Gender Differences in Film Conversations

Ankit Saxena
University of Pennsylvania
Philadelphia, PA 19104

Abstract

Film dialogues can provide insight into gender-based patterns and biases that are prevalent in the filmmaking industry. In this paper, we compare gender-based differences in film dialogues using simple statistical tools. Then we examine structural differences in conversations between women and men in terms of quantity and tone using sentiment analysis tools in Python.

1. Introduction

Films are an important form of communication for filmmakers to convey their vision to people. While each film has many aspects, from its music to the set design, dialogues and conversations are one of the most important ways to communicate a film's message and tone. Therefore, gender stereotypes represented in dialogues and conversations have the power to set certain expectations in the minds of viewers. Prolonged exposure to certain gender stereotypes in films can profoundly affect our attitudes and beliefs (Fearing, 1947).

Previously, people have studied differences in gender in terms of dialogues through methods like the Bechdel–Wallace test which checks if a movie has at least two women in it who talk to each other about something besides a man. A more scientific approach used by scholars is to understand the gender-based differences in quantity of dialogues in films (Schofield and Mehr, 2016; Sap et al., 2017).

In this paper, we hope to uncover such gender gaps in film dialogues and conversations using a series of hypothesis tests. Our results revealed statistically significant differences in both the quantity of dialogues in films as well as qualitative aspects like tone.

We began by quantifying the dialogues between women and men in films and identified gaps in the number of sentences, words, and characters in films. We then shifted our focus to conversations and identified gaps in the sequence of dialogues between women and men. Finally, we compared certain qualitative factors like the use of pronouns and tone of dialogues between women and men.

2. Data Description

Our dataset came from the Cornell Movie-Dialogs Corpus (Danescu-Niculescu-Mizil et al., 2011), a collection of dialogues from 617 film scripts. The dataset contains 304,713 individual utterances of dialogues which translate to 220,579 conversational exchanges between 10,292 characters.

All movies with fewer than five votes on the Internet Movie Database (IMDb) were not included. The pairs of characters that interacted were identified using simple data processing heuristics. Finally, the gender of characters of a film were determined by equating them with the gender of the actress or actor performing the role, which was taken from IMDb.

Simple exploration of the dataset revealed that overall, there are fewer women dialogues in films compared to men. The same trend is observed in Figure 1 if we group the data in terms of decade of film released. We also hypothesized that the proportion would be

closer to 0.5 with increase in release year, but that was not the case.

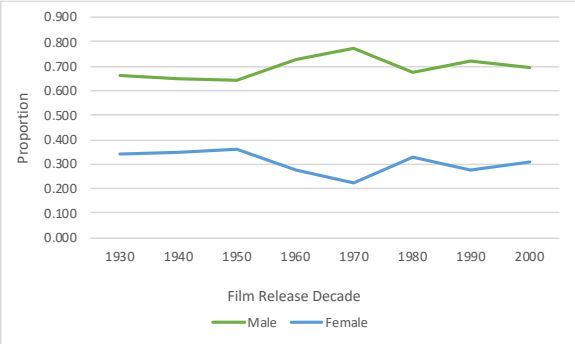

Figure 1: Proportion of dialogues for each gender grouped by the decade of the film's release

3. Experiments

3.1. Gender Gap in Dialogue Quantity

We hypothesized that women would have fewer dialogues compared to men in films and were able to confirm that through a series of t-tests.

When testing for differences in the average length of each dialogue by calculating character count per dialogue, we discovered that there is a statistically significant difference between the average dialogue length for men compared to women. On average, men had longer dialogues compared to women, which can be seen in Table 1 and Figure 2.

