

# Was man heute noch sagen kann – Eine empirische Studie zu Slurs unter Embeddings

Carlotta Schneeberger

14.-16. Mai 2026

# Inhalt

- ① Hintergrund
- ② Experiment
- ③ Ergebnisse
- ④ Fazit

# Inhalt

① Hintergrund

② Experiment

③ Ergebnisse

④ Fazit

# Slurs

- “words that disparage people on the basis of their membership of a certain group” (Nunberg 2018: 238)
- Gruppenbezogene Beleidigungen wie das N-Wort, nicht allgemeine Beleidigungen wie die heutige Benutzung von ‘Idiot’
- Gruppen basierend auf z.B. Herkunft, Religion, Sexualität, Geschlecht
- Abwertung aufgrund von Verhalten das als inhärent für eine Gruppe gesehen wird
- Stereotype, die als positiv oder negativ dargestellt werden, aber immer entmenschlichen (Kirk-Giannini 2019: 3) und Gruppe als minderwertig sehen (Jeshion 2013: 232)

# Slurs

Darum: **content warning**

Als Inhalt dieser Arbeit werden Slurs notwendigerweise erwähnt (aber nicht benutzt) und tauchen auch in den Experiment-Stimuli auf.

Ich heiÙe die Benutzung von Slurs nicht gut und werde sie in den meisten Fällern zensieren.

# Literatur

- Unklar durch welchen Prozess die Abwertung zustande kommt
- Wortbedeutung (Croom 2015, Foster 2020, Hom & May 2013)
- Pragmatische Ebene (Camp 2018, Cepollaro 2015, McCready 2010, Nunberg 2018)

# Wortbedeutung

- Paraphrase eines Slurs: *X ist Teil der Gruppe G (hat darum Charakteristiken  $c_1, c_2, \dots$ ) und ist darum verachtenswert*
- Zählen stereotypische Eigenschaften als Teil der wörtlichen Bedeutung

# Wortbedeutung

- Paraphrase eines Slurs: *X ist Teil der Gruppe G (hat darum Charakteristiken  $c_1, c_2, \dots$ ) und ist darum verachtenswert*
- Zählen stereotypische Eigenschaften als Teil der wörtlichen Bedeutung
- Problem: Embeddings wie Verneinungen, die Entailments canceln
- “Carlotta ist keine \*\*\*” sollte demnach keine abwertende Wirkung mehr haben
- Hat es aber, Abwertung projiziert aus Embeddings heraus

# Pragmatische Ebene

- Projektion aus Verneinungen (Fragen, Konditional-Konstruktionen, ...) ist ein Merkmal für eine zusätzliche Bedeutungsebene wie Presuppositions und Conventional Implicatures (CI) (Chierchia & McConnell-Ginet 1990, Potts 2005)
- Paraphrase eines Slurs: *X ist Teil der Gruppe G*
- Durch einen pragmatischen Mechanismus wird zusätzlich ausgedrückt: *Gruppe G ist verachtenswert*
- Diese zweite Ebene kann von z.B. Verneinungen nicht ohne weiteres angesteuert werden

# Pragmatische Ebene

Carlotta ist keine \*\*\*  $\rightsquigarrow$  Gruppe G ist verachtenswert

- Presuppositions
  - Es ist nicht der Fall, dass meine Tante hat ein Haustier hat  $\rightsquigarrow$  Ich habe eine Tante
- CIs
  - Es ist nicht der Fall, dass meine Tante, die in einem Haus wohnt, ein Haustier hat  $\rightsquigarrow$  Meine Tante wohnt in einem Haus
- schwer zu differenzieren, CIs sind im Gegensatz zu Presuppositions Sprecher\*in-Orientiert, nicht Subjekt-Orientiert

presupposition is a troubled one. The beginning of the seventies, in particular, witnessed such fierce debates over this subject that even today some authors will never use the p-word without enclosing it in scare quotes, if indeed they are prepared to use it at all. However, I don't see there is any

