TRAVEL, GRIEF AND FIGHT: COMPUTATIONAL TRACING OF FEMALE WORKS IN THE DUTCH EAST INDIES LITERATURE

Amalia Putri Astari
Philosophy Faculty, Universität zu Köln
aastari@smail.uni-koeln.de

ABSTRACT

In the canon of the Dutch Indies literature, the works of female writers are sometimes overlooked. Mainly when it appears in the early period of the Dutch colonial period, the central issue and genre in most of the literary works are frequently written by male authors through a masculine colonial perspective. This paper explores the works of female authors in the period of the Dutch East Indies (1800-1942) using two computational tools. The aim is to examine the most common themes that are appeared in the female authors' works throughout this long period. Using two computational tools, Stylo and Antconc, this paper reveals genre closeness and the most common themes that are appeared in twelve Dutch East Indies female novels which are collected from dbnl.org. Based on the cluster analysis, this paper divides the novels into four different clusters. More over the computational analysis also shows the three main themes; exploration, disadvantages of being a repressed group, and fighting against the oppressed. The use of extensive digital data in this research gives an alternative for literary researchers to find an alternative way of doing their research during this pandemic time, which makes the physical visit to the library limited. It also gives an alternative in the literature field to combine this discipline with computational tools.

KEYWORDS: Indies literature, female literature, computational literary studies, digital humanities.

INTRODUCTION

The existence of Dutch East Indies literary genre is not as simple as it seems to be. There is a lot more to it than meets the eye until this literary genre can be defined as a literature that people understand now, ”the Dutch-language literature of colonial and post-colonial Indonesia from the seventeenth century to the present day that includes Dutch, Indo-European and Indonesian authors ”(KITLV definition of Dutch East Indies literature).

The discussions about the definition of the Dutch East Indies literature used to be obscure. In 1972 Rob Nieuwenhuys published his Oost-Indische Spiegel, in which he defined for the first time the new literary genre: Indische literatuur. His work is without any doubt, the first guideline to the canon of the Indies literature. This guideline refers to two main characteristics of Indies literature. Firstly, this literature consists of the works of Dutch writers who had lived in the Dutch East Indies and specifically refers to the limited time and space. Secondly, this literature manifested a slightly distinguishing characteristic due to the complete absence of a literary tradition in the Indies environment (Praamstra, 257:1997). This statement
has thus eliminated the works of Indonesian writers, although their works are written in Dutch and they tell stories about The Netherlands Indies. Nieuwenhuys’ definition of Indies literature has been severely criticized over the years. Despite the discussion and the introduction of the new genre *Indische literatuur*, this literature still had to struggle to be included in the Dutch literary history based on Ton Anbeek’s guideline to Dutch literature history written in 1990 (Bel, 195:1992). This literature remains, like Termorshuizen (249-256:1990) states in his article as “an exotic stepchild in the history of literature.”

The criticisms and discussions led to a broader definition of Indies literature. This literature even also revolves not only around the VOC and Dutch Indies eras but also includes the postcolonial discourse (KITLV definition of Dutch East Indies literature). Although the broader definition includes most of the literature rotates around Indies themes, the discourse of Indies authors during the period 1900-1945 embraces mostly the works of male authors. In Rob Nieuwenhuys Indies literature guideline (1978 : 255-323), most of the female authors are placed on the last chapter of the book with the title, ‘Not to be Forgotten’. Only one female author, Székelyl-Lulof, is mentioned in earlier in the fifth chapter, and the other four, Maria Dermoût, Beb Vuyk, Hella Haasse, and Aya Zikken are mentioned briefly at the end of the book. Hence it shows, while the Indies literature is sometimes absent in the history of Dutch literature, the female Indies authors are completely neglected both in the canon of East Indies literature and the Dutch literature.

On the 1st of September 2000, Vilan van de Loo launched a website that gives fresh air to the publication of the East Indies female authors name list, damescompartiment.nl. She arranges the name from A to and she succeeded to record seventy names on her website. Among those names, man can see not only Dutch authors but also Dutch-Indo author as well as native Indonesian female authors, Kartini and Soewarsih Djopoeptio. Van de Loo arranges this list based on the broader definition of Indies literature that involves an unlimited range period of the books published, and the nationality background of the author. Two red lines connect these authors are, first of all, the language they use in their works, Dutch, and furthermore the theme they use in the story, The Netherlands East Indies.