	Men	*Women*
Mean	56.34358	51.774756
Variance	12482.0802	3493.3473
Observations	169844	70892
Hypothesized Mean Difference	0	
df	228182	
t Stat	13.039502	
P(T<=t) one-tail	3.7653E-39	
t Critical one-tail	1.6448603	
P(T<=t) two-tail	7.5306E-39	
t Critical two-tail	1.95997438	

Table 1: Results from t-test for differences in average number of characters per dialogue for men and women

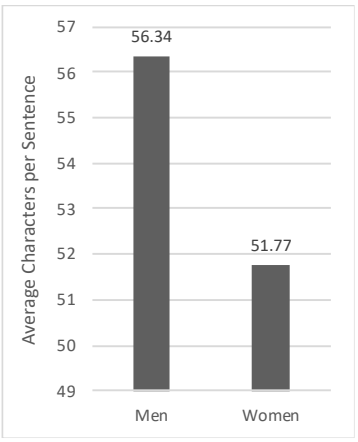

Figure 2: Average characters per sentence for men and women

When testing for the length of dialogues by calculating the number of words per sentence, we observed a similar pattern. On an average, men tend to have a higher number of words in their dialogues compared to women which can be seen in Table 2 and Figure 3.

	Men	*Women*
Mean	10.73211889	9.99709417
Variance	271.1650168	127.09889
Observations	169844	70892
Hypothesized Mean Difference	0	
df	190359	
t Stat	12.62525503	
P(T<=t) one-tail	7.92535E-37	
t Critical one-tail	1.644861632	
P(T<=t) two-tail	1.58507E-36	
t Critical two-tail	1.959976447	

Table 2: Results from t-test for differences in average number of words per dialogue for men and women

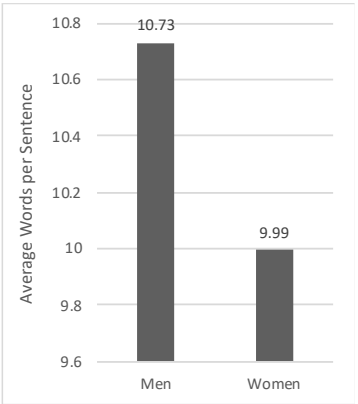

Figure 2: Average words per sentence for men and women

3.2. Gender Gap in Conversations

Very few characters in films engage in monologues. Most characters engage in conversations with other characters. Conversations are a crucial aspect of communication in films because they can reveal important information about the emotional and power dynamics between two or more characters which cannot be understood by analyzing individual dialogues in isolation.

We codified the film dialogues into arrays of sequential conversations using simple heuristics of character interactions. Using these arrays, we were able to identify which gender tends to start or end conversations, which could give linguists cues into the power dynamics between men and women.

For this analysis, we removed the conversations for which we did not have information on the gender of either the conversation starter of conversation finisher. That left us with 55,639 from 515 films.

For conversation between a man and a woman, we saw a higher number of men who start and end conversations compared to women – both overall in our dataset and on an average per film – which can be seen from Table 3 and Table 4 respectively.

	Conversation Starters	Conversation Finishers
Women	17,005	16,081
Men	38,634	39,558

Table 3: Total number of conversations where men and women started or finished a conversation

	Conversation Starters	Conversation Finishers
Women	33	31
Men	75	77

Table 4: Average number of conversations where men and women started or finished a conversation per film

Also, when calculating the total number of conversations as well as the average number of conversations between gender pairs (women-women, women-men, and men-men) in our reduced dataset of 515 movies, we observed that conversations between men and men exceeded those of both women and men as well as women and women.

	Total Conversations	Conversations per Film
Women-Women	832	1.6
Women-Men	8,383	16.3
Men-Men	8,800	17.1

Table 5: Average and total number of conversations between different gender pairs

3.3. Sentiment Analysis

By measuring the sentiment of dialogues, we could objectively quantify the tone of different characters. We used sentiment labels from the Python-based sentiment analysis tool Valence Aware Dictionary and sEntiment Reasoner or VADER (Hutto and Gilbert, 2014) to obtain scores for positive, negative, neutral, and composite value of each dialogue. VADER uses a sentiment lexicon which is a list of lexical features (e.g., words) which are generally labelled according to their semantic orientation like positive or negative.

First, when analyzing for negative sentiment of dialogues for different genders, we observed that there is no statistical difference between women and men, which can be seen in the t-test results in Table 6.

