

CATs are Fuzzy PETs: A Corpus and Analysis of Potentially Euphemistic Terms

Martha Gavidia, Patrick Lee, Anna Feldman, Jing Peng
{gavidiam1, leep6, feldmana, pengj}@montclair.edu

LREC 2022



**MONTCLAIR STATE
UNIVERSITY**

Euphemisms

Euphemisms are mild or indirect expressions that are used in place of more unpleasant or offensive ones.

passed away instead of **died**

armed conflict instead of **war**

dismissed instead of **fired**

John was dismissed from his job.

euphemistic

The teacher dismissed the class.

literal

Related Work

Politeness

Euphemisms are related to the language of politeness and politeness has been a central concern in pragmatic theory. (Grice, 1975; Leech, 1983; Lakoff, 1979; Brown et. al 1987).

Danescu-Niculescu-Mizil et al.,
2013

Politeness and social power

Rababah, 2014

Translatability of euphemisms

Chavez and Gerosa, 2021;
Morrisey and Kirakowski 2013,
Jain et al., 2018; Tallyn et., al,
2018;
Toxtili et al., 2018

Chatbots and politeness

Related Work cont.

Computational Approaches

There is not much computational work on recognizing and interpreting euphemisms.

Felt and Riloff, 2020

X-phemisms and Sentiment Analysis

Magu Luo, 2018; Zhu et al., 2021;
Zhu and Bhat, 2021

Euphemisms as code words
Zhu- BERT

Kapron-King and Xu, 2021

Gender study; Source of euphemisms

Corpus Creation

PETs Dataset

PET	real_meaning	topic	source
tinkle	urinate	bodily functions	-
addict	drug user	substances	-
family jewels	testicles	physical/mental attributes	Kapron-King, A., OED
accident	peeing your pants	bodily functions	-
juvenile detention center	jail for children	employment	-
aging	getting old	physical/mental attributes	Gormandy White, Mary

PET	Variations
lose your lunch	lose/lost my lunch, lose/lost her lunch, lose/lost his lunch, lose/lost their lunch, lose/lost our lunch
senior citizen	senior citizens
lay off	laid off, laying off

Sensitive/Taboo Topics

Taboo/Sensitive Topic	Count	PET Examples
death	214	pass away, late, put to sleep
sexual activity	68	make love, sex worker, sleep with
employment	246	lay off, dismissed, downsize
bodily functions	51	accident, pass gass, time of the month
politics	279	freedom fighter, correctional facility, pro-life
physical/mental attributes	387	aging, overweight, disabled
substances	115	weed, substance abuser, getting clean

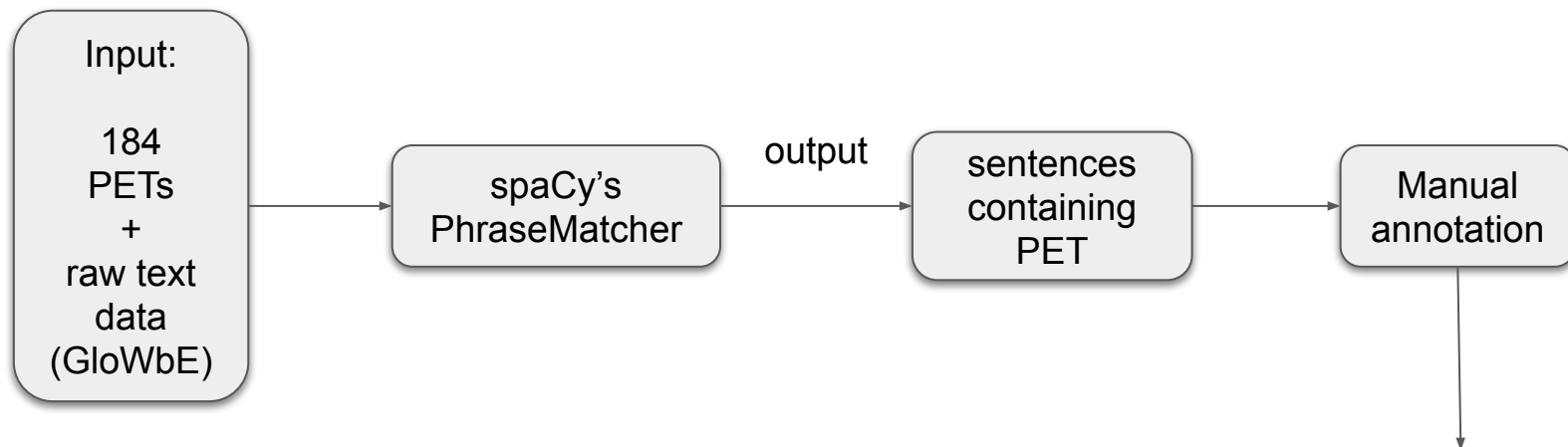
Table 3: Sensitive/Taboo Topics with PET examples

