Analysis in R: Partition

Abhilasha Sawlani, Akshar Katariya, Jahnavi Selvakumar

08/03/2021

## Title: Reading Beyond the (Radcliffe) Line: The Syncretic Worlds of Partition Novels

### Introduction and Research Question

Why did the Partition of the Indian subcontinent take place? Historians and literary scholars alike have struggled with this question for over seven decades, focussing on the high politics of big players like the British, the Congress and the Muslim League. One of the commonly held assumptions is that religious differences played a key role in fomenting communal sentiments, leading to widespread riots in the years leading up to Partition. This project aims to reorient the scholarly gaze by focussing on literary texts — mainstream political machinations lie at the margins of these works as they focus on the lived experience of Partition and everyday negotiation of religious differences. What might these texts convey about the sentiments associated with various religions in the popular consciousness? What kinds of collectives (India, Pakistan, Bangladesh or a more fluid notion of watan) are perceived positively or negatively? In addition, sentiment analysis might also point us towards other causal factors that, in the perception of the common people populating these novels, precipitated the Partition.

### Research Hypothesis

Partition literature presents a syncretic worldview quite different from the anti-Muslim rhetoric that permeates discussions about the reasons and consequences of Partition today. Consequently, it can be argued that irreconcilable religious differences were not the primary cause for Partition.

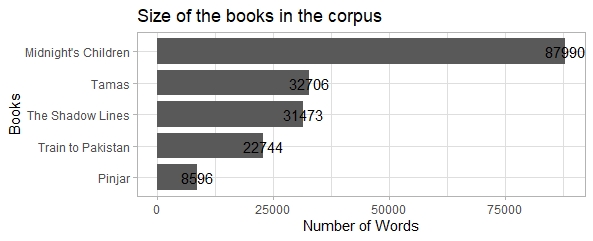
To prove this hypothesis, we will attempt to examine the sentiments associated with popular markers of Muslimness in particular. Tangentially, we will attempt to explore who might have been deemed the villain of the Partition story in the popular consciousness — in other words, which powerful entities are surrounded with negative sentiments in the novels. In the process, we will also attempt to understand the sentiments surrounding the ideas of nationhood, Partition and freedom.

### Corpus Description

We selected five novels about the Partition written by authors of Indian origin. These include — Bhisham Sahni’s *Tamas* (1973), Khushwant Singh’s *Train to Pakistan* (1956), Amrita Pritam’s *Pinjar* (1950), Amitav Ghosh’s *The Shadow Lines* (1988), and Salman Rushdie’s *Midnight’s Children* (1981). Since these novels are temporally dispersed, they might allow us a glimpse of the changing trends in the genre. Moreover, the writers hail from different religious backgrounds which might allow us a more balanced perspective on the religious politics depicted in their works. A possible shortcoming of the selected corpus could be that two of the texts — *Tamas* and *Pinjar* — appear in translation, possibly diluting the meanings found in the original texts. The Urdu word ‘watan’, for instance, carries a different emotional valence than its English counterparts, ‘nation’ or ‘country’.

### Summary Paragraph

The top ten words in the words are all common nouns with the exception of the word ‘Singh’(6th most frequent). This could be due to the fact that the Partition and its notable associated events took place near Punjab where ‘Singh’ is a common surname and a middle name. The word ‘mother’(2nd most frequent) appears more frequently than ‘father’(10th most frequent). This difference can be accredited to the prevalence of chants like ‘Bharat Mata’ which translate into ‘Mother India’ These chants were especially prevalent during the mid 1940s owing to the heights of the independence movement during that period. The topic of Partition was key in that movement. Within the corpus, there are visually significant outliers with respect to the length of these books. *Midnight’s Children* has the highest number of words and is 3.7 times the length of *Tamas* — the second longest book in our corpus. *Pinjar* has the shortest length and is 2.6 times shorter than *Train to Pakistan* — the second shortest book. This variation can induce a bias which might skew our data, and, with that, the analysis following from the data. To account for this, we have looked at the intra-corpus data as well to check for the ‘influence’ of those outliers and have taken it into consideration for our inferences.



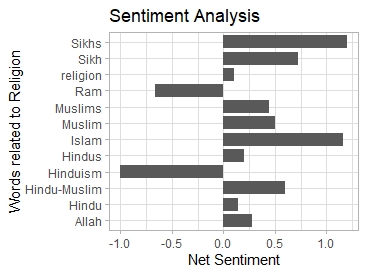
We see a moderate correlation between the books in our corpus. Even the texts with the highest temporal gap between the years of publication, namely, 32 years between *Train to Pakistan* and *The Shadow Lines* yield a correlation of 0.49 with statistical significance. The majority of the correlations between the books are in the range of 0.4 - 0.65 owing to the similarity in the plots of these books as they are all based on the Partition. The correlation can also be attributed to the common nationality of the authors, and the short range of the years of publication between the books.

