

Final Presentation

Vedika Ahuja



My Project

Objective: Explore how the media uses the words lesbian, gay, bisexual, transgender, and queer (LGBTQ), and how the usage and meaning of the words have changed over time.

Sub-goals:

1. Determine where various words related to sexual orientation and gender identity (stud, butch, androgynous, gay, lesbian, etc.) are located on the cultural dimensions of race, class and gender in the news media
2. Study how the description, meaning, and context of LGBTQ groups and related words have changed over the last 20 years in the media
3. Detect linguistic shifts in the meaning of LGBTQ words over time

My Project

1. Determine where various words related to sexual orientation and gender identity (stud, butch, androgynous, gay, lesbian, etc.) are located on the cultural dimensions of race, class and gender in the news media
2. Study how the description, meaning, and context of LGBTQ groups and related words have changed over the last 20 years in the media
3. Detect linguistic shifts in the meaning of LGBTQ words over time

Corpus

- The Corpus of Contemporary American English (COCA)
- Contains more than one billion words of text (20 million words each year 1990-2019) from five genres: spoken, fiction, popular magazines, newspapers and academic texts
- Specifically all news articles from 1990 to 2015
- A total of 3,648 documents use the words lesbian, gay, bisexual, transgender or queer

Methodology

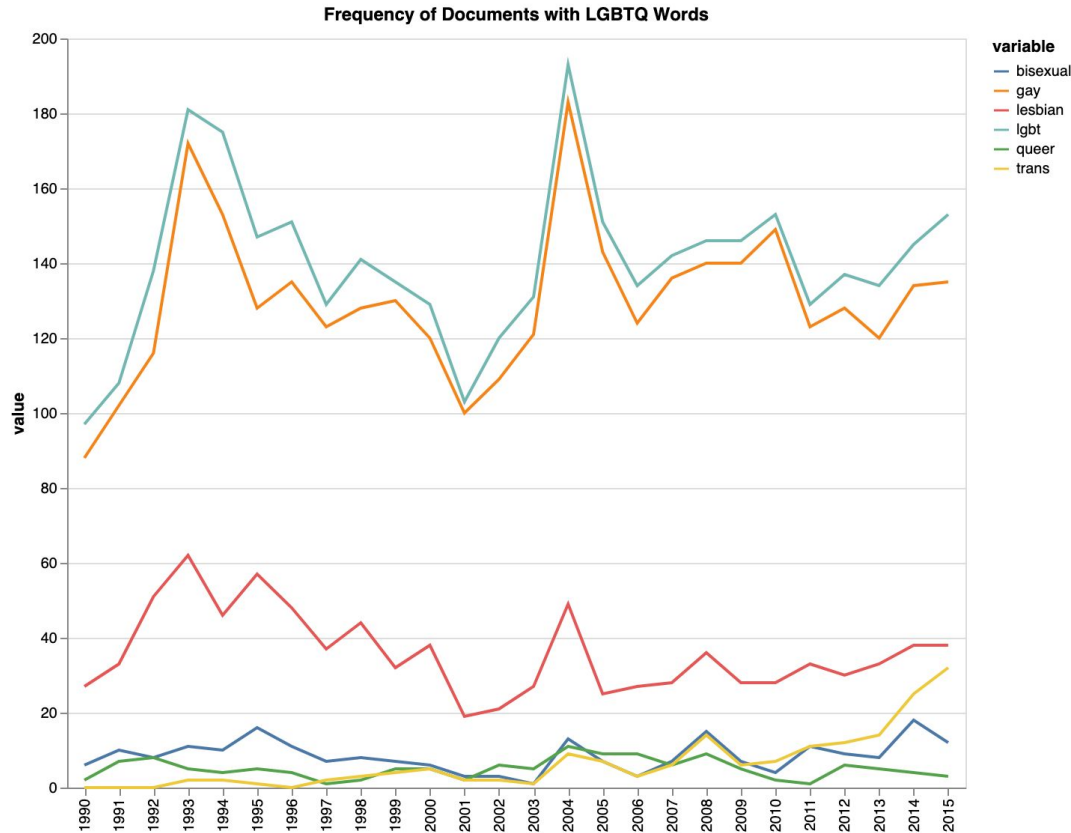
- Find similar words to LGBTQ words over time
 - Using word2vec word embedding models
- Classify and categorize LGBTQ words along cultural dimensions of race, class, and gender
 - Used word2vec embedding models to position LGBTQ words relative to other words representing the cultural dimensions of gender, race, and class
 - Approach from Kozlowski, Austin, Matt Taddy, James Evans. 2019. "The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings."
- Identify semantic change of LGBTQ words over time
 - Using approach piloted at Stanford by William Hamilton, Daniel Jurafsky and Jure Leskovec, used in HW 7

