a) and mapping fill text to explanatory background information (Sharma et al., 2023b). We hope that our method of construing templates as semantic predicates can contribute to this body of work.

In the social computing space, another line of research focuses on understanding how memes originate (Morina and Bernstein, 2022) and spread across platforms (Zannettou et al., 2018). These treat meme templates as discrete tokens. We model template semantics, which have the granularity to enable analysis of variation and social meaning. Qu et al. uses CLIP to understand how memes evolve as they spread. While they use the text in comments to model the high-level concepts indexed by particular variants, we use the fill text of memes to model low-level template semantics.

9 Conclusion

In this paper, we analyze memes as a form of language subject to the same kinds of sociolinguistic variation as other modalities, such as written language and speech. We propose a new approach to understanding meme semantics, taking advantage of the multimodal structure of memes to learn semantic representations of templates from an unlabeled dataset. We use this method on a large dataset of memes scraped from Reddit, and demonstrate that it yields coherent, visually diverse clusters of semantically similar memes. We make these clusters and the code publicly available for future research. Finally, we use these clusters to study language variation and change in subreddits. We show that variations between meme template are socially meaningful and memes often share usage patterns with the textual language that has been studied in the past. We find that memes can be rich resources for understanding social language use.

10 Ethical considerations

The data used in this work was collected from Reddit in 2021 and is publicly available. To preserve the right to be forgotten, we release only the post IDs of the posts we used and the labels from the semantic clustering process. There may be offensive, hateful, or sexual messages present in the memes and comments in this dataset.

The models we trained are also publicly available. We use them only to better understand the semantics of memes. We do not train any generative models, and warn against training generative models on the data without careful consideration of how to mitigate the toxic, offensive, or otherwise harmful outputs that might be generated.

11 Limitations

We note several limitations to this work. First, we only study memes posted to Reddit meme communities, which are topically-focused and primarily English-speaking. One should be cautious in extrapolating these results to other settings in which memes are used. However, our data pipeline and models are platform agnostic—the semantic clusters can be generated from any set of memes. By making our code and models public, we hope to encourage other researchers to replicate and extend our analysis to other settings.

The meme clustering pipeline is also imperfect. As we note in the paper, the visual clustering is overly aggressive, resulting in the same base image being split into multiple template clusters. Although we show that the semantic clustering step mostly addresses this issue, improving the visual clustering could yield more precise analysis. Additionally, there are edge cases where a small visual modification changes the semantic meaning (e.g., sometimes the order of panels is reversed). The data pipeline does not always identify these visual differences.

Finally, the time series analyses are limited by data sparsity in earlier years—this is due in part to a smaller Reddit user-base, but also because many images have been deleted or removed since they were first posted. Though it is unlikely that this natural decay is systematic in a way that would significantly bias our estimates, it nonetheless reduces the precision of our analysis.

Acknowledgments

The research reported in this article was supported by funding from the National Science Foundation (Graduate Research Fellowship DGE-2146752 and grant IIS-1942591) and from the Volkswagen Foundation.

We thank Nicholas Tomlin, Lucy Li and Isaac Bleaman for helpful suggestions and feedback.

References

Eduardo G. Altmann, Janet B. Pierrehumbert, and Adilson E. Motter. 2011. Niche as a Determinant of Word Fate in Online Groups. *PLOS ONE*, 6(5):e19009.

David Bamman, Jacob Eisenstein, and Tyler Schnoebelen. 2014. Gender identity and lexical variation in social media. *Journal of Sociolinguistics*, 18(2):135–160.

Jason Baumgartner, Savvas Zannettou, Brian Keegan, Megan Squire, and Jeremy Blackburn. 2020. The pushshift reddit dataset. *ArXiv*, abs/2001.08435.

Su Lin Blodgett, Lisa Green, and Brendan O'Connor. 2016. Demographic dialectal variation in social media: A case study of African-American English. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 1119–1130, Austin, Texas. Association for Computational Linguistics.

Mary Bucholtz and Kira Hall. 2005. Identity and interaction: A sociocultural linguistic approach. *Discourse Studies*, 7(4-5):585–614.

Kathryn Campbell-Kibler. 2009. The nature of sociolinguistic perception. *Language Variation and Change*, 21(1):135–156.

Kenneth Ward Church and Patrick Hanks. 1990. Word association norms, mutual information, and lexicography. *Computational Linguistics*, 16(1):22–29.

Cristian Danescu-Niculescu-Mizil, Robert West, Dan Jurafsky, Jure Leskovec, and Christopher Potts. 2013. No country for old members: User lifecycle and linguistic change in online communities. In *Proceedings of the 22nd International Conference on World Wide Web*, WWW '13, pages 307–318, New York, NY, USA. Association for Computing Machinery.

Marco Del Tredici and Raquel Fernández. 2017. Semantic variation in online communities of practice. In *IWCS 2017 - 12th International Conference on Computational Semantics - Long papers*.

Marco Del Tredici and Raquel Fernández. 2018. Semantic Variation in Online Communities of Practice. *arXiv:1806.05847 [cs]*.

Dorottya Demszky, Devyani Sharma, Jonathan Clark, Vinodkumar Prabhakaran, and Jacob Eisenstein. 2021. Learning to recognize dialect features. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 2315–2338, Online. Association for Computational Linguistics.

Penelope Eckert. 2008. Variation and the indexical field. *Journal of Sociolinguistics*, 12(4):453–476.

Jacob Eisenstein. 2015. Systematic patterning in phonologically-motivated orthographic variation. *Journal of Sociolinguistics*, 19(2):161–188.

Jacob Eisenstein, Brendan O'Connor, Noah A. Smith, and Eric P. Xing. 2010. A latent variable model for geographic lexical variation. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 1277–1287, Cambridge, MA. Association for Computational Linguistics.

Howard Giles and Peter F. Powesland. 1975. Accommodation Theory. In *Speech Style and Social Evaluation*, Speech Style and Social Evaluation, pages 232–239. Academic Press, Oxford, England.

William L. Hamilton, Jure Leskovec, and Dan Jurafsky. 2016. Diachronic word embeddings reveal statistical laws of semantic change. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1489–1501, Berlin, Germany. Association for Computational Linguistics.

Jack Hessel, Ana Marasovic, Jena D. Hwang, Lillian Lee, Jeff Da, Rowan Zellers, Robert Mankoff, and Yejin Choi. 2023. Do androids laugh at electric sheep? humor "understanding" benchmarks from the new yorker caption contest. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 688–714, Toronto, Canada. Association for Computational Linguistics.

Dirk Hovy and Christoph Purschke. 2018. Capturing regional variation with distributed place representations and geographic retrofitting. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 4383–4394, Brussels, Belgium. Association for Computational Linguistics.

EunJeong Hwang and Vered Shwartz. 2023. MemeCap: A Dataset for Captioning and Interpreting Memes. ArXiv:2305.13703 [cs].

Dan Jurafsky, Victor Chahuneau, Bryan R. Routledge, and Noah A. Smith. 2014. Narrative framing of consumer sentiment in online restaurant reviews. *First Monday*.

Daniel Jurafsky and James H. Martin. 2009. *Speech and Language Processing (2nd Edition)*. Prentice-Hall, Inc., USA.

Andres Karjus, Richard A. Blythe, Simon Kirby, and Kenny Smith. 2020. Communicative need modulates competition in language change.

Douwe Kiela, Hamed Firooz, Aravind Mohan, Vedanuj Goswami, Amanpreet Singh, Pratik Ringshia, and Davide Testuggine. 2021. The Hateful Memes Challenge: Detecting Hate Speech in Multimodal Memes. ArXiv:2005.04790 [cs].

Gunther R. Kress and Theo Van Leeuwen. 2001. *Multimodal Discourse: The Modes and Media of Contemporary Communication*. Arnold; Oxford University Press.

William Labov. 1963. The Social Motivation of a Sound Change. WORD, 19(3):273–309.

Chen Liu, Gregor Geigle, Robin Krebs, and Iryna Gurevych. 2022. FigMemes: A dataset for figurative language identification in politically-opinionated memes. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 7069–7086, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach.

Li Lucy and David Bamman. 2021. Characterizing English variation across social media communities with BERT. *Transactions of the Association for Computational Linguistics*, 9:538–556.

Brian MacWhinney. 1989. Competition and lexical categorization. In Roberta Corrigan, Fred R. Eckman, and Michael Noonan, editors, *Current Issues in Linguistic Theory*, volume 61, page 195. John Benjamins Publishing Company, Amsterdam.

Matej Martinc, Petra Kralj Novak, and Senja Pollak. 2020. Leveraging contextual embeddings for detecting diachronic semantic shift. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 4811–4819, Marseille, France. European Language Resources Association.

Shreyash Mishra, S. Suryavardan, Parth Patwa, Megha Chakraborty, Anku Rani, Aishwarya Reganti, Aman Chadha, Amitava Das, Amit Sheth, Manoj Chinnakotla, Asif Ekbal, and Srijan Kumar. 2023. Memotion 3: Dataset on Sentiment and Emotion Analysis of Codemixed Hindi-English Memes. ArXiv:2303.09892 [cs].

Burt L. Monroe, Michael P. Colaresi, and Kevin M. Quinn. 2017. Fightin' Words: Lexical Feature Selection and Evaluation for Identifying the Content of Political Conflict. *Political Analysis*, 16(4):372–403.

Durim Morina and Michael S. Bernstein. 2022. A web-scale analysis of the community origins of image memes. *Proc. ACM Hum.-Comput. Interact.*, 6(CSCW1).