### Data Visualization 1

#### Sentiment Analysis for Word Clusters

The ‘two nation theory’ that motivated the Partition of the subcontinent put forward the notion that Hindus and Muslims constituted two exclusive, irreconcilable communities that could not coexist. Such a theory ran counter to the lived experience of centuries of coexistence where communities intermingled in relationships fluctuating between amity and occasional hostility. To understand the nature of Hindu-Muslim relations depicted in these novels, we manipulated the NER sentiment analysis code to inspect the nature of sentiments surrounding particular clusters of words. These words were segregated into different categories like religion, nations, language, space and actors.

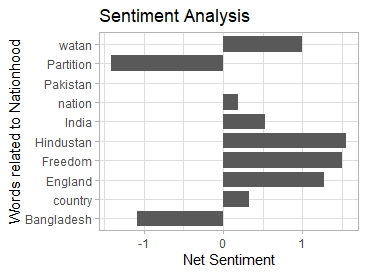
Under the religion category, words like Sikh/Sikhs, Muslim/Muslims, Hindu/Hindus have a net positive sentiment. In addition, the words Allah and Islam are also surrounded by significantly positive sentiment. The hyphenated word Hindu-Muslim also has positive sentiment, perhaps gesturing towards phrases like Hindu-Muslim harmony or unity, often invoked by secular groups in the context preceding Partition. The positive sentiment around this hyphenated word also gestures towards the hyphenated identities in pre-Partition India. On the other hand, the sentiment around the word religion, albeit positive, is significantly lower (0.1) when compared with clusters like Muslim/Muslims(0.5, 0.44), Sikh/Sikhs(0.72, 1.2). The word Hinduism, in particular, has a net negative sentiment (-1).



This might suggest that while religious groups or individuals in particular were deemed in a positive light, the increasing institutionalization of religion was perceived negatively. This institutionalization was furthered by the activities of reformist movements like the Arya Samaj that wished to purge Hinduism of all “alien” or Islamic influences. In a similar vein, the word Ram (-0.6) is surrounded by negative sentiment, perhaps due to the invocation of chants such as “Jai Shri Ram” during scenes of mob violence or “Hey Ram” during moments of despondence (The radar plot for the word inches significantly towards disgust).

The analysis of word clusters associated with religious spaces yields a similar insight. The exclusive religious spaces such as temple and mosque are surrounded by negative sentiment whereas the word tomb is associated with positive sentiment. Tombs or dargahs or shrines dedicated to Muslim (often Sufi) saints were syncretic spaces visited by both communities; the positive sentiment associated with them might gesture towards their inviting, peaceful status as opposed to spaces like temples or mosques that were often the primary sites of Partition violence. In the radar plot, the word tomb greatly inclines towards joy. On the other hand, the sentiment analysis of language words (Hindi, Urdu, Hindustani, Bengali) did not yield significant insights; the word Hindustani has positive sentiments surrounding it but it remains ambiguous in meaning — it could refer either to the language that syncretically combined Hindi and Urdu or to the collective identity of the people of the subcontinent.

Under the category of nations, the word watan was found to have a greater positive sentiment (1) as compared to nation (0.18) or country (0.32). As opposed to the more territorially-defined idea of the nation, the Urdu word ‘watan’ is supposed to refer to an affective connection with the land in Partition literature, and does not always intersect with the idea of religious nationalism. The words India (0.5) and Hindustan (1.5) have a positive sentiment while the net sentiment around Pakistan is 0, suggesting perhaps the equal presence of proponents of both pro and anti-Pakistan sentiments (The radar plot for Pakistan inches significantly towards disgust). The word Hindustan, as scholars like Manan Ahmed Asif have suggested, initially referred to a more inclusive conception of collective identity than the word India [[1]](#footnote-29); the high positive sentiment around Hindustan, coupled with the significantly negative sentiment around the word Partition (-1.4) suggests the prioritization of a free (freedom - 1.5) but unified collective in the novels.

