

**Comparative Analysis of Gender-Related Words in Parallel Corpora of *Bumi Manusia***

Lalitia Apsari  
Department of Descriptive Linguistics, University of Indonesia,  
West Java, Indonesia  
lalitia.apsari@ui.ac.id

**Abstract**

In literary criticism, the representation of gender in text production from the reader's perception has been a long and continuous debate. This comparative analysis using parallel corpora is a study to break down and compare the two languages' attempts in delivering the similar semantic goals. The main focus is on gender-related words in the *Bumi Manusia* novel (international title: *This Earth of Mankind*), a fictional novel by Pramoedya Ananta Toer, in both Bahasa Indonesia and English as translated by Maxwell Lane. Quantitative analysis supported with qualitative detailing are used to detect and analyze ATTITUDE-conveying terms as specified in the Appraisal Theory. Preliminary evaluation indicates that there is a difference in patterns between the Bahasa Indonesia text and the English text which convey each language's distinction in representing gender and identity. This also poses a question of how much is lost in the translation in terms of semantic values. This paper can be scaled into both synchronic and diachronic comparative analyses of general language comprehension using a notably bigger parallel corpora.

*Keywords:* ATTITUDE, Appraisal Theory, literary translation, corpus linguistics, gender representation

**Introduction**

There are many ways of studying the representation of gender, particularly in literary works. These studies are mainly an attempt to identify any inclination or formation of patterns of gender portrayal. This study utilizes both quantitative and qualitative analyses to examine the gender-related words and their interaction with the surrounding words in parallel corpora of *Bumi Manusia*. *Bumi Manusia* is the first of the four books that comprise Pramoedya Ananta Toer's Buru Quartet. As one of the classic Indonesian literary works, the book has been translated into many languages, including an international version in English under the title *This Earth of Mankind*, and is considered as one of the most significant literary works by an Indonesian author.

**Objectives and Research Questions**

A previous study on gender representation has been conducted focusing only on ATTITUDE analysis surrounding gender-related words in one literary work written in Indonesian (Apsari, 2017). This study focuses on a similar theme but using parallel corpora of a literary work written in Bahasa Indonesia (Indonesian) and its translation in English to detect the similarities and differences in delivering the same semantic messages. It is an attempt to showcase how quantitative analysis supported with qualitative findings can be significantly useful to examine literary works from the linguistic aspect. The text is analyzed using a computational approach and studied qualitatively to extract traces of inclination on how femininity or masculinity is

discursively performed within the text as well as cross-language with certain limitations. The use of parallel corpora poses a question on how the semantic values in Indonesian is delivered in English and whether there are differences that cause any loss in semantic values targeted in the Indonesian text by scrutinizing how particular tokens are used, their interaction with the gender-related words, and the sentiments formed according to the ATTITUDE aspect of Appraisal Theory (Martin & White, 2005). However, this paper does not include the characters in the analysis nor provide any generalization on male or female character in the book.

### Theory

This paper uses Appraisal Theory as the basis theory and methodology of qualitative analysis related to the ATTITUDE values (Martin & White, 2005). The term ‘appraisal’ was first coined into psychology by Magna Arnold (1960) and is defined as the process through which the significance of a situation for an individual is determined. Appraisal brings about the attraction or aversion while emotion is associated with this “felt tendency towards anything intuitively appraised as good (beneficial), or away from anything intuitively appraised as bad (harmful).” (Arnold, 1960). The ATTITUDE outlines the specifics of the appraisal being expressed, the APPRECIATION values (evaluation of phenomena, reaction or composition), and the JUDGEMENT values (social or ethical appraisal), and the AFFECT values (emotional state).

### Methodology

For a typical corpus-based linguistic analysis, the most significant method is to search the corpus in question for the occurrence of certain strings (i.e., words or phrases). This study involves parallel corpora of the novel *Bumi Manusia* (Toer, 1980), its original text in Indonesian as written by Pramoedya Ananta Toer and the international version *This Earth of Mankind* as translated into English by Maxwell Lane.