# Pragmatische Ebene

- Um Presuppositions und CIs auseinander zu halten gibt es ein weiteres Embedding
- CIs projizieren aus indirekter Rede heraus, Presuppositions nicht (werden geplugged (Karttunen 1973))
  - Megan hat gesagt meine Tante hat ein Haustier, aber das stimmt gar nicht, ich habe gar keine Tante  $\rightarrow$  Ich habe eine Tante
  - #Megan hat gesagt meine Tante, die in einem Haus wohnt, hat ein Haustier, aber das stimmt gar nicht, meine Tante wohnt in einer Wohnung  $\rightsquigarrow$  ?Meine Tante wohnt in einem Haus
- Megan hat gesagt Carlotta ist eine **\*\*\***, aber das stimmt nicht, ich hab nichts gegen Gruppe G  $\rightsquigarrow$  ?

# Pragmatische Ebene

- Einen Slur zu sagen ist beleidigend, Verneinungen etc. ändern daran nichts, aber ist es immer noch beleidigend wenn man einen Slur zitiert bzw. wird es so aufgefasst als würde man als zitierende Person die Perspektive teilen?
- Kein Konsens in der Literatur
- Experimente bis jetzt widersprüchliche Ergebnisse (Panzeri 2016, Cepollaro u. a. 2024)
- Grad der Abwertung untersucht, nicht wer 'Verantwortlich' ist (Zitierende\*r oder Zitierte\*r)

# Inhalt

① Hintergrund

② Experiment

③ Ergebnisse

④ Fazit

# Experiment

## Verhalten sich Slurs wie Presuppositions oder wie CIs?

Ist man der Meinung, dass man als berichtende Person auch die Verantwortung für einen zitierten Slur trägt, dann folgt man einer CI-Analyse von Slurs.

- Slurs unter mehreren Embeddings, um zu sehen ob sie noch herausprojizieren
- Fragen (um auszuschließen dass die Abwertung auf der Wortebene ist), indirekte Rede (um zwischen Presupposition und CI zu unterscheiden)
- zusätzlich zu Slurs auch Presuppositions und CIs mittesten für den direkten Vergleich (Tonhauser u. a. 2018)
  - Presupposition (*stop*, possessives)
  - CI (NRRC, *damn*)
  - Slur (*n\*gger*, *f\*ggot*)

# Experiment

- online Experiment auf PCIBex (Zehr & Schwarz 2018)
- Demo-Link: <https://farm.pcibex.net/r/FRwcdL/>
- 50 VPs auf Prolific, Zustimmung zu 'harmful content'
- Experimentelle Items: trigger in verschiedenen lexikalischen Kontexten (z.B. Cupcakes)

# Experiment

- Block 1: inference task für Frage-Embedding
- Ratings von 0-100
- 'ja'-Antwort zeigt projektion (Abwertung nicht auf der Wortebene)
- Filler: lex. Kontext instantiiert ohne Trigger, projizieren nicht  
"Megan asks: Did Mark bring the cupcakes? - Would you conclude that Mark brought the cupcakes?"

Megan asks:  
'Did that faggot bring the cupcakes?'

Would you conclude that Megan  
views gay people negatively?

no  yes

# Experiment

- Block 2: acceptability task für indirekte Rede-Embedding
- Ratings von 0-100
- 'unnatürlich'-Antwort zeigt projektion (Abwertung nicht auf der Wortebene)
- Filler: natürlich ("Megan says: Katelyn said X. And then she adds: But  $\neg X$ "), unnatürlich ("Megan says: Katelyn said X. And then she adds: But  $\neg$ Kately said X")

Megan says:  
'Katelyn said that nigger is a stuntman.'  
And then she adds:  
'But there's nothing bad about black people.'

