

Who was the most polite F.R.I.E.N.D? - Predicting Conversational Politeness Scores of Characters on a TV Show

Abhinav Gupta

December 14, 2021

Abstract

The politeness level classifier put forth by Mizil et al [1] is one of the most popular and robust computational frameworks for identifying linguistic aspects of politeness. This can also be further extended to study the relationships between politeness and social behavior. In this work, we leverage this politeness classifier, which is now part of the Cornell Conversational Analysis Toolkit ([ConvoKit](#)), to obtain politeness features and markers of the popular [F.R.I.E.N.D.S corpus](#). This corpus consists of all the conversations that occurred over 10 seasons of Friends, a popular American TV sitcom that ran in the 1990s. We use the ConvoKit politeness classifier to predict the politeness scores of the show’s characters and find the most polite character¹. Furthermore, we also study the relation of politeness with social factors such as the gender and social status of all six F.R.I.E.N.D.S characters.

Literature Review

The authors of [1] put forth a computational framework for identifying linguistic aspects of politeness. They train a classifier on a new corpus of various annotated requests from wikipedia/stack-exchange which achieves close to human performance. They also further use their framework to study relationships between politeness and social power and claim that there is a negative correlation between politeness and power. They also explore the relation with other social factors such as gender and community membership.

Zhang et al [2] investigate how politeness strategies are tied to future trajectories of conversations. The politeness package tool put forth by Yeomans et al [3] helps extract politeness markers in the English language. This paper extends and further improves the work done in [1]. Aubakirova and Bansal [4] study the use of neural networks to help understand politeness in natural language requests and perform politeness classification using deep learning. Using machine learning tools and frameworks to classify conversational data can be very challenging. Hoffmat et al [5] evaluate the challenges faced by such ML classifiers to predict politeness and propose future directions for improvement.

¹[Link to our code](#), which further elucidates on how we trained the classifier on the wikipedia corpus and used it to predict labels for the FRIENDS corpus. The zip file contains two jupyter notebooks, one for identifying politeness strategies and one for using the pretrained model on the FRIENDS dataset.

It may be difficult to classify politeness for indirect requests. For instance, the FRIENDS corpus consists of many indirect sarcastic comments such as *Shouldn't you be telling me where Monica is?* instead of *Can you please tell me where Monica is?*. Clark and Schunk [6] evaluate politeness scores for such indirect requests. Burke and Kraut [7] study the impact of politeness on desired outcomes, such as increased reply rate on social media. They propose a machine learning model of linguistic politeness that can work as a 'politeness checker'. Cheng et al [8] study the causes of trolling behavior in online forums and whether such impolite behavior is innate or situational. They conclude that the mood of the interlocutor and the discussion context affect the politeness of the conversation. Watts studies how we can model linguistic politeness effectively. [9].

In this work, we use the code made available by Mizil et al [1] for our politeness classification task on the FRIENDS corpus in order to identify characters with the most number of politeness characters and study whether such pretrained models for the wikipedia corpora can be generalized to real-life conversational data.

Approach and Experiments

Our primary objective is to leverage the politeness classifier [1] for the FRIENDS corpus and compute politeness scores for each character for a few seasons of the show. To this end, we first train the politeness classifier on the wikipedia corpus. We first annotate the corpus with politeness strategies by adding dependency parses and strategy information, as put forth in [1]. The politeness strategies for this corpus are shown in Figure 1.