Pramoedya Ananta Toer was born on February 6, 1925. His father, who was politically active against the Dutch, was a headmaster and a man of some social prominence. Pramoedya Ananta Toer was known to write corrosive effects of colonialism on human relations which led him to a long imprisonment. He was denied pens or any kind of paper in prison, so Pramoedya told the story of Minke, the hero of the Buru Quartet thus the hero of *Bumi Manusia*, to his fellow prisoners every night. In this way, the author was playing on a long Indonesian tradition of oral storytelling. The English translation of *Bumi Manusia* was by an Australian diplomat, Maxwell Lane, who was posted to Jakarta in the early 1980s. It was that translation from the Malay-based Indonesian language that showed the English-reading world Pramoedya Ananta Toer’s spare but emotive writing style (Perlez, 2006).

### Material

The novel *Bumi Manusia* and its translation *This Earth of Mankind* are the parallel corpora comprises of more than 200,000 tokens that require the use of a small-scale computational linguistic method as well as a manual analysis to detect patterns within the linguistic evidence. For a more focused research goal on the words surrounding gender-related words, this study conducts an analysis using selected words of *wanita*, *perempuan*, *dara*, and *gadis* for the female gender and *lelaki*, *pemuda*, and *pria* for the male gender in Indonesian. From the initial observation, the parallel words in English found in the novel are *woman/women*, *female*, *girl*, and *maiden* for female gender and *man/men*, *male*, and *boy* for male gender. Gender-related pronouns are not considered for the analysis of both

languages because the interaction with their surrounding words is mainly to portray a particular person and less about gender representation.

### Quantitative and Qualitative Approaches

This paper utilizes corpus linguistics as the main approach for the analysis. Corpus linguistics is considered as an empirical method of linguistic analysis and description using real-life examples of language data stored in corpora as the starting point (Crystal, 1992; Jackson, 2007). From this approach, concordances and collocations of the focus words are examined using the Key-Word In Context (KWIC) method (Sinclair, 2005) which is useful for hypothesis testing and generation. The ATTITUDE analysis utilizes the SentiWordNet (Baccianella, Esuli, & Sebastiani, 2010) to detect the ATTITUDE-related words in the English corpus and combining it with a semi-manual approach for the Indonesian corpus<sup>1</sup>.

The concordances and collocations from the KWIC samples are annotated using a combination of rule-based computational linguistic approaches and manual labelling to support the categorization and polarity of a word or phrase based on the sentence and to spot keywords based on common-sense knowledge (Moilanen & Pullman, 2007; Subrahmanian & Reforgiato, 2008; Liu, Lieberman & Selker, 2003). The attitudinal expressions consist of ‘app’ for APPRECIATION; ‘nor’, ‘cap’, and ‘ten’ for NORMALITY, CAPACITY, TENACITY respectively in the SOCIAL ESTEEM group; ‘ver’ and ‘prop’ for VERACITY and PROPRIETY respectively in the SOCIAL SANCTION group; all of JUDGEMENT (Martin & White, 2005); and eight AFFECT categories (Lövheim, 2011). ‘+’, ‘-’, and ‘ø’ indicate positive, negative, and undecided polarity, respectively.

This study uses the R software (version 3.3.3) (R Core Team, 2017) for the quantitative analysis and visualizations as well as Gephi software (Bastian, Heymann, & Jacomy, 2009) for network graph visualizations. Association measurements to detect collocations are based on the likelihood ratio (Manning & Schütze, 1999) as implemented in the ‘quanteda’ package (Benoit & Nulty, 2017) in R. The network graph visualization uses ForceAtlas2 algorithm to portray the interaction between the gender-related words and their surrounding words (Jacomy et al., 2014).

### Literature Review

Particular words are phrases that are usually used to deliver certain semantic goals, particularly those related to gender representation and construction. In the field of literature, numerous gender-related words emerge to convey somewhat different purposes in its products. Therefore, gender definitions are adaptable and malleable concepts. Because the concepts may vary culturally, socially, or generationally (Baker, 2003), it should be noted that this study using *Bumi Manusia* should not be related to the current gender representation happening in reality. However, with further and contextual studies, this might be relevant to gender representation within the related timeframe. The ATTITUDE system plays a vital part of this research which is acts as the basis framework in constructing the interactions between the words that will be compared in Indonesian and English. This framework will help to construe the semantic themes surrounding the focus words which potentially represent the author’s perspective and the translator’s take in illustrating a certain type of gender within the story.