Using the list from ‘Dames Compartiment’, tracing the works of Indies female authors becomes more convenient. Unfortunately, the global pandemic situation also affected researchers’ access to trace books at the library. The limitation of personal attendance makes it more difficult for most of the researcher to freely request books or read books in the reading room. This drives researchers to be more creative and to find alternatives in conducting literature research. The most possible resource of data we can easily find is available on the internet.

In the Dutch Studies and history fields, the website dbnl.org (De digitale bibliothek voor de Nederlandse Letterkunde) is one of the most promising Dutch text website collections. This website is launched in 1999 as an initiative of the Maatschappij der Nederlandse Letterkunde and as of 2020 is maintained by a collaboration of the Nederlandse Taalunie, Flemish Libraries and Royal Library of the Netherlands. This website contains thousands of literary text and it has a filter menu to make it easier for the user to look up the collections. From the seventy names of female Dutch authors, DBNL only has the collections of eight female authors. The novels are, *Vrouwen lief en leed onder de tropen* (1892), Therese Hoven;
The twelve novels are available in digital PDF format and downloaded for free from the dbnl.org website. I then converted the files into txt format to adjust to the two computational tools I use for the analysis; Stylo and Antconc.

Stylo package (Eder et al. 2013) provides multiple analyses in the field of computational stylistics, including authorship attribution, genre clustering, stylochronometry, etc. This package’s main function stylo() implements cluster analysis that reveals the closeness of the writing style between one text to another. Stylometric studies works as simple as this, one piece of software, in this case Stylo, extracts the meaning from the mass of words in the texts, arranging, loading texts and visualizing the results by drawing graphs. The stylometrists have begun to produce statistical methods of their own, the most important method is Burrow’s Delta whose method I use for clustering the data using Stylo. Meanwhile, Antconc is a free computational tool that allows you to examine several possibilities in texts. This tool includes seven tools; concordance tool (to show word search results in context); concordance plot tool (to show the position of the searched words in the target text); file view tool (to investigate more detail of the text of individual files); cluster/N-Grams (to find common expressions in a corpus); collocates (to show the collocates of a searched words that allow you to investigate patterns in using of the word); word list (to count all the words in the corpus and present them in an ordered list); Keyword list (to show which words are unusually frequent in the texts). This paper uses only four tools to analyse the twelve texts.

Several researchers have attempted to classify the literature genre or identify the primary characteristic of the literary movement in the Dutch Indies period. The Mirror of the Indies (1982) by Rob Nieuwenhuys, provides the reader with the most in-depth classification of the Dutch Indies Literature. Unfortunately, there is no specific discussion of female authors. Instead, they appear in the last chapter of the book with a title, Not to be Forgotten. In 2000, Vilian van de Loo launched her Website damescompartiment.nl and gathered information about women’s author in the Dutch East Indies literature. In this Website, she publishes mostly a short biography and the works of each author but not explicitly classified the themes or genre. However, there has been some research that has used computational tools to classify topics in extensive collections of journals and political texts (Riddel, 2012; Block, Newman, 2011). In the field of Dutch literature, there are several attempts have been made to explore the features of a literary movement using computational tools (Naber, Boot, 2019; Kao, Jurafsky, 2015). However, the computational research in classifying the female writers of the Dutch East Indies literature has never been done before.

This paper aims to reveal the similarities of the writing style between the twelve novels with clustering the genre of these novels. This clustering analysis works with the help of Stylo
that able to do a distance fast reading. Moreover, departing from the different clusters, the computational tool Antconc can see closer to the text and help the process of a closer reading of the texts to conclude the main themes appear in the twelve novels.

ANALYSIS

The function \texttt{stylo()} is currently the primary tool in the package. The function is signified to allow users to automatically process a digital corpus from a specified working directory and to deliver a variety of stylometric interpretations from multivariate statistics to estimate and visualize stylistic similarities between data texts. This function presents analytical analyses. This function will commonly be used to compose a most-frequent-word (MFW) list for the entire corpus. It will obtain the frequencies of the MFWs in the individual texts to generate an initial pattern of words in the rows by each individual texts in the columns (see figure 1) (Eder, Rybicki, Kestemont, 2017:7).

