

# Literature survey

## LNS project

Team – papa ka para

Aditya Raghuvanshi & Utsav Shekhar

## The politeness Package: Detecting Politeness in Natural Language by Michael Yeomans, Alejandro Kantor, Dustin Tingley

Politeness is a universal dimension of human communication, The politeness package is designed to make it easier to detect politeness in English natural language, by quantifying relevant characteristics of polite speech, and comparing them to other data about the speaker, For example, researchers may want to know whether politeness is associated with some situational or trait-level covariate (text as description). Or researchers might want to know whether people respond differently to polite rather than impolite language (text as treatment).

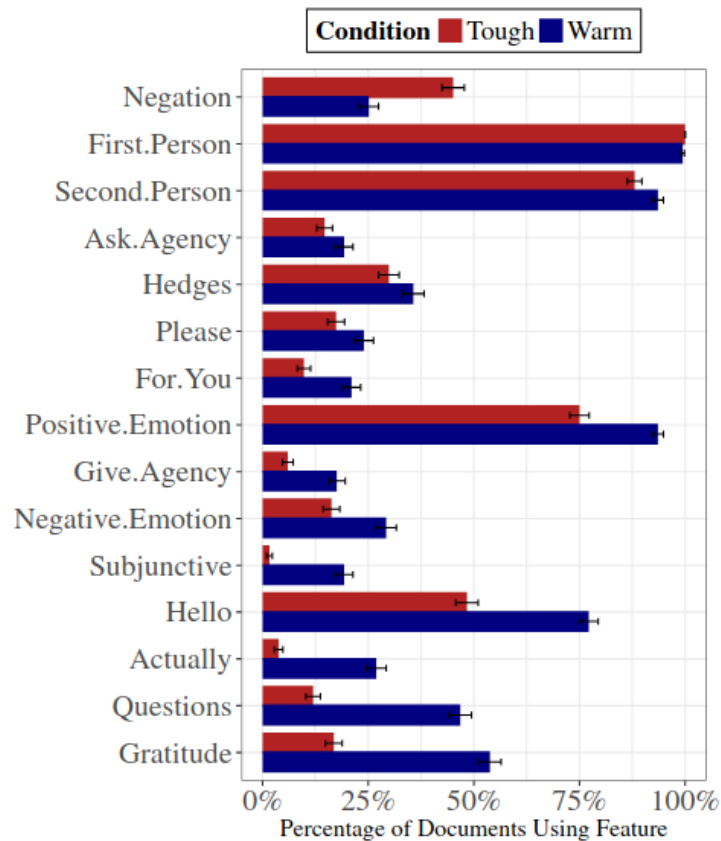
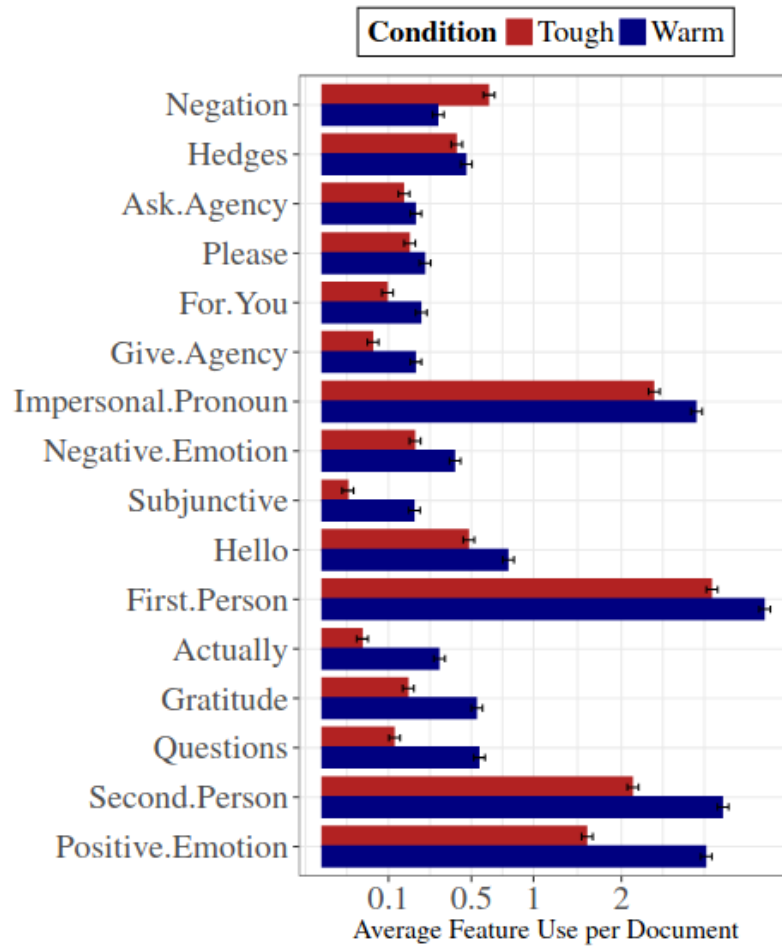
### Politeness workflow