APPRECIATION, as one of the systems in ATTITUDE by which evaluations are made of products and processes. It comprises the evaluation related to appearance or physical traits

---

<sup>1</sup> This paper’s author is also currently conducting a study on a scalable system to detect sentiment-conveying terms in Indonesian based on dependency length and parsing.

by value. Although it falls under the general heading of aesthetics, APPRECIATION also involves a non-aesthetic category of 'social valuation'. The attitudinal sub-system of JUDGEMENT involves evaluation of human behaviors by referencing a set of institutionalized norms which can be differentiated between personal judgements of admiration or criticism (SOCIAL ESTEEM) and moral judgements of praise or condemnation (SOCIAL SANCTION) (Iedema, Feez, & White, 1994; Martin & White, 2005). In general, a JUDGEMENT value can be realized explicitly (as tokens) or implicitly (inscribed). Due to the abstract nature of inscribed terms, this study considers only the explicit tokens as the analysis data. AFFECT is related to emotional response and disposition and is typically realised through mental processes of reaction. The AFFECT values are typically realized through mental processes of reaction and attributive relation of the types. They provide one of possibly the most obvious ways that a speaker--which in this case are gender-related words--can adopt a stance towards some phenomenon. These values also provide the resources by which the speaker can indicate how that phenomenon affected them emotionally in order to appraise that phenomenon in affectual terms.

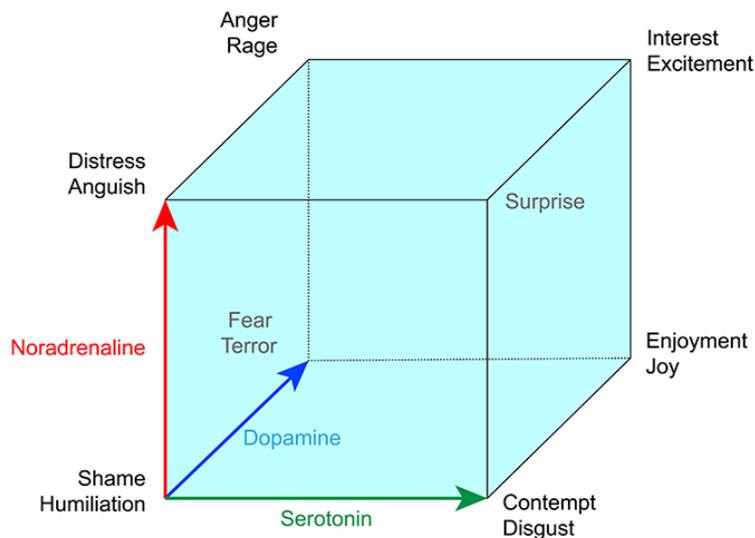


Figure 1. Lövheim cube of emotion

This paper uses a dimensional model of emotion which is described by Lövheim (2011) as the Lövheim cube of emotion (Figure 1), as a grounding system in building a contextual lexicon for ATTITUDE analysis, particularly on the AFFECT values.

Overall, the direction of the opinion (i.e., whether positive or negative) which in this case is represented by the text from the novel *Bumi Manusia*, is sometimes referred to as semantic orientation. Esuli and Sebastiani (2006) define the problem as having three different aspects: (i) determining the text's subjectivity (i.e., whether the text is factual in nature or whether it expresses an opinion on its subject matter); (ii) determining the text's polarity, or deciding if a given subjective text expresses a positive or negative opinion on its subject matter; and (iii) determining the strength of the text's polarity (i.e., deciding whether the positive opinion expressed by a text on its subject matter is weakly positive, mildly positive, or strongly positive).

### Findings

A preliminary analysis includes mapping the dispersion of the focus words which then followed by a comparative ATTITUDE analysis between the Indonesian and English corpus using the KWIC method. The preliminary quantitative analysis for the overall

## COMPARATIVE ANALYSIS OF GENDER-RELATED WORDS IN

comparison between the focus words in Indonesian and English shows similarities in many aspects with a distinct difference in a few aspects.