Methodology

- Find similar words to LGBTQ words over time
 - Using word2vec word embedding models
- Classify and categorize LGBTQ words along cultural dimensions of race, class, and gender
 - Used word2vec embedding models to position LGBTQ words relative to other words representing the cultural dimensions of gender, race, and class
 - Approach from Kozlowski, Austin, Matt Taddy, James Evans. 2019. "The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings."
- Identify semantic change of LGBTQ words over time
 - Using approach piloted at Stanford by William Hamilton, Daniel Jurafsky and Jure Leskovec, used in HW 7

Descriptive Statistics:

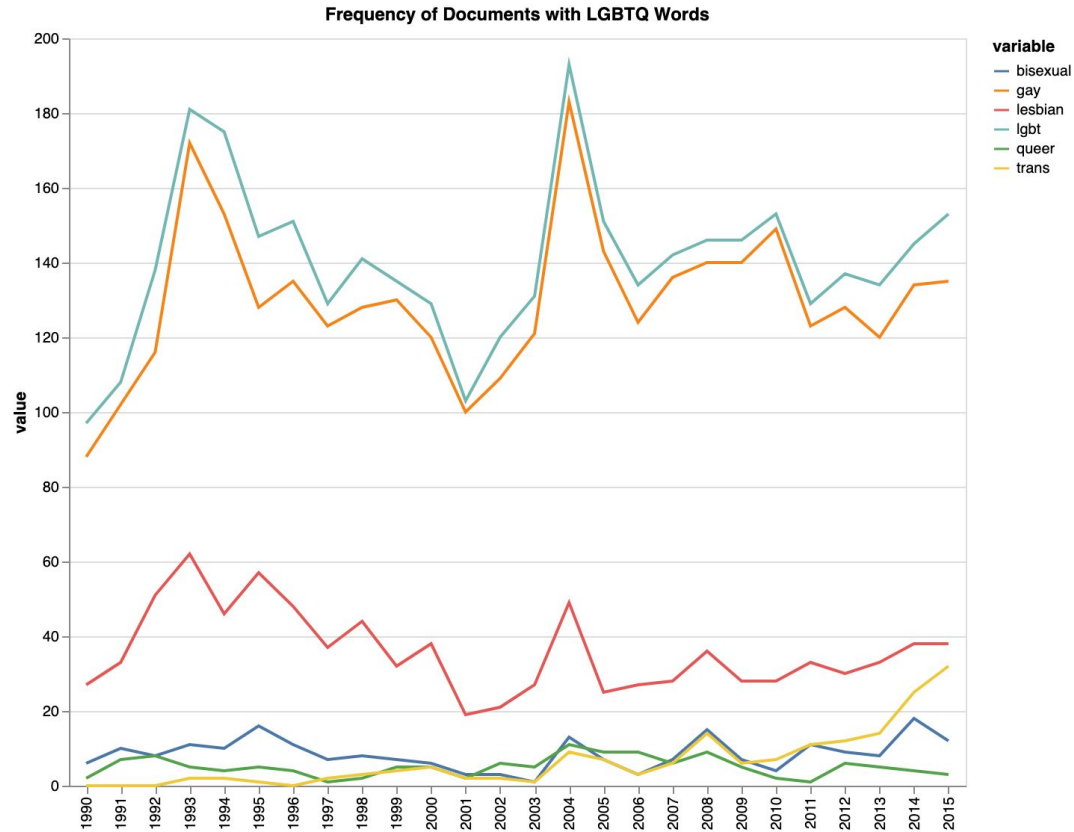
Usage of LGBTQ words over time



- The usage of gay makes up majority of LGBT mentions in the news
- Big spike in new articles with “gay” in it during the early 1990s, due to the AIDS epidemic
- Also a spike in 2004, when same-sex marriage first became legal (in Massachusetts)
- The usage of transgender has steadily increased since 2009
- “Queer” is the least used word in the news