This paper examines the twelve texts with Delta methods. The workflow consists of producing cluster analysis tress between most-frequent-word frequencies in the texts compared (Rybicki, 2014; Eder, 2015). This method was proposed by John Burrows to discover the stylistic difference between two or more texts by comparing the relative recurrences of very common words (Burrows, 2002:267).

First and foremost, the texts are converted to UTF-8 (aka Unicode) format. After that, the files are renamed in the following naming syntax: Author Name_Publication Year_title. However, despite the careful corpus preparation, the experiment performed a systematic error in the corpus. From twelve files, only eleven files are clustered properly (see table 1) meanwhile the other one, ‘Koloniaaltje’ (1920) by Marie van Zeggelen is corrupted.

After the final files are arranged, the function \texttt{stylo()} runs the GUI box on the screen. Graphical User Interface (GUI) enables the user to control the script’s main parameter without having to interfere with the actual coding system. This tool renders an easy shortcut for a beginner humanities scholars who want to explore digital tools. GUI provides several tabs. In this paper, I use only several functions of tools provided in this Interface. Firstly, I set the input text as a plain text as the format of the files are UTF-8 and ticked the point Dutch as a setting language (Figure 2). The second setting I changed is the tab FEATURES. This tab provides the feature of the texts appear in the corpora. This tab is also crucial for the analysis because here the user can set the minimum and Maximum Most Frequent Words. The user can also set the increment of the analyses. For the analyses in this paper, I set four times different MFW Settings; min:100, max:100; min:100, max:2000; min:100, max:3000; min:100, max:4000; min:100, max:5000 (Figure 3). The changes of the maximum MFW affects the position of clustering between the twelve novels. I also set the pronunciation deletion to make sure that the words counted as MFW are those word classes except pronouns. The last thing I set for this analysis is the Statistic. As mentioned before, I use Cluster Analysis with Classic Delta distance (see Figure 4) to create a graph with the row and column measurements just like in Figure 1.

These GUI parameters settings result in the findings in the Figure 5,6,7,8, and 9. From the most minimal MFW culling setting (See Figure 5), the CA classified the eleven novels in two different clusters. Kartini and Hoven’s writings were identified as the most different writing styles than the other nine novels. Three novels written by De Wit share the stylistic
similarities while one novel by Zeggelen separates itself from the other two novels. The increase from 1000-2000 MFW changes the graphic (Figure 6). Interestingly, with more counting of MFW, the novels by Zeggelen are grouped in one cluster while De Wit’s novels are separated into two different clusters. The finding of these settings classified the novels into three main clusters. The two novels from Kartini and Hoven have not changed the cluster position and identified as the most distinctive writing styles as the other nine. In figure 7 and 8, the most noticeable difference is seen in the position of Soewarsih’s novel. With 4000 MFW maximum words, this book identified itself in the same cluster as Hoven’s novel, next to Kartini’s book. In 5000 MFW settings, the finding remains the same. It divided the eleven novels into two main clusters. The first cluster includes several subclusters. while the second cluster only arranges two different clusters. Based on these CA findings, the clusters are divided into four different clusters (see table 2).

Based on this cluster grouping, the research workflow moves to Antconc tool. Using Antconc, I first examine the clusters one by one by the world list option in Antconc. The results present the list of the most frequent words from each cluster. To designate the main object of this analysis, I eliminate articles, prepositions, and conjugations. The word classes I collect are only verbs, nouns, and adjectives. I limited the most frequent words only the tenth most frequent words on every cluster. Table 2 shows the top ten most frequent words from each cluster. The word class of the most frequent words varies from verb, noun, and adjective.

The four clusters have however one similarity: the word ‘vrouw’ (=woman) appears as one of the most common words in all four clusters. To look closer to the context of the text, I use ‘concordance’ tool to see the words appear before and after the word ‘vrouw’. In the first cluster, the word ‘vrouw’ combines mostly with adjective; jonge vrouw (=young woman), slechte vrouw (= bad woman), lieve vrouw (=lovely woman), vrije vrouw (=free woman), Mohammedaansche vrouw (= Moslem woman), Oostersche vrouw (=Eastern woman), goede vrouw (=good woman), wijze vrouw (=wise woman). The most hits of the word ‘vrouw’, appears mostly in Lulofs’ work with 124 word hits, and 25 hits in De Wit’s Gods goocheelaartjes, and 6 hits in De Wit’s Orpheus in de dessa.

