

# Humor in Word Embeddings: Cockamamie Gobbledegook for Nincompoops

Limor Gultchin

Genevieve Patterson Nancy Baym  
Adam Tauman Kalai

Nathaniel Swinger

Presenters:

Siba Smarak Panigrahi and Sohan Patnaik

Reading Session III  
Kharagpur Data Analytics Group  
IIT Kharagpur

March 14, 2021

# Important Notes

- **W**e expect you all to have certain queries regarding the presentation, certain intricate doubts maybe... **Please put them in the chat of this meeting.** We will address them all at the end!
- **T**o ensure smooth flow of the presentation, we need you all to **keep your microphones and video turned off.**
- **F**inally, a **Google Form** [\[Link\]](#) will be released for feedback but most importantly, we ask you to put up name of any AI-related paper to be presented in the upcoming sessions!

**Remember, if we can convert it into a presentation, we are indeed gonna present it!**

# Outline

- Important Notes
- Introduction
- Relevant Features in Humor Theories
- Data
- Linear Regression : Identifying Mean Humor Direction
- CLustering: Sense of Humor Embeddings
- Predicting individual differences in word humor
- Conclusion

# Introduction

- This paper aims at detecting and generating humor for AI systems.
- Word Embeddings, which are vector representations of words are shown to capture the humor in the words.
- Different humor theories have also been proposed in order to distinguish among the words, and what kind of humor is proposed by the word.
- Crowdsourcing approach was also used to generate dataset for this task.

Now let us look at each aspect in detail. You will see that an amazingly simple approach has been used to find the mean humor direction. Have patience!

**This Presentation is Rated-R**

The 216 top-rated words, sorted in order of mean humor ratings (funniest first), used in the main crowdsourcing experiment:

asshattery, clusterfuck, douchebaggery, poppycock, craptacular, cockamamie, gobbledegook, nincompoops, wanker, kerfuffle, cockle pickers, pussyfooting, tiddlywinks, higgledy piggledy, kumquats, boondoggle, doohickey, annus horribilis, codswallop, shuttlecock, bejeezus, bamboozle, whakapapa, artsy fartsy, pooper scoopers, fugly, dunderheaded, dongles, didgeridoo, dickering, bacon butties, woolly buggers, pooch punt, twaddle, dabbawalas, goober, apeshit, nut butters, hoity toity, glockenspiel, diktats, mollycoddling, pussy willows, bupkis, tighty whities, nut flush, namby pamby, bugaboos, hullabaloo, hoo hah, crapola, jerkbaits, batshit, schnitzels, sexual napalm, arseholes, buffoonery, lollygag, weenies, twat, diddling, cockapoo, boob tube, galumphing, ramrodded, schlubby, poobahs, dickheads, fufu, nutjobs, skedaddle, crack whore, dingbat, bitch slap, razzmatazz, wazoo, schmuck, cock ups, boobies, cummerbunds, stinkbait, gazumped, moobs, bushwhacked, dong, pickleball, rat ass, bootlickers, skivvies, belly putter, spelunking, faffing, spermatogenesis, butt cheeks, blue tits, monkeypox, cuckolded, wingnuts, muffed punt, ballyhoo, niggly, cocksure, oompah, trillion dong, shiitake, cockling, schlocky, portaloos, pupusas, thrust reverser, pooja, schmaltzy, wet noodle, piggeries, weaner, chokecherry, tchotchkes, titties, doodad, troglodyte, nookie, annulus, poo poo, semen samples, nuttied, foppish, muumuu, poundage, drunken yobs, yabbies, chub, butt whipping, noobs, ham fist, pee pee, woo woo, squeegee, flabbergasted, yadda yadda, dangdut, coxless pairs, twerps, tootsies, big honkin, porgies, dangly, guffawing, wussies, thingies, bunkum, wedgie, kooky, knuckleheads, nuttin, mofo, fishmonger, thwack, teats, peewee, cocking, wigwams, red wigglers, priggish, hoopla, poo, twanged, snog, pissy, poofy, newshole, dugong, goop, whacking, viagogo, chuppah, fruitcakes, caboose, cockfights, hippocampus, vindaloo, holeshoot, hoodoo, clickety clack, backhoes, loofah, skink, party poopers, civvies, quibble, whizzy, gigolo, bunged, whupping, weevil, spliffs, toonie, gobby, infarct, chuffed, gassy, crotches, chits, proggy, doncha, yodelling, snazzy, fusarium, bitty, warbled, guppies, noshes, dodgems, lard, meerkats, lambast, chawl

# Relevant Features in Humor Theories

- **Relief Theory:** The relief theory offers a physiological explanation, which argues that laughter functions as a pressure relief for the nervous system, sometimes as a means to address taboo subjects that cause such stress.
- **Incongruity Theory:** Incongruity theory explains humor as a reaction to moments of uncertainty, in which two non-related issues which seem to be unfitting are juxtaposed in one sentence or event, but are then resolved.
  - This theory could be seen as an explanation of the intuitive importance of punch lines in common jokes, where the “set up” builds an expectation that the “punch line” violates.