Descriptive Statistics:

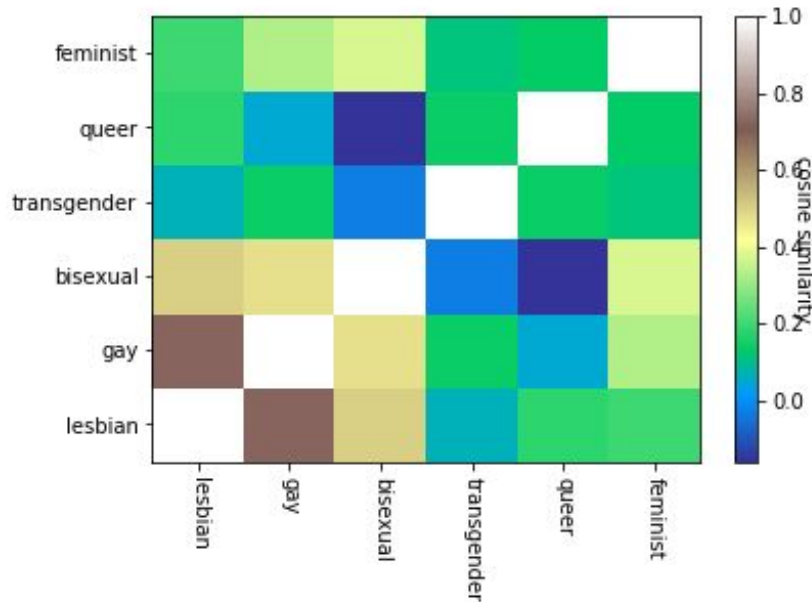
Usage of LGBTQ words over time



- The usage of gay makes up majority of LGBTQ mentions in the news
- Big spike in new articles with “gay” in it during the early 1990s, due to the AIDS epidemic
- Also a spike in 2004, when same-sex marriage first became legal (in Massachusetts)
- The usage of transgender has steadily increased since 2009
- “Queer” is the least used word in the news

Word Vector Similarities

- Queer is the most different from other LGBT words, though it is the closest to “lesbian.”
- Transgender is located far from lesbian and bisexual

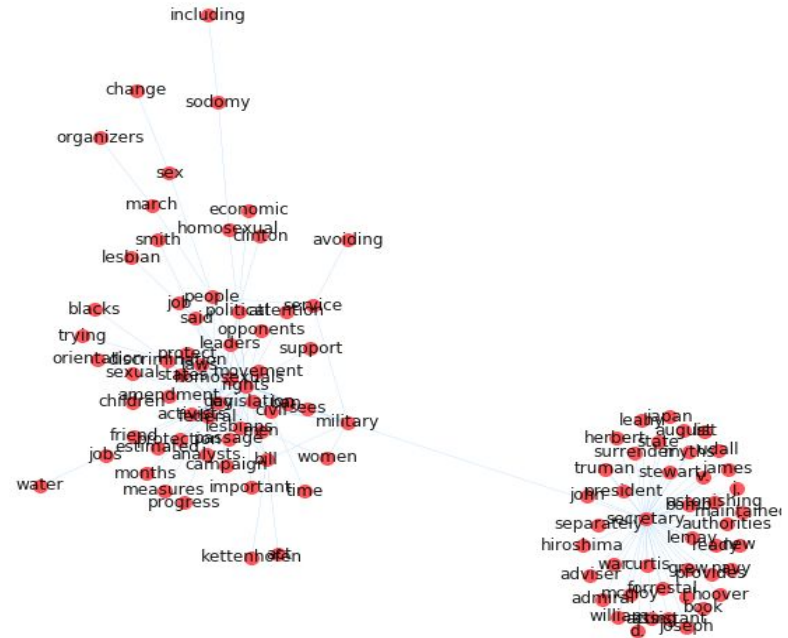


Semantic Network Analysis - Transgender

Method: Compare the subnetworks around transgender in the corpus from 1990-1995 to the corpus from 2010-2015.