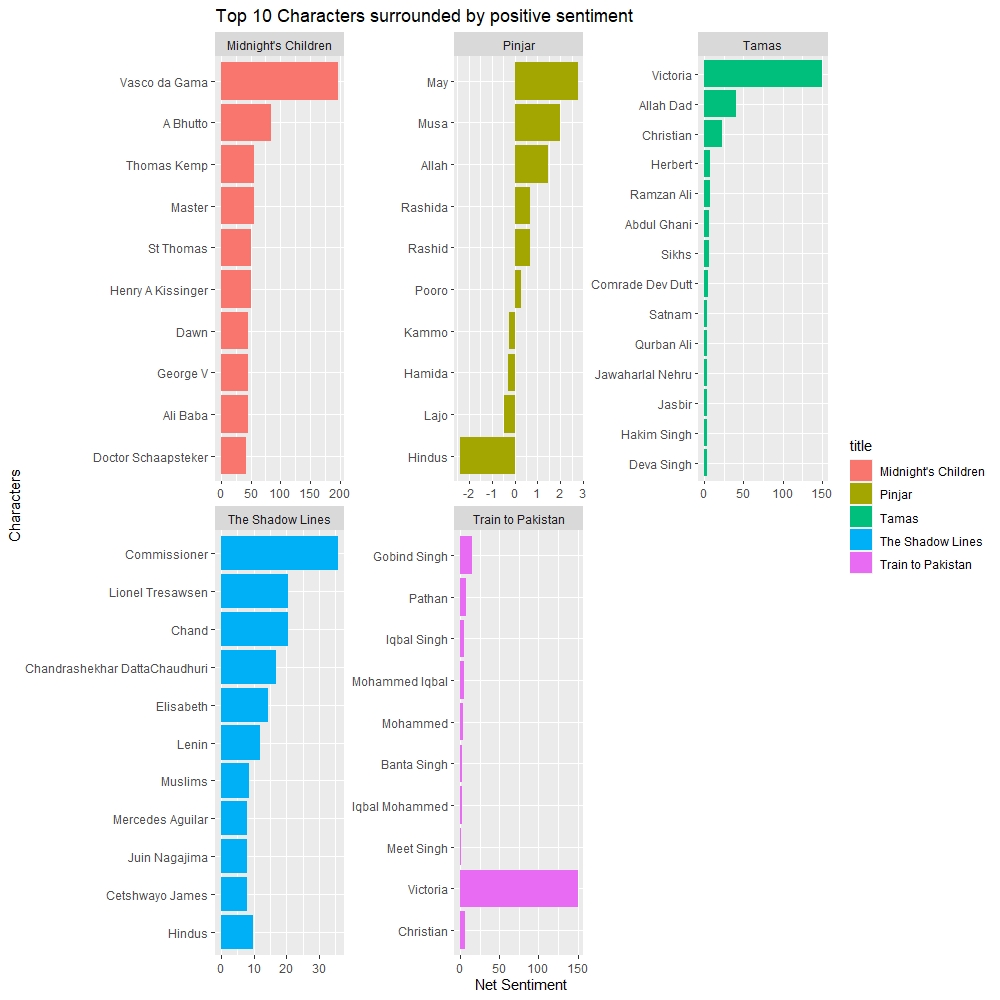


That assumption seems to be challenged by the graph plotting the sentiments around major political actors during the time of Partition. The graph points towards a significantly positive sentiment around entities like the Muslim League (3.03, as opposed to the 0.56 for Congress), Muhammad Ali Jinnah (4) and Mountbatten (1.42), popularly identified with the Partition plan. Mahatma Gandhi and Jawaharlal Nehru, on the other hand, are markedly surrounded by negative sentiment (-0.50 and -0.9 respectively). In the absence of the granular context in which conversations about these entities take place in the novels (who is leading these conversations and when?), it becomes difficult to surmise the dominant perception about them. The negative sentiment around Nehru and Gandhi, however, may be explained by the predominant sense of betrayal (for accepting the Partition plan) that these figures came to be associated with in the popular mind in the aftermath of Partition.

### Data Visualization 2

#### NER

Next, we ran the named entity recognition analysis to examine which characters and locations in the novels were surrounded by positive or negative sentiments. Based on the hypothesized syncretic nature of these texts, we assumed that characters with a Muslim name would be surrounded by positive sentiments. That assumption was proven correct when we plotted the list of top ten characters surrounded by positive sentiments in the novels — several of them had evidently Muslim names such as Ali Baba (45) in *Midnight’s Children*; Ramzan Ali (7), Qurban Ali (4) and Abdul Ghani (6) in *Tamas*. Mythical and religious Muslim figures like Allah (1.45) and Musa (2) were also surrounded by positive sentiment in *Pinjar*.



The sentiment analysis for characters in *Pinjar* revealed an interesting insight about abduction and forced religious conversion of women during Partition. Pooro, a Hindu woman surrounded by positive sentiments (0.28) is abducted by Rashid and thenceforth referred to as Hamida which is surrounded by negative sentiments (-0.28). Her abductor Rashid, depicted as a well-meaning and caring man in the novel, is surrounded by positive sentiments (0.66) suggesting that the negativity associated with Hamida is owing to the forced nature of her conversion rather than a negativity surrounding Muslimness itself. This is further testified by the negative sentiments surrounding Lajo (-0.46), another abducted Hindu woman. Perhaps it is the institutionalization of otherwise lived and hybrid religious practices and an imposition of that religion onto individual bodies (particularly that of women) that is perceived in a negative light in the novels.

