What makes us laugh? Investigations into Automatic Humor Classification

Taradheesh Bali *
International Institute of
Information Technology
Hyderabad, India
taradheesh.bali
@research.iiit.ac.in

Vikram Ahuja *
International Institute of
Information Technology
Hyderabad, India
vikram.ahuja
@research.iiit.ac.in

Late Prof. Navjyoti Singh
International Institute of
Information Technology
Hyderabad, India
navjyoti@iiit.ac.in

Abstract

Most scholarly works in the field of computational detection of humour derive their inspiration from the incongruity theory. Incongruity is an indispensable facet in drawing a line between humorous and non-humorous occurrences but is immensely inadequate in shedding light on what actually made the particular occurrence a funny one. Classical theories like Script-based Semantic Theory of Humour and General Verbal Theory of Humour try and achieve this feat to an adequate extent. In this paper we adhere to a more holistic approach towards classification of humour based on these classical theories with a few improvements and revisions. Through experiments based on our linear approach and performed on large data-sets of jokes, we are able to demonstrate the adaptability and show componentizability of our model, and that a host of classification techniques can be used to overcome the challenging problem of distinguishing between various categories and sub-categories of jokes.

1 Introduction

Humor is the tendency of particular cognitive experiences to provoke laughter and provide amusement. Humor is an essential element of all verbal communication. Natural language systems should be able to handle humor as it will improve user-friendliness and human-computer interaction. Humour has been studied for a number of years in computational linguistics in terms of both humour generation (Ritchie and Masthoff, 2011), (Stock and Strapparava, 2006) and detection, but no such work has been done to create a classification of humor. Humor Detection has been approached as a classification problem by (Mihalcea and Strapparava, 2005). Classification of humour is a very dif-

ficult task because even theoretically there is not much consensus among theorists regarding what exactly humour is? Even if there were a specific theory as to what are the categories of humor, the sense of humour varies from person to person and therefore giving its types is even more difficult. Consensus is yet to be achieved regarding the categorization of humour (Attardo et al., 1994). To achieve this difficult feat of classification we try to answer the most basic question of Why do we laugh on a joke?. What factors motivate us. This is the most novel thing that only we are trying to achieve as of now. First of all the possible types of humor can be virtually infinite. Some researchers reduce humor to just one, or a few types, for example, incongruity (Ruch and Carrell, 1998). Since there are infinite possible types, there is a continued lack of any generally accepted taxonomy of humor, thus it may be classified according to different purposes. These classifications may often overlap. For instance the joke: A clean desk is a sign of a cluttered desk drawer can be labeled as a sarcastic joke as well as a wordplay joke/pun(antonyms).

We are trying to formulate the problem of determining different types of humor as a traditional classification task by feeding positive and negative datasets to a classifier. The data-set consists of one liners jokes of different types collected from many jokes websites, multiple subreddits and multiple twitter handles.

In short, our contributions can be summarized as follows:

- We present a theoretical framework which also provides the base for the task of computational classification of a vast array of types of jokes into categories and sub-categories

- We present a comparative study of a wide range of topic detection methods on large

---

* Both authors have contributed equally towards the paper (names in lexicographic sequence).

1

data sets of one-liner jokes.

- We analyze jokes based on the theme that they expresses and the emotion that they evoke.

The remainder of the paper is structured as follows. Section 2 provides an overview of related work and their shortcomings. Section 3 presents the framework proposed. Section 4 presents the dataset along with some pre-processing steps. Section 5 presents the various experiments conducted on the data set. Section 6 discusses the results, while Section 7 concludes the paper.

2 Related Work

Research in humour is a field of interest pertaining not only to linguistics and literature but neuroscience and evolutionary psychology as well. Research in humor has been done to understand the psychological and physiological effects, both positive and negative, on a person or groups of people. Research in humor has revealed many different theories of humor and many different kinds of humor including their functions and effects personally, in relationships, and in society.