Articles in Epoch 1

- Only 5 articles mentioned “transgender” in my corpus during this time
- These articles have are related to two topics of words:
 - The military
 - Trans and lgbt civil rights and activism

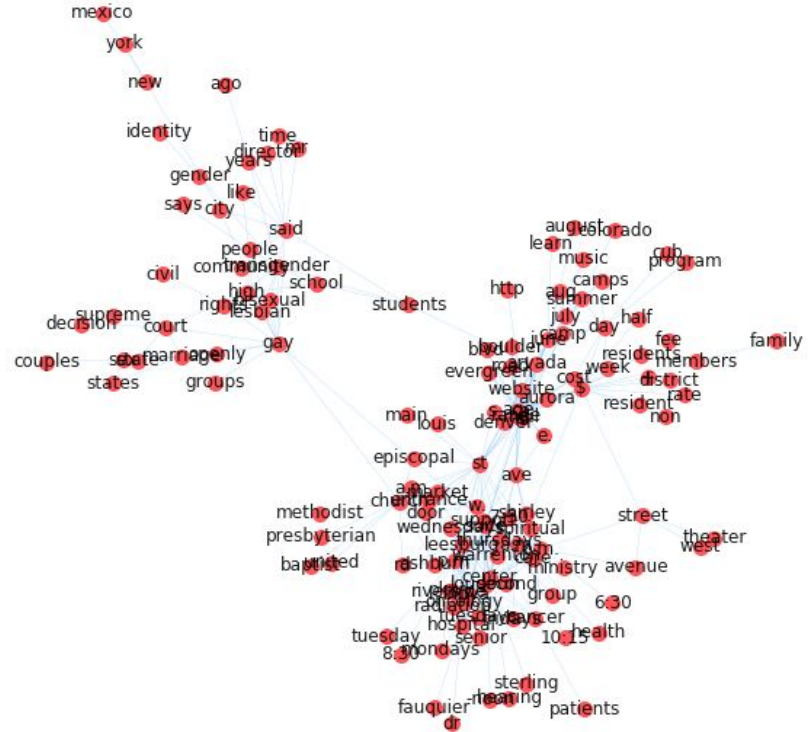


Semantic Network Analysis - Transgender

Method: Compare the subnetworks around transgender in the corpus from 1990-1995 to the corpus from 2010-2015.

Articles in Epoch 5

- 101 articles mentioned “transgender” in my corpus during this time
- 4 parts of the network:
 - Civil rights
 - Art and music
 - Health Care
 - Religion

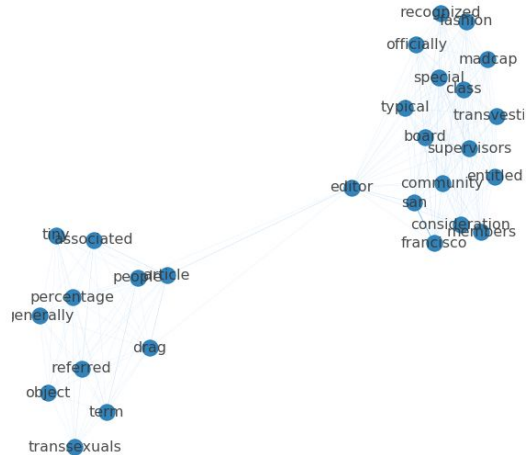


Semantic Network Analysis - Transgender

Nodes within 1 or 2 network steps of 'transgender'

Epoch 1

- Not related to any other communities in the LGBTQIA groups



Epoch 5

- Closely related to gay and closer to lesbian



Similar Words Over Time - Gay

- Biggest difference has to do with marriage equality (equality, licenses, religious)
- Also in 2010-2015 mentions of women and abortion are not as prominent