Dong Nguyen, Laura Rosseel, and Jack Grieve. 2021. On learning and representing social meaning in NLP: a sociolinguistic perspective. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 603–612, Online. Association for Computational Linguistics.

Asaf Nissenbaum and Limor Shifman. 2017. Internet memes as contested cultural capital: The case of 4chan's /b/ board. *New Media & Society*, 19(4):483–501.

Pamela Perniss. 2018. Why We Should Study Multimodal Language. *Frontiers in Psychology*, 9.

Jingnong Qu, Liunian Harold Li, Jieyu Zhao, Sunipa Dev, and Kai-Wei Chang. 2022. DisinfoMeme: A Multimodal Dataset for Detecting Meme Intentionally Spreading Out Disinformation.

Yiting Qu, Xinlei He, Shannon Pierson, Michael Backes, Yang Zhang, and Savvas Zannettou. 2023. On the Evolution of (Hateful) Memes by Means of Multimodal Contrastive Learning. In *2023 IEEE Symposium on Security and Privacy (SP)*, pages 293–310. IEEE Computer Society.

Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. 2021. Learning transferable visual models from natural language supervision.

Alex Rosenfeld and Katrin Erk. 2018. Deep neural models of semantic shift. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 474–484, New Orleans, Louisiana. Association for Computational Linguistics.

Shivam Sharma, Atharva Kulkarni, Tharun Suresh, Himanshi Mathur, Preslav Nakov, Md. Shad Akhtar, and Tanmoy Chakraborty. 2023a. Characterizing the entities in harmful memes: Who is the hero, the villain, the victim? In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, pages 2149–2163, Dubrovnik, Croatia. Association for Computational Linguistics.

Shivam Sharma, Ramaneswaran S, Udit Arora, Md. Shad Akhtar, and Tanmoy Chakraborty. 2023b. MEMEX: Detecting explanatory evidence for memes via knowledge-enriched contextualization. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 5272–5290, Toronto, Canada. Association for Computational Linguistics.

Sameer B. Srivastava, Amir Goldberg, V. Govind Manian, and Christopher Potts. 2018. Enculturation Trajectories: Language, Cultural Adaptation, and Individual Outcomes in Organizations. *Management Science*, 64(3):1348–1364.

Ian Stewart, Stevie Chancellor, Munmun De Choudhury, and Jacob Eisenstein. 2017. #Anorexia, #anarexia, #anarexyia: Characterizing online community practices with orthographic variation. In *2017 IEEE International Conference on Big Data (Big Data)*, pages 4353–4361.

S.A. Tagliamonte. 2006. *Analysing Sociolinguistic Variation*. Key Topics in Sociolinguistics. Cambridge University Press.

Kohtaro Tanaka, Hiroaki Yamane, Yusuke Mori, Yusuke Mukuta, and Tatsuya Harada. 2022. Learning to evaluate humor in memes based on the incongruity theory. In *Proceedings of the Second Workshop on When Creative AI Meets Conversational AI*, pages 81–93, Gyeongju, Republic of Korea. Association for Computational Linguistics.

William Timkey and Marten van Schijndel. 2021. All bark and no bite: Rogue dimensions in transformer language models obscure representational quality. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 4527–4546, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

V. A. Traag, L. Waltman, and N. J. van Eck. 2019. From Louvain to Leiden: guaranteeing well-connected communities. *Scientific Reports*, 9(1):5233. Number: 1 Publisher: Nature Publishing Group.

Peter Trudgill. 1986. *Dialects in Contact*. Number 10 in Language in Society. B. Blackwell, Oxford, UK ; New York, NY, USA.

Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Remi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest, and Alexander Rush. 2020. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 38–45, Online. Association for Computational Linguistics.

Savvas Zannettou, Tristan Caulfield, Jeremy Blackburn, Emiliano De Cristofaro, Michael Sirivianos, Gianluca Stringhini, and Guillermo Suarez-Tangil. 2018. On the origins of memes by means of fringe web communities. In *Proceedings of the Internet Measurement Conference 2018*, IMC '18, page 188–202, New York, NY, USA. Association for Computing Machinery.

Dongyu Zhang, Minghao Zhang, Heting Zhang, Liang Yang, and Hongfei Lin. 2021. MultiMET: A multimodal dataset for metaphor understanding. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 3214–3225, Online. Association for Computational Linguistics.

Justine Zhang, William L. Hamilton, Cristian Danescu-Niculescu-Mizil, Dan Jurafsky, and Jure Leskovec. 2017. Community Identity and User Engagement in a Multi-Community Landscape. ArXiv:1705.09665 [physics].

Qing Zhang. 2005. A Chinese yuppie in Beijing: Phonological variation and the construction of a new professional identity. *Language in Society*, 34(3):431–466.

Jian Zhu and David Jurgens. 2021a. Idiosyncratic but not arbitrary: Learning idiolects in online registers reveals distinctive yet consistent individual styles. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 279–297.

Jian Zhu and David Jurgens. 2021b. The structure of online social networks modulates the rate of lexical change. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 2201–2218.

A Details on visually clustering templates

A.1 Preprocessing

One common meme layout that would caused issues in the template clustering step was a text frame around the source image, where there is a border around a source image, as well as some text above or below (see Figure 9a for an example).

For each image, we use a rectangular kernel to detect potential text patches, replace those patches with the background color, and finally identify the bounding box for the remaining source image without any excess borders. Figure 9 walks through the steps visually.

A.2 Image hashing

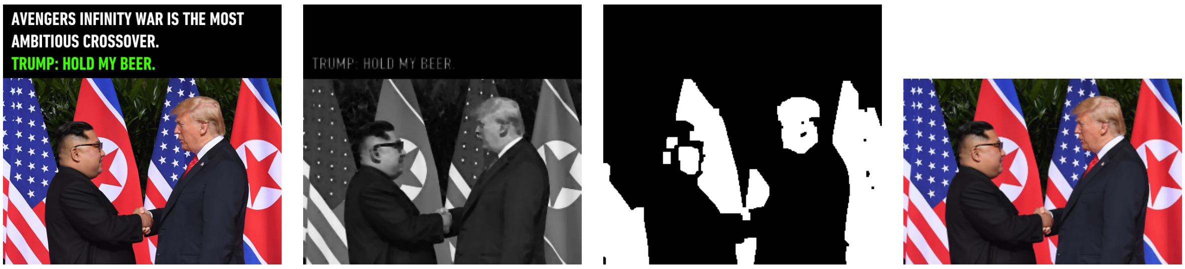
We create a 64-bit perceptual hash for each preprocessed image in the dataset. The preprocessed images are only used for the hashing step; all other steps use the original image. Figure 10 shows examples of images whose preprocessed versions have the same hash.

A.3 Hash clustering

We first compute pairwise Hamming distance between all the hashes. Then, we discard any pairs with a Hamming distance greater than 10. Then we construct a network of hashes, where edges of the graph are calculated as $11 - d_{ij}$ for Hamming distance d_{ij} between the ith and jth hashes before finally using the Leiden algorithm to cluster hashes. Figure 11 shows the top 18 most heavily represented hash clusters, with 4 sampled images from each.

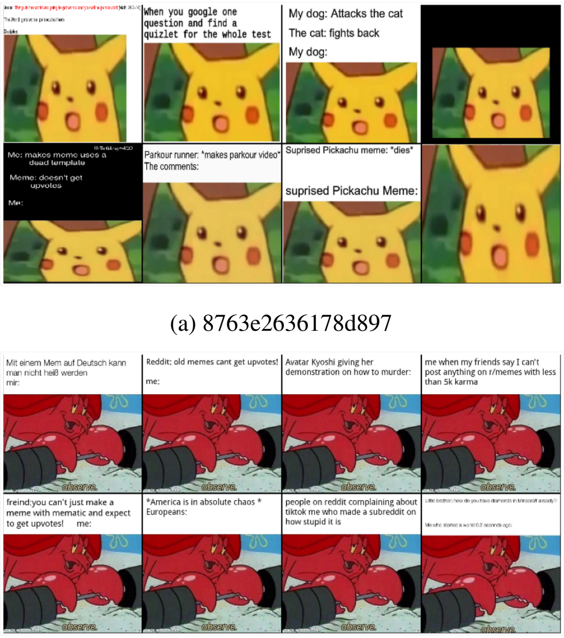
We use the Leiden algorithm with the Constant Potts Model (CPM) as the quality function; we use a density of 1.0, but experiments with other density values (0.01, 0.1, 10) yielded qualitatively similar or worse results. The algorithm results in aggressively split clusters, where each cluster is coherent, but there are some memes that share a base template but are split between two clusters (e.g. Winnie the Pooh appears twice in Figure 11, among others).

We find that these duplicate hash clusters are often merged when we generate the semantic clusters.



(a) Original image (b) Remove text (c) Remove text artifacts (d) Trim excess borders

Figure 9: Example of the image preprocessing steps, described in Section 2.1


When you google one question and find a quizlet for the whole test
My dog: Attacks the cat
The cat: fights back
My dog:
Me: makes meme uses a dead template
Meme: doesn't get upvotes
Me:
Parkour runner: *makes parkour video*
The comments:
Suprised Pickachu meme: *dies*
suprised Pickachu Meme:
Mit einem Mem auf Deutsch kann man nicht heiß werden
mir:
Reddit: old memes cant get upvotes!
me:
Avatar Kyoshi giving her demonstration on how to murder:
me when my friends say I can't post anything on r/memes with less than 5k karma
freind:you can't just make a meme with mematic and expect to get upvotes! me:
*America is in absolute chaos *
Europeans:
people on reddit complaining about tiktok me who made a subreddit on how stupid it is
observe.