Historically, humour has been synonymous with laughter but major empirical findings suggest that laughter and humour do not always have a one-to-one association. For example, Non-Duchenne laughter (Gervais and Wilson, 2005). At the same point of time it is also well documented that even though humour might not have a direct correlation with laughter it certainly has an influence in evoking certain emotions as a reaction to something that is considered humorous (Samson and Gross, 2012). Through the ages there have been many theories of humour which attempt to explain what humor is, what social functions it serves, and what would be considered humorous. Though among the three main rival theories of humour, incongruity theory is the more widely accepted as compared to relief[1] and superiority[2] theories, it is necessary but not sufficient in containing the scope of what constitutes humour.

---

[1]Relief theory maintains that laughter is a homeostatic mechanism by which psychological tension is reduced.(2018)

[2]The general idea behind Superiority Theory is that a person laughs about either misfortunes of others (so called schadenfreude) as laughter expresses feelings of superiority over them or over a former state of ourselves.(2016)

Script Semantic Theory of Humour (SSTH): In his book Raskin (Raskin, 2012) divulges the concept of semantic scripts. Each concept expressed by a word which is internalized by the native speaker of a language, is related to a semantic script via some cognitive architecture to all the surrounding pieces of information. Thereafter, he posits that in order to produce the humor of a verbal joke, the following 2 conditions must be met

- "The text is compatible, fully or in part, with two different (semantic) scripts

- The two scripts with which the text is compatible are opposite. The two scripts with which the text is compatible are said to overlap fully or in part on this text."

Humor is evoked when a trigger at the end of the joke, the punch line, causes the audience to abruptly shift its understanding from the primary (or more obvious) script to the secondary, opposing script.

General Verbal Theory of Humour (GVTH): The key idea behind GVTH are the 6 levels of independent Knowledge Resources (KRs) defined by (Attardo and Raskin, 1991). These KRs could be used to model individual jokes and act as the distinguishing factors in order to determine the similarity or differences between types of jokes. The KRs are ranked below in the order of their ability to '*determine*'/restrict the options available for the instantiation of the parameters below them:

1. Script Opposition (SO)

2. Logical Mechanism (LM)

3. Situation (SI)

4. Target (TA)

5. Narrative Strategy (NS)

6. Language (LA)

Owing to the use of Knowledge Resources GVTH has a much higher coverage as a theory of humour as compared to SSTH, but there still are a few aspects where GVTH comes up short. In prior sections we have established that humour has a direct correlation with the emotions that it evokes. In a similar manner emotions also act as a trigger to a humorous event. In such said events because the reason for inception of the humorous content lies

with the post-facto realization/resolution of the incongruity caused by the emotion rather than the event itself applying script opposition is out of line. For example, fear, a negative emotion that can stem as a result of some incongruity in the expected behaviour of our surroundings. Our primary emotion to such a situation is fear. Even so, the result of this incongruity caused in our emotional state, which incipiently was caused by the incongruity in our physical surroundings, can lead to humour. It must be noted that the trigger here is neither the situation nor any LM or script opposition, but the emotional incongruity.

Correspondingly, humour can also prompt itself in form of meta-humour just as emotions do. For example, one way to appreciate a bad joke can be the poorness of the joke. Another major point of contention in GVTH is Logical Mechanism. Here, logical does not stand for deductive logic or strict formal logicality but rather should be understood in some looser quotidian sense rational thinking and acting or even ontological possibility.

In his paper (Krikmann, 2006) correctly points out that in SSTH and GVTH both, Raskins concept of script is merely a loose and coarse approximation, borrowed from cognitive psychology which attempts to explain what actually happens in human consciousness. Such scripts encapsulate not only direct word meanings, but also semantic information presupposed by linguistic units as well as the encyclopaedic knowledge associated to them. Even so, in order to explain certain instances where direct or indirect script opposition is missing we need to inject an inference mechanism and a script elaborator to the current cognitive model, which would work off of the pre-existing script and ones that are newly formed through the inference mechanism. These two features become indispensable as, it is not always the case that opposing scripts are readily available to us.