How natural is the sentence in blue given the previous report?

completely unnatural  completely natural

# Inhalt

① Hintergrund

② Experiment

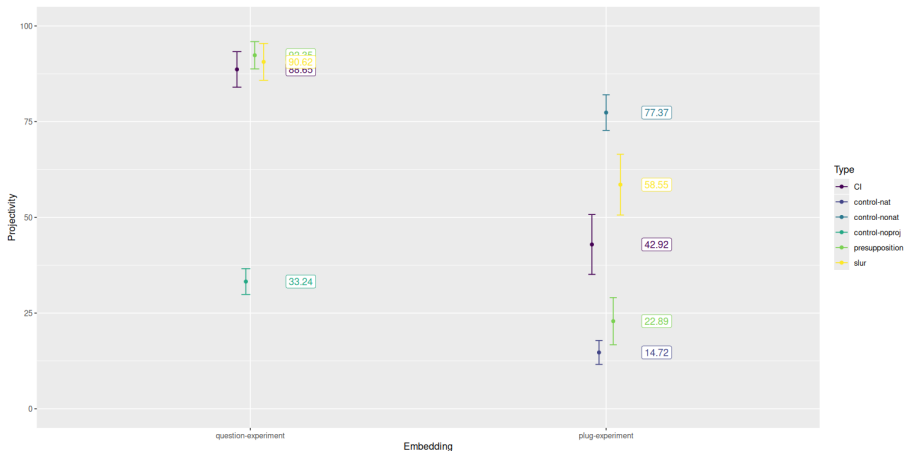
③ Ergebnisse

④ Fazit

# Erwartungen

- Wortebene: Slurs im Fragen-Block alle niedrige Ratings wie Filler, nicht wie Presuppositions und CIs
- CI-Theorie: Slurs wie CIs und unnatürliche Filler im indirekte Rede-Block
- Presupposition-Theorie: Slurs wie Presupposition und natürlicher Filler im indirekte Rede-Block

## Analyse



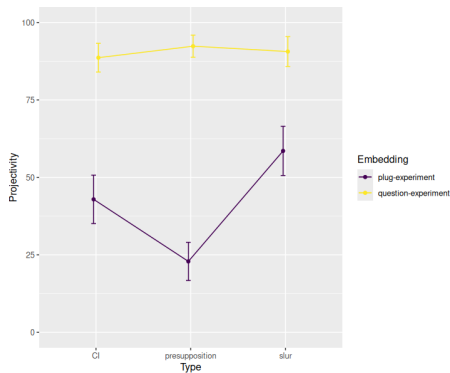
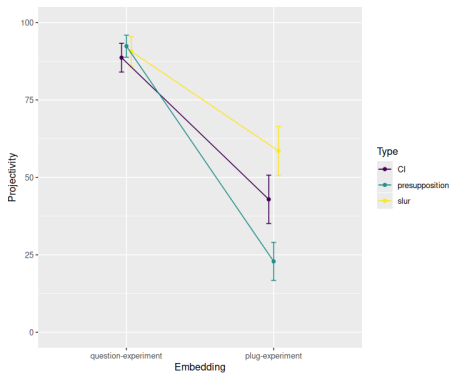
**Abbildung:** Mittelwerte der (transformierten) Ratings mit 95% Konfidenzintervall der Types in beiden Experimentblöcken

# Analyse

- linear mixed effect models und post-hoc tests
- Fragen-Experiment: alle Types signifikant unterschiedlich zum Filler ( $p < 0.001$ ) aber nicht zu einander ( $p \approx 1$ )
- indirekte Rede: nur Presuppositions unterscheiden sich nicht signifikant vom natürlichen Filler ( $t(36.7) = -2.388$ ,  $p = 0.22$ )
- Slurs verhalten sich nicht wie Presuppositions
- Slurs sind aber auch marginal signifikant unterschiedlich zu CIs ( $t(48.9) = -3.111$ ,  $p = 0.0311$ ).