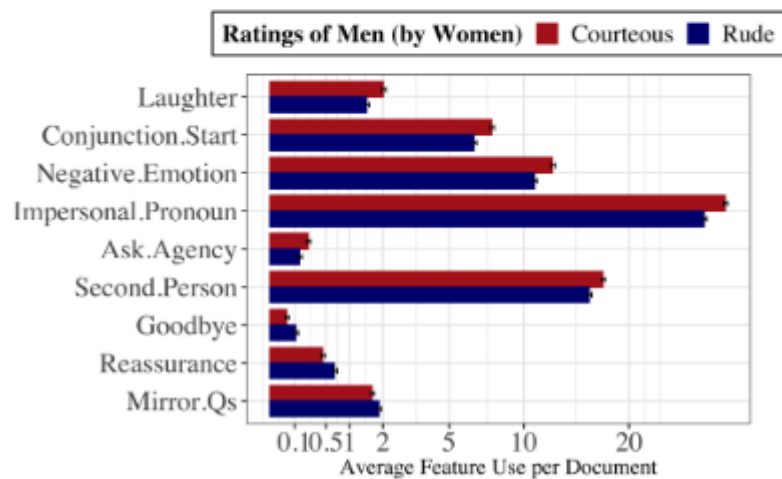
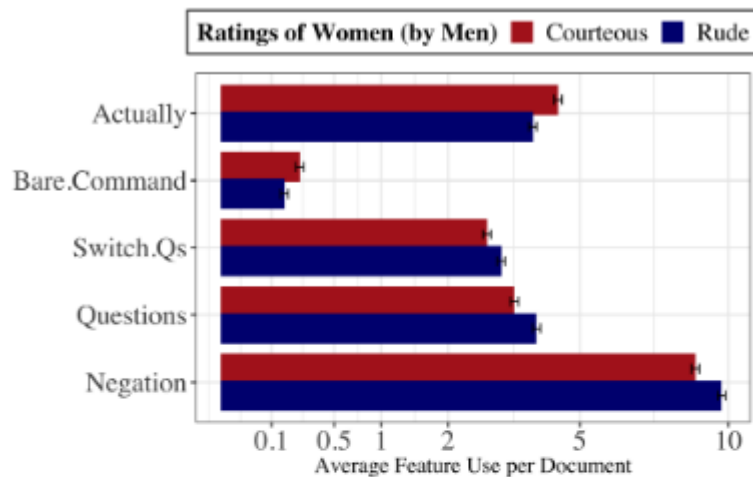
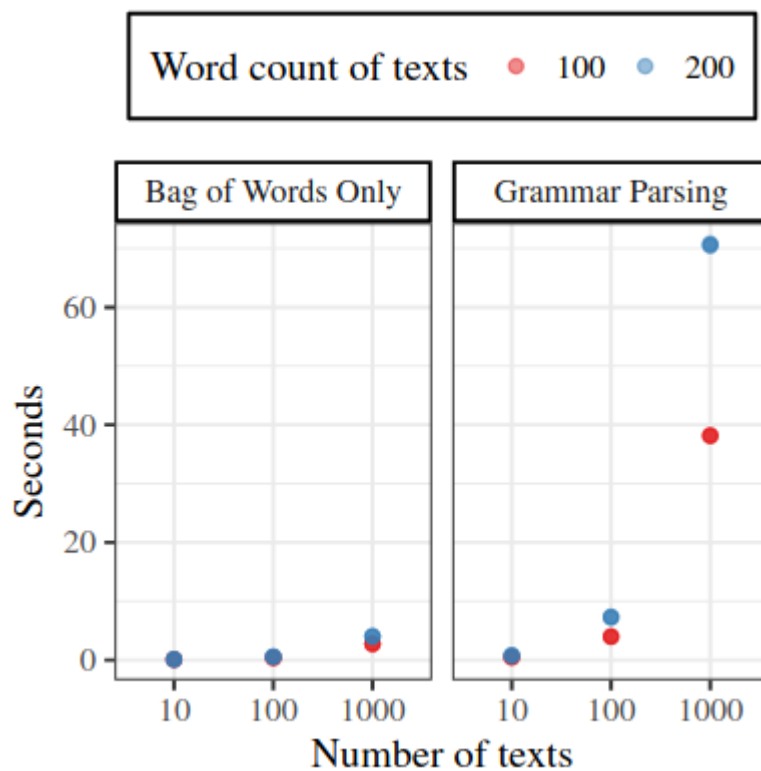
The politeness package provides functions to identify politeness markers in natural language, graphically compare these to covariates of interest, develop a supervised model to detect politeness in new documents, and inspect high- and low-politeness documents. Table 1 summarizes the main functions of the polite