3 Proposed Framework

Having Script Opposition as the only derivative bedrock behind the start of a humorous event proves deleterious in SSTH and GVTHs ability to be able to adapt to different kinds of incongruities. Further, due to the inability of GVTH to accommodate emotions at any level, uncertainty surrounding Logical Mechanism with its really vague identity, and the order of the Knowledge resources instigate us to diverge from SSTH and

GVTH as the foundation for our computational setup. Rather, in order to address such shortcomings we have kept the structure of our theory to be much more consequence driven.

Having an approach solely derived from the existing types of humour, would be subject to changes and alterations with the addition of every new type of humor and will add the limitation of the model being either too rigid, which might lead to overfitting while performing computational analysis or can lead to a model which becomes unstable as it is unable to sustain new types after more and more changes. In preference to this we proceed with caution keeping in mind the scope of this problem, drawing from the successes of the previous theories such as SSTH and GVTH with a more holistic approach in mind.

From the outset, Attardo and Raskin (Script theory revis(it)ed: joke similarity and joke representation model) had their features focused towards recognizing the distinguishing parameters of various degrees of similarity among jokes. In a similar manner we recognize three major marked characteristics which are reflected across all types of jokes, viz.

1. **Mode (*Modus Operandi*) :** Each joke whether verbal, textual or graphic has a way in which it is put across to the respective audience. This *mode* of delivery of a joke can be (but not always) decided upon by the performer of the humorous act. The mode can be a matter of conscious choice or the spontaneous culmination of a dialogue. Different situations might warrant for different modes of delivery leading to varied effects after the humour behind the joke is resolved. For example, the delivery of joke can be sarcastic, where the speaker might want to retort to someone in a conversation or it can be deadpan, where the triviality of speakers reaction becomes the source of humour. As compared to SSTH and GVTH which investigate the reason behind the incongruity (incongruity being the single source of humour) in the scripts or situations in such scenarios, we embrace incongruity as one of the many mechanisms that can be possible and keep the scope open for all categories which encompass far greater types of humour including and not being limited to juxtaposition of opposing scripts.Thus, the tools that are at the

disposal to bring about variations in the mode become more than mere language based artifacts like puns, alliterations etc. The mode can be based on the phonetics of the words such as in a limerick.

Two unique sub-categories that can be addressed here which would otherwise cause problems in SSTH and GVTH, due to their structure of logical mechanism are Anti-Humour and Non-Sequitur. Both are unconventional forms of humour and posit a stringent challenge to such theories. Non-Sequitur is difficult to accommodate even for GVTH due to its reliance on Logical Mechanisms. While all the jokes which follow any sort of logical structure could have been classified according to GVTH due to LM, Non-Sequitur does not follow any logical structure whatsoever. The entire point of a non-sequitur is that it is *absurd* in its reason and it also makes *no sense* according to semantics or meaning. The case with anti-humour could not be more different as it is not a play on the logical structure of the normal conversation but on that of the joke. Hence, as we have also mentioned in the criticisms section, there does not exist a mechanism in the previous theories to deal with such second order humour and meta-jokes.

2. **Theme :** Each joke through the use of its language and the subject matter conveys a feeling or an emotion along with it. As we have discussed at lengths in the previous sections emotion plays a very important role in a humorous event. It can by itself spur a new thread for a joke as well as act as the conclusive feeling that we get along with the humorous effect. For example, the feeling of disgust on hearing a joke about a gross situation or thing. Hence, the function that the *'theme'* of a joke can serve is, as a pointer towards the overall affect the joke has during its delivery and after its resolution. In this way we are able to tackle the aspects of a humorous event which are content and language dependent.

3. **Topic :** Most jokes have some central element, which can be regarded as the butt of the joke. This element is the key concept around which the joke revolves. It can be based on stereotypes, such as in blonde jokes or can be based off of a situation such as *'__walks into a bar'*. As can be observed in the latter case it is mostly but not always the case that the central element be single object or a person. The *'__walks into a bar'* might further lead to a topic or a situation which ends up with the punchline being on the *'dumb blonde'* stereotype. Hence, a single joke can therefore, without such restrictions on its definition can have multiple topics at the same time. Also by not restricting ourselves to only stereotypes about things, situations and beings we can also play with cases where the topic is the stereotype of a particular type of joke itself, leading to humour about stereotypes of humour. For example, a joke about a bad knock knock joke.