(a) 8763e2636178d897

(b) 9465a9596e1a66bc

Figure 10: Examples of groups of meme instances with the same perceptual hash when processed.

B Details on the semantic clusters

B.1 Training details

The RoBERTa and CLIP models were fine-tuned using the Huggingface transformers library (Wolf et al., 2020). The RoBERTa model was fine-tuned on an NVIDIA A5000 GPU for 3 epochs (27 hours) with a learning rate of 1e-6. The CLIP model was fine-tuned with mixed-precision on 8 NVIDIA A6000 GPUs for 3 epochs (27 hours) with a learning rate of 4e-6. There was no hyperparameter tuning.

B.2 Calculating edge weights.

To prepare the data for clustering, we constructed a weighted adjacency matrix by keeping only the top 10 nearest neighbors for each template embedding.

Figure 11: Sample images from the 18 most common perceptual hash clusters; each row contains two clusters with four sampled images.

Model	# Clusters	Avg. Size
RoBERTa	784	8.7
CLIP	657	10.4
CLIP-diff	617	11.1
Concat.	685	10.0

Table 2: Count and average sizes of semantic clusters generated from each embedding model.

We calculated the weight as

$$w_a(b) = \lambda^{r_a(b)},$$

where $w_a(b)$ is the weight of the edge between templates b and a, $r_a(b)$ is the rank of the cosine similarity between a and b, and λ is a discount factor (we set this to 0.9).

We chose to weight by a function of ranked similarity instead of cosine similarity directly because we found the cosine similarity was often low even for the embeddings of semantically equivalent templates, resulting poor recall and many small clusters. Using the weighted ranking, we get more templates per cluster without introducing too many false positives.

B.3 Outputs

There are 784 semantic clusters generated from the RoBERTa embeddings—table 2 shows statistics for cluster sizes for all of the models. We include more extensive examples of the clusters in appendix D.

The distribution of clusters is highly skewed. Figure 12 shows the distribution of cluster sizes in our Reddit dataset, for the RoBERTa embeddings. The largest 103 clusters account for 50% of the posts the dataset. The largest 10 account for 12%—this distribution is shown in Figure 13. The median semantic cluster contains 2,587 posts; the smallest one in the dataset contains 122.

Similarly, posts in r/memes account for 38% of all posts in the dataset. Accordingly, most of our analyses have employed stratified sampling or techniques to normalize by size. Figure 14 shows that the top three most commonly occurring subreddits are all generic meme subreddits; other more topically or geographically localized ones follow.

C Model evaluation details

Human annotators were presented with pairs of images with the following instructions:

> You will be looking at pairs of memes;

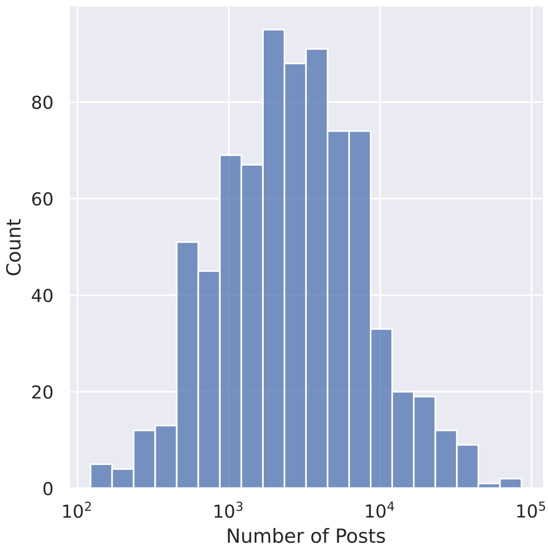


Figure 12: Distribution of semantic cluster coverage

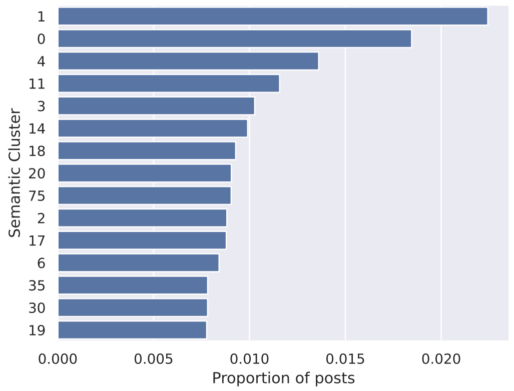


Figure 13: Proportion of posts using the 15 most common clusters

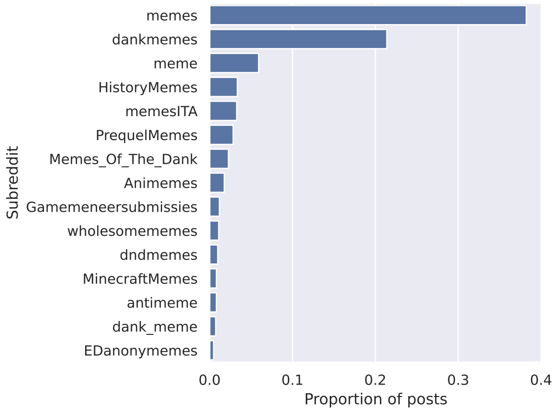


Figure 14: Proportion of posts in the 15 most common subreddits

Model	Krippendorff's α
RoBERTa	0.77
CLIP	0.76
CLIP-diff	0.67
Concat.	0.76

Table 3: Krippendorff's α for the human evaluations, stratified by model.

> for each pair, you will be answering two yes/no questions.
>
> 1. is this pair semantically similar (can you conceivably copy / paste the text of one into the other with minor changes and have it still make sense)
> 2. is this pair visually similar (do they have the same characters, art style, etc? e.g. two harry potter memes. If the layout is the same but the characters are different, you should mark it as not visually similar)

The annotators were three student employees whose job responsibilities included such annotations. They were recruited at a large public university in the United States and paid above the minimum wage for the geographic region they're in. They were informed that their annotations would be used to evaluate different model outputs.

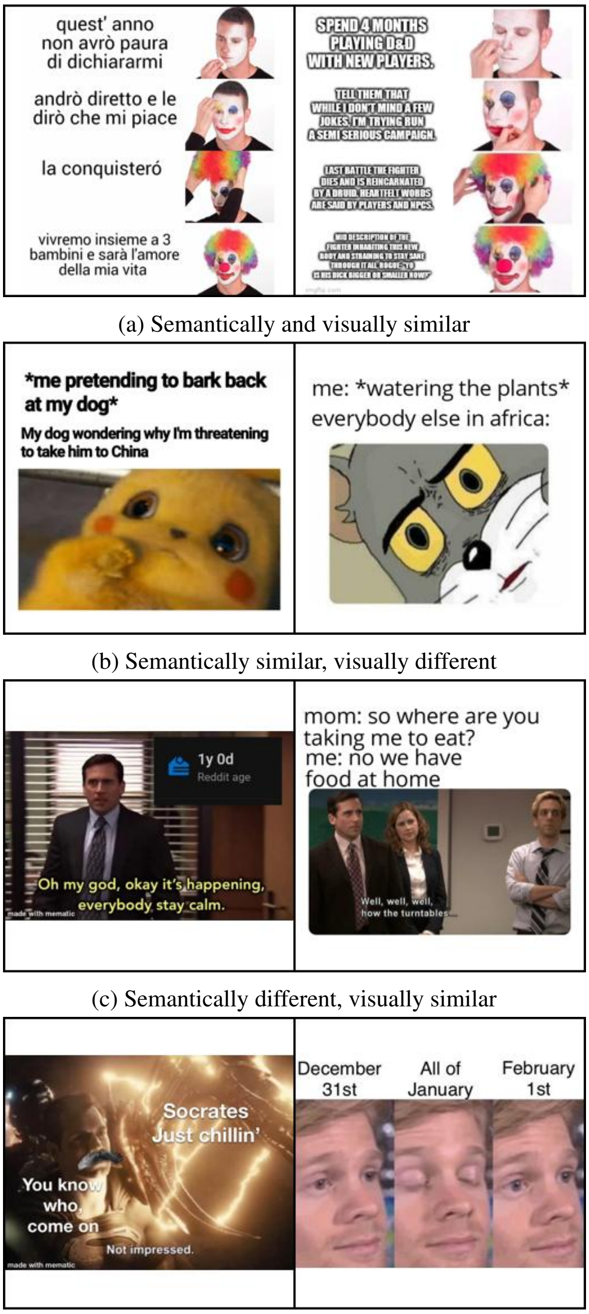
C.1 Agreement

For the RoBERTa and CLIP models, all three annotators performed evaluations. For the CLIP-diff and Concat. models, which were evaluated later, two of the three annotators performed evaluations after we had verified high agreement on RoBERTa and CLIP. When calculating the Krippendorff's α for each of the three models, the agreement is similar across all models with the exception of CLIP-diff, which had slightly lower agreement. Table 3 shows the agreement scores for each.

C.2 Examples and further discussion

Figure 15 includes examples of some of the image pairs presented to annotators, as well as the expected judgments for semantic and visual similarity.