Detecting politeness features The function `politeness()` takes in an `n`-length vector of texts, and returns an `n`-by-`f` data.frame, with `f` columns corresponding to the number of calculated features. We note that our package does not perform any spell- or grammar-checking on the text- though these kinds of errors can degrade the fidelity of the information in the text. Instead, we recommend that users do this on their own- either in R, using another package like `hunspell`, or else using other software or by hand. There are 36 features in total (see Table 2), but some user options affect the number of features that are returned. For one, if a part-of-speech tagger is not used (by setting `parser = "none"`) then some features that depend on these tags will not be calculated, as detailed in Table 2. Additionally, you may use the default setting `drop_blank =`

TRUE which will remove features that were not detected in any of the texts. For example, consider the following texts. `library(politeness) texts <- c("Hello, you", "You") politeness()` with `parser="none"` (default) will identify that both texts contain the feature `Second.Person` and that only the first contains `Hello`.

Plotting politeness features `politenessPlot()` combines the politeness feature matrix with another measure of substantive interest that might covary with the prevalence of some of the politeness features- in particular, a ground-truth measure of politeness from annotation or assignment to treatment. This function generates a horizontal bar plot (with 95% confidence intervals) of feature prevalence among two groups of documents. The function can handle any kind of politeness metric (counts, binary, or averaged), but that distinction must be made in the initial call to `politeness()`. Below, we plot the politeness features from a feature count matrix in Figure 1, and a binary feature matrix in Figure 2. The order of the bars themselves are sorted automatically, and determined by calculating the variance-weighted log odds of each feature with respect to the binary covariate. Many covariates are continuous. By default, this package treats the top and bottom terciles of that distribution as binary categories, with the middle tercile dropped. Users can also create their own categories beforehand, and enter those labels in place of the covariate. Often some features are not meaningful for further analysis – either because they are too rare in the data, or else because they do not meaningfully covary with the covariate of interest. Users have two options to exclude these from the plot. First, the `drop_blank` parameter can remove rare features – it takes a number between 0 and 1, which determines a cut-off based on prevalence. Specifically, all features which appear in less than this proportion of texts are excluded from the plot. To include all features, leave this value at 0.