On inspection of the aforementioned categories we can clearly observe that unlike GVTH giving a hierarchical structure to these metrics is unsustainable. This works in our favour as we get rid of establishing problematic dependencies like ontological superiority for each category. Instead, we provide a flatter approach where a joke can be bred out of various combinations from each category and belong to multiple sub-categories at the same time.

The culmination of our work towards creating computationally detectable entities leads us to recognizing a sub-set in each of the categories that we have defined above. In the coming sections we venture towards testing our theoretical framework in real-life scenarios extracted through various social-media. Table 1 provides a catalogue of the sub-categories that we detect in each category.

4 Dataset

- **Topic Detection :** For the task of topic detection in Jokes we mined many jokes websites and collected their tags and considered those our topics. We have restricted our Jokes to the following categories: Animal, Blonde, Fat, Food, Profession, Kids, Marriage, Money, Nationality, Sports, News/politics, Police/military , Technology, Height, Men/Women, Celebrities/Pop Culture, Travel, Doctor, Lawyer, God/religion, Pick up lines, school, party, Walks into a bar, Yo-mama. Most of the

Table 1: Computationally Detectable Characteristics

Categories	Sub-Categories
Mode	Sarcastic
	Exaggeration/Hyperbole
	Phonetics Assisted
	Semantic Opposites
	Secondary Meaning
Theme	Dark Joke
	Gross Joke
	Adult/Sexual Joke
	Insults
Topics	Animal, Blonde, Fat, Food, Profession, Kids, Marriage, Money, Nationality, Sports, News/politics, Police/military, Technology,Height,Men/ Women, Celebrities, Pop Culture, Travel, Doctor, Lawyer, God/religion, Pick up lines, School, Parties, Walks into a bar, Yo-mama

jokes websites had the above topics as common topics. We mined nearly 40,000 one liners jokes belonging to these 25 categories for the use of Topic Detection. Since they were collected automatically, it is possible to have noise in the dataset.

- **Sarcastic Jokes :** For the task of Sarcasm Detection we mined Sarcastic jokes(positive) from reddit and other jokes websites which had sarcasm tags in it. For negative data we considered data under tags other than Sarcasm and manually verified the jokes. We created a dataset of 5000 jokes with 2500 belonging to the the positive set and and equal amount of negative instances and manually verified them

- **NSFW Jokes :** These are the types of jokes which are most famous on the online media.These types of jokes are mainly associated with heavy nudity, sexual content, heavy profanity and adult slangs. We collected multiple one liner jokes from subreddit /r/dirtyjokes and took jokes from various jokes websites with tags NSFW, dirty, adult and sexual. We created a dataset of 5000 jokes with 2500 belonging to the positive instances and equal number of negative

instances verified manually.

- **Insults :** These kinds of jokes mainly consists mainly of offensive insults directed someone else or towards the speaker itself. (Mendrinos, 2004) Typical targets for insult include individuals in the show's audience, or the subject of a roast. The speaker of an insult joke often maintains a competitive relationship with the listener. We collected multiple jokes from the subreddit /r/roastme and after manual verification we had 2000 jokes of positive instances and for negative instances we manually created a dataset of 2000 one liner jokes.

- **Gross :** A joke having to do with disgusting acts or other things people might find grotesque. We extracted 500 jokes various jokes website which had a "gross" category/tag in it. We selected equal number of non gross jokes from the above datatset. After manual verification we had a total of 1000 jokes in this category, 500 belonging to both positive and negative sets.

- **Dark Humor :** It's a form of humor involving a twist or joke making the joke seen as offensive, harsh, horrid, yet the joke is still funny. We collected multiple jokes from subreddit /r/darkjokes as well as as many jokes websites containing the tag Dark Humor. After removing duplicates we had a dataset of 3500 dark jokes. For negative samples we randomly selected 3500 jokes from the jokes websites which did not contain Dark Humor in their tags and manually verified them.