The limitation of our approach became evident in the sentiment analysis graph for characters in *Train to Pakistan*. Four similar sounding character names (all with positive sentiments) appeared — Iqbal Singh, Mohammed Iqbal, Mohammed and Iqbal Mohammed. Sensing another narrative of religious conversion, we searched for those names in the corpus and came across the following sentence:

“He did not have to say what Iqbal he was. He could be a Muslim, Iqbal Mohammed. He could be a Hindu, Iqbal Chand, or a Sikh, Iqbal Singh. It was one of the few names common to the three communities. In a Sikh village, an Iqbal Singh would no doubt get a better deal, even if his hair was shorn and his beard shaved, than an Iqbal Mohammed or an Iqbal Chand. He himself had few religious feelings.”

Given the ambiguity surrounding the religion of a non-religious man, it would be erroneous to infer the religious affiliation of a character based on their name alone. A character’s visual appearance (is their hair long like a Sikh or shorn like a Muslim?) influences the way they are perceived by other characters in the novel and thereby the sentiments surrounding them. The absence of a granular understanding of context makes all cursory inferences myopic. In addition, several Partition novels make a distinction between the peaceful residents of the syncretic home (village) and the rioters — the outsiders who always come from elsewhere. This explains why quite a few Muslim characters can be found even in the graph plotting the characters with the highest negative sentiments. In short, any inference based only on the sentiments surrounding characters belonging to a particular religion would be insufficient (without contextual knowledge) to indicate the nature of the social world depicted in the novel — syncretic or exclusionary.

The graphs plotting sentiments around locations were similarly inconclusive due to intra-corpus variations. The sentiment around Lahore (-2.5) is negative in *Pinjar* and *Midnight’s Children* (-12.66) and positive in *Tamas* (23) and *Train to Pakistan* (0.50). The sentiment around Punjab is negative in *Pinjar* (-16) and positive in *Train to Pakistan* (11). The sentiment around India is positive throughout the corpus and that around Pakistan positive in all the books except *Tamas*. Any meaningful inference based on sentiments surrounding a particular location would require contextual information; the narrative space accorded to moments of violence or moments of pre-Partition harmony rooted in a locale vary across the titles thereby influencing the sentiments associated with particular locations. Our initial assumption that spaces across the border would be imbued with positive sentiment due to pre-Partition syncretism could not be sufficiently validated using NER.

### Conclusion

Sentiment analysis conducted on a cluster of words enabled us to prove our hypothesis partly. While particular individuals belonging to a religion (Muslim/Muslims, Sikh/Sikhs, Hindu/Hindus) are associated with a positive sentiment, institutionalized and exclusionary forms of religion (Hinduism, mosques and temples) are perceived in a negative light in these novels. This is further borne out by the discussion about religious conversion in the section covering named entity recognition.

The existence of syncretic social worlds in the novels could not be conclusively established owing to the absence of contextual information about everyday inter-religious interactions. How much narrative space is accorded to such interactions as opposed to the violence of Partition, though it greatly influenced our data, also remained unclear. The absence of this granular perspective further prevented us from arriving at a clear understanding of how dominant political actors of the time were perceived. Novels, by their nature, are polyphonic entities, often staging a dialogue between diverse discourses; consequently, it is difficult to make certain inferences without taking sufficient cognizance of this complexity.

### Reflection

While Voyant allowed us to glean the frequency of occurrence of particular words and word clusters, sentiment analysis and NER enabled us to determine, to a certain extent, the meanings that particular words and entities might be imbued with in a text. The sentiment found around those entities gestures towards the ways in which they are treated in a text — whether they are perceived by other characters (or by the narrator or author) positively or negatively. In other words, as compared with Voyant, sentiment analysis takes into account the contexts in which words appear rather than just the frequency of their appearance.

However, without a close reading of the texts, a nuanced recording of the contexts in which conversations occur and characters interact, even sentiment analysis only offers a partial understanding of phenomena in the text. Moreover, such an analysis does not account for differences in the formal structure of narratives — the narratorial consciousness, the perspective from which the story is told and that might influence the sentiments found around particular entities. It also doesn’t account for the treatment of time and thereby fails to yield insights about how and why the sentiments around particular characters or events might change over time. Do the sentiments around Muslim or Hindu characters change as the narrative draws closer to Partition? At what point does religious harmony transform into communal conflict? How and where do characters from different religious groups interact in the texts? Questions such as these require human intervention and cannot be answered adequately using the bird’s eye view that sentiment analysis provides.

1. Asif, Manan Ahmed.*The Loss of Hindustan: The Invention of India.* Harvard University Press, 2020. [↑](#footnote-ref-29)