# Interaktion

- signifikante Interaktion von Type:Embedding
- indirekte Rede beeinflusst die drei Types verschieden stark



# Inhalt

① Hintergrund

② Experiment

③ Ergebnisse

④ Fazit

# Fazit

## Gov. Andrew Cuomo uses 'n-word' to make point about derogatory terms against Italians

**Abbildung:** New York Gov. Andrew Cuomo zitiert einen Slur aus der New York Times (<https://edition.cnn.com/2019/10/15/politics/andrew-cuomo-racial-slur>)

- wenn wir nur auf die Unterschiede zu den nicht-projizierenden Fillern schauen verhalten sich Slurs wie CIs in der Literatur beschrieben sind
- $\neq$  Abwertung auf Wortebene,  $\neq$  Abwertung als Presupposition
- Projizieren zum speaker-level  $\rightarrow$  Abwertung wird der sprechenden Person zugeschrieben  
(wenn sich nicht aktiv davon abgegrenzt wird, Level der Beleidigung sinkt z.B. mit Anführungszeichen (Cepollaro u. a. 2024))

# Fazit

- allerdings passen sie nicht komplett zu den CI-Items
  - Unterschied von Slurs und CIs hauptsächlich ausgehend von den NRRCs, näher an Presuppositions
  - passt zu Harris & Potts (2009), dass CIs nicht zwingend speaker-oriented sind (NRRCs mehr als *expressives*)
  - größeres Spektrum für CIs, wo auch Slurs zugehören
- Keine Slurs benutzen **und** auch nicht unzensiert zitieren weil es im allgemeinen eine befürwortende Wirkung hat (im paradigmatischen Kontext)










# Ausblick

Schreibt mir gerne, wenn ihr Fragen oder relevante Beispiele habt!

▶ [carlotta.schneeberger@hhu.de](mailto:carlotta.schneeberger@hhu.de)

Danke :-)

# Literatur I

-  Camp, Elisabeth. 2018. A dual act analysis of slurs. *Bad words: Philosophical perspectives on slurs* 1.
-  Cepollaro, Bianca. 2015. In defence of a presuppositional account of slurs. *Language Sciences* 52. 36–45.
-  Cepollaro, Bianca, Simone Sulpizio, Claudia Bianchi & Isidora Stojanovic. 2024. Slurs in quarantine. *Mind & Language* 39(3). 381–396.
-  Chierchia, Gennaro & Sally McConnell-Ginet. 1990. *Meaning and grammar: An introduction to semantics*. MIT press.
-  Croom, Adam M. 2015. The semantics of slurs: A refutation of coreferentialism. *Ampersand* (2). 30–38.
-  Foster, Jennifer. 2020. Beyond 'Neutral Counterparts': Towards an Overlap Theory of Derogatory Terms.
-  Harris, Jesse A. & Christopher Potts. 2009. Perspective-shifting with appositives and expressives. *Linguistics and Philosophy* 32(6). 523–552.  
<https://doi.org/10.1007/s10988-010-9070-5>.
-  Hom, Christopher & Robert May. 2013. Moral and semantic innocence. *Analytic Philosophy* 54(3). 293–313.
-  Jeshion, Robin. 2013. Slurs and stereotypes. *Analytic Philosophy* 54(3). 314–329.

# Literatur II

-  Karttunen, Lauri. 1973. Presuppositions of compound sentences. *Linguistic inquiry* 4(2). 169–193.
-  Kirk-Giannini, Cameron Domenico. 2019. Slurs Are Directives. *Philosophers' Imprint* 19. 1–28.
-  McCready, Elin. 2010. Varieties of conventional implicature. *Semantics and Pragmatics* 3, Article 8. 1–57.
-  Nunberg, Geoffrey. 2018. The social life of slurs. *New work on speech acts*. 237–293.
-  Panzeri, Francesca. 2016. L'offensività degli slurs. Un'indagine sperimentale/The offensiveness of slurs: an experimental investigation. *Rivista Italiana di Filosofia del Linguaggio*.
-  Potts, Christopher. 2005. *The Logic of Conventional Implicatures*. Oxford University Press.
-  Tonhauser, Judith, David I Beaver & Judith Degen. 2018. How projective is projective content? Gradience in projectivity and at-issueness. *Journal of Semantics* 35(3). 495–542.
-  Zehr, Jérémy & Florian Schwarz. 2018. PennController for Internet Based Experiments (IBEX). <https://doi.org/10.17605/OSF.IO/MD832>.