4.1 Data Preprocessing

The content of user created jokes on Twitter and Reddit can be noisy. They could contain elements like @RT, links, dates, ID's, name of users, HTML Tags and hashtags to name a few. To reduce the amount of noise before the classification task , the data is subjected to the following pre processing tasks.

- **Tokenization :** In a raw post, terms can be combined with any sort of punctuation and hyphenation and can contain abbreviations, typos, or conventional word variations. We use the NLTK tokenizer package to extract tokens from the joke by removing stop words,

Table 2: My caption

Sarcastic Joke	I asked my North Korean friend how it was there? He said he couldn't complain.
Exaggeration/Hyperbole	You know what, we need a huge spoon to take care of this. Guy who invented shovels
Phonetics Assisted	Coca Cola went to town, Diet Pepsi shot him down. Dr. Pepper fixed him up, Now we are drinking 7up.
Semantic Opposites	Humpty Dumpty had a great fall - and a pretty good spring and summer , too .
Secondary Meaning	Those who like the sport fishing can really get hooked
Dark Jokes	Why don't black people go on cruises? They are not falling for that one.
Gross Joke	Q: Why did the skeleton burp? A: It didn't have the guts to fart.
Adult/Sexual Joke	Does time fly when you're having sex or was it really just one minute?
Insults	You are proof that evolution can go in reverse.

punctuation, extra white space and hashtags and removing mentions, i.e., IDs or names of other users included in the joke and converting to lowercase.

- **Stemming :** Stemming is the process of reducing words to their root (or stem), so that related words map to the same stem or root form. This process naturally reduces the number of words associated with each document, thus simplifying the feature space. We used the NLTK Porter stemmer in our experiments.

5 Experiment

We performed various experiments on our dataset. For the evaluation we randomly divided our dataset into 90% training and 10% testing. All the experiments were conducted 10 fold and the final performance is reported by averaging the result.

- **Topic Detection :** There are a wide variety of methods and variables and they greatly affect the quality of results. We compare results from three topic detection methods on our dataset to detect topics of these jokes. We use LDA, Naive Bayes and SVM along with lexical and Pragmatic features and compared their results. We also augment the used approaches by boosting proper nouns and then, recalculating the experiment results on the same dataset. The boosting techniques that we have used are duplication proper nouns.

This boosting technique was chosen keeping in mind the need to give priority to the tweet semantic.

- **Sarcastic :** We treat sarcasm detection as a classification problem. After pre-processing the data we extracted n-grams more precisely, unigrams and bigrams from the dataset and then were added to the feature dictionary. Along with this we used brown clustering which helped us to put similar kinds of words in same cluster. Along with these features we also took sentiment values of the different parts of joke(here 3) as a feature because there is usually a great difference in sentiment scores in different part of a sarcastic joke or a tweet. Using these lexical as well as pragmatic features as in (González-Ibánez et al., 2011) we train a logistic regression and a SVM to distinguish between sarcastic jokes from non sarcastic jokes.

- **Exaggeration :** These are types of statements that represents something as better or worse than it really is. They can create a comical effect when used appropriately. For eg: In the joke "You grandma is as old as mountains", the intensity of the statement is increased by using phrase like "as old as". We detect such intense phrases in jokes to categorize under this category by getting sentiment score of every token. Individual sentiment score of every token in phase as well the

combined sentiment score will be in positive range to generate an exaggeration effect.

- **Antonyms/Semantic Opposites :** An antonym is one of a pair of words with opposite meanings. Each word in the pair is the antithesis of the other. We use the antonym relation in WORDNET among noun, adjectives and verbs and used approach similar to (Mihalcea and Strapparava, 2005)

- **Phonetic Features :** Rhyming words also create a joke. For instance the joke - Coca Cola went to town, Diet Pepsi shot him down. Dr. Pepper fixed him up, Now we are drinking 7up creates a comical effect due the fact that town and down , up and 7up are rhyming words. Similar rhetorical devices play an important role in wordplay jokes, and are often used in. We used CMU Pronunciation Dictionary to detect rhyming words

- **Secondary Meaning :** These are the types of the jokes where we find that there is semantic relation among words in a jokes and that relation could be in a form located in, part of, type of, related to, has, etc. For eg: In the joke "Those who like the sport fishing can really get hooked" comical effect is created due to the relation between "hook" and "fishing". In order to detect these relations in a joke we are using Concept Net (Speer et al., 2017). It is a multilingual knowledge base, representing words and phrases that people use and the common-sense relationships between them. So, using concept net we are able to give a used in relationship between hook and fishing. We are going upto three levels to detect secondary relationship between different terms in a joke.

