

## **A Century of Media Representations of Muslims and Chinese Minorities in the Philippines (1870s-1970s)**

Cruz, Frances Antoinette (University of Antwerp, University of the Philippines Diliman)

Ortuño Casanova, Rocío (University of Antwerp, Universidad de Alcalá de Henares)

### Abstract

The postcolonial era has often been characterized by consolidating multi-ethnic and multi-religious groups brought together under colonial regimes. In the Philippines, these struggles continue to be seen in the ambivalent treatment of ethnic or religious minorities such as Chinese and Muslims or Moros, particularly in nation-building. This issue mirrors situations in other former colonies, which ended up excluding some ethnic-cultural groups from the national identity – such as in Cuba, where nationalist discourse left out the significant Chinese population (López-Calvo 2009, 18). Our hypothesis is that differences between regime and language, and how different regimes failed to result in significant changes of how minorities were represented in discourse before 1972, paved the way to ongoing and unresolved issues of national belonging, lending methodological possibilities and further empirical support for the analysis of identities over time in large corpora.

One of the key differences in colonial treatment of the Indio on the one hand, and the Chinese and Moro on the other, lies mainly in the intersection of religious conversion and (proto)racial ontologies of the Spanish colonists. Spanish representations centered on a ‘bifurcated image’ of the Indios as either “simple children of nature who would be receptive to tutelage in civilization and Christianity” (Fredrickson 2002, 36), or outrightly hostile, and therefore worthy of violent countermeasures. Whereas the Moros were treated as Muslims, who in Europe were believed to have “been exposed to the gospel and rejected it” (ibid., 37) they were also treated as hostile Indios. On the other hand, the Chinese were tolerated as both foreign as well as ““an entrepreneurial minority” – the kind of group that is likely to be deeply resented and readily turned into a scapegoat when conditions are unstable and times are hard.” (ibid., 92).

The present study thus utilizes print media from three time periods – the Spanish colonial period (1872 - December 10 1898), the American colonial period (December 11 1898 – July 4 1946), and Independence (July 5 1946-1972). The total corpus consists of 5,397,708 words, with the English, Filipino and Spanish subcorpora comprising roughly 37.39%, 10.54% and 52.07% of the total. Timewise, the texts from the Spanish Colonial Era, the American Colonial Era and Independence comprise of 10.74%, 67.75% and 21.50% respectively.

We apply techniques such as word frequency extraction, collocation analysis and word embeddings, supporting them with close readings of relevant excerpts identified by quantitative tools to determine time-bound and language-bound semantic fields associated with Muslims and Chinese in the Philippines. We further make use of word embeddings to

provide a ‘macro’ lens for identifying common themes in the corpora before closely examining term occurrences contextually. Here, we apply Word2Vec (Mikolov et al. 2013), which involves the identification of semantically-similar words through the generation of a multi-dimensional vector space from a large corpus of text. In order to visualize the data, this paper makes use of t-distributed stochastic neighbor embedding (t-SNE) (van der Maaten and Hinton 2008), a machine learning algorithm, which reduces the results gathered from Word2Vec into two or three dimensions.

We first performed a layout analysis on the texts using Transkribus’ print block detection module (Kahle et al. 2017) and OCRed the texts using a customized text recognition model prior to pre-processing. The model was based on a training set of 157 manually annotated pages that included samples from each of the target languages. The character error rate for the model used in the study is 1.94% (training set) and 1.87% (validation set).

Due to the presence of multilingual periodicals, a combined stop word list was built from the main languages in the study (Filipino, English, and Spanish) and applied to the dataset. Subsequently, the appearance of words related to identities was studied through examining word frequencies and word collocations appearing in each sub-corpus through Gephi and AntConc (Anthony 2019). The collocations were identified on the basis of the mutual information score that measures the probability of a key word occurring next to a collocate, relative to the number of times each word occurs in total (Stubbs 1995). As mutual information scores tend to be higher for infrequently occurring collocates, the frequency floor for this corpus was set to three occurrences and above. Due to the multilingual nature of the corpus, collocations were identified for all possible morphological forms of the target word across languages.

Due to the corpus size, all time periods and languages were entered as input for the measurement of similarity in Word2Vec. For ‘Moro/s’, associations included notably negatively charged terms, such as ‘outlaws’ and ‘insurgents’, but also words with implications along the binary dimensions of savagery/civilization and rural/urban. The results for the Chinese terms revealed associations with other nationalities, professions ‘campesino’ (farmer), ‘arquitecto’ (architect), and otherness (‘dayuhan’ (foreigner), ‘kani(-)kanilang’ (their own)).

Both the quantitative and qualitative approaches revealed that while different discourses developed with regard to Moros and Chinese in the Philippines, their signifiers retained a relatively stable core with shifting meanings depending on the nature of the political regime.

Particularly striking about the Spanish corpus was their treatment of the Chinese with regard to the latter’s sincerity in religion, while discourse on Moros was strongly militarized. During the American Period, citizenship was recognized based on a binary of Filipino vs. alien (Chu 2021) which implied both a legal and ethnic form of othering. While Moros were assumed to be citizens of the Philippines, the American objective of socializing Moros into their administration only casted Moros as ‘good’ if they were compliant with American directives.

The independence period presented an acceptance of legal statuses of both groups but did not significantly progress beyond stereotyped attributes. The appropriation of selected features of

an American form of ‘official’ Filipino nationalism during independence, was the result of a “widespread educative process [in the Philippines] [...] that changed the public perception of things, events, peoples and places” (Quibuyen 2008, 378), including minorities. As far as the representation of minorities living in the Philippines goes, visions that transcend the colonial character of representation are only beginning to scratch the surface of possibility.

## References

- Anthony, Laurence. 2019. *AntConc* (Version 3.5.8). Tokyo, Japan: Waseda University. <https://www.laurenceanthony.net/software>.
- Fredrickson, George M. 2002. *Racism: A Short Introduction*. Princeton: Princeton University Press.
- Kahle, Philip; Colutto, Sebastian; Hackl, Günter; and Günter Mühlberger. 2017. *Transkribus - A Service Platform for Transcription, Recognition and Retrieval of Historical Documents*. 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), pp. 19–24. <https://doi.org/10.1109/ICDAR.2017.307>.
- Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient Estimation of Word Representations in Vector Space. *ArXiv:1301.3781v3 [Cs.CL]*, 1–12.
- López-Calvo, Ignacio. 2009. *One World Periphery Reads the Other: Knowing the “Oriental” in the Americas and the Iberian Peninsula*. Newcastle Upon Tyne: Cambridge Scholars Publishing.
- Quibuyen, Floro. 2008. *A Nation Aborted*. Quezon City: Ateneo de Manila University Press.
- Stubbs, Michael. 1995. Collocations and Semantic Profiles: On the Cause of Trouble with Quantitative Studies. *Functions of Languages* 2 (1): 23–55.