

## Comparison of Entities in Excerpts from the Novel *Perahu Kertas* and Its Translation *Paper Boats*

Dea Yuanita Nasution<sup>✉</sup>, Harni Kartika Ningsih

Universitas Indonesia, Indonesia

### Article Info

History Articles  
Received:  
10 October 2023  
Accepted:  
07 November 2023  
Published:  
30 December 2023

Keywords:  
entities, field of  
experience,  
translation, novel,  
taxonomic relations

### Abstract

A text is a product of the cognitive representation of its producer, including literary and translation texts. To identify the influence of this cognitive background, this research aims to explore shifts in the field of experience in the translation of novels by identifying participants constructing the field. The study utilizes a qualitative method with the IDEATION framework from Systemic Functional Linguistics, specifically taxonomic relations theory. A parallel corpus is used, extracted from excerpts of the Indonesian novel "Perahu Kertas" and its English translation titled "Paper Boats." The investigation is conducted using the AntConc software as a corpus processing instrument. The data for this research includes nouns categorized into concrete, abstract, and metaphoric entities. The results reveal two crucial findings. Firstly, everyday concrete entities dominate in both data. Secondly, there is a shift in the translation process of selected nouns, indicated by differences in the types and frequencies of entities between the two datasets. Besides the differences in the field of experience backgrounds of each text producer and the use of translation procedures, the two findings also suggest systematic differences in each language that influence the translator's choices and decisions underlying these shifts. This research provides a deeper understanding of the translation process and its impact on the representation of entities in the literary context.

<sup>✉</sup> Correspondence address:  
Jl. Margonda Raya, Kampus Baru Depok, Jawa Barat, Indonesia  
E-mail: deayuanitanasution3434@gmail.com

## INTRODUCTION

Literary works are a product of human cognition that reflect human's experiences. The human experiences then influence the construction of lexical elements built in the work. Each lexical element built in a literary work can determine various effects, sensations, and meanings that define the literary work. In the context of translating literary works, shifts in patterns of lexical elements between an original literary work and its translated version can undoubtedly be identified. This is because the creators of the source text (ST) and the target text (TT) are different individuals with distinct cognitive backgrounds. This discussion is often referred to as social cognition, a dimension explaining how individuals or groups of text makers produce a text (Putri, et al., 2022).

The literary work analyzed in this research is a novel. Although translation of literary works is considered the most challenging type of translation (Newmark, 1988), literary works require dynamic meanings. This dynamic nature provides a broad space for the translators to interpret meanings within the novels. This is reinforced by Landers (2001), who states that the translation of literary works requires the translator to consistently engage in the creative process. In the context of literary translation, a translator is required to make choices and decisions, collectively shaping the translation product (Landers, 2001). The broad scope for interpretation and authority held by a translator in determining their translation allows significant shifts of meanings between the ST and the TT. Therefore, novels are the well-suited type of literary work to provide the necessary data for this research.

A theoretical framework which may assist in revealing shifts of meanings is the field of linguistics, specifically Systemic Functional Linguistics (SFL) (Halliday and Matthiessen, 2014). SFL is the theory used to understand how a text constructs its meaning in a given context. SFL provides a framework for discussing how human experience (field of experience) influences the description built within a discourse, namely

Ideation (Martin, 1992; Martin and Rose, 2007). Ideation further explains that in all languages, the field of experience is built from a process involving people, objects, places, and qualities constructed behind the grammar, more specifically in each clause. Ideation itself is divided into three parts: taxonomic relations, nuclear relations, and activity sequences. Taxonomic relations are often used to examine the dynamics of pronouns in scientific texts, allowing to determine if a text contains a sense of everyday or specialized discourse. Additionally, the analysis using this framework is typically employed to facilitate researchers in examining the relationships between lexical elements that construct the field of experience and the field of activity within a discourse.