### APPRECIATION Values

Overall findings between the focus words in Indonesian and English show a few differences in terms of cross-language sentiment polarity in terms of APPRECIATION values (Table 1). There is a distinct semantic value emerges within the text in *Bumi Manusia* where the female and male genders are portrayed almost equally through APPRECIATION values whereas the ratio of interaction number between female and male in the English corpus is near to 2:1. It also applies in the negative sentiment category where the ratio in the Indonesian and English corpus is around 4:3 and almost 4:1 respectively. There is a potential significant difference of semantic goals in describing an equal evaluation of female and male genders as initially portrayed by Pramoedya Ananta Toer.

Table 1

*Percentage of word interaction in terms of polarity for APPRECIATION values*

APPRECIATION	Female (ID)	Female (ENG)	Male (ID)	Male (ENG)
Positive	79.31%	85.53%	62.07%	46.15%
Negative	14.94%	9.41%	20.69%	34.62%
Undecided	5.75%	7.06%	17.24%	19.23%

Looking at the word usage within the language, the limitation of female gender-related words *woman/women* in the English corpus as opposed to the options of *perempuan* and *wanita* in Indonesia causes the nodes to interact differently (Figure 2). The network graph of APPRECIATION values based on modularity class for female gender-related words in the Indonesian corpus shows that *perempuan* interacts only with the aesthetic category of 'social valuation'. This can be identified by the presence of the words related to appearance: *muka* (face), *bibir* (lips), *mata* (eyes) as well as words like *berdandan* (wearing make-up), *dibedaki* (powdered), *indah* (lovely), and *berhias rapi* (well-dressed). On the other hand, *wanita* interacts with both aesthetic and non-aesthetic categories of 'social valuation'. In the English corpus, there is no particular interaction in the aesthetic category because *woman/women* interact interchangeably with both categories.

## COMPARATIVE ANALYSIS OF GENDER-RELATED WORDS IN

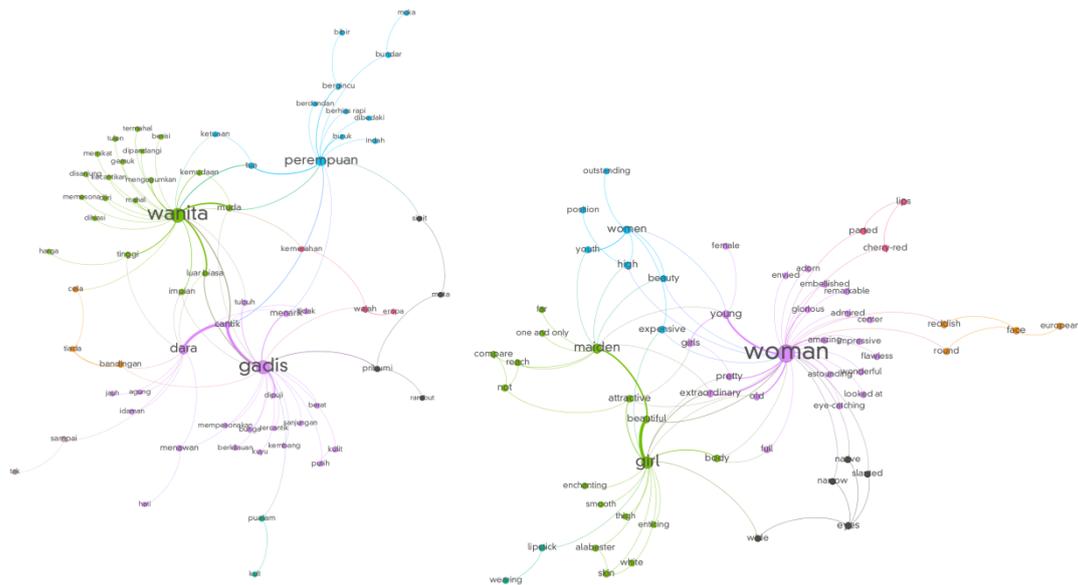


Figure 2. Network graph of APPRECIATION values based on modularity class for female gender-related words in the Indonesian (left) and English (right) corpus

The specific interaction is also identified in the male gender network even though not as clear and distinct. The word *pemuda*, which means ‘young man’ in English, in the Indonesian corpus, plays a different role with *lelaki muda*, which also means ‘young man’. The word *pemuda* interacts only with positive sentiments such as *gagah* (stout) and *ganteng* (handsome) whereas *lelaki muda* interacts with all words regardless of the sentiment. In the English corpus, the word *young man* could not accurately reflect how *pemuda* is being portrayed because it functions interchangeably with *young man* as *lelaki muda*.