GloWbE

The Corpus of Global
Web-based English

- 1.9 billion words of text
- 20 different English speaking countries
- Optimal source since euphemisms are cultural and geographical
- We use a portion of the US Dialect of English

Corpus Creation Overview



keyword	edited_text	is_euph	category	type	euph_status
droppings	When I was a kid, I found a duck wrapped up in some fishing line. I got the line untangled, but its wing was hurt so I took it home with me. "He laughed quietly. " By the time I got there, I was covered with bird <droppings>. Mom about freaked."		1 bodily functions	droppings	always_euph
disadvantaged	I believe Samsung and its lawyers are very frustrated with Judge Koh because of the two preliminary injunctions she recently granted. Whatever the reason for Samsung's release of material deemed inadmissible evidence may be, it's very unusual for a party to a litigation to claim at the very outset that it is <disadvantaged>. I quoted Samsung's statement to the press further above, and what I find shocking is the final sentence: "Fundamental fairness requires that the jury decide the case based on all the evidence."		0 employment	disadvantaged	sometimes_euph

Corpus Details

Euphemism Corpus	
Total Sentences	1, 965
Euphemistic Sentences	1, 382
<i>with always euphs</i> PET	777
<i>with sometimes euphs</i> PET	605
Literal Sentences	583
Avg. Word Count	65 words
Avg. Char. Count	373 char.
Total PETs	129
<i>always euphs</i>	71
<i>sometimes euphs</i>	58

Euphemistic and Literal Usages of PETs

PET	Label	Example
weed	<i>euphemistic</i>	You will want to stop off at the medical marijuana dispensary for a supply of fireworks, alcohol, personal weaponry and dope-<weed>. Then, fill a glass or pop a top or load a bong or whatever one does, to get along these days.
	<i>literal</i>	In some ways, cultivating for <weed >control is almost a lost art. Herbicides seemed to work so well for so long that many farmers abandoned mechanical means of control.
disabled	<i>euphemistic</i>	No no no no. I'm in the same situation- <disabled >, chronic pain, artist, no "visible disability" (even when I'm in my chair), and nobody understands that it takes us longer to do *everything*. I'm honestly surprised you even humored your neighbor this far!
	<i>literal</i>	They claim there is no network or storage capability in these machines, clearly this is not true. These features may be <disabled >or only available to administrators who service the equipment, but in any event the TSA @ @ @ @ @ @ @ @ @ @ @ problems. As to the veterans out there who work for the TSA, I share your frustration
between jobs	<i>euphemistic</i>	I would still donate food and clothing for people in need but at least I would know that it was my choice and it was being used for it's intended purpose. I applied for temporary assistance when I was <between jobs >for a month to support my family. We had no savings or income and we were denied because I had made too much money the previous year.
	<i>literal</i>	The more new people you meet, the more your chances of finding out about a great job increases. Then if you hear back from multiple places, you'll have choices and who wouldn't want to be able to choose <between jobs >rather than grasping at the first one that comes along.

Experiments

Experiment 1: roBERTa

- We hypothesize that replacing PETs with their literal meanings should produce a shift in sentiment and offensiveness, since the usage of euphemisms makes our speech less emotionally charged.
- roBERTa base model (Liu et al., 2019; Barbieri et al., 2020)

*Just from my personal observations, among **low-income** kids, those with a strong home life tend to do better.*



*Just from my personal observations, among **poor** kids, those with a strong home life tend to do better.*

Model	Label	% Change
Sentiment	Neutral	-2.6%
	Positive	-11.3%
	Negative	54.6%
Offensiveness	Not-Offensive	-6.6%
	Offensive	30.0%

Table 5: Changes in sentiment/offensiveness scores after replacement of euphemism