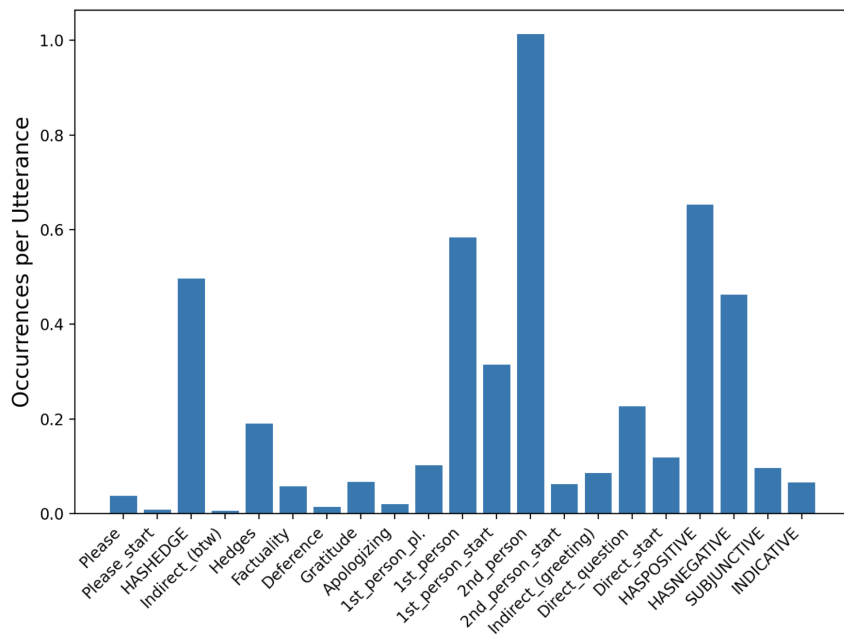


Figure 1: A glimpse of the overall use of politeness strategies in the wikipedia corpus. We annotate the wikipedia with such politeness strategies before training our politeness classifier on it.

We then subsequently use the ConvoKit politeness classifier and train it on the annotated wikipedia

corpus. In order to evaluate its training, we sample out a small subset of 100 utterances from the corpus which acts as our test dataset. We evaluate the performance of our trained classifier on this test set and are able to successfully predict politeness with an average accuracy of 74%, as shown in Figure 2. We can then use this pre-trained classifier and test it on the FRIENDS corpus to predict politeness for each character’s utterance. This has been further elaborated in our jupyter notebook [here](#).

	precision	recall	f1-score	support
False	0.70	0.76	0.73	46
True	0.78	0.72	0.75	54
accuracy			0.74	100
macro avg	0.74	0.74	0.74	100
weighted avg	0.74	0.74	0.74	100

Figure 2: Before testing the FRIENDS corpus, we first train the classifier for predicting politeness labels on the wikipedia dataset. We check these predictions on a test set (also from the wikipedia corpus) and obtain an accuracy of 74% as shown in this classification report.

Finding politeness labels for the FRIENDS corpus

Now that we have a pretrained model of the politeness classifier, we use this to find politeness labels for dialogues uttered by each character on FRIENDS. First, we only consider the first season of the TV show. We segregate out all dialogues from the first season and apply our trained politeness classifier to obtain the politeness scores for each utterance. From these results, we compute the number of polite and impolite utterances for each of the six individual characters. We compute a ‘politeness score’ as simply the percentage of polite utterances of the total number of utterances by that particular character. This has been further elucidated in our code.

$$\text{Politeness Score} = \frac{\text{number of polite utterances}}{\text{total number of utterances}} * 100 \quad (1)$$

Our results are shown in Figure 3. We can see that Rachel Green has the highest ratio of polite utterances in this season, with 50 polite statements out of 852. This does seem like a rather low number, as this would imply that the other $(852-50) = 802$ dialogues used by Rachel are impolite. However, due to the nature of the wikipedia corpus (which mostly consists of requests), the classifier is only trained to understand something as polite with specific politeness strategies. Hence, if a dialogue is not necessarily impolite, the classifier would still label it as impolite as it assumes it is being used in wikipedia’s context. Therefore, it would be smarter to think of impolite as *neither polite nor rude* and not as *impertinent*.

Specific Examples of Polite and Impolite Utterances

We now extract out the text of a few utterances in order to better comprehend what is and what isn’t labelled as polite.