### JUDGEMENT Values

In terms of polarity in JUDGEMENT analysis, the ratio between female and male genders in both languages shows only a slight difference (Table 2) with a note that the percentage of the negative sentiment for the female gender in the Indonesian corpus is notably higher than the percentage in the English corpus.

Table 2

*Percentage of word interaction in terms of polarity for JUDGEMENT values*

JUDGEMENT	Female (ID)	Female (ENG)	Male (ID)	Male (ENG)
Positive	55.62%	67.38%	67.65%	71.43%
Negative	38.75%	24.82%	29.41%	22.45%
Undecided	5.62%	7.8%	2.94%	6.12%

The challenge in identifying ATTITUDE terms, particularly JUDGEMENT values, is that several words are inscribed. However, in this novel, the inscribed terms have an explicit trace, therefore, the identification for this part can be done semi-manually. The striking difference for this aspect is that no terms are found for the male gender in the Indonesian

## COMPARATIVE ANALYSIS OF GENDER-RELATED WORDS IN

corpus related to NORMALITY whereas there is a 10.25% portion found in the English corpus (Table 3). Related to VERACITY, the percentage ratio between the female and male genders in the Indonesian corpus is in contrast with that in the English corpus. The percentage related to PROPRIETY in the Indonesian corpus is also more distant in ratio than that in the English corpus.

Table 3

*Percentage of word interaction in terms of category for JUDGEMENT values*

JUDGEMENT	Female (ID)	Female (ENG)	Male (ID)	Male (ENG)
NORMALITY	9.38%	17.73%	-	10.25%
CAPACITY	39.38%	29.79%	47.06%	38.78%
TENACITY	19.37%	17.73%	2.94%	8.16%
VERACITY	4.37%	6.38%	8.82%	4.08%
PROPRIETY	27.5%	28.37%	41.18%	38.78%

As opposed to the specific interaction found in the APPRECIATION analysis, the words *wanita* and *perempuan* as well as *pemuda* do not show any significant difference in terms of JUDGEMENT values.

### AFFECT Values

The ratio between female and male genders in both languages does not show any significant difference in terms of polarity in the AFFECT analysis (Table 4). However, there are a few interesting findings based on the categories of AFFECT.

Table 4

*Percentage of word interaction in terms of polarity for AFFECT values*

AFFECT	Female (ID)	Female (ENG)	Male (ID)	Male (ENG)
Positive	32.62%	32.76%	21.95%	33.3%
Negative	61.7%	61.21%	68.29%	62.5%
Undecided	5.67%	6.03%	9.76%	4.17%

Table 5 shows that no terms are found related to the emotional category of SHAME/HUMILIATION for the male gender in the English corpus. There is also no vague terms such as emotion that is related to numbness for that specific gender. There are terms found for the two categories in the Indonesian corpus albeit the small number. There are notable differences between the focus words which might indicate a difference of each gender's portrayal. The female gender-related words in the Indonesian corpus interact mostly and almost equally in the DISTRESS/ANGUISH, CONTEMPT/DISGUST, and INTEREST/EXCITEMENT categories whereas there is a notable difference in the English corpus. This also happens in the male gender-related words in the Indonesian corpus which interact mostly and almost equally in the DISTRESS/ANGUISH, ANGER/RAGE, and

## COMPARATIVE ANALYSIS OF GENDER-RELATED WORDS IN

INTEREST/EXCITEMENT categories whereas the ones in the English corpus interact mostly in the DISTRESS/ANGUISH category followed by JOY/ENJOYMENT. There are several terms (particularly verbs) in the Indonesian corpus that play a specific role to portray an act of ANGER/RAGE such as *menikam* (stab), *memaksa* (force), and *membenci* (hate) that are not found surrounding the male gender-related words in the English corpus.