We found that CLIP-based methods would include clusters that were visually similar but semantically different. Figure 16 shows an example of a


quest' anno non avrò paura di dichiararmi
andrò diretto e le dirò che mi piace
la conquisteró
vivremo insieme a 3 bambini e sarà l'amore della mia vita
SPEND 4 MONTHS PLAYING D&D WITH NEW PLAYERS.
TELL THEM THAT WHILE I DON'T MIND A FEW JOKES, I'M TRYING RUN A SEMI SERIOUS CAMPAIGN.
LAST BATTLE THE FIGHTER DIES AND IS REINCARNATED BY A DRUID. HEARTFELT WORDS ARE SAID BY PLAYERS AND NPCS.
MID DESCRIPTION OF THE FIGHTER INHABITING THIS NEW BODY AND STRAINING TO STAY SANE THROUGH IT ALL. ROGUE: "YO IS HIS DICK BIGGER OR SMALLER NOW?"
me pretending to bark back at my dog
My dog wondering why I'm threatening to take him to China
me: *watering the plants*
everybody else in africa:
1y 0d
Reddit age
Oh my god, okay it's happening, everybody stay calm.
made with mematic
mom: so where are you taking me to eat?
me: no we have food at home
Well, well, well, how the turntables...
Socrates
Just chillin'
You know who, come on
Not impressed.
made with mematic
December 31st
All of January
February 1st

(a) Semantically and visually similar

(b) Semantically similar, visually different

(c) Semantically different, visually similar

(d) Semantically and visually different

Figure 15: Example meme pairs for annotation

Model	**Precision**	**Visual-adjusted**
All		
RoBERTa	**0.78**	**0.44**
CLIP	0.65	-0.09
CLIP-diff	0.69	0.18
Concat.	0.70	0.30
Common		
RoBERTa	0.81	**0.53**
CLIP	0.80	0.06
CLIP-diff	**0.84**	0.47
Concat.	0.80	0.42
Random		
RoBERTa	**0.67**	**0.23**
CLIP	0.46	-0.22
CLIP-diff	0.45	-0.13
Concat.	0.51	0.17

Table 4: Comparison of cluster quality for different embedding models, stratified by the random and common samples. CLIP-based models yield clusters that are biased towards visual features.

cluster generated with CLIP embeddings that contains meme templates that have different semantic functions, but almost all contain characters from the *Star Wars* franchise. This cluster had an visually adjusted precision score of -2.08.

The failure case of over-indexing on visual similarity is not restricted to creating semantically incoherent clusters with visual similarity. The CLIP-diff and, to a greater extent, Concat embeddings were good at surfacing less-used variants of templates. However, they do so at the expense of splitting into several stylistically delineated semantic clusters, which is detrimental to our desired analysis on variation. Figure 17 shows how templates from the large RoBERTa cluster for declarative templates are divided into several Concat clusters.

D Further semantic cluster examples

Figures 18,19,20,21 contain more examples of semantic clusters generated using the different embedding models. Content warning: though we tried to filter for toxic speech, these memes may still contain offensive content.

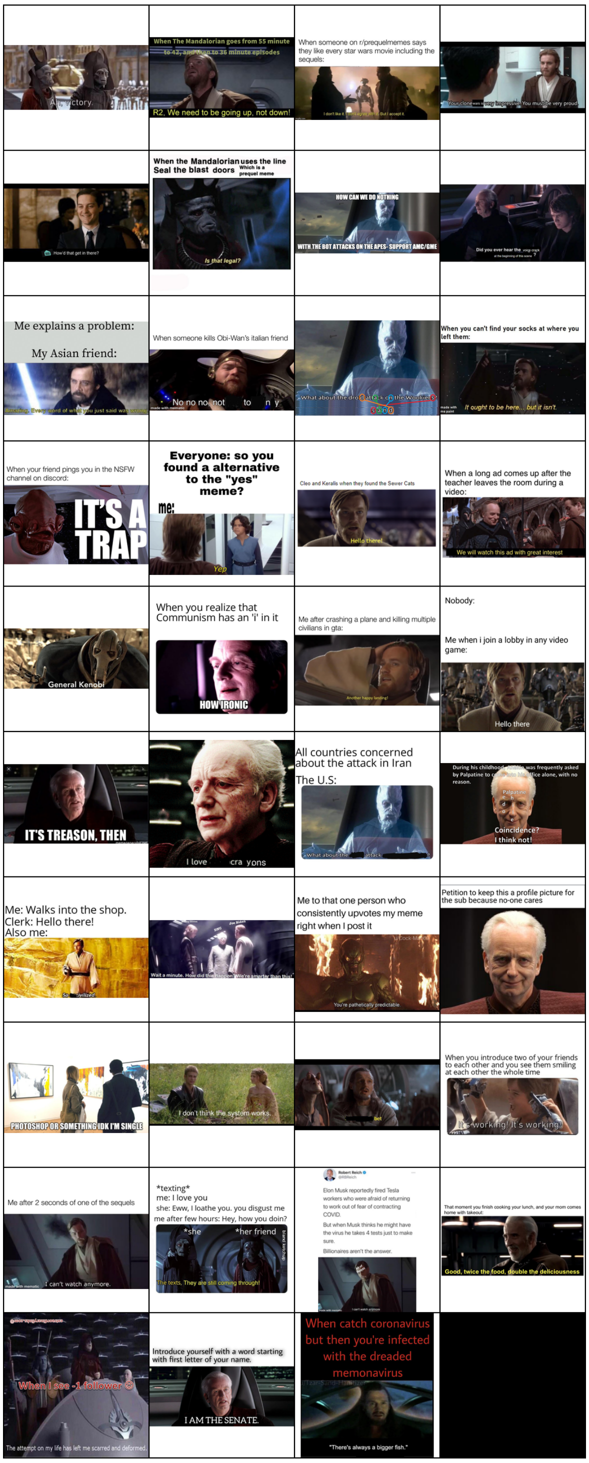
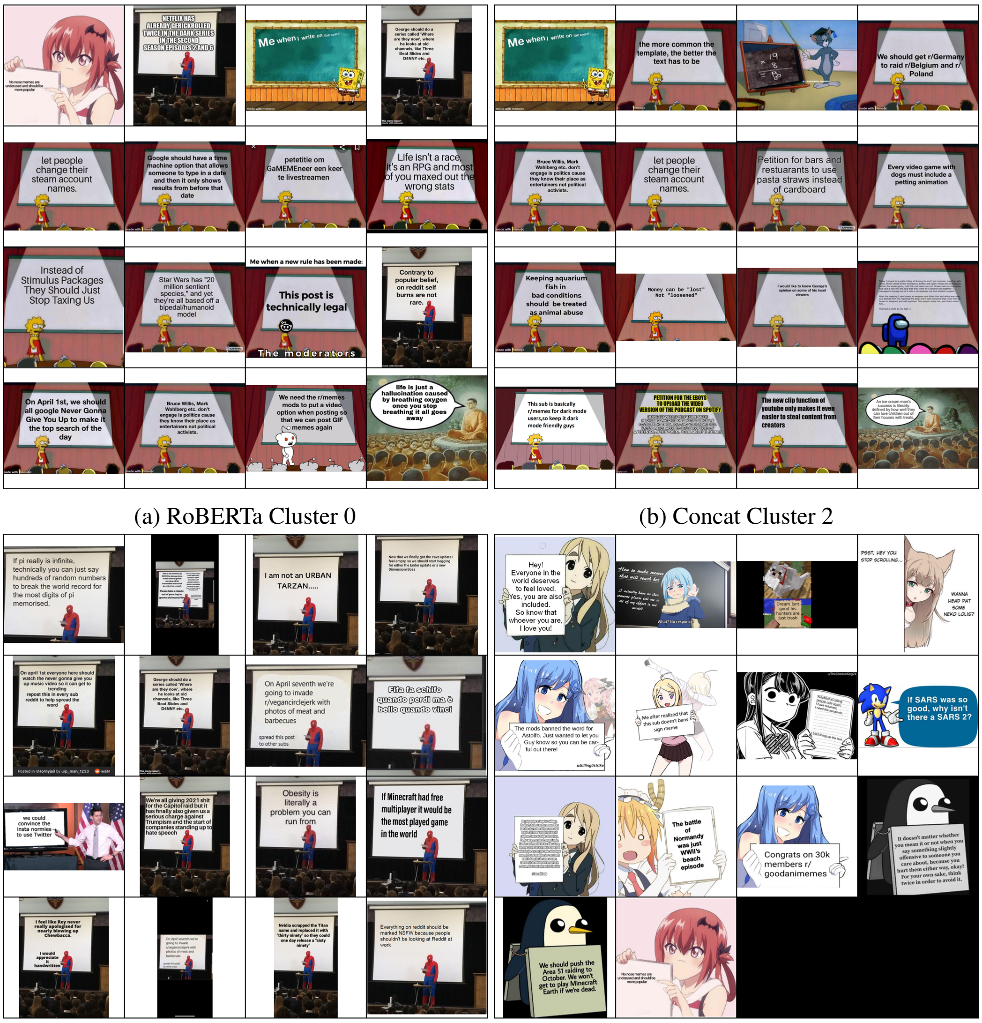