Several studies have employed the taxonomic relations framework to analyze a discourse. Grahmayanuri (2017), for instance, discusses semantic discourse analysis in *Parpokatan Orja* (a traditional discussion about the presentation of customary ceremonies in South Tapanuli) using taxonomic relations by Martin (1992) and Martin and Rose (2007). Unlike the study of Grahmayanuri (2017), which analyzes types of taxonomic relations including repetition, synonym, contrast, part, and class, this current research uses taxonomic relations to analyze types and shifts of entities in the case of literary translation. Research on the analysis of taxonomic relations in translations has also been conducted by Lubis and Khairuna (2019). In the article, they selected the English translations of ten surahs of the Quran from one of the popular Quran translation applications. The framework used is still the same, referring to the taxonomic relations framework by Martin and Rose (2007). The difference is that Lubis and Khairuna's (2019) article also analyzed the types of taxonomic relations based on patterns of lexical elements found, such as repetition, synonym, contrast, part, and class. Meanwhile, this current study focuses more on analyzing and comparing the movements of entities within the discourse.

Similar research on taxonomic relations in the translation of the Quran has been conducted by Iqbal and Grahmayanuri (2019), specifically

in Surah Ar-Rahman. However, similar to previous articles, this study also analyzes the types of taxonomic relations based on patterns of lexical elements. Lubis, et al (2019) have also published a research related to the analysis of taxonomic relation types found in the English translation of Surah Al-Baqarah. The framework and analysis conducted in this article are still the same, analyzing types of taxonomic relations based on patterns of lexical elements from Martin and Rose (2007). A recent and slightly different study that the researcher found regarding the use of the taxonomic relations framework in discourse analysis is an article by Lubis (2020). This study attempts to analyze the realization of repetition dominating in Surah Al-Baqarah. However, unlike the previous studies, the realization of repetition analyzed in this article uses the logogenetic pattern (following Halliday, 1992) to determine the density of the text in Surah Al-Baqarah.

From various similar studies related to taxonomic relations, it is found that previous studies generally discuss taxonomic relations that focus on the patterns formed among lexical elements in a discourse. It means that mostly the previous studies only analyze the frequency of repetition, synonym, contrast, class, and part within the discourse. Meanwhile, the scope of taxonomic relations also addresses the presence of entity elements (the overall depiction of people and objects) participating in building the field of experience and field of activity within the text. Considering previous research, this study addresses a gap by focusing on the complete picture of people and objects participating in building the field of experience and the field of activity in the text. It also integrates discourse analysis with corpus linguistics and the use of corpus processing software. Additionally, previous studies did not use the taxonomic relations framework to compare the movement of entities between the source and target texts. Therefore, this research aims to fill these gaps by connecting them to the translation domain.

This study focuses on the types of entities constructing the field of experience and the field

of activity within the text. It also compares the movements of entities in excerpts from the novel "Perahu Kertas" as the source text corpus (ST) and its translated novel "Paper Boats" as the target text corpus (TT). The novel was chosen because it is considered to have the potential of revealing abundant data regarding entity shifts that support this research. Created from the renowned uniqueness of its author's ideas, the novel contains many cultural markers and eccentric lexis. This richness contributes to the novel's philosophical values and meanings due to the inclusion of many words that are not commonly heard in daily language. Therefore, it is highly likely that there are many shifts in entities resulting from the translator's personal choices and decisions in building her best translation.

Martin and Rose (2007) state that taxonomic relations consist of relationships between lexical elements that depict the chain relationship among those lexical elements and portray the entire picture of people and objects participating in elaborating the field of experience and field of activity in discourse. The chain relationship among lexical elements in a text is illustrated by the presence of repetition, synonym, and contrast, each of which is further divided into its types. The presentation of the entire picture of people and objects participating in elaborating the field of experience and the field of activity is analyzed by the type of concrete, abstract, and metaphoric entities, each of which is also divided into types. The field of experience refers to how human experience as the producer of discourse is constructed in a text, while the field of activity refers to the 'sense' or nuances built within the text. If referring to taxonomic relations as the complete picture of people and objects, taxonomic relations can be used to identify the types of entities to observe participants (nouns) involved in constructing both fields of the text. The types of entities are detailed in the table adapted from Martin and Rose (2007) as follows:

**Table I.** Kinds of Entities

| Kinds of Entities |                      | Definition   | Examples  |
|-------------------|----------------------|--|---|
| <i>Concrete</i>   | <i>Everyday</i>      | Referring to concrete, tangible nouns commonly used in everyday life and understood by the common people | Man, girlfriend, face, hands, apple, house, hill    |
|                   | <i>Specialized</i>   | Referring to concrete nouns that are tangible but typically only known by those involved in the field.   | Mattock, lathe, gearbox                             |
| <i>Abstract</i>   | <i>Technical</i>     | Referring to abstract concepts generally understood only by professionals engaged in the field.          | Inflation, metafunction, gene                       |
|                   | <i>Institutional</i> | Referring to terms typically used exclusively within one institution.                                    | Offence, hearing, applications, violation, amnesty, |
|                   | <i>Semiotic</i>      | Referring to language features.  | Question, issue, letter, extract                    |
|                   | <i>Generic</i>       | Referring to dimensions of terms that have classes and parts.  | Colour, time, manner, way, kind, class, part, cause |
| <i>Metaphoric</i> | <i>Process</i>       | Referring to nouns resulting from the nominalization of verbs.   | Relationship, marriage, exposure, humiliation       |
|                   | <i>Quality</i>       | Referring to nouns resulting from the nominalization of adjectives.                                      | Justice, truth, integrity, bitterness, security     |

This article aims to explore lexical patterns in the translation of a literary work using the IDEATION framework (Martin, 1992). In support of Colina's theory (2015), this exploration is necessary to assess the success of a translation in terms of the degree of resemblance between the original literary work and its translated version. More specifically, the analysis in this research utilizes the taxonomic relations framework to further examine the comparison of the entities in a source text (ST) and a target text (TT). The objectives are formulated into two questions:

1. What entities are found and dominant in the excerpt from the novel "Perahu Kertas" (ST) and the excerpt from the novel "Paper Boats" (TT)?

What changes occur in entities between the excerpt from the novel "Perahu Kertas" (ST) and the excerpt from the novel "Paper Boats" (TT)?

## METHODOLOGY

This research adopts a qualitative method with a discourse analysis design and corpus linguistic software as a tool to trace data frequencies. The discourse analysis design used in this study is the Ideational meaning system of Systemic Functional Linguistics, specifically taxonomic relations (Martin, 1992; Martin & Rose, 2007). This framework is employed to identify the types of entities found in the selected novel excerpts. Meanwhile, the corpus linguistic tool used is AntConc version 3.5.9 (2020). The tool is employed to sift through data by examining collocations and word frequencies in the corpus.

The data source for this research is the novel "Perahu Kertas" by Dewi Lestari, first published in 2009 and translated into English by Tiffani Tsao in 2017, with the title "Paper Boats". "Perahu Kertas" was chosen as the data source of this research due to its popularity. Besides being a work by Dewi Lestari, a renowned author with many other works, this novel is her first bestseller,

leading to its adaptation into a film with the same title. The novel contains numerous philosophical values and cultural markers due to the presence of unfamiliar words in the everyday Indonesian language, the mixture of the Dutch language, and the main character's obsession with Roman mythology and literature.

Considering the substantial number of words in both "Perahu Kertas" and its translation "Paper Boats," data sampling was conducted to select several excerpts to serve as the corpus. The resulting corpus from both data sources includes the ST corpus (in Indonesian) containing Chapter 1 and 2 of the novel "Perahu Kertas" and the TT corpus (in English) containing Chapter 1 and 2 of

the translated novel "Paper Boats". Chapter 1 and 2 of each novel were selected as the corpus to be analyzed in this study because these initial chapters contain character introductions and the storyline, making it likely to find numerous nouns for examination. In these first two chapters, the story unfolds around two main characters, Kugy and Keenan, who initially are in different countries but later reunite and begin their journey. These initial chapters also provide insights into character introductions, settings, and cultural markers, which are sought as data in this research. The following is detailed information about the two corpora analyzed using the AntConc software.