Lulof’s novel, Koeli, tells a story about the life of Njai, a designation of a female concubine in The Dutch East Indies. In this book, appear fourteen times the term ‘vrije vrouw’. This term refers to Njai character who does not have any longer contract with any Dutch men. Lulofs pictures it as a phase of freedom when women can finally free from the trap of concubine. In the second cluster, the word ‘vrouwen’ is mostly related to the verb ‘boogen’ (=to bow to someone). This action happens in the relation between Dutch East Indies native women with Dutch people. The books in the cluster 2 shows the depiction of women as an inferior group against Dutch people. Moreover, in the cluster 2, the word ‘man’ also relates several times with the word ‘boogen’. The word ‘man’ can be translated in two different ways. Firstly, it means a male and the second one means a person in general. The context of the word ‘man’ in this cluster is mostly referred to the first meaning. Thus, both male and female native characters in this cluster are associated with the action ‘boogen’, to show a respect towards Dutch people.

The opposite appears in the cluster 3. The word ‘vrouwen’ in the two novels of this group associates mostly with the sentences that tell stories about female native Dutch East
Indies people who do hard working labour, for instance; vrouwen aan den arbeid (=women were working), vrouwen aan het rijst stampen (=women were stomping rice), vrouwen bezig met het morgen maal (= women were busy with preparing breakfast), vrouwen de harde werkers zijn op het eiland (=women are the most hardworking people on the island), vrouwen snijden de dieren in stukken (=women are cutting the animals into pieces), vrouwen dansten, met messen gevochten (=women were dancing and fighting with knife). These activities show how women can lead the society and stand for themselves without relying on men’s figures. Moreover, the word combination of ‘vrouw’ also mentions in this cluster in the words; vrouwenarbeid (=women labour) and vrouwenbeweging (=women’s movement). However, the word ‘vrouwen’ that associates with Dutch women only appears one time in, ‘Hollandsche vrouwen; lachen (= Dutch women are laughing). The reference of native women dominates the stories in the worjs of cluster 3.

In the last cluster, the word ‘vrouw’ appears mostly after the adjective ‘Javaansche’ and ‘Europesche.’ Mostly in Kartini’s and Soewarsih’s work, comparations between Javanese women and European women are mentioned a lot. It is not a surprise since both of the writers share the same background, a native Indonesian background. In Hoven’s novel, the adjectives of ‘Javaansche’ and ‘Europesche’ do not expose that much. Nonetheless, the similarity between all these three books is the frequent appearance of the word combinations; vrouwenarbeid (=working labor); vrouwenbestaan (women’s existence); vrouwenbeweging (=women’s movement); vrouwenemancipatie (= women’s emancipation) and vrouwen wereld (=female world). These associations of women, show that most of the issues that reveal in the stories represent the woman’s consciousness of movement and emancipation. Furthermore, it becomes even more emotional on how in this cluster, the word women’s are also associated with the word ‘leed’ (=sorrow) and vrouwenhart (=women’s heart). The fight of the women’s issues in the stories also often leads to an emotional contemplation of the characters in these books. Moreover, compare to the other clusters, cluster 4 also raises an important issue in the family life. The word ‘echgenoot’ (=spouse) appears 110 times in Hoven’s, 291 times in Kartini’s and 128 times in Soewarsih’s novel. These three books construct several problematic stories of being housewives. These issues do not appear significantly in other clusters. Cluster 4 mentions to such degree most of the female issues in the Dutch East Indies.

Besides the word ‘women’, words that associate with a male are also arise in the most frequent word’s hits. The most exciting finding appears in the first place of cluster 1, the word ‘Toewan’.This word is a salutation used by native characters in the stories to call Dutch male characters. With the concordance plot tool in Antconc, it appears that most of the contexts in the use of the word ‘toewan’ are related to power legitimation of Dutch colonization power towards native Indies people in the Dutch East Indies. The fact that the salutation of Toewan resembles significantly in cluster 1, shows a huge contrast of colonial power issues raised in cluster 1 compared to cluster 4. However, in cluster 4, the other male word’s association comes from the use of the word ‘vader’ (=father). The word comes mostly from Kartini’s book and it is influenced by Kartini’s adoration of her father, the Regency Chief of Jepara. Furthermore, also in Soewarsih’s book, the role of the father is mentioned as the role of the guard who is mostly responsible for his daughter; to pick up his daughter from the station; give permission for a daughter to go somewhere. Different from the use of the word ‘Toewan’, the word ‘vader’
appears to be more of a daughter respect towards her father rather than a colonial power relation represents from the salutation of ‘Toewan’.