Table 5

*Percentage of word interaction in terms of category for AFFECT values*

AFFECT	Female (ID)	Female (ENG)	Male (ID)	Male (ENG)
DISTRESS/ANGUISH	17.73%	24.14%	21.95%	31.25%
CONTEMPT/DISGUST	19.15%	18.97%	9.76%	10.42%
INTEREST/EXCITEMENT	19.15%	18.97%	19.51%	14.58%
JOY/ENJOYMENT	14.89%	16.38%	9.76%	20.83%
ANGER/RAGE	13.48%	8.62%	21.95%	14.58%
FEAR/TERROR	4.96%	4.31%	2.44%	4.17%
SHAME/HUMILIATION	4.26%	3.45%	4.88%	-
SURPRISE	1.42%	0.86%	7.32	4.17%
Undecided	4.96%	4.31%	2.44%	-

A special interaction between the female gender-related words in both languages can be seen in the network graph of Figure 3. There are differences in the variety of interactions between the words in the Indonesian corpus and those in the English corpus. The positive interaction regarding INTEREST/EXCITEMENT with words such as *cinta* (love), *suka* (like), and *dipandangi* (gazed upon) in the Indonesian corpus are much more diverse than in the English corpus that only uses *love*.



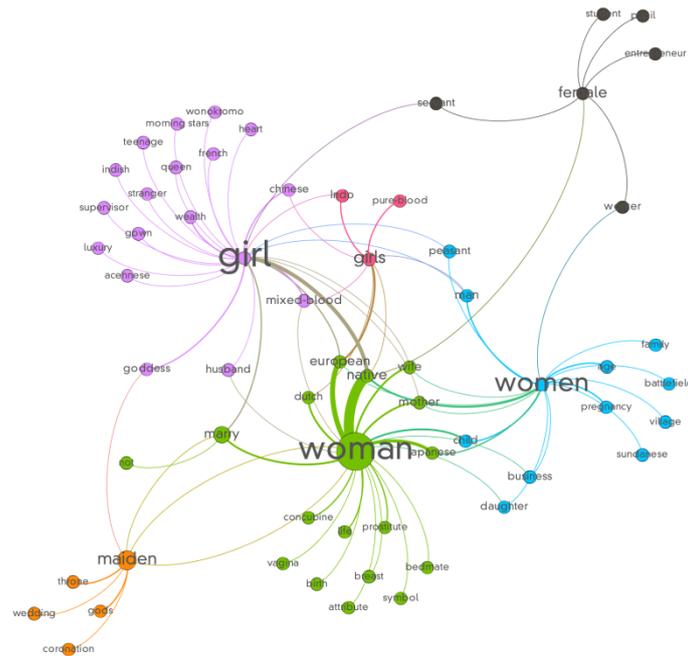


Figure 4. Network graph of persona based on modularity class for female gender-related words in the Indonesian (above) and English (below) corpus

The words *dara* and *maiden* interact with words that place female gender in the highest position such as *kekasih dewa* (god’s lover), *dewi* (goddess), *ratu* (queen), *penobatan* (coronation), and *tahta* (throne). Between female and male gender-related words in both languages, the words connecting the gender nodes include *pribumi* (native), *eropa* (European), *indo* (Indonesian), *pekerja* (worker), *ibu* (mother). These words are expected due to the general theme of the *Bumi Manusia* novel.

### Discussion

From the quantitative analysis and qualitative findings, it is identified that generally Pramoedya Ananta Toer did not differentiate the valuation between female and male genders based on the ATTITUDE analysis as part of the Appraisal Theory and the persona mapping, with the exception of the word *dara* in Indonesian and its English translation *maiden* that has a distinct semantic theme surrounding it. However, there are notable disparities between the Indonesian text and its English translation on how certain gender-related words interact with their surrounding in creating specific gender portrayals. This disparity sheds a light on the initial research question that focuses on the differences in achieving particular semantic goals in the Indonesian and English text of *Bumi Manusia*, and supply profound information on the scale of the possible loss in semantic values targeted in the original text.