**Table II.** Detail Information of the Corpus

|                               | <b>ST Corpus</b> | <b>TT Corpus</b> |
|-------------------------------|------------------|------------------|
| <b>Language</b>               | Indonesian       | English          |
| <b>Number of Words</b>        | 4.326            | 5.686            |
| <b>Number of <i>Token</i></b> | 4.415            | 5.870            |
| <b>Number of <i>Types</i></b> | 1.595            | 1.521            |

The data in this study consist of lexical elements categorized into one of the entities according to Martin and Rose (2007), namely concrete everyday, concrete specialized, abstract technical, abstract institutional, abstract semiotic, abstract generic, metaphoric process, and metaphoric quality. Additionally, the data extracted from each corpus are limited, considering only nouns with a minimum frequency of two, which will be considered as data in this research. In other words, nouns that appear only once in the corpus will not be selected. This is done to better define the entities that truly build the text through repetition.

The data processing steps begin technically with the acquisition of data sources. First, each novel is downloaded in .pdf format, then

converted into .txt format using AntFileConverter. After that, the selected data samples, namely Chapter 1 and Chapter 2 from each novel, are transferred to Notepad++ with UTF-8 encoding to ensure that all entered characters can be read, and the format remains intact. Afterwards, the ST and TT corpus are arranged in parallel based on sentences and cleaned of unnecessary parts such as page numbers, headers, footers, and so on. The prepared corpora, in .txt format, are alternately uploaded to AntConc to observe the frequency of the most common words and their collocations in each corpus. The data selection is performed by sorting the nouns that fall into one of the entity types as outlined in the taxonomic relations framework mentioned earlier.

AntConc  
File Edit Settings Help

Target Corpus  
Name: temp  
Files: 1  
Tokens: 5870

KWIC Plot File Cluster N-Gram Collocate Word Keyword Wordcloud  
Types 1521/1521 Tokens 5870/5870 Page Size 100 hits 1 to 100 of 1521 hits

|    | Type    | Rank | Freq | Range |
|----|---------|------|------|-------|
| 44 | school  | 42   | 19   | 1     |
| 45 | them    | 42   | 19   | 1     |
| 46 | when    | 42   | 19   | 1     |
| 47 | if      | 47   | 18   | 1     |
| 48 | this    | 47   | 18   | 1     |
| 49 | time    | 47   | 18   | 1     |
| 50 | mother  | 50   | 17   | 1     |
| 51 | who     | 51   | 16   | 1     |
| 52 | bandung | 52   | 15   | 1     |
| 53 | by      | 52   | 15   | 1     |
| 54 | ll      | 52   | 15   | 1     |
| 55 | not     | 52   | 15   | 1     |
| 56 | your    | 52   | 15   | 1     |
| 57 | d       | 57   | 14   | 1     |
| 58 | go      | 57   | 14   | 1     |
| 59 | into    | 57   | 14   | 1     |
| 60 | other   | 57   | 14   | 1     |
| 61 | re      | 57   | 14   | 1     |
| 62 | an      | 62   | 13   | 1     |
| 63 | father  | 62   | 13   | 1     |

Korpus Paper Boats (Bab 1 & Ba

Search Query  Words  Case  Regex  
Start  Adv Search

Sort by Frequency  Invert Order

**Figure I.** Display of Corpus Processing Using AntConc

From the numerous nouns that have been sorted, the researcher then categorizes them into several tables. The first table is created to compare and observe the frequency shifts among similar nouns (those with equivalent meanings) between the ST and TT corpus. Subsequently, the next two tables are generated to categorize nouns according to their respective entity types found in both the ST and TT corpus. Finally, the overall frequency and entity type shifts between the ST and TT corpora are accumulated into a table comparing the number of nouns based on their entity types. Through this data analysis, the researcher can observe the shifts and relationships among lexical elements, especially nouns of people and objects participating in building the field of experience in the text, and how the

translator selects and maintains the translation through the entities that appear.