The raising issues of colonial power appear not only from the word ‘Toewan’ but also expresses in the use of the word ‘arme’ (= poor) and ‘klein’ (=small), bruine (= brown) to call native Indies people. Especially in cluster 1 and 2, the adjectives ‘donker’ (=dark) are used before the nouns that are related to native people and their residence, donkere kampong (=native village), donkere hoofd (=native heads), donkere meisjes (=native girls) and donkere kindertjes (=native kids). These expressions are used by the narrators in the stories to describe the native community in the Dutch East Indies. However, these adjectives do not appear frequently in the cluster 3 and 4.

The stories of the Dutch Indies society and environment also appear frequently in cluster 1, 2, and 3. The most repeated words in these three clusters are regularly related to activities that are happened in the natural landscape and adjectives which express the exploration and admiration of natural areas. For example in the word ‘stond’ (=stood), in cluster 1 and ‘water’ in cluster two. Both words have the concordances that are related to the representation of the natural landscape in the Dutch East Indies. Moreover, nature and landscape descriptions such as; Standing alone in the darkness; stood in the silence; stood next to the volcano; stood above the smoky evening red sky; stood in the kampong; stood the wing; stood the volcano; stood under the little cloud; stood under the sun, lead the characters in the stories (see cluster 1) to do a self-contemplation about weakness, solitude, and death. In cluster 3, however, the substantive ‘dag’ (= day), nuances more positive expression in the texts. It mentions, for example, how the day brings another light and hopes for the characters.
Figures and Tables

Fig. 1 Row and Column Measurements in Stylo CA function

Fig. 2 stylo() launched GUI, first tab appeared: Input Language
Fig. 3 GUI Features Tab

Fig. 4 GUI Statistics Tab
Fig. 5 Cluster Analysis Maximum MFW 100

Fig. 6 Cluster Analysis Maximum MFW 2000

Fig. 7 Cluster Analysis Maximum MFW 3000
Fig. 8 Cluster Analysis Maximum MFW 4000
Table 1 Final Corpus

| Title                                      | Author            | Year of Publication |
|--------------------------------------------|-------------------|--------------------|
| Vrouwen lief en leed onder de tropen       | Therese Hoven     | 1892               |
| Orpheus in de dessa                        | Augusta de Wit    | 1903               |
| De gouden Kris                             | Marie C. van Zeggelen | 1908              |
| Onderworpenen                              | Marie C van Zeggelen | 1910              |
| Door duisternis tot licht                  | Kartini           | 1911               |
| Natuur en mensen in Indie                  | Augusta de Wit    | 1914               |
| Bij het hart van Indië                     | Marie C. van Zeggelen | 1926              |
| De Wit Gods goochelaartjes                 | Augusta de Wit    | 1934               |
| Koeli                                      | Székely-Lulofs    | 1932               |
| Het laatste huis van de wereld             | Beb Vuyk          | 1941               |

Fig. 9 Cluster Analysis Maximum MFW 5000
Table 2 Final Clustering

| Cluster 1                                                                 | Cluster 2                                      | Cluster 3                                                                 | Cluster 4                                                                 |
|---------------------------------------------------------------------------|------------------------------------------------|---------------------------------------------------------------------------|---------------------------------------------------------------------------|
| De Wit_Orpheus in de dessa; De Wit_Gods goochelaarthes; Lulofs_Koelie     | Zeggelen_Onderworpen; Zeggelen_De gouden Kris; Zeggelen _Bij het hart van Indie | Vuyk_Het laatste huis; De Wit_ Natuur en menschen in Indie               | Kartini_Door duisternis; Hoven_Vrouwen lief en leed onder de tropen; Soewarsih_Buiten het gareel |

Table 2 The Most Frequent Words

| Cluster 1                                    | Cluster 2                                      | Cluster 3                                      | Cluster 4                                      |
|----------------------------------------------|------------------------------------------------|-----------------------------------------------|------------------------------------------------|
| toewan, stond, vrouwen, oogen, alleen, weg, kon, hoofd, wist, alles       | had, groote, huis, hart, menschen, land, donkere, man, vrouwen, keek. | heeft, komen, water, man, vrouwen, dag, jaar, volk, staan. | kunnen, vrouwen, leven, goed, vrouw, vader, leed, zien, weet hoop. |