Words Most Similar to Gay From 1990-1995

```
cocaW2V_1.most_similar('gay')
```

```
[('lesbian', 0.9127849340438843),  
 ('black', 0.8849095106124878),  
 ('civil', 0.8727728128433228),  
 ('rights', 0.8678826093673706),  
 ('anti', 0.8632321357727051),  
 ('human', 0.8596915006637573),  
 ('men', 0.8425957560539246),  
 ('community', 0.8386508226394653),  
 ('abortion', 0.8160275220870972),  
 ('women', 0.8047685623168945)]
```

Words Most Similar to Gay From 2010-2015

```
cocaW2V_5.most_similar('gay')
```

```
[('sex', 0.9113506078720093),  
 ('civil', 0.8692562580108643),  
 ('human', 0.8567376732826233),  
 ('equality', 0.8484565019607544),  
 ('licenses', 0.8195568919181824),  
 ('religious', 0.8189222812652588),  
 ('lgbt', 0.8178043365478516),  
 ('bisexual', 0.8148080110549927),  
 ('gays', 0.8106257915496826),  
 ('interracial', 0.8050644397735596)]
```

Similar Words Over Time - Gay

- Biggest difference has to do with marriage equality (equality, licenses, religious)
- Also in 2010-2015 mentions of women and abortion are not as prominent

Words Most Similar to Gay From 1990-1995

```
cocaW2V_1.most_similar('gay')
```

```
[('lesbian', 0.9127849340438843),  
 ('black', 0.8849095106124878),  
 ('civil', 0.8727728128433228),  
 ('rights', 0.8678826093673706),  
 ('anti', 0.8632321357727051),  
 ('human', 0.8596915006637573),  
 ('men', 0.8425957560539246),  
 ('community', 0.8386508226394653),  
 ('abortion', 0.8160275220870972),  
 ('women', 0.8047685623168945)]
```

Words Most Similar to Gay From 2010-2015

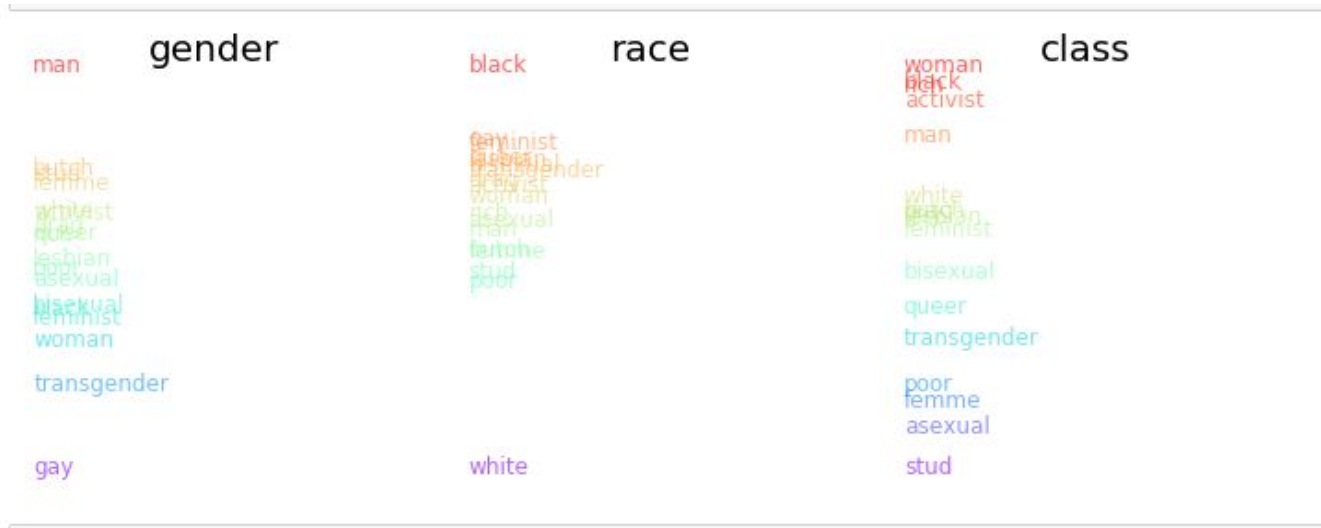
```
cocaW2V_5.most_similar('gay')
```

```
[('sex', 0.9113506078720093),  
 ('civil', 0.8692562580108643),  
 ('human', 0.8567376732826233),  
 ('equality', 0.8484565019607544),  
 ('licenses', 0.8195568919181824),  
 ('religious', 0.8189222812652588),  
 ('lgbt', 0.8178043365478516),  
 ('bisexual', 0.8148080110549927),  
 ('gays', 0.8106257915496826),  
 ('interracial', 0.8050644397735596)]
```