The highlights of the analysis include:

- The number of terms interacting with female gender-related words are prominently higher than those interacting with male-gender related words across all aspects.
- Both female and male gender-related words interact mostly with CAPACITY values represented by words such as *educated*, *triumph*, *capable*, *captaincy*, and *capable* which are expected due to the general theme of the novel.
- Both gender-related words also interact mostly with DISTRESS/ANGUISH, CONTEMPT/DISGUST, INTEREST/EXCITEMENT, JOY/ENJOYMENT, and ANGER/RAGE which are expected as well due to the general theme of the novel.
- Based on the APPRECIATION analysis, the use of *perempuan* (woman) and *pemuda*

(young man) serve particular purposes. The word *perempuan* interacts only with the aesthetic category of ‘social valuation’ and the words surrounding *pemuda* suggest a portrayal of a strong and capable young man. In line with the Appraisal Theory (Martin & White, 2005), it can be inferred that *perempuan* is not intended to be associated with meanings other than appearance such as ‘significant’ and ‘harmful’, and *pemuda* serves a certain representation due to the categorization of its interacting words based on APPRECIATION values. Reflecting the research question, these certain portrayals of female and male genders are not achieved through the portrayal in the English version because the words *woman/women* and *young man* function more generally.

- The author uses *dara* and *maiden* to portray the highest form of female gender. This can be observed from their relation with words like *goddess* and *queen* in terms of persona. Among other words, *dara* and *maiden* are also associated the most with words with positive sentiments. This construction through Appraisal analysis relates to the theory on how we associate emotions and opinion to certain linguistic aspects such as specific words or syntactic patterns (Biber & Finegan 1989, Hunston 2011, Stein 1995, Martin & White, 2005)
- There is an inclination of polarity scaling for the female gender-related words in both languages where the words *dara* or *maiden* and *gadis* or *girl/girls* are noticeably interacting with more positive sentiments than *perempuan*, *wanita* or *woman/women*. This suggests that the words serve different semantic values in accordance with the second point from Esuli and Sebastiani (2006) in deciding if a given subjective text expresses a positive or negative opinion on its subject matter.
- There is also an inclination of deterministic value in terms of sentiment polarity which can be seen from the lower percentage of ‘undecided’ sentiment in the Indonesian corpus as opposed to the higher percentage in the English corpus throughout most of the analyses. This can also be an additional evidence to indicate a possible loss of semantic value in the translated text.

### Limitations

The main challenge of this research is to determine the terms constructing gender identity description which are sometimes characterized not only explicitly (tokens), but also implicitly (inscribed). The tokens, even though are explicitly mentioned, still pose an issue of ambiguity. The computational resources for the English language are adequate to disambiguate the meanings but for the Indonesian language, that is still considered under-resourced in terms of computational methods as this still requires manual-labelling. That being said, extracting the inscribed meaning is still a paramount challenge for all languages. The second limitation is the lack of language resources in semantic scaling which refrains this study to dive into the third point of Esuli and Sebastiani’s (2006) aspect which focuses on the strength of the text’s polarity.

### Recommendation

This paper can be used as a case-study in developing a comprehensive and annotated open dictionary, particularly for the Indonesian language that is compatible and congruent with libraries in any programming language to advance linguistic analyses. This can also include a consideration to explore the area of semantic scaling to provide a deeper comprehension of the text’s semantic orientation. Potential development of this study can also aim to expand the framework by incorporating dependency-based approaches such as dependency length and parsing to detect sentiment-conveying terms by examining the

relationship between the headword and its dependents (Ferrer-i-Cancho, 2004; Hudson, 2010; Liu, 2009).

### Conclusion

This study is an attempt to demonstrate how quantitative analysis and qualitative findings can provide a systematic analytical framework to analyze gender representation and possible identity construction in literature. Using parallel corpora of the novel *Bumi Manusia* written in Indonesian and its English translation *This Earth of Mankind*, this analysis combines approaches from the basic systemic functional linguistics framework (Halliday, 1994) and the Appraisal Theory, particularly the ATTITUDE types (Martin & White, 2005) with the Lövheim cube of emotion (Lövheim, 2011) to support the AFFECT analysis.

This study shows that there are prominent differences in the number of terms interacting with the female gender-related words which are greater than those with the male gender-related words in both languages. However, in terms of ratio, there is no significant difference or contrast based on both sentiment polarity and ATTITUDE types between female and male gender, potentially suggesting that the author's intention to illustrate an equal gender representation. The analyses also indicate that particular gender-related words such as *dara*, *perempuan* and *pemuda* in Indonesian and *maiden* in English interact with their surrounding words in specific ways. It also relates to the disparity of gender-related vocabulary between the two languages. Due to its specific ways of interaction, the words *dara*, *perempuan*, and *pemuda* in the Indonesian corpus manage to form distinct semantic themes that are different from the others as opposed to the words *woman/women* and *man/men* in the English corpus which function more generally. This poses a further question and possibly further studies of whether the translation process of a literary work can benefit from this study, particularly in studying the lexical and semantic relationship and networks of the original work in order to grasp a deeper understanding of gender (or beyond gender) representation.