Ah, victory.
When The Mandalorian goes from 55 minute to 42, and then to 36 minute episodes
R2, We need to be going up, not down!
When someone on r/prequelmemes says they like every star wars movie including the sequels:
I don't like it. I don't agree with it. But I accept it.
Your clone wars is very impressive. You must be very proud.
How'd that get in there?
When the Mandalorian uses the line Seal the blast doors Which is a prequel meme
Is that legal?
HOW CAN WE DO NOTHING
WITH THE BOT ATTACKS ON THE APES- SUPPORT AMC/GME
Did you ever hear the voice crack at the beginning of this scene ?
Me explains a problem:
My Asian friend:
Amazing. Every word of what you just said was wrong
When someone kills Obi-Wan's italian friend
No no no, not to n y
made with mematic
What about the droid attack on the Wookiees?
sand
When you can't find your socks at where you left them:
made with ms paint
It ought to be here... but it isn't.
When your friend pings you in the NSFW channel on discord:
IT'S A TRAP
Everyone: so you found a alternative to the "yes" meme?
me:
Yep
Cleo and Keralis when they found the Sewer Cats
Hello there!
When a long ad comes up after the teacher leaves the room during a video:
We will watch this ad with great interest
General Kenobi
When you realize that Communism has an 'i' in it
HOW IRONIC
Me after crashing a plane and killing multiple civilians in gta:
Another happy landing!
Nobody:
Me when i join a lobby in any video game:
Hello there
IT'S TREASON, THEN
I love cra yons
All countries concerned about the attack in Iran
The U.S:
What about the [redacted] attack [redacted]?
During his childhood, Anakin was frequently asked by Palpatine to come into his office alone, with no reason.
Palpatine
Coincidence? I think not!
Me: Walks into the shop.
Clerk: Hello there!
Also me:
So uncivilized.
DNC
Joe Biden
Wait a minute. How did this happen? We're smarter than this!
Me to that one person who consistently upvotes my meme right when I post it
u/Cock-Master
You're pathetically predictable.
Petition to keep this a profile picture for the sub because no-one cares
PHOTOSHOP OR SOMETHING IDK I'M SINGLE
I don't think the system works.
[redacted] bet
When you introduce two of your friends to each other and you see them smiling at each other the whole time
It's working! It's working!
Me after 2 seconds of one of the sequels
made with mematic
I can't watch anymore.
texting
me: I love you
she: Eww, I loathe you. you disgust me
me after few hours: Hey, how you doin?
*she
*her friend
bland ketchup
The texts, They are still coming through!
Robert Reich
@RBReich
Elon Musk reportedly fired Tesla workers who were afraid of returning to work out of fear of contracting COVID.
But when Musk thinks he might have the virus he takes 4 tests just to make sure.
Billionaires aren't the answer.
made with mematic
I can't watch anymore
That moment you finish cooking your lunch, and your mom comes home with takeout:
Good, twice the food, double the deliciousness
When I see -1 follower 😊
The attempt on my life has left me scarred and deformed.
Introduce yourself with a word starting with first letter of your name.
I AM THE SENATE.
When catch coronavirus but then you're infected with the dreaded memonavirus
u/Tzar-Sand-Hanitizer
"There's always a bigger fish."

Figure 16: This Concat cluster has a low visual-adjusted precision.


NETFLIX HAS ALREADY GERICKROLLED TWICE IN THE DARK SERIES IN THE SECOND SEASON EPISODES 2 AND 6
Me when I write on the board
let people change their steam account names.
Google should have a time machine option that allows someone to type in a date and then it only shows results from before that date
petetitie om GaMEMEneer een keer te livestreamen
Life isn't a race, it's an RPG and most of you maxed out the wrong stats
Instead of Stimulus Packages They Should Just Stop Taxing Us
Star Wars has "20 million sentient species," and yet they're all based off a bipedal/humanoid model
Me when a new rule has been made:
This post is technically legal
The moderators
Contrary to popular belief, on reddit self burns are not rare.
On April 1st, we should all google Never Gonna Give You Up to make it the top search of the day
Bruce Willis, Mark Wahlberg etc. don't engage is politics cause they know their place as entertainers not political activists.
We need the r/memes mods to put a video option when posting so that we can post GIF memes again
life is just a hallucination caused by breathing oxygen once you stop breathing it all goes away
the more common the template, the better the text has to be
We should get r/Germany to raid r/Belgium and r/Poland
Petition for bars and restuarants to use pasta straws instead of cardboard
Every video game with dogs must include a petting animation
Keeping aquarium fish in bad conditions should be treated as animal abuse
Money can be "lost" Not "loosened"
This sub is basically r/memes for dark mode users,so keep it dark mode friendly guys
PETITION FOR THE EBOYS TO UPLOAD THE VIDEO VERSION OF THE PODCAST ON SPOTIFY
The new clip function of youtube only makes it even easier to steal content from creators
If pi really is infinite, technically you can just say hundreds of random numbers to break the world record for the most digits of pi memorised.
I am not an URBAN TARZAN.....
On april 1st everyone here should watch the never gonna give you up music video so it can get to trending repost this in every sub reddit to help spread the word
On April seventh we're going to invade r/vegancirclejerk with photos of meat and barbecues
spread this post to other subs
Fifa fa schifo quando perdi ma è bello quando vinci
we could convince the insta normies to use Twitter
We're all giving 2021 shit for the Capitol raid but it has finally also given us a serious charge against Trumpism and the start of companies standing up to hate speech
Obesity is literally a problem you can run from
If Minecraft had free multiplayer it would be the most played game in the world
I feel like Rey never really apologized for nearly blowing up Chewbacca.
Nvidia scrapped the Titan name and replaced it with "thirty ninety" so they could one day release a "sixty ninety"
Everything on reddit should be marked NSFW because people shouldn't be looking at Reddit at work
Hey! Everyone in the world deserves to feel loved. Yes, you are also included. So know that whoever you are, I love you!
PSST, HEY YOU STOP SCROLLING...
WANNA HEAD PAT SOME NEKO LOLIS?
The mods banned the word for Astolfo. Just wanted to let you Guy know so you can be careful out there!
Me after realized that this sub doesn't bans sign meme
if SARS was so good, why isn't there a SARS 2?
The battle of Normandy was just WWII's beach episode
Congrats on 30k members r/ goodanimemes
It doesn't matter whether you mean it or not when you say something slightly offensive to someone you care about, because you hurt them either way, okay? For your own sake, think twice in order to avoid it.
We should push the Area 51 raiding to October. We won't get to play Minecraft Earth if we're dead.

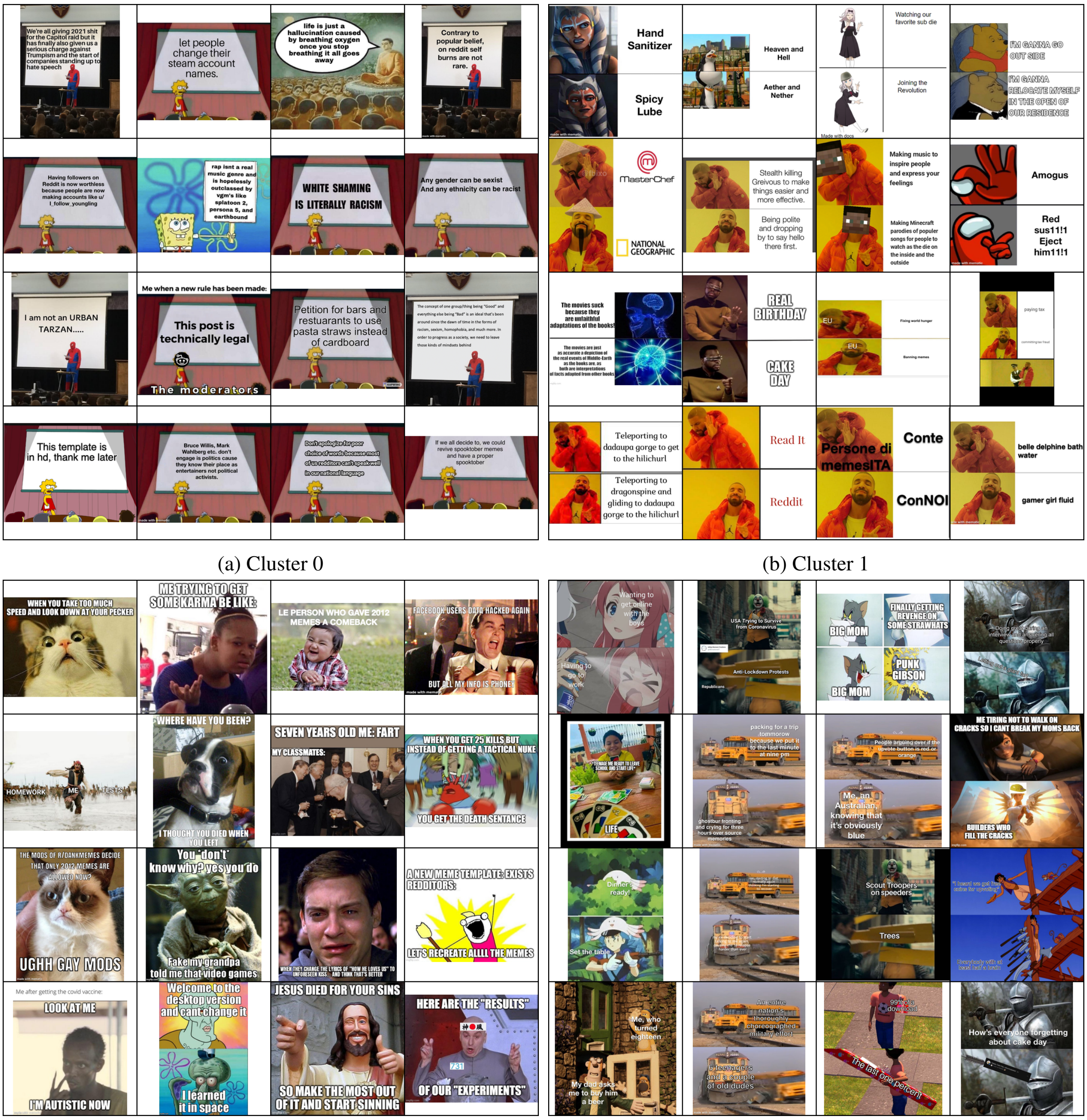
(a) RoBERTa Cluster 0

(b) Concat Cluster 2

(c) Concat Cluster 31

(d) Concat Cluster 134

Figure 17: Templates that are clustered together by RoBERTa appear in stylistically delineated semantic clusters in the Concat clusters. Displayed are a sample of up to 16 templates from each cluster.