Results

Projecting Words onto Cultural Dimensions

- Gay is the most feminine word in the LGBT lexicon, more so than “woman”!
- Transgender is a close second - interesting because transgender is a gender neutral word, but this could mean the media discusses trans women more than trans men or non-binary folks
- All LGBT words are slightly closer to black than white
- Transgender is closer to poor than rich



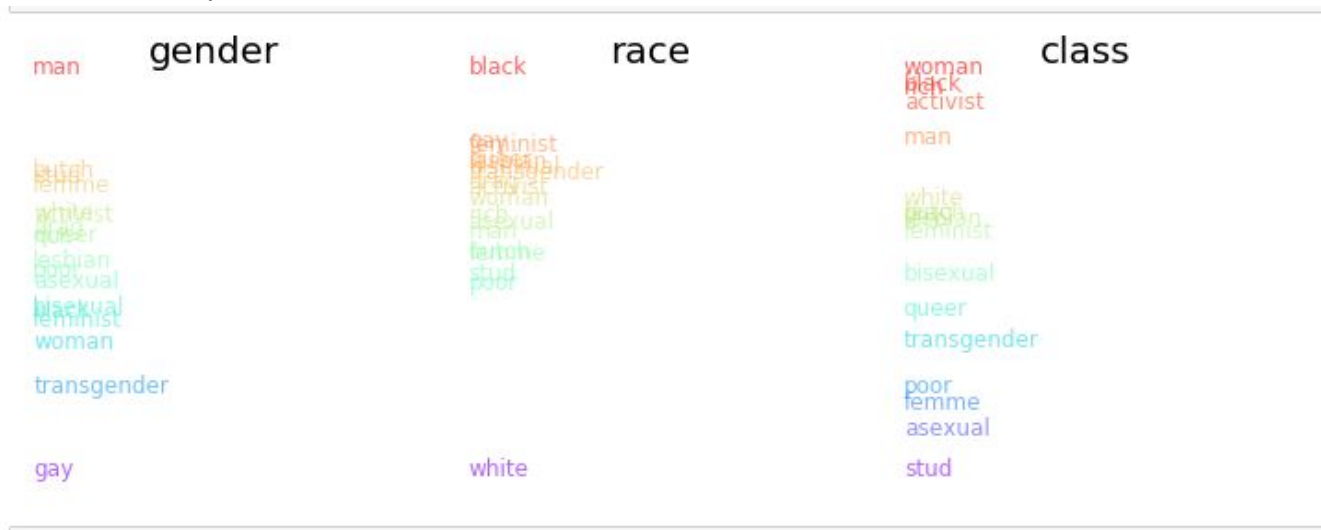
Projecting Words onto Cultural Dimensions

-
- man gender black race woman class
- black
- woman
- activist
- man
- white
- lesbian
- feminist
- bisexual
- queer
- transgender
- poor
- femme
- asexual
- stud
- gay
- white
- poor
- feminist
- lesbian
- queer
- feminist
- femme
- stud
- poor
- white
- black
- bisexual
- asexual
- gender
- man
- woman
- transgender
- gay
- lesbian
- queer
- feminist
- femme
- stud
- poor
- white
- black
- bisexual
- asexual