Table 3 The concordance of the MFW’s

| Cluster 1                                      | Cluster 2                                      | Cluster 3                                      | Cluster 4                                      |
|-----------------------------------------------|------------------------------------------------|-----------------------------------------------|------------------------------------------------|
| toewan; toewan besar; toewan controleur; toewan dokter. | Had; Aangenomen; achtergelaten; Nederlander had gelijk | Heeft; achtergelaten; aanhoudend; aan bled; beloofd; gevecht; soldaten. | Kunnen; Apprecieren; Bereiken; Bewijzen; Bijdragen; Blijven; |
| Stond; | Groote; | Komen; | Vrouwen; |
|---|---|---|---|
| Stond alleen in de donkerheid; stond de stilte; stond daar de vulkaan; stond boven het rookerig avondrood; stond in de kampong; stond de vleugel; stond de vulkaan; Stond de kleine wolk; stond de zon. | donkere ogen; groote heer; groote huis; kantoor; verbaazing | Spoedig komen; komen brengen; komen Javanen; zij komt. | Vrouwen; |

| Vrouwen; | Huis; | Water; | Leven; |
|---|---|---|---|
| Mohammedaansche vrouw; wijze vrouw; slechte vrouw; lieve vrouw; vrije vrouw; Oostersche vrouw; goede vrouw. | huis der Hollanders; vrouw liepen om het huis; van haar huis. | water, aarbeving, ziekte en dood; groene schemering in het snelle water af te koelen; water brak en troebel; springt in het water; water en terugkaast van een blauwe, onbarmhartige hemel. | echtgenooten |

| Oogen; | Hart; | Man; | Goed; |
|---|---|---|---|
| haar oogen; zwarte oogen; scheeve chineesche oogen; roode oogen; | hart hart kloptte; heftig; hart kloppen van | man als slavin gebonden; zwakte; kleinen man. | Voorbeeld; dag goed vooruit; goed voor dochtertje. |
| neergeslagen oogen; uitpuilende oogen; sloeg haar ogen neer. | verwachting; hart is angstig; hart sneller; hart rouwt; hart raakte. |  |
|---|---|---|
| Alleen; Ik ben alleen, alleen de donkere huid; de koelie, glad als een pisangblad. | Menschen; bruine menschen; arme menschen; blanke menschen. | Vrouwen’ aan den arbeid; aan het rijst stampen; Hollandsche vrouwen; lachen vrouwen bezig met het morgen maal; vrouwen dansten met messen gevochten; vrouwen de harde werkers zijn op zij eiland; vrouwen snijden de dieren in stukken. | Vrouw; Javaansche vrouw; Europesche vrouw; gelukkig te; vrouw in het bijzonder; vrouw moet uitstaan. |
| Weg; Geen weg; weer weg; ging zij weg; geluk op je weg; Dat alles weg; Weg te loopen. | Land; mijn land; haar land, zuiveren land. | Dag; de dag begint met een lichte verkleuring; de volgende dag begint het beraad opnieuw; de dag groeien de hoopen; dag opnieuw; 165 hits | Vader; vader aan mijne zijde; Arme, lieve, oude vader; vader die hen van het station afhaalde; vader’s toestemming; vader vindt het onderwijs het beste 368 hits |
| Kon; | Kon een mensch hier leven; kon een mensch leven van zo goed als niets. | Donkere; donkere hoofd; donkere kampong; donkere kindertjes; donkere massa; donkere meisjes; donkere oogen. | Jaar; jaar geleden; verleden jaar | Leed; dragen jonge vrouw; leed pinjnen; besparen; leed en nadenken; leed de gevoelens der vrouwen. |
|------|-------------------------------------------------------------------|-----------------------------------------------------------------|---------------------------------|---------------------------------|
| Hoofd; | zijn hoofd; hoofd en dood, hoofd en neergeslagen oogen; hoofd gebogen; omgewend hoofd; haar oog; groot en angstig; hoofd mandoer. | Man; Man wist; man boog | Volk; immigranten volk uit voortdurend besmette streken; | Zien; zijne familie mag haar nog niet zien; moeder natuur te zien; dankbaarheid te zien, hoe haar man haar eert en waardeert; Het is heerlijk om te zien, hoe er leven komt in de takken. Hoe de goede ziel steeds vooruitgaat; buitenwereld te zien; vielzijdig zien, leerde haar toegevendheid; Hij kann me niet zien lijden; onze lieven te zien lijden |
| Wist; | het geluk wist; wist dat de pijn in zijn hart; wist dat hij er | Vrouwen; Boog; vrouwen en kinderen | Immigratie; de sterke immigratie die van oudsher uit Java hierheen is; een | Weet; ik weet dat, ... |
CONCLUSION