We're all giving 2021 shit for the Capitol raid but it has finally also given us a serious charge against Trumpism and the start of companies standing up to hate speech
let people change their steam account names.
life is just a hallucination caused by breathing oxygen once you stop breathing it all goes away
Contrary to popular belief, on reddit self burns are not rare.
Having followers on Reddit is now worthless because people are now making accounts like u/ I_follow_youngling
rap isnt a real music genre and is hopelessly outclassed by vgm's like splatoon 2, persona 5, and earthbound
WHITE SHAMING IS LITERALLY RACISM
Any gender can be sexist And any ethnicity can be racist
I am not an URBAN TARZAN.....
Me when a new rule has been made:
This post is technically legal
The moderators
Petition for bars and restuarants to use pasta straws instead of cardboard
This template is in hd, thank me later
Bruce Willis, Mark Wahlberg etc. don't engage is politics cause they know their place as entertainers not political activists.
Don't apologize for poor choice of words because most of us redditors can't speak well in our national language
If we all decide to, we could revive spooktober memes and have a proper spooktober
Hand Sanitizer
Spicy Lube
Heaven and Hell
Aether and Nether
Watching our favorite sub die
Joining the Revolution
I'M GANNA GO OUT SIDE
I'M GANNA RELOCATE MYSELF IN THE OPEN OF OUR RESIDENCE
MasterChef
NATIONAL GEOGRAPHIC
Stealth killing Grievous to make things easier and more effective.
Being polite and dropping by to say hello there first.
Making music to inspire people and express your feelings
Making Minecraft parodies of populer songs for people to watch as the die on the inside and the outside
Amogus
Red sus11!1 Eject him11!1
The movies suck because they are unfaithful adaptations of the books!
REAL BIRTHDAY
CAKE DAY
EU
EU
Teleporting to dadaupa gorge to get to the hilichurl
Teleporting to dragonspine and gliding to dadaupa gorge to the hilichurl
Read It
Reddit
Persone di memesITA
Conte
ConNO!
belle delphine bath water
gamer girl fluid
WHEN YOU TAKE TOO MUCH SPEED AND LOOK DOWN AT YOUR PECKER
ME TRYING TO GET SOME KARMA BE LIKE:
LE PERSON WHO GAVE 2012 MEMES A COMEBACK
FACEBOOK USERS DATA HACKED AGAIN
BUT ALL MY INFO IS PHONEY
HOMEWORK
ME
WHERE HAVE YOU BEEN?
I THOUGHT YOU DIED WHEN YOU LEFT
SEVEN YEARS OLD ME: FART
MY CLASSMATES:
WHEN YOU GET 25 KILLS BUT INSTEAD OF GETTING A TACTICAL NUKE
YOU GET THE DEATH SENTANCE
THE MODS OF R/DANKMEMES DECIDE THAT ONLY 2012 MEMES ARE ALLOWED NOW?
UGHH GAY MODS
You 'don't' know why? yes you do
Fake my grandpa told me that video games
A NEW MEME TEMPLATE EXISTS REDDITORS:
LET'S RECREATE ALLLL THE MEMES
Me after getting the covid vaccine:
LOOK AT ME
I'M AUTISTIC NOW
Welcome to the desktop version and cant change it
I learned it in space
JESUS DIED FOR YOUR SINS
SO MAKE THE MOST OUT OF IT AND START SINNING
HERE ARE THE "RESULTS"
OF OUR "EXPERIMENTS"
Wanting to get online with the boys
Having to go to work
USA Trying to Survive from Coronavirus
Anti-Lockdown Protests
Republicans
BIG MOM
FINALLY GETTING REVENGE ON SOME STRAWHATS
BIG MOM
PUNK GIBSON
LIFE
packing for a trip tommorow because we put it to the last minute at nine pm
ghostbur fronting and crying for three hours over source memories
People arguing over if the upvote button is red or orange
Me, an Australian, knowing that it's obviously blue
ME TIRING NOT TO WALK ON CRACKS SO I CANT BREAK MY MOMS BACK
BUILDERS WHO FILL THE CRACKS
Dinner's ready!
Set the table
Scout Troopers on speeders
Trees
Me, who turned eighteen
My dad asks me to buy him a beer
An entire nation's thoroughly choreographed military effort
a teenager's download
The last one percent
How's everyone forgetting about cake day

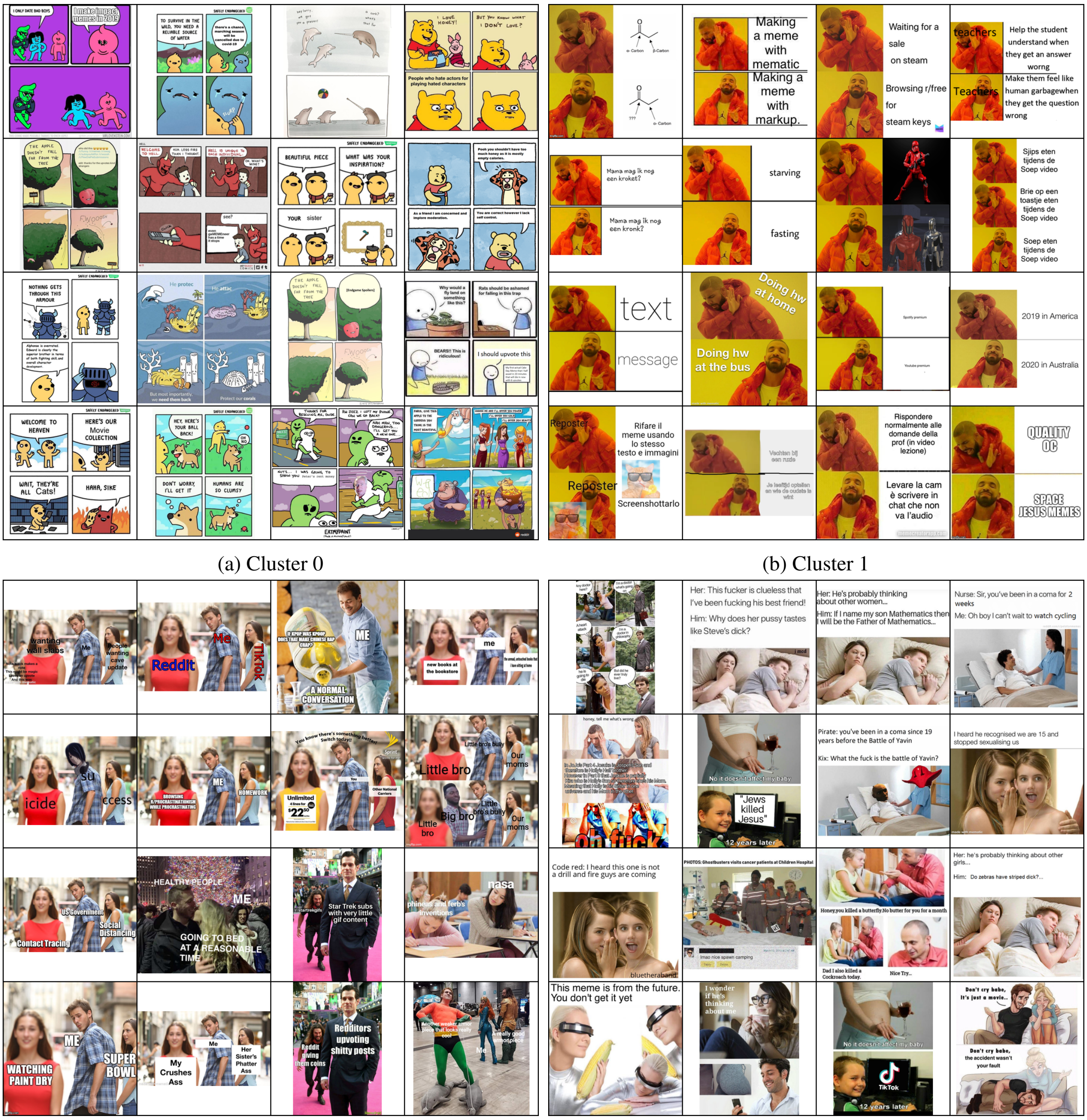
(a) Cluster 0

(b) Cluster 1

(c) Cluster 10

(d) Cluster 30

Figure 18: Samples from RoBERTa clusters


I make impact memes in 2016
HE'S OUR Movie COLLECTION
WELCOME TO HEAVEN
WAIT, THEY'RE ALL Cats!
HAHA, SIKE
DON'T WORRY I'LL GET IT
HUMANS ARE SO CLUMSY
NOTHING GETS THROUGH THIS ARMOUR
He protec
He attac
But most importantly, we need them back
Protect our corals
I should upvote this
Making a meme with mematic
Making a meme with markup.
Waiting for a sale on steam
Browsing r/free for steam keys
teachers
Help the student understand when they get an answer worng
Teachers
Make them feel like human garbagewhen they get the question wrong
Mama mag ik nog een kroket?
starving
fasting
Sjips eten tijdens de Soep video
Brie op een toastje eten tijdens de Soep video
Soep eten tijdens de Soep video
text
message
Doing hw at home
Doing hw at the bus
2019 in America
2020 in Australia
Reposter
Rifare il meme usando lo stesso testo e immagini
Screenshottarlo
Rispondere normalmente alle domande della prof (in video lezione)
Levare la cam e scrivere in chat che non va l'audio
QUALITY OC
SPACE JESUS MEMES
Reddit
TikTok
ME
A NORMAL CONVERSATION
me
icide
su
ccess
HOMEWORK
Little bro
Our moms
Big bro
HEALTHY PEOPLE
GOING TO BED AT A REASONABLE TIME
Contact Tracing
US Government
Social Distancing
Star Trek subs with very little gif content
nasa
phineas and ferb's inventions
WATCHING PAINT DRY
SUPER BOWL
My Crushes Ass
Her Sister's Phatter Ass
Redditors upvoting shitty posts
Her: This fucker is clueless that I've been fucking his best friend!
Him: Why does her pussy tastes like Steve's dick?
Her: He's probably thinking about other women...
Him: If I name my son Mathematics then I will be the Father of Mathematics...
Nurse: Sir, you've been in a coma for 2 weeks
Me: Oh boy I can't wait to watch cycling
No it doesn't affect my baby
"Jews killed Jesus"
12 years later
Pirate: you've been in a coma since 19 years before the Battle of Yavin
Kix: What the fuck is the battle of Yavin?
I heard he recognised we are 15 and stopped sexualising us
Code red: I heard this one is not a drill and fire guys are coming
Her: he's probably thinking about other girls...
Him: Do zebras have striped dick?...
This meme is from the future. You don't get it yet
I wonder if he's thinking about me
TikTok
Don't cry babe, it's just a movie..
Don't cry babe, the accident wasn't your fault