Results

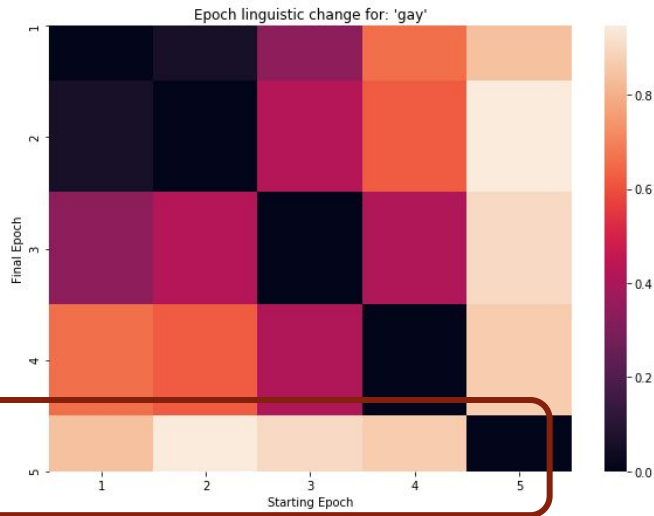
Projecting Words onto Cultural Dimensions

- Gay is the most feminine word in the LGBT lexicon, more so than “woman”!
- Transgender is a close second - interesting because transgender is a gender neutral word, but this could mean the media discusses trans women more than trans men or non-binary folks
- All LGBT words are slightly closer to black than white
- Transgender is closer to poor than rich

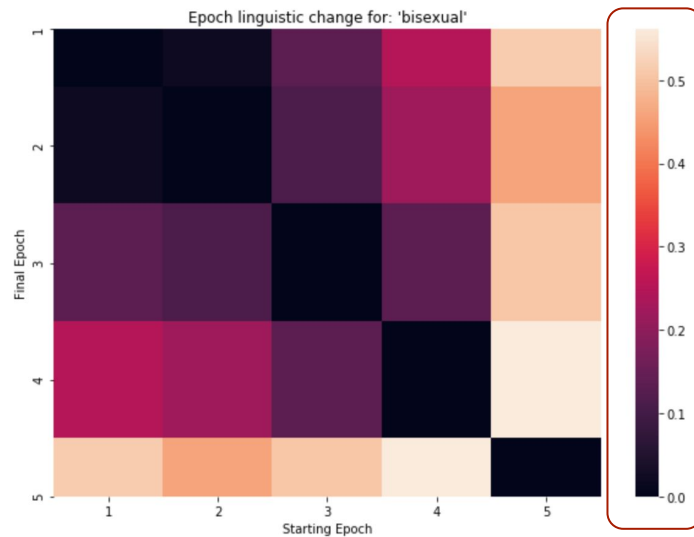


Results - Epoch Linguistic Change - Gay

- The meaning of “gay” changes the most in the last period (2010-2015) compared to every other period
- Very little change in the meaning/usage between 1990 and



- The semantic meaning of bisexual changed the least and the most slowly of all the LGBTQ words



Conclusion

- Out of all LGBTQ issues, the media discusses “gay” issues the most
 - There was a spike in the usage of gay and lesbian in the early 90s with the AIDS epidemic, and all LGBTQ words in the mid 2000s with greater discussion of same-sex marriage
- On cultural dimensions:
 - Gay is the farthest away from “man”
 - Transgender, queer, bisexual are all closer to poor
- The change in the semantic network surrounding transgender indicates more inclusion by the media of trans issues as part of LGBT issues, and generally reflects more court cases about transgender rights
- The words located near lesbian, gay, and bisexual have shifted over time from being more about civil liberties to more of equality and marriage
- The semantic meaning of gay has changed the most of any LGBTQ words over time, while lesbian has changed the least

Conclusion

- Out of all LGBTQ issues, the media discusses “gay” issues the most
 - There was a spike in the usage of gay and lesbian in the early 90s with the AIDS epidemic, and all LGBTQ words in the mid 2000s with greater discussion of same-sex marriage
- On cultural dimensions:
 - Gay is the farthest away from “man”
 - Transgender, queer, bisexual are all closer to poor
- The change in the semantic network surrounding transgender indicates more inclusion by the media of trans issues as part of LGBT issues, and generally reflects more court cases about transgender rights
- The words located near lesbian, gay, and bisexual have shifted over time from being more about civil liberties to more of equality and marriage
- The semantic meaning of gay has changed the most of any LGBTQ words over time, while lesbian has changed the least