### References

- Apsari, L. 2017. Attitude Analysis Surrounding Gender-Related Words in *Lelaki Harimau*. In *International Journal of Languages, Literature and Linguistics* Vol. 3, No. 3.
- Arnold, M. 1960. *Emotion and Personality*, New York: Columbia University Press.
- Baccianella, S., Esuli, A., & Sebastiani, F. 2010. SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. In *LREC* (Vol. 10, pp. 2200-2204).
- Benoit, K. & Nulty, P. 2017. "Quanteda: Quantitative Analysis of Textual Data". R package version: 0.9.9-24. URL <https://github.com/kbenoit/quanteda>
- Biber, D. & Finegan, E. 1989. Styles of stance in *English: Lexical and grammatical marking of evidentiality and affect*. *Text* 9: 93-124
- Crystal, D. 1992. *An Encyclopedic Dictionary of Language and Languages*. Oxford, UK: Blackwell.
- Esuli, A. & Sebastiani, F. 2006. SentiWordNet: A publicly available lexical resource for opinion mining. In *Proceedings of 5th International Conference on Language Resources and Evaluation (LREC)*, pp. 417-22. Genoa, Italy
- Ferrer-i-Cancho, R. 2004. Euclidean distance between syntactically linked words. *Physical review E*, 70, 056135.
- Halliday, M. A. K. 1994. *An introduction to functional grammar*. London: Edward Arnold

## COMPARATIVE ANALYSIS OF GENDER-RELATED WORDS IN

- Hudson, R. 2010. *An Introduction to Word Grammar*. Cambridge University Press, Cambridge.
- Hunston, S. 2011. *Corpus Approaches to Evaluation: Phraseology and Evaluative Language*. New York: Routledge
- Iedema, R., Feez, S. & White, P. R. R. 1994. *Media Literacy, Sydney, Disadvantaged Schools Program*. NSW Department of School Education.
- Jackson, H. 2007. *Key Terms in Linguistics*. London: Continuum.
- Jacomy, M., Venturini, T., Heymann, S., & Bastian, M. 2014. ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PloS one*, 9(6), e98679.
- Liu, H. 2009. *Dependency Grammar: From Theory to Practice*. Science Press, Beijing.
- Manning, C. & Schütze, H. 1999. *Foundations of Statistical Natural Language Processing*. Cambridge, MA: MIT Press.
- Martin, J. R. R. 1997. 'Analysing Genre: Functional Parameters', in *Genres and Institutions: Social Processes in the Workplace and School*, Christie, F. & Martin, J.R. (eds), London, Cassell: 3-39.
- Martin, J. R., & White, P. R. R. 2005. *The Language of Evaluation: Appraisal in English*. Palgrave, London, UK.
- Moilanen, K. & Pulman, S. 2007. Sentiment Composition. *Proceedings of the Recent Advances in Natural Language Processing International Conference*, pp. 378-382.
- Lövheim, H. 2011. "A New Three-dimensional Model for Emotions and Monoamine Neurotransmitters." *Medical Hypotheses* 78.2: 341-48.
- Perlez, J. 2006. *Pramoedya Ananta Toer, 81, Indonesian Novelist, Dies*. New York Times, Retrieved October 18, 2017
- R Core Team. 2017. R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. Vienna, Austria. URL <https://www.R-project.org/>
- Sinclair, J. M. 2005. *Trust the Text: Language, Corpus and Discourse*. London: Routledge.
- Stein, D. 1995. Subjective meanings and the history of inversions in English. In *Subjectivity and Subjectivisation: Linguistic Perspectives*, ed. D Stein, S Wright, pp. 129-50. Cambridge: Cambridge University Press
- Toer, P. A. 1980. *Bumi Manusia*. Indonesia: Hasta Mitra.
- Yu, C. S. and Lin, Y. W., Differentiating Strategy of online banking Service Quality, *Proceedings of the 7th Annual Global Information Technology Management Association World Conference*, CD-Format, Orlando, Florida, USA, June 11-13 2006.