	Name	Polite	Impolite	Total	Politeness Percentage
0	Monica Geller	32	850	882	3.63
1	Joey Tribbiani	26	617	643	4.04
2	Ross Geller	34	942	976	3.48
3	Rachel Green	50	802	852	5.87
4	Phoebe Buffay	29	624	653	4.44
5	Chandler Bing	30	804	834	3.6

Figure 3: The number of polite and impolite utterances by each character in the first season of FRIENDS. Rachel has the most polite dialogues, with 5.87% of her utterances being labelled as polite. It is important to note here that this does not imply that she is being impertinent or disrespectful the other 94% of the time. It simply means that she is noticeably and considerably very polite 5.87% of the time. Chandler has the least number of utterances that are polite.

1. *Daddy, I just... I can't marry him! I'm sorry. I just don't love him. Well, it matters to me!*

Speaker: Rachel Green

Result: polite, probability estimates = 0.5495772703902778

2. *Ross, could you come and help me with the spaghetti, please?*

Speaker: Monica Geller

Result: polite, probability estimates = 0.5055128748584624

3. *Phoebe, could you do me a favour? Could you try this on? I just wanna make sure it fits.*

Speaker: Joey Tribbiani

Result: polite, probability estimates = 0.5634086002071629

We can see that all of these utterances have been labelled as polite, which makes sense, as a major politeness strategy used by the classifier is to look for words such as *please* and *sorry*.

4. *Joey, stop hitting on her! It's her wedding day!*

Speaker: Monica Geller

Result: impolite, probability estimates = 0.08390555318208881

5. *No I don't, to hell with her, she left me!*

Speaker: Ross Geller

Result: impolite, probability estimates = 0.07222739997069622

6. *So Rachel, what're you, uh... what're you up to tonight?* Speaker: Ross Geller

Result: impolite, probability estimates = 0.06131959914245545

The three utterances above are examples of those that have been labelled as *impolite*. We can see, particularly from example 6, that what Ross is saying isn't exactly insolent or discourteous. It's just a classic example of everyday dialogue, where Ross is asking Rachel what her plans for that

night are. But it isn't exactly a *polite request* so the classifier labels it as *impolite*. As pointed out earlier, it would be better to think of *impolite* as *not a polite request*.

Politeness Labels from more seasons

Season 5

Furthermore, we also predict politeness labels for the characters from other seasons - season 5 and season 9. The code takes a long time to run and apply the politeness strategies to each individual utterance, hence the scope of this paper has been limited to just three seasons in the interest of time. However, one tangible and substantial result would be to analyze the trend of politeness labels of each character across different seasons and relate it to the social status and the character development in the show.

First, we apply our pretrained classifier on all utterances from season 5 and show our results in Figure 4. We can see that Rachel and Chandler, once again, have the highest and lowest ratio of polite dialogues respectively.

	Name	Polite	Impolite	Total	Politeness Percentage
0	Monica Geller	31	883	914	3.39
1	Joey Tribbiani	36	906	942	3.82
2	Ross Geller	44	826	870	5.06
3	Rachel Green	47	871	918	5.12
4	Phoebe Buffay	31	785	816	3.8
5	Chandler Bing	26	922	948	2.74

Figure 4: The number of polite and impolite utterances by each character in the fifth season of FRIENDS. Rachel triumphs again while Chandler's sarcasm causes his score to be the lowest of all.

Here are a few specific examples of Polite and Impolite Utterances from the fifth season.

7. *Hey! Joey, would you mind giving me and Ross a hand moving his couch?*

Speaker: Rachel Green

Result: polite, probability estimates = 0.5907886723768967

8. *! Good thing we have that, 'Not in New York' rule.*

Speaker: Chandler Bing

Result: polite, probability estimates = 0.6093129379637341

9. *I'm so sorry that you had to find out this way. I'm sorry, but it's true, I love him too.*

Speaker: Monica Geller

Result: polite, probability estimates = 0.6771867490987232

These are a few specific instances of utterances labelled as *impolite* from season 5.