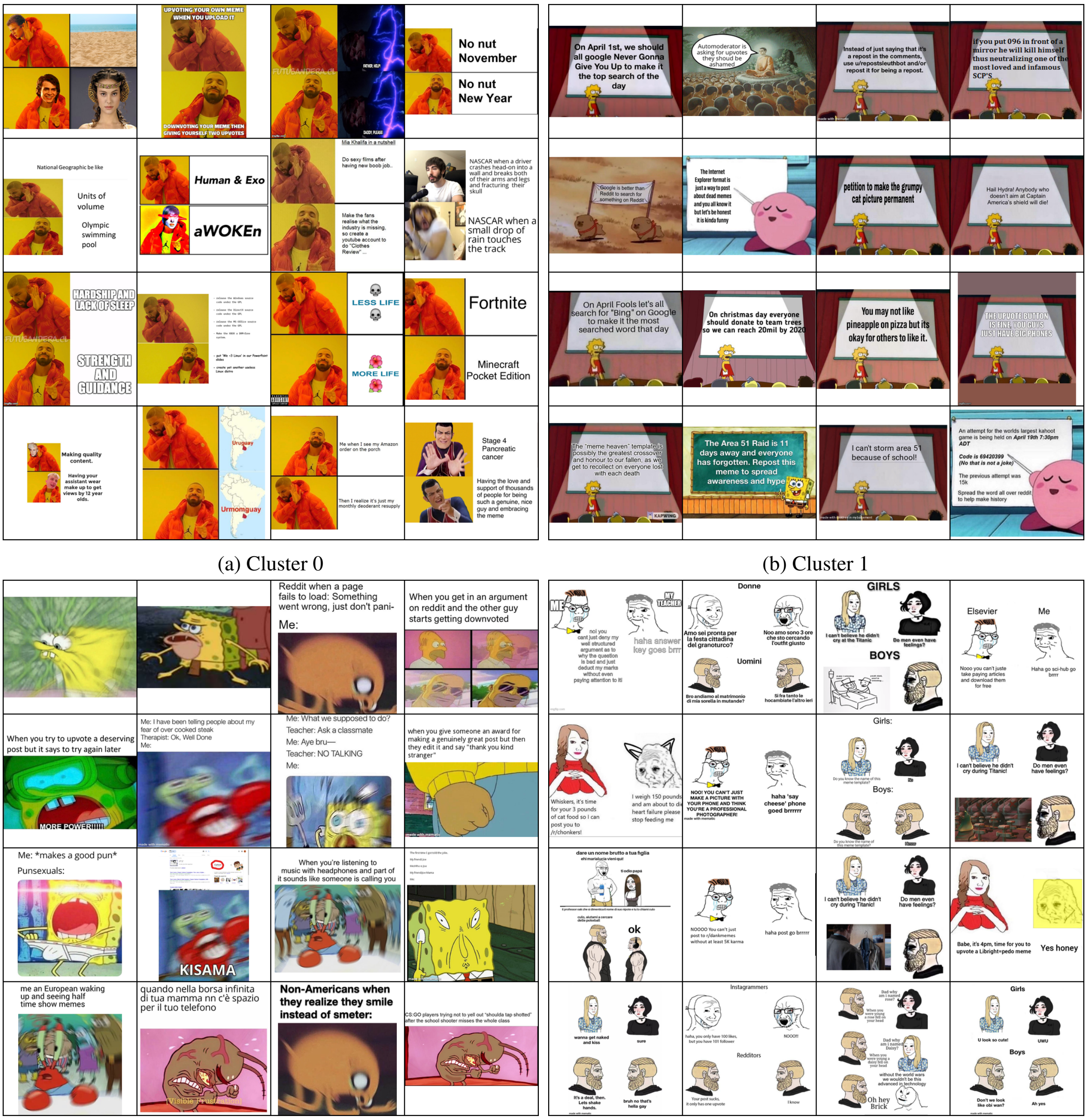
(a) Cluster 0

(b) Cluster 1

(c) Cluster 36

(d) Cluster 23

Figure 19: Samples from CLIP clusters


UPVOTING YOUR OWN MEME WHEN YOU UPLOAD IT
DOWNVOTING YOUR MEME THEN GIVING YOURSELF TWO UPVOTES
FATHER, HELP!
DADDY, PLEASE!
No nut November
No nut New Year
National Geographic be like
Units of volume
Olympic swimming pool
Human & Exo
aWOKEn
Mia Khalifa in a nutshell
Do sexy films after having new boob job..
Make the fans realise what the industry is missing, so create a youtube account to do "Clothes Review" ...
NASCAR when a driver crashes head-on into a wall and breaks both of their arms and legs and fracturing their skull
NASCAR when a small drop of rain touches the track
HARDSHIP AND LACK OF SLEEP
STRENGTH AND GUIDANCE
- release the Windows source code under the GPL
- release the DirectX source code under the GPL
- release the MS Office source code under the GPL
- Make the XBOX a DRM-free system.
- put 'We <3 Linux' in our PowerPoint slides
- create yet another useless Linux distro
LESS LIFE
MORE LIFE
Fortnite
Minecraft Pocket Edition
Making quality content.
Having your assistant wear make up to get views by 12 year olds.
Uruguay
Urmomguay
Me when I see my Amazon order on the porch
Then I realize it's just my monthly deoderant resupply
Stage 4 Pancreatic cancer
Having the love and support of thousands of people for being such a genuine, nice guy and embracing the meme
On April 1st, we should all google Never Gonna Give You Up to make it the top search of the day
Automoderator is asking for upvotes they shoud be ashamed
Instead of just saying that it's a repost in the comments, use u/repostsleuthbot and/or repost it for being a repost.
if you put 096 in front of a mirror he will kill himself thus neutralizing one of the most loved and infamous SCP'S
Google is better than Reddit to search for something on Reddit
The Internet Explorer format is just a way to post about dead memes and you all know it but let's be honest it is kinda funny
petition to make the grumpy cat picture permanent
Hail Hydra! Anybody who doesn't aim at Captain America's shield will die!
On April Fools let's all search for "Bing" on Google to make it the most searched word that day
On christmas day everyone should donate to team trees so we can reach 20mil by 2020
You may not like pineapple on pizza but its okay for others to like it.
THE UPVOTE BUTTON IS FINE, YOU GUYS JUST HAVE BIG PHONES
The "meme heaven" template is possibly the greatest crossover and honour to our fallen, as we get to recollect on everyone lost with each death
KAPWING
The Area 51 Raid is 11 days away and everyone has forgotten. Repost this meme to spread awareness and hype
I can't storm area 51 because of school!
An attempt for the worlds largest kahoot game is being held on April 19th 7:30pm ADT
Code is 69420399 (No that is not a joke)
The previous attempt was 15k
Spread the word all over reddit to help make history
Reddit when a page fails to load: Something went wrong, just don't pani-
Me:
When you get in an argument on reddit and the other guy starts getting downvoted
When you try to upvote a deserving post but it says to try again later
MORE POWER!!!!!
Me: I have been telling people about my fear of over cooked steak
Therapist: Ok, Well Done
Me:
Me: What we supposed to do?
Teacher: Ask a classmate
Me: Aye bru—
Teacher: NO TALKING
Me:
when you give someone an award for making a genuinely great post but then they edit it and say "thank you kind stranger"
Me: *makes a good pun*
Punsexuals:
KISAMA
When you're listening to music with headphones and part of it sounds like someone is calling you
me an European waking up and seeing half time show memes
quando nella borsa infinita di tua mamma nn c'è spazio per il tuo telefono
[Visible Frustration]
Non-Americans when they realize they smile instead of smeter:
CS:GO players trying not to yell out "shoulda tap shotted" after the school shooter misses the whole class
ME
MY TEACHER
no! you cant just deny my well structured argument as to why the question is bad and just deduct my marks without even paying attention to it!
haha answer key goes brrr
Donne
Amo sei pronta per la festa cittadina del granoturco?
Noo amo sono 3 ore che sto cercando l'outfit giusto
Uomini
Bro andiamo al matrimonio di mia sorella in mutande?
Si fra tanto le hocambiate l'altro ieri
GIRLS
I can't believe he didn't cry at the Titanic
Do men even have feelings?
BOYS
Elsevier
Me
Nooo you can't juste take paying articles and download them for free
Haha go sci-hub go brrrr
Whiskers, it's time for your 3 pounds of cat food so I can post you to /r/chonkers!
I weigh 150 pounds and am about to die heart failure please stop feeding me
NOO! YOU CAN'T JUST MAKE A PICTURE WITH YOUR PHONE AND THINK YOU'RE A PROFESSIONAL PHOTOGRAPHER!
haha 'say cheese' phone goed brrrrrr
Girls:
Do you know the name of this meme template?
No
Boys:
Do you know the name of this meme template?
I know
I can't believe he didn't cry during Titanic!
Do men even have feelings?
dare un nome brutto a tua figlia
ehi marialucia vieni qui!
ti odio papà
il professor oak che si dimentica il nome di suo nipote e tu lo chiami culo
culo, aiutami a cercare delle pokeball
ok
NOOOO You can't just post to r/dankmemes without at least 5K karma
haha post go brrrrr
Babe, it's 4pm, time for you to upvote a Libright=pedo meme
Yes honey
wanna get naked and kiss
sure
It's a deal, then. Lets shake hands.
bruh no that's hella gay
Instagrammers
haha, you only have 100 likes, but you have 101 follower
NOOO!!!
Redditors
Your post sucks, it only has one upvote
I know
Dad why am i named rose?
When you were young a rose fell on your head
Dad why am i named Daisy?
When you were young a daisy fell on your head
without the world wars we wouldn't be this advanced in technology
Oh hey Brick
Girls
U look so cute!
UWU
Boys
Don't we look like obi wan?
Ah yes