The findings of the two computational analyses using Stylo and Antconc explicate how a distant reading tool can unhitch the way for humanities scholars to examine a large number of literary corpora. It also signifies that a successful digital humanities experiment requires an excellent corpus preparation. Working with initially twelve digital books, in the end, this research can only analyze eleven books due to technical file error. From these eleven books, Stylo arranges the eleven books in four different clusters that have a range of closeness from one to another. The results show how cluster 1 and 2 share the same characteristic of the most frequent words. Both clusters are focus on how the characters explore the environment and society of the Dutch East Indies. The stories of the books in these two clusters also appear to be more cognizant of the colonial power relations in the Dutch East Indies society. It is shown by the use of adjectives to describe native people that are told to have lower prosperity than Dutch people. Cluster 1 and 2 also expose how women are represented as the repressed group compared to male characters. Cluster 3 and 4, however, present more awareness of women’s issues. It signifies by the most frequent words regarding female issues and the position of the female in family and society in the Dutch East Indies.
REFERENCES

Reference to a journal publication:
Bel, Jacqueline (1992). Receptie-onderzoek en literatuur-geschiedenis: De Indische roman als broeierige vrijplaats van de Nederlandse literatuur. Indische Letteren, Jaargang 7, 1992, pp. 195-203.  
Block, S., Newman, D. (2011). What, Where, When, and Sometimes Why: Data Mining Two Decades of Women’s History Abstracts. Journal of Women’s History 23, no.1. 81-109. Naber, F., Boot, P., (2019). Exploring the features of naturalist prose using LIWC in Nederlab. Journal of Dutch Literature. 50-70. Praamstra, Olf (1997). ‘De omstreden bloei van de Indisch-Nederlandse letterkunde. Een afbakening van het corpus’. Tijdschrift voor Nederlandse Taal- en Letterkunde, Vol. 113, 1997, pp. 257-274. Riddell, A.B. (2012). How to Read 22,198 Journal Articles: Studying the History of German Studies with Topic Models Termorshuizen, Gerard (1990), ‘De Indische bellettrie: een exotisch stiefkind binnen de literatuurgeschiedenis’. In: Traditie en progressie. Handelingen van het 40ste Nederlands Filologencongres (’s-Gravenhage 1990) 249-256.

Reference to a book:
Burrows, John. (2002). ‘Delta: A Measure of Stylistic Difference and a Guide to Likely Authorship’, Literary and Linguistic Computing 17.3 (2002), 267–86. Eder, Maciej; Rybicki, Jan; Kestemont, Mike. (2017) ‘Stylo’: a package for stylometric analyses. Computational Stylistics Group. Eder, Maciej. (2013). Mind your corpus: systematic errors in authorship attribution. Literary and Linguistic Computing, 28(4): 603–14. Eder, Maciej. (2015). Visualization in stylometry: some problems and solutions. Submitted to Literary and Linguistic Computing. Nieuwenhuys, Rob. 1978. Oost-Indische Spiegel. Amsterdam: Em. Querido's uitg. Rybicki, Jan. (2014). Visualizing literature: artistic statistics. In Bleinert, M., Curylo-Klag, I., Kucala, B. (eds), Art in Literature, Literature in Art. Krakow Wydawnictwo Uniwersytetu Jagiellon skiego, pp.135–46.

Reference to a web source:
Van de Loo, V., (2000) Dames compartiment. http://damescompartiment.nl (1st Augustus, 2020) De Digitale Bibliotheek voor de Nederlandse Letteren (DBNL). https://dbnl.org (3rd Augustus, 2020)

Acknowledgment
This work was supported by Lisa Maskell Fellowship, Gerda Henkel Stiftung, Germany.