- **Dark Humor :** It is a comic style that makes light of subject matter that is generally considered taboo, particularly subjects that are normally considered serious or painful to discuss such as death. Some comedians use it as a tool for exploring vulgar issues, thus provoking discomfort and serious thought as well as amusement in their audience. Popular themes of the genre include violence, discrimination, disease, religion and barbarism. Treating it as a classification problem, we extracted unigrams from the dataset. We also

Table 3: Topic Detection

Classifier	Accuracy
LDA	59%
Naive Bayes	63%
SVM	72%
SVM + Proper Noun Boosting	**76%**

Table 4: Sarcastic Jokes

Results	
Features	**Acc.**
Logistic Regression (LR)	68%
LR + (1,2) grams	71%
LR + (1,2) grams + Brown Clustering	71.5%
LR + (1,2) grams + Brown Clustering + Sentiment Scores	75.2%
SVM + Sentiment Scores + N garms	77%

extracted sentiment scores of the sentence because of the hypothesis that dark humor tends to have a very negative sentiment score throughout the joke. We then compared the accuracies of classification techniques such as SVM and Logistic Regression.

- **Adult Slangs/Sexual Jokes :** These types of jokes are most famous on the internet.After pre-processing we extracted unigrams and bigrams. To detect these types of jokes we used a slang dictionary called Slang SD (Wu et al., 2016). It contains over 90,000 slang words/phrases along with their sentiment scores. We used these features and compared accuracies of classification methods such as SVM and Logistic Regression.

- **Gross :** Treating the problem of detecting Gross Jokes as a classification problem, unigrams are extracted after pre-processing. We kept a list of top 100 gross words according to their tf-idf score. This feature indicated the presence of gross words. Along with this we also maintain sentiment scores because of the

Table 5: Dark Jokes

Results	
Features	Accuracy
Logistic Regression (LR)	59%
LR + Sentiment Scores	63%
SVM + Sentiment Scores	64%

7

Table 6: Adult Slang/Sexual Jokes

Results	
Features	Accuracy
Logistic Regression (LR)	71%
LR + (1,2)grams + Slang SD	85%
SVM + (1,2)grams + Slang SD	88%

Table 7: Gross Jokes

Results	
Features	Accuracy
Logistic Regression (LR)	56%
LR + Common Gross Words + Sentiment	65%
SVM + Common Gross Words + Sentiment	67%

hypothesis that gross jokes tends to have a negative sentiment. Using all these features we compare accuracies using SVM and Logistic Regression.

- **Insults:** After pre-processing we are extracting unigrams and bigrams from the dataset. Along with this we are creating a list of insulting words using top 100 words according to their Tfidf score. Along with this we calculated semantic scores of each of the joke and used these features in a Naive Bayes Classifier and a SVM.

6 Analysis

In Tables 3, 4, 5 ,6 , 7 and 8 we can see results of our classifiers. We see that SVM has a better accuracy in all the cases than Naive Bayes and Logistic Regression. In the case of Topic Detection, Proper noun boosting increases the accuracy furthermore. In the case of sarcasm detection, we see the sentiment scores as well as unigrams and bigrams given to a SVM gave the best possible result. In the case of detection of dark humor we see that there is significant increase in in accuracy as sentiment values are introduced. These maybe because of the fact