Based on these theories, the authors considered 6 features of word humor.

1. **Humorous sound:** Certain words such as **bugaboo** or **razzmatazz** are funny regardless of their meaning. This feature is related to the incongruity theory in that an unusual combination of sounds that normally do not go together can make a word sound funny.
2. **Juxtaposition/Unexpected incongruity:** Certain words are composed of terms which are otherwise completely unrelated. For example, there is little relation between the **hippocampus** part of the brain and the mostly unrelated words **hippo** and **campus**. This feature is motivated by incongruity theory.
3. **Sexual connotation:** some words are explicitly sexual (such as sex) or have sexual connotations (such as **thrust reverser**). The fact that some people find those words funny can be explained by the 'relief theory', which suggests humor is a venting mechanism for social taboos.
4. **Scatological connotation:** Some words have connotations related to excrement to varying degrees, such as **nincompoop** or **apeshit**. The justification of this feature is similar to the sexual connotation feature above.
5. **Insulting words:** in the context of word humor, the superiority theory suggests that insulting words may be humorous to some people.
6. **Colloquial words:** Extremely informal words such as **crapola** can be surprising and provide relief in part due to their unusually informal nature.

# Data

1. **Engelthaler-Hill Dataset:** This dataset provides mean ratings on 4997 words, each rated by approximately 35 raters on a scale of 1-5. It was further segregated by means on gender and age.
2. **Crowdsourcing Studies:**
  - a. A subset of 120k most frequent words and phrases were taken from GNEWS
    - i. Each string was shown to at least three different participants in three different sextuples.
    - ii. 80,062 strings were not selected by any participant, consistent with EH's finding that the vast majority of words are not found to be funny.
  - b. The 8k study applied the same procedure (except without a “none are humorous” option) to the 8,120 words that were chosen as the most humorous in a majority of the sextuples they were shown.
  - c. To filter down to 216 words, we selected the words with the greatest fraction of positive labels.
  - d. To avoid near duplicate words, amongst words with the same stem, the word with the greatest score was selected. For instance, among the four words **wank**, **wanker**, **wankers**, and **wanking**, only **wanker** was selected for the third round as it had the greatest fraction of positive votes.



They reduce 216 to most humorous word/phrase with the following approach:

- Show 36 sextuples to all the people to select the best humor word amongst them
- Per participant, we have 36 best humor words for each participant
- Again show these 36 word/phrases randomly in 6 sextuples
- Now, we have 6 best humor words for each participant
- Kays, show them in one sextuple and lo and behold - we have the best one for each participant

Total - 43 clicks per participant!

### **Rating association:**

- The single funniest word: 3
- The five other words: 2
- The remaining 30 words: 1
- Rest were: 0

## Feature Annotation Study:

- a) 1500 words were chosen from the original 120,000 words by including the 216 words discussed above plus random words (half from 8k study and half from 120k study).
- b) Participants were asked to annotate six features discussed earlier namely **humorous sound, juxtaposition, colloquial, insulting, sexual connotation, and scatological connotation**.
  - a) Each feature was given a binary “yes/no” annotation, and results were averaged over at least 8 different raters per feature/word pair.
  - b) In each task, each rater was given the definition of a single feature and was presented with 30 words.
  - c) A small number of workers were eliminated due to poor performance on gold-standard words.

# Linear Regression : Identifying Mean Humor Direction

1. **Data:** 4997 EH words + their humor ratings
2. **Approach:** Fitting a simple least squares linear regression for predicting the EH mean humor rating for each word (The embedding of each word is 300 dimensional and that serves as the feature vector for regression).

Embedding name	Corr.	Corr. (using CV)
GNEWS (word2vec)	0.721	$0.675 \pm 0.003$
WebSubword (fastText)	0.767	$0.729 \pm 0.002$
WebFast (fastText)	0.767	$0.730 \pm 0.002$
WebGlove (gloVe)	0.768	$0.730 \pm 0.002$

Table 1. Correlation coefficients compared across embeddings using least-squares regression to predict EH humor ratings. Correlations with 10-fold cross-validation and confidence intervals also reported.