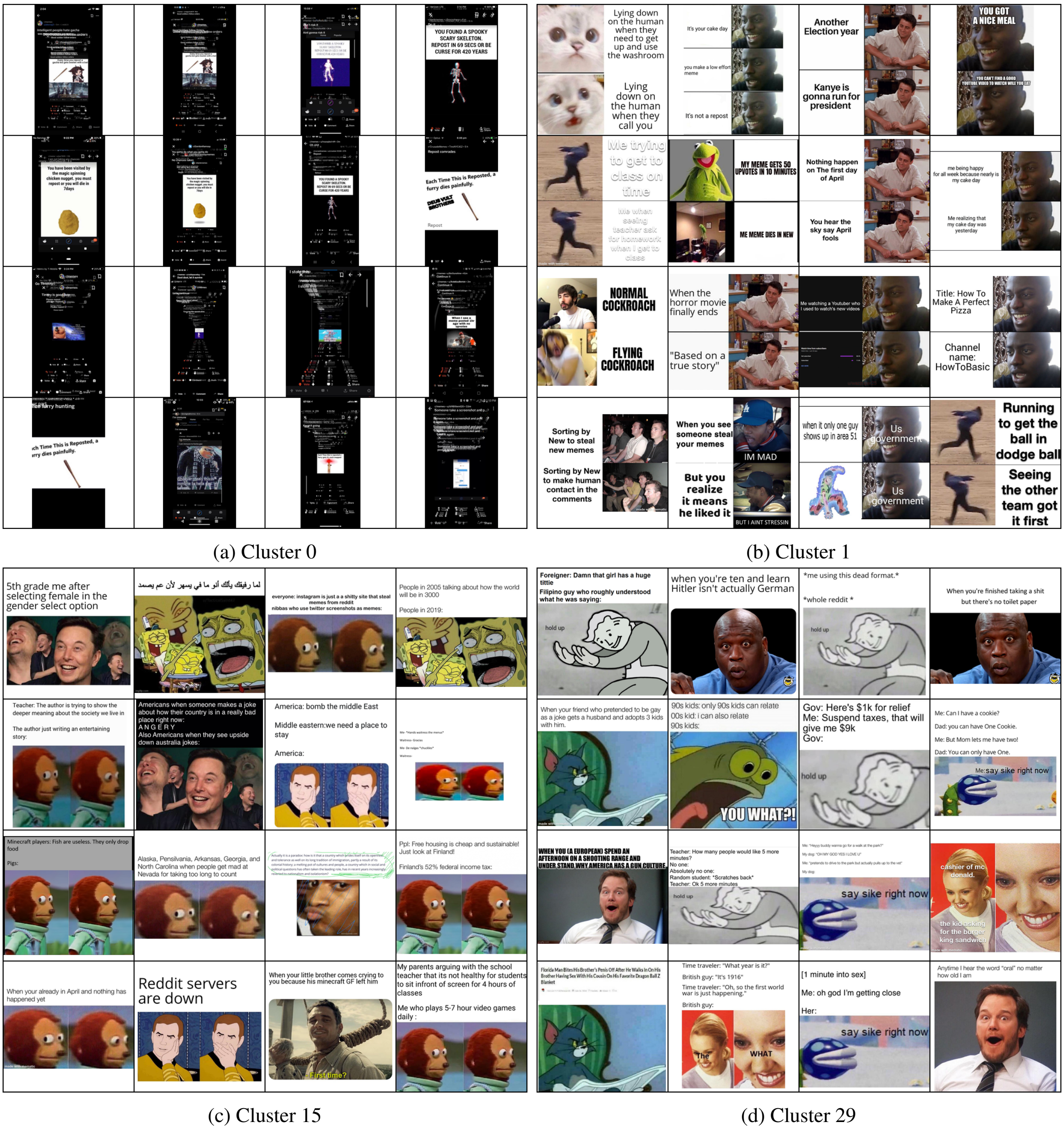
(a) Cluster 0

(b) Cluster 1

(c) 32

(d) Cluster 22

Figure 20: Samples from CLIP-diff clusters


YOU FOUND A SPOOKY SCARY SKELETON. REPOST IN 69 SECS OR BE CURSE FOR 420 YEARS
You have been visited by the magic spinning chicken nugget. you must repost or you will die in 7days
Each Time This is Reposted, a furry dies painfully.
Lying down on the human when they need to get up and use the washroom
Lying down on the human when they call you
It's your cake day
you make a low effort meme
It's not a repost
Another Election year
Kanye is gonna run for president
YOU GOT A NICE MEAL
YOU CAN'T FIND A GOOD YOUTUBE VIDEO TO WATCH WHILE YOU EAT
Me trying to get to class on time
Me when seeing teacher ask for homework when I get to class
MY MEME GETS 50 UPVOTES IN 10 MINUTES
ME MEME DIES IN NEW
Nothing happen on The first day of April
You hear the sky say April fools
me being happy for all week because nearly is my cake day
Me realizing that my cake day was yesterday
NORMAL COCKROACH
FLYING COCKROACH
When the horror movie finally ends
"Based on a true story"
Title: How To Make A Perfect Pizza
Channel name: HowToBasic
Sorting by New to steal new memes
Sorting by New to make human contact in the comments
When you see someone steal your memes
IM MAD
But you realize it means he liked it
BUT I AINT STRESSIN
when it only one guy shows up in area 51
Us government
Running to get the ball in dodge ball
Seeing the other team got it first
5th grade me after selecting female in the gender select option
لما رفيقك يألك أنو ما في يسهر لأن عم يصمد
everyone: instagram is just a shitty site that steal memes from reddit
People in 2005 talking about how the world will be in 3000
People in 2019:
Teacher: The author is trying to show the deeper meaning about the society we live in
The author just writing an entertaining story:
Americans when someone makes a joke about how their country is in a really bad place right now:
ANGERY
Also Americans when they see upside down australia jokes:
America: bomb the middle East
Middle eastern:we need a place to stay
America:
Minecraft players: Fish are useless. They only drop food
Pigs:
Alaska, Pensilvania, Arkansas, Georgia, and North Carolina when people get mad at Nevada for taking too long to count
Ppl: Free housing is cheap and sustainable! Just look at Finland!
Finland's 52% federal income tax:
When your already in April and nothing has happened yet
Reddit servers are down
When your little brother comes crying to you because his minecraft GF left him
First time?
My parents arguing with the school teacher that its not healthy for students to sit infront of screen for 4 hours of classes
Me who plays 5-7 hour video games daily :
Foreigner: Damn that girl has a huge tittie
Filipino guy who roughly understood what he was saying:
hold up
when you're ten and learn Hitler isn't actually German
me using this dead format.
*whole reddit *
When you're finished taking a shit but there's no toilet paper
When your friend who pretended to be gay as a joke gets a husband and adopts 3 kids with him.
90s kids: only 90s kids can relate
00s kid: i can also relate
90s kids:
YOU WHAT?!
Gov: Here's $1k for relief
Me: Suspend taxes, that will give me $9k
Gov:
Me: Can I have a cookie?
Dad: you can have One Cookie.
Me: But Mom lets me have two!
Dad: You can only have One.
Me: say sike right now
WHEN YOU (A EUROPEAN) SPEND AN AFTERNOON ON A SHOOTING RANGE AND UNDER STAND WHY AMERICA HAS A GUN CULTURE.
Teacher: How many people would like 5 more minutes?
No one:
Absolutely no one:
Random student: *Scratches back*
Teacher: Ok 5 more minutes
say sike right now
cashier of mc donald.
the kid asking for the burger king sandwich
Florida Man Bites His Brother's Penis Off After He Walks In On His Brother Having Sex With His Cousin On His Favorite Dragon Ball Z Blanket
Time traveler: "What year is it?"
British guy: "It's 1916"
Time traveler: "Oh, so the first world war is just happening."
British guy:
The
WHAT
[1 minute into sex]
Me: oh god I'm getting close
Her:
say sike right now
Anytime I hear the word "oral" no matter how old I am

(a) Cluster 0

(b) Cluster 1

(c) Cluster 15

(d) Cluster 29

Figure 21: Samples from CLIP-diff + RoBERTa clusters