## The romance of Urdu is lost in translation.

This was an article we found online in which the author was referring to the elements of romance and sootherness that Urdu have. The author regrets that he is unaware of the beautiful language of Urdu and not able to comprehend it as the local people do. His floral descriptions of the Lucknow of his childhood only intensify my regret. Apparently, even the tongawalas of Lucknow spoke polished and courteous Urdu.

Examples of the same in other languages :

1. Listening to The Magic Flute, he says, is a completely different experience if you know German
2. reading War and Peace in the original Russian offers nuances that are lost in the English translation.
3. America's exuberance, for instance, isn't quite captured in English.
4. Punjabi enterprise and risk-taking isn't fully expressed in the crude bravado of its language.
5. Mumbaiiya Hindi captures the capitalist matlabi Mumbaikar but fails to represent Mumbai's generous spirit that swells, it seems, with every monsoon flood and natural catastrophe that hits this city.
6. Italian barely conveys the depths of Italy's contribution to the arts, particularly opera.

But ....

Urdu... the cadences and sighs of this beautiful language seem to encompass the courtly graces of Islamic-Indian culture which flourished in Awadh. Urdu is the language of romance; it is the language of poetry, of diplomacy.

The way in which author describes what he feels about Urdu

"I can't help think that if Juliet had uttered her "Good Night" and "Parting is such sweet sorrow" in Urdu, the whole Montague-Capulet quarrel would have been sorted out. When my untutored ear listens to Urdu shairi, what I enjoy is the relish with which the poem is told and the wah wahs with which it is appreciated. When I listen to a Farida Khanum ghazal, I don't understand the

lyrics, but my heart hears the longing. The Mughals were many things but they were, above all, bon vivants. They epitomized the title of my column. They lived the good life by surrounding themselves with decorative things, aromatic cuisine, soulful music and heartfelt poetry. All of this is encompassed in the exquisite nuances of their language.”

Like Bengalis, who take pride in their language, lovers of Urdu are uniformly fanatical, believing Urdu to be the mother lode, the fairest language of them all.

Citations - <https://www.livemint.com/Opinion/B9PWW3liEjY59D4Y1IGYkL/The-romance-of-Urdu-is-lost-in-translation.html>

## Politeness in language

According to this paper “Linguistic politeness can be defined as the ways in which language is employed in conversation to show consideration for the feelings and desires of one’s interlocutors, to create and uphold interpersonal relationships (so-called politic behavior), and to comply with the rules for what society or one’s culture considers appropriate behavior.”

Their initial research was based on contemporary theoretical pragmatics and sociology and on the cooperative principle, speech act theory, and the notion of face in particular. Their research challenged theories like they described a universal feature of languages and cultures, as well as their ideas to capture politeness.

### Proposed alternatives

The Discursive method, undoubtedly the most popular alternative, emphasises how members of the studied group use the term in conversation and views politeness as something that is formed via speech.

A lot of focus is being placed on language issues that are generally covered in the literature on politeness in conversation analysis. Some of the themes covered by contemporary study into politeness date back decades, despite the radical changes the subject as a whole has seen and the ongoing discussion. Youngsters learning to be polite through language, cross-cultural pragmatics, conceptualising the idea of a face, honorifics, and the interaction of gender with politeness are a few examples. Impoliteness, the function of politeness in

language evolution, and diachronic changes to politeness cultures and forms itself are subjects that have only lately or sporadically been studied.

Reference - <https://www.oxfordbibliographies.com/view/document/obo-9780199772810/obo-9780199772810-0161.xml#:~:text=Linguistic%20politeness%20can%20be%20defined,what%20society%20or%20one's%20culture>