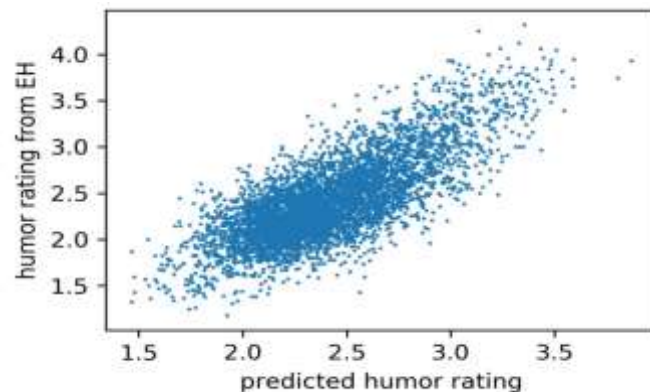


Figure 1. Humor rating predictions from a linear regression model trained on the WebSubword embedding vectors of words from the EH dataset, plotted against the EH mean humor ratings. See Table 1 for comparison with other embeddings.

# Clustering Sense of Humor Embeddings

*Table 2.* Raters were clustered based on their sense-of-humor embeddings, i.e., the average vector of the words they rated funny. For each cluster, the single words that differentiate each rating cluster are shown together with scores on six humor features and demographics. Note that the demographics differences emerged from clustering the sense-of-humor vectors—demographics were not used during clustering.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
word 1	gobbledegook	tootsies	clusterfuck	whakapapa	dickheads
word 2	kerfuffle	fufu	batshit	codswallop	twat
word 3	hullabaloo	squeegee	crapola	dabbawalas	cocking
word 4	razzmatazz	doohickey	apeshit	pooja	titties
word 5	gazumped	weenies	fugly	spermatogenesis	asshattery
word 6	boondoggle	muumuu	wanker	diktats	nuttet
word 7	galumphing	thingies	schmuck	annulus	dong
word 8	skedaddle	wigwags	arseholes	chokecherry	wanker
word 9	guffawing	weaner	dickheads	piggeries	cockling
word 10	bamboozle	peewee	douchebaggery	viagogo	pussyfooting
sound	<b>1.11</b>	1.02	0.97	1.02	0.90
scatological	0.80	0.99	<b>1.15</b>	0.89	1.14
colloquial	0.95	1.00	<b>1.14</b>	0.87	1.02
insults	0.86	0.90	<b>1.23</b>	0.84	1.12
juxtaposition	0.89	0.86	0.99	1.10	<b>1.13</b>
sexual	0.81	0.91	0.99	1.00	<b>1.25</b>
female %	70.3%*	57.5%	53.8%	52.4%	35.2%*
mean age	38.6	37.4	42.3*	37.2	34.7*

**Bolded numbers** represent the highest mean value obtained for that feature in any of the clusters.

\*statistically significant difference with  $p\text{-value} < 10^{-6}$ .

# Predicting Individual Differences in Word Humor

1. The average of embeddings of the words a person finds funny, captures their sense of humor in vector form.
2. Two random people  $p_1$  and  $p_2$  with funniest words  $w_1, w_2$  (rating of 3) were chosen respectively, with a condition that  **$p_1$  rated  $w_2$  as 0** and  **$p_2$  rated  $w_1$  as 0** (Remember the ratings assigned).
3. Given the **two sets of 35 other words (the training words)** each participant rated positively, and given words  $\{w_1; w_2\}$  (test words), the goal was to **predict which participants rated  $w_1$  or  $w_2$  funniest**.
4. **For example**, the **training sets** might be {bamboozle, bejeezus, ..., wigwams, wingnuts} and {batshit, boobies, ..., weaner, whakapapa} and **test set** to match might be the words {poppycrack, lollygag}.
5. **To test the embedding training word vectors were averaged and similarity with the test words was computed**. In particular, if  $\mathbf{r}_1$  and  $\mathbf{r}_2$  are the two training word vector averages and  $\mathbf{v}_1$  and  $\mathbf{v}_2$  are the corresponding test word vectors, then  $\mathbf{r}_1$  is matched to  $\mathbf{v}_1$  if and only if,

$$\begin{aligned}\|\vec{r}_1 - \vec{v}_1\|^2 + \|\vec{r}_2 - \vec{v}_2\|^2 &< \|\vec{r}_1 - \vec{v}_2\|^2 + \|\vec{r}_2 - \vec{v}_1\|^2 \\ 0 &< (\vec{r}_1 - \vec{r}_2) \cdot (\vec{v}_1 - \vec{v}_2)\end{aligned}$$

In the **easy version**, a pair of people was chosen who have disjoint training sets of positively rated words, indicating distinct senses of humors.

In the **hard version**, they used as training words only the five words each person rated 2.

Know-your-audience test	Success rate
Easy: disjoint sets of 35 training words.	78.1%
Normal: 35 training words.	68.2%
Hard: 5 training words.	65.0%

# Conclusion

- Multiple aspects of humor as covered in various humor theories can be captured in the word embeddings
- The meaning of humor varies from person to person
- We can find an individual's sense of humor (represented as humor vector) - can help to find their sense of humor in new, unrated words (see the tip below)
- Upon clustering humor ratings of multiple demographic groups, different humor preferences emerge across the different groups.