Table 8: Insult Jokes

Features	**Accuracy**
Naive Bayes + (1,2) grams	72%
SVM + (1,2) grams	72%
SVM + insulting words sentiment values +(1,2) grams	79%

the sentiment values in the negative instances are opposites to what it is in positive instances. This result is expected because dark jokes tend to have negative sentiment values. In case of adult slang detection we are getting a very good accuracy as soon as a slang dictionary is introduced. In detection of gross jokes, the accuracy is increased as soon as sentiment and common gross words are introduced. In short,we find that sentiment values prove to be a very important feature in detection of various sub categories. We are also able to detect intense phrases which lead to exaggeration as well as jokes in which there is some kind of a semantic relation among different terms. Using these subcategories we have covered a lot in our ground in categorization of jokes. The results that we achieve act as binary indicators for each subcategory in our experiment, thus giving multiple tags according to topic, theme and mode to a joke, making our approach more extensive and unique as compared to our counterparts.

7 Future Work

Given the constraints of the scope of our paper as well as our research we have tried to assimilate as many sub-categories as possible to include as a part of our computational framework, but at the same point of time we also make an ambitious yet modest assumption that it is still possible to add a few more sub-categories. As our model is versatile enough to handle the addition of such subcategories seamlessly, the only impediment would the the feasibility of the effort and availability of the computational tools for them to be integrated. With the addition of more and diverse data the model can be made more robust and accurate as well. In future, the framework can also be extended to distinguish between humorous and non-humorous events, allowing us to use the complete tool on various types of data, such as, movie or television show scripts to detect the occurrences of various types of humour and hence, giving birth to a more holistic classification of said media.

References

Salvatore Attardo, Donalee Hughes Attardo, Paul Baltes, and Marnie Jo Petray. 1994. The linear organization of jokes: analysis of two thousand texts. *Humor-International Journal of Humor Research*, 7(1):27–54.

Salvatore Attardo and Victor Raskin. 1991. Script the-

ory revis (it) ed: Joke similarity and joke representation model. *Humor-International Journal of Humor Research*, 4(3-4):293–348.

Wikipedia contributors. 2018. Theories of humor — wikipedia, the free encyclopedia. [Online; accessed 12-March-2018].

Matthew Gervais and David Sloan Wilson. 2005. The evolution and functions of laughter and humor: A synthetic approach. *The Quarterly review of biology*, 80(4):395–430.

Roberto González-Ibánez, Smaranda Muresan, and Nina Wacholder. 2011. Identifying sarcasm in twitter: a closer look. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: Short Papers-Volume 2*, pages 581–586. Association for Computational Linguistics.

Arvo Krikmann. 2006. Contemporary linguistic theories of humour. *Folklore*, 33(2006):27–58.

James Mendrinos. 2004. *The Complete Idiot's Guide to Comedy Writing*. Penguin.

Rada Mihalcea and Carlo Strapparava. 2005. Making computers laugh: Investigations in automatic humor recognition. In *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pages 531–538. Association for Computational Linguistics.

John Morreall. 2016. Philosophy of humor. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*, winter 2016 edition. Metaphysics Research Lab, Stanford University.

Victor Raskin. 2012. *Semantic mechanisms of humor*, volume 24. Springer Science & Business Media.

Graeme Ritchie and Judith Masthoff. 2011. The standup 2 interactive riddle builder. In *Proceedings of the Second International Conference on Computational Creativity*. Universidad Aut´ onoma Metropolitana, Mexico City.

Willibald Ruch and Amy Carrell. 1998. Trait cheerfulness and the sense of humour. *Personality and Individual Differences*, 24(4):551–558.

Andrea C Samson and James J Gross. 2012. Humour as emotion regulation: The differential consequences of negative versus positive humour. *Cognition & emotion*, 26(2):375–384.

Robert Speer, Joshua Chin, and Catherine Havasi. 2017. Conceptnet 5.5: An open multilingual graph of general knowledge. In *AAAI*, pages 4444–4451.

Oliviero Stock and Carlo Strapparava. 2006. Laughing with hahacronym, a computational humor system. In *PROCEEDINGS OF THE NATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE*, volume 21, page 1675. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999.

Liang Wu, Fred Morstatter, and Huan Liu. 2016. Slangsd: Building and using a sentiment dictionary of slang words for short-text sentiment classification. *arXiv preprint arXiv:1608.05129*.