	Men	*Women*
Mean	0.080082893	0.080076836
Variance	0.029460338	0.030722288
Observations	170003	70904
Hypothesized Mean Difference		0
df		130268
t Stat		-0.00777617
P(T<=t) one-tail		0.496897794
t Critical one-tail		1.644865324
P(T<=t) two-tail		0.993795589
t Critical two-tail		1.959982195

Table 6: Results from t-test for differences between men and women for average negative words per sentence calculated by VADER

When performing a similar test for differences in neutral language between men and women, we found that men tend to use

more neutral language on an average, which can be observed in Table 7.

	Men	*Women*
Mean	0.802602795	0.793920639
Variance	0.062747633	0.067037161
Observations	170003	70904
Hypothesized Mean Difference		0
df		128876
t Stat		-7.572464148
P(T<=t) one-tail		1.84317E-14
t Critical one-tail		1.644865451
P(T<=t) two-tail		3.68634E-14
t Critical two-tail		1.959982392

Table 6: Results from t-test for differences between men and women for average neutral words per sentence calculated by VADER

Performing a similar test for positive language showed that women tend to use more positive words compared to men, which can be seen in Table 7.

	Men	*Women*
Mean	0.117291124	0.125988308
Variance	0.044517764	0.048983119
Observations	170003	70904
Hypothesized Mean Difference	0	
df	127219	
t Stat	-8.910460909	
P(T<=t) one-tail	2.5735E-19	
t Critical one-tail	1.644865605	
P(T<=t) two-tail	5.147E-19	
t Critical two-tail	1.959982632	

Table 7: Results from t-test for differences between men and women for average positive words per sentence calculated by VADER

Finally, performing a t-test for VADER scores for compound language revealed that women tend to have a higher score than men denoting an overall more positive sentiment, which can be seen in Table 8. The compound score is a metric that calculates the sum of all the lexicon ratings which have been normalized between -1 (most extreme negative) and +1 (most extreme positive).

	Men	*Women*
Mean	0.034171998	0.046549683
Variance	0.11978839	0.11600472
Observations	170003	70904
Hypothesized Mean Difference	0	
df	134706	
t Stat	-8.090324452	
P(T<=t) one-tail	2.99977E-16	
t Critical one-tail	1.644864939	
P(T<=t) two-tail	5.99953E-16	
t Critical two-tail	1.959981595	

Table 8: Results from t-test for differences between men and women for average compound words per sentence calculated by VADER

3.4. Experiments in Progress

- Advanced emotional analysis using LIWC dictionary.
- Breaking down values by film genres.
- Controlling for screenwriter gender.
- Relationship between IMDb score, box office and proportion of male-to-female dialogues.
- Relationship between end-credit sequence and dialogue quantity, separately for men and women.
- Most common vocabulary
- Checking for 1st personal pronouns (I, me, we) that denote more self-centered dialogues versus 2nd and 3rd person pronouns (him, she, they) that denote more outward-centered dialogues – both for men and women.

3.5. Conclusion

Using simple numerical analyses as well as more advanced sentiment analyses, we were able to discover interesting differences in the quantity and quality of dialogues between men and women. This study does not account for non-binary genders in films and is an area that should be explored in future studies.

Quantifying such differences provides scientific evidence that there is a gender gap in the representation of men and women in films, at least through their dialogues. Studios should take note and try to close that gap to avoid misrepresenting genders in their films.

4. References

[Danescu-Niculescu-Mizil et al., 2012a] Cristian Danescu-Niculescu-Mizil, Justin Cheng, Jon Kleinberg, Lillian Lee. 2012a. *You had me at hello: How phrasing affects memorability*. In Proceedings of ACL, pages 892–901.

[Fearing, 1947] Franklin Fearing. *The Annals of the American Academy of Political and Social Science* Vol. 254, The Motion Picture Industry (Nov., 1947), pp. 70-79.

[Hutto and Gilbert, 2014] CJ Hutto and Eric Gilbert. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Eighth International AAAI Conference on Weblogs and Social Media.

[Schofield and Mehr, 2016; Sap et al., 2017] Alexandra Schofield, Leo Mehr. *Gender-Distinguishing Features in Film Dialogue*. NAACL-CLfL (2016).