**Tip:** So next time you want to impress someone with your own “quirky” humor - be sure to get the humor embeddings of the person (and average them)



The 216 top-rated words, sorted in order of mean humor ratings (funniest first), used in the main crowdsourcing experiment:

asshattery, clusterfuck, douchebaggery, poppycock, craptacular, cockamamie, gobbledegook, nincompoops, wanker, kerfuffle, cockle pickers, pussyfooting, tiddlywinks, higgledy piggledy, kumquats, boondoggle, doohickey, annus horribilis, codswallop, shuttlecock, bejeezus, bamboozle, whakapapa, artsy fartsy, pooper scoopers, fugly, dunderheaded, dongles, didgeridoo, dickering, bacon butties, woolly buggers, pooch punt, twaddle, dabbawalas, goober, apeshit, nut butters, hoity toity, glockenspiel, diktats, mollycoddling, pussy willows, bupkis, tighty whities, nut flush, namby pamby, bugaboos, hullabaloo, hoo hah, crapola, jerkbaits, batshit, schnitzels, sexual napalm, arseholes, buffoonery, lollygag, weenies, twat, diddling, cockapoo, boob tube, galumphing, ramrodded, schlubby, poobahs, dickheads, fufu, nutjobs, skedaddle, crack whore, dingbat, bitch slap, razzmatazz, wazoo, schmuck, cock ups, boobies, cummerbunds, stinkbait, gazumped, moobs, bushwhacked, dong, pickleball, rat ass, bootlickers, skivvies, belly putter, spelunking, faffing, spermatogenesis, butt cheeks, blue tits, monkeypox, cuckolded, wingnuts, muffed punt, ballyhoo, niggly, cocksure, oompah, trillion dong, shiitake, cockling, schlocky, portaloos, pupusas, thrust reverser, pooja, schmaltzy, wet noodle, piggeries, weaner, chokecherry, tchotchkes, titties, doodad, troglodyte, nookie, annulus, poo poo, semen samples, nuttied, foppish, muumuu, poundage, drunken yobs, yabbies, chub, butt whipping, noobs, ham fist, pee pee, woo woo, squeegee, flabbergasted, yadda yadda, dangdut, coxless pairs, twerps, tootsies, big honkin, porgies, dangly, guffawing, wussies, thingies, bunkum, wedgie, kooky, knuckleheads, nuttin, mofo, fishmonger, thwack, teats, peewee, cocking, wigwams, red wigglers, priggish, hoopla, poo, twanged, snog, pissy, poofy, newshole, dugong, goop, whacking, viagogo, chuppah, fruitcakes, caboose, cockfights, hippocampus, vindaloo, holeshoot, hoodoo, clickety clack, backhoes, loofah, skink, party poopers, civvies, quibble, whizzy, gigolo, bunged, whupping, weevil, spliffs, toonie, gobby, infarct, chuffed, gassy, crotches, chits, proggy, doncha, yodelling, snazzy, fusarium, bitty, warbled, guppies, noshes, dodgems, lard, meerkats, lambast, chawl



Table 5. The ten words rated funniest in our study, their female/male mean rating discrepancy (if significant), and some features of these words.

	FM	sound	juxtaposition	colloquial	insulting	sexual	scatological
asshattery			✓	✓	✓	✓	✓
clusterfuck	M		✓	✓		✓	✓
douchebaggery				✓	✓		
poppycock	M	✓	✓		✓	✓	
craptacular					✓		✓
cockamamie	F				✓	✓	
gobbledegook	F	✓					
nincompoops	F	✓	✓	✓	✓		✓
wanker	M			✓	✓	✓	
kerfuffle	F	✓		✓			

Table 6. 10 most highly annotated words in each humor feature category.

Colloquial	Insulting	Juxtaposition	Scatological	Sexual	Funny sounding
dissing	fuckers	party poopers	dog poop	foreskins	lollygag
wee lad	douche	hippocampus	dung	scrotum	gobbledegook
clusterfuck	dickheads	pooch punt	pooper scoopers	nudism	ballyhoo
twat	nincompoops	port potties	poo poo	nuttid	tiddlywinks
mofo	asshat	bacon butties	urination	blue tits	higgledy piggledy
fugly	wussies	sexual napalm	apeshit	boobies	schlubby
flimflam	twat	pickleball	poo	pussyfooting	hobnobbing
woo woo	crack whore	pussyfooting	butt cheeks	crotches	hoo hah
crack whore	rat ass	boob tube	crapola	vibrators	didgeridoo
nutjobs	smartass	jobholders rose	diaper clad	masturbating	poppycock

Thank You!