Experiment 2- Annotation Task

keyword	edited_text	is_euph	interpretation	confidence
special needs	While the Hearing Officer never mentioned NCLB in his ruling, accountability, charting educational progress, and the need to teach reading skills was central to his @ @ @ @ @ @ @ @ @ @ Child Left Behind in special education litigation-I use the definitions of reading, scientifically based reading research, diagnostic reading assessment, and essential components of reading programs from the law, and appropriate training of qualified staff. Children with <Special Needs> Left Behind One issue of The Beacon: Journal of Special Education Law and Practice focused on these issues.			
disabled	I was homeschooled up until 5th grade, which is when I started attending a public elementary school, followed by a public high school for grades 7-12. I couldn't help but feel like I was in a school for the mentally <disabled> for most of my public school experience. Grammar, geography, the ability to do any sort of math without a calculator, foreign language, artistic talent, intelligent discussion, and the ability to effectively read, were all not very important whatsoever throughout my eight years of public education.			
pass away	Your story has touched my heart and I wanted to say thank you for bravely sharing Ewan's beautiful story with us. I had a baby sister <pass away> from similar problems and I have a taste of what that level of grief is like. My heart goes out to you and your family in this time @ @ @ @ @ @ @ @ @ @ on so many of us and he will not be forgotten.			

Inter-rater Agreement

- Observed Agreement
- Krippendorf's alpha

$$\alpha = 1 - \frac{D_o}{D_e}$$

1.00 - perfect reliability

0.00 - absence of reliability

Inter-rater Reliability Results	
Annotators	5
Cases	500
Decisions	2500
Average observed percent agreement	71.74%
Krippendorf's alpha	0.415

Disagreement Examples

Varying Interpretations

Below are example texts where annotators differed in their labels, and had significantly varying interpretations of the PET.

PET	text	annotator1	annotator2	annotator3	annotator4
freedom fighter	The Palestinian woman is unlike @ @ @ @ @ @ @ @ @ @ be the daughter, wife, sister or mother of the prisoners, the dead, the injured. She is a stone thrower, a <freedom fighter>. She is an inspiration for every woman as she smiles, works hard, revolts, rebels and dreams despite all what all difficulties she does through.	Label: 0 Interpretation: someone with strong beliefs and actions	Label: 0 Interpretation: actively opposing suffering under occupation	Label: 0 Interpretation: a person who fights for freedom	Label: 1 Interpretation: uses violence to achieve political goals
collateral damage	This is very very bad; that we have a CIC that will relieve field commanders for doing the right thing. I guess we are supposed to consider our Benghazi dead as <collateral damage >to BO's foreign policy.	Label: 1 Interpretation: injuries that were not intended	Label: 0 Interpretation: a talking point	Label: 1 Interpretation: unintended deaths due to war	Label: 0 Interpretation: damage as consequence

Disagreement Examples

Commonly Accepted Terms (CATs)

Below are example texts where annotators differed in their labels, and the PET in question could be considered a CAT.

PET	text	annotator1	annotator2	annotator3	annotator4
disabled	Your life testifies to the fact that life with disability can be abundantly fulfilling. Here in Nigeria, my organization AAD INITIATIVE is working toward that-fulfilling life for the <disabled>, and hopefully we shall get there.	Label: 0 Interpretation: disabled	Label: 1 Interpretation: those who have a disability, patronizing	Label: 0 Interpretation: physically or mentally impaired	Label: 0 Interpretation: broad term for people with disability
pro-choice	Since there is no agreement, the Pro-Choice >position (not a belief) understands reality. The decision to give birth is exclusively a woman's- whether or not it is legal.	Label: 1 Interpretation: pro abortion rights	Label: 0 Interpretation: commonly accepted name of a sociopolitical movement	Label: 1 Interpretation: pro-abortion	Label: 0 Interpretation: a term that represents people who know that women should make choices for their body

CATs are Fuzzy PETs: A Corpus and Analysis of Potentially Euphemistic Terms

Martha Gavidia, Patrick Lee, Anna Feldman, Jing Peng
{gavidiam1, leep6, feldmana, pengj}@montclair.edu



**MONTCLAIR STATE
UNIVERSITY**

This material is based upon work supported by the National Science
Foundation under Grant No. 1704113

Our corpus and dataset can be found at
https://github.com/marsgav/euphemism_project