10. *Yeah, I think you should keep looking!*

Speaker: Monica Geller

Result: impolite, probability estimates = 0.17651734781647002

11. *Ooh, Atlantic City! Oh, that's a great plan! Who's plan was that?*

Speaker: Phoebe Buffay

Result: impolite, probability estimates = 0.09146143134539647

As already discussed, this doesn't imply that these utterances are disrespectful. It's just that they are not considerably and formally polite.

Season 9

And finally, we use our pretrained classifier for predicting politeness labels for the penultimate season. The results (in the form of a pandas dataframe) are displayed in Figure 5.

	Name	Polite	Impolite	Total	Politeness Percentage
0	Monica Geller	22	887	909	2.42
1	Joey Tribbiani	41	823	864	4.75
2	Ross Geller	44	835	879	5.01
3	Rachel Green	38	866	904	4.2
4	Phoebe Buffay	41	759	800	5.12
5	Chandler Bing	47	871	918	5.12

Figure 5: The number of polite and impolite utterances by each character in the ninth season of FRIENDS. Phoebe and Chandler both have the highest politeness scores, while Monica has the lowest.

Here are a few specific examples, which one may relate to if they've watched this season of FRIENDS!

12. *I'm sorry. I, I told you I was in Tulsa because I wanted to spend the night with Monica and I, I didn't know . . . I didn't think you'd understand.*

Speaker: Chandler Bing

Result: polite, probability estimates = 0.5860504803214516

13. *I'm sorry. That's the one thing I can't do. I promised I'd be with Monica.*

Speaker: Chandler Bing

Result: polite, probability estimates = 0.5495772703902778

One might remember this hilarious scene when Ross orders everything from room service when he stays at a hotel with Chandler. He does so very politely with *please* and *thank you*!

14. *Hi, this is Ross Geller in suite 206. It seems you forgot a couple of things. Could you have some complimentary toiletries sent up to my room? Thank you! Ok. Toothbrush, toothpaste, razor,*

mouthwash, deodorant, dental floss, band aids, shaving cream, after shave... and I feel like I am forgetting something... Is there anything else you have that I haven't asked for already? Yeah, go ahead, send up some tampons.

Speaker: Ross Geller

Result: polite, probability estimates = 0.9585009163942574

Limitations and Issues

Our classifier is also prone to a few mistakes, where it may not be able to identify sarcasm or rhetorical comments. For instance, here is an utterance by Phoebe in season 1.

15. *Thanks. Okay, now go away so we can talk about you.*

Speaker: Phoebe Buffay

Result: polite, probability estimates = 0.6953336770291142

Phoebe is clearly being rude here and asking someone to go away so they can gossip about that person. But the classifier gives this a high politeness score, mostly due to the word *thanks*. This is a good example of why classifiers trained on wikipedia corpuses may not generalize very well to real-life conversational data. Here is another sarcastic example from Chandler, in season 5.

16. *I'm sorry, I thought you were Joey's other grandmother. I've done it again.*

Speaker: Chandler Bing

Result: polite, probability estimates = 0.5860504803214516

And another one from Ross when he was presenting his work at the conference in Barbados!

17. *Yep! I'd like to thank you guys for coming down here to complain about the rain and ruin my career!*

Speaker: Ross Geller

Result: polite, probability estimates = 0.8710572795917456

Hence, we can see that the politeness classifier is unable to accurately label such utterances. The best way to combat this issue is to train it for real-life conversational data, such as the FRIENDS corpus, by manually annotating and identifying apt politeness strategies, which would be different from those found on wikipedia and reddit corpora.

Relationship between politeness and social status

Mizil et al [1] also use their politeness computational framework to analyze relationships between politeness and social power and obtain a negative correlation between politeness and power on Wikipedia. We can notice from our results that in the first season, Rachel has a very high politeness score, which is when she used to work at the Central Perk coffee shop. In the penultimate season, when Rachel has risen and is working as a buyer at Bloomingdale's, her politeness score falls down to 4.2. This is also noticed with Monica's scores, where her politeness scores fall from 3.63 in season 1 to 2.42 in season 9, after she has been promoted multiple times to the head chef of her restaurant.

Similarly, we notice that Chandler's scores in the first season are very low, which is when he had a reputable job in *statistical analysis and data reconfiguration*. However, he later moves to

advertising, and his scores rise up. All these instances may potentially also indicate a negative correlation between politeness and social status, but this is inconclusive due to the lack of data. This would be better comprehensible with politeness scores for all the seasons, which is a promising future direction for this study.

Conclusion and Future Work

In this work, we leverage the politeness classifier formulated in [1] to obtain politeness features and markers of the FRIENDS corpus and subsequently label each character’s utterance as polite or impolite. We compute the politeness scores for all characters for three seasons and successfully identify the most polite and impolite character for that particular season. We also look at specific examples from the corpus and their predicted labels to qualitatively analyze the performance of the classifier and identify limitations.

This work only considers three seasons in the interest of time as the code is computationally expensive for the FRIENDS corpus, which consists of over 67,000 utterances. One future direction for this work is to compute the politeness labels for all the characters in all the seasons and plot these scores on a graph to identify any trends. This would also help us further concretely analyze the relationship between social status and politeness and whether there exists any negative correlation.

We also identify certain limitations of this classifier, where it incorrectly classifies sarcastic comments as *polite*. This shows the inability of such classifiers which are pretrained on online forum corpora to generalize to real-life conversational datasets. Hence, a promising future direction for this study is to annotate the FRIENDS corpus with politeness strategies and then train the classifier, which would potentially yield stronger results. But overall, our classifier is able to identify most of the polite utterances and give us an excellent idea of all the characters on the show and their politeness scores across multiple seasons.

References

- [1] Cristian Danescu-Niculescu-Mizil et al. “A computational approach to politeness with application to social factors”. In: *Proceedings of ACL*. 2013.
- [2] Justine Zhang et al. “Conversations Gone Awry: Detecting Early Signs of Conversational Failure”. In: *Proceedings of ACL*. 2018.
- [3] Michael Yeomans, Alejandro Kantor, and Dustin Tingley. “The politeness Package: Detecting Politeness in Natural Language”. In: *R J.* 10 (2018), p. 489.
- [4] Malika Aubakirova and Mohit Bansal. “Interpreting Neural Networks to Improve Politeness Comprehension”. In: *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*. Austin, Texas: Association for Computational Linguistics, Nov. 2016, pp. 2035–2041. DOI: [10.18653/v1/D16-1216](https://doi.org/10.18653/v1/D16-1216). URL: <https://aclanthology.org/D16-1216>.
- [5] Erin R. Hoffman, David W. McDonald, and Mark Zachry. “Evaluating a Computational Approach to Labeling Politeness: Challenges for the Application of Machine Classification to Social Computing Data”. In: *Proc. ACM Hum.-Comput. Interact.* 1.CSCW (Dec. 2017). DOI: [10.1145/3134687](https://doi.org/10.1145/3134687). URL: <https://doi.org/10.1145/3134687>.
- [6] Herbert H. Clark and Dale Schunk. “Polite responses to polite requests”. In: *Cognition* 8 (1980), pp. 111–143.

- [7] Moira Burke and Robert Kraut. “Mind Your Ps and Qs: The Impact of Politeness and Rudeness in Online Communities”. In: *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*. CSCW '08. San Diego, CA, USA: Association for Computing Machinery, 2008, pp. 281–284. ISBN: 9781605580074. DOI: [10.1145/1460563.1460609](https://doi.org/10.1145/1460563.1460609). URL: <https://doi.org/10.1145/1460563.1460609>.
- [8] Justin Cheng et al. “Anyone Can Become a Troll: Causes of Trolling Behavior in Online Discussions”. In: *Proceedings of CSCW*. 2017.
- [9] Richard J. Watts. *Politeness*. Key Topics in Sociolinguistics. Cambridge University Press, 2003. DOI: [10.1017/CB09780511615184](https://doi.org/10.1017/CB09780511615184).