## RESULTS AND DISCUSSION

This research reveals several findings, one of which is the frequency shift of nouns between the ST and TT corpus. This finding was obtained through the process of calculating selected words or nouns as data. These nouns were then compared by arranging them in parallel according to their meanings. The arrangement of nouns in the table was also done based on the frequency, namely from the most frequent to the least frequent nouns. The following table provides an example of the comparison and frequency shift of nouns found between the ST and TT corpus:

**Table III.** Example of the Comparison and Frequency Shift of Nouns Found in the ST and TT corpus:

| No. | ST      | Frequency | TT                  | Frequency |
|-----|---------|-----------|---------------------|-----------|
|     | Orang   | 19        | People/person       | 7         |
|     | Buku    | 17        | Book/books          | 15        |
|     | Air     | 16        | Water               | 9         |
|     | Anak    | 14        | Child/kids/children | 10        |
|     | Mata    | 13        | Eye/eyes            | 7         |
|     | Rumah   | 12        | House               | 7         |
|     | Pintu   | 10        | Door                | 11        |
|     | Tahun   | 10        | Year/years          | 10        |
|     | Telepon | 10        | Phone               | 7         |
|     | Kota    | 9         | City                | 7         |
|     | Kamar   | 8         | Room/bed            | 17        |
|     | Laut    | 8         | Sea                 | 6         |
|     | Mama    | 8         | Mom/mother          | 25        |
|     | Mobil   | 7         | Car                 | 4         |
|     | Sekolah | 7         | School              | 19        |

Table III above presents 15 examples of data out of a total of 106 noun types found in the ST corpus and 93 noun types found in the TT corpus after applying the data sampling technique. From this table, it is evident how noun shifts occur in the translation process. One of the most significant shifts is directly indicated by the noun that appears most frequently in the ST corpus, namely the noun "orang", which is translated as "people" or "person" in the TT corpus. The noun "orang," which initially appeared 19 times in the ST corpus, then changed after the translation process, appearing only 7 times in the TT corpus. This finding also indicates that the frequency changes of translated noun equivalents are dynamic. There are instances where the frequency decreases from ST to TT, and others where it increases significantly, as shown by the nouns "kamar" and "mama" with equivalents "room" or "bed" and "mom" or "mother" in English. Additionally, some show the same frequency in both ST and TT, such as the noun "tahun" with equivalents "year" or "years" in English.

This phenomenon is, of course, caused by the translation process itself. The translation process generally involves decision-making by the translator to achieve the translation's purpose (skopos). Colina (2015) explains that translation is the transfer of text from one language to

another with a significant degree of resemblance to the source text. Furthermore, Colina (2015) clarifies that the term "degree of resemblance" refers to the general similarity between the source and target languages, not detailed similarities such as structural or semantic equivalence. This is because no two languages in the world are entirely identical. Additionally, other factors, such as the use of specific translation strategy theories, are believed to contribute to the dynamic shifts in noun frequency.

Another finding after categorizing nouns according to their respective entity types is the difference in the types and frequencies of entities found between the ST and TT corpus. Tables IV and V below provide examples of the categorization of nouns based on their entity types. From the overall analysis of data in the ST and TT corpus, a total of 182 data were obtained from categorizing nouns based on their entity types. However, only a maximum of ten examples of nouns from each entity type are displayed in the two tables below to optimize space in this article. The total of 182 data is divided into ST and TT data, with the number of data coincidentally being the same between ST and TT. There is a total of 91 data in ST, consisting of 68 data classified as concrete everyday entities, 13 data classified as abstract generic entities, 8 data classified as metaphoric

process entities, and 2 data classified as metaphoric quality entities. Meanwhile, from the 91 data found in STa, 72 data fall into concrete everyday entities, 12 data into abstract generic entities, and 7 data into metaphoric process

entities. The list of nouns in both tables below is also arranged based on their frequency in the ST corpus, namely from the most to the least frequent. The detailed data is completely displayed in Tables IV and V below:

**Table IV.** Examples of Entities Found in the ST Corpus:

| No | Concrete |             | Abstract  |               |          | Metaphoric |            |            |
|----|----------|-------------|-----------|---------------|----------|------------|------------|------------|
|    | Everyday | Specialized | Technical | Institutional | Semiotic | Generic    | Process    | Quality    |
| 1  | Orang    | x           | x         | x             | x        | Tahun      | Bacaan     | Kebiasaan  |
| 2  | Buku     |             |           |               |          | Hari       | Pandangan  | Legalisasi |
| 3  | Air      |             |           |               |          | Sore       | Penampilan |            |
| 4  | Anak     |             |           |               |          | Waktu      | Penekanan  |            |
| 5  | Mata     |             |           |               |          | Arah       | Pertanyaan |            |
| 6  | Rumah    |             |           |               |          | Detik      | Pelukis    |            |
| 7  | Pintu    |             |           |               |          | Jam        | Penulis    |            |
| 8  | Telepon  |             |           |               |          | Menit      | Tulisan    |            |
| 9  | Kota     |             |           |               |          | Minggu     |            |            |
| 10 | Kamar    |             |           |               |          | Pagi       |            |            |

**Table V.** Examples of Entities Found in the TT Corpus:

| No | Concrete |             | Abstract  |               |          | Metaphoric |              |         |
|----|----------|-------------|-----------|---------------|----------|------------|--------------|---------|
|    | Everyday | Specialized | Technical | Institutional | Semiotic | Generic    | Process      | Quality |
| 1  | School   | x           | x         | x             | x        | Time       | Reading      | x       |
| 2  | Mother   |             |           |               |          | Years      | Attention    |         |
| 3  | Father   |             |           |               |          | Week       | Departure    |         |
| 4  | Books    |             |           |               |          | Afternoon  | Entrance     |         |
| 5  | Room     |             |           |               |          | Day        | Expression   |         |
| 6  | Door     |             |           |               |          | Age        | Preparation  |         |
| 7  | Face     |             |           |               |          | Minutes    | Relationship |         |
| 8  | Water    |             |           |               |          | Morning    |              |         |
| 9  | Bed      |             |           |               |          | Night      |              |         |
| 10 | Mom      |             |           |               |          | Centuries  |              |         |



From both tables above, the most striking observation that can be immediately inferred is the difference in the types of entities between the ST and TT corpus. In the ST corpus, four types of entities were found, namely concrete everyday entities, abstract generic entities, metaphoric process entities, and metaphoric quality entities. No concrete specialized, abstract technical, abstract institutional, and abstract semiotic entities were found at all in the ST corpus. Meanwhile, there are only three types of entities found in the TT corpus, namely concrete everyday entities, abstract generic entities, and metaphoric process entities. Not a single noun falls into the categories of concrete specialized, abstract technical, abstract institutional, abstract semiotic, and metaphoric quality entities in the TT corpus. This certainly indicates a shift or change, both in terms of entity types and the number of nouns found in each entity type. Again, this occurs because the data source used in this study is a translation. As stated by Hoed (2006), translation is the process of transferring a written message from ST to TT using a different language from the ST. This definition further clarifies that there is never a translation process from ST to TT without any shifts, as the language systems used are already different.

Additionally, the nouns in both tables above are arranged in sequence according to the frequency of nouns that appear most frequently in

both the ST and TT corpus. This is to facilitate the researcher in comparing the shifts in the appearance of nouns in each corpus. Therefore, from Table IV above, it can be concluded that the nouns "orang" (people), "tahun" (year), "bacaan" (reading), and "kebiasaan" (habit) are the most frequently appearing nouns in the ST corpus based on their respective entity types. Meanwhile, from Table V, it can be concluded that the nouns "school," "time," and "reading" are also the most frequently appearing nouns in the TT corpus based on their respective entity types. This certainly reveals new findings that clearly indicate shifts in noun types and entities between the ST and TT corpus. These shifts are undoubtedly caused by various factors, such as the translator's decision-making in determining the translation, the use of specific translation procedures, and the differences in language systems between Indonesian and English. A concrete depiction of the process of noun and entity shifts will be further explained in Table VII below.

After calculating and categorizing the whole data, the next step is to calculate the data according to their entity types to observe the comparison of frequencies and shifts in entity types between the ST and TT corpus. The overall comparison and shift in noun frequencies based on their entity types are displayed in the following table:

**Table VI.** Comparison of Entity Frequencies between the ST and the TT Corpus:

| No    | ST                    | Frequency | Percentage | TT                    | Frequency | Percentage |
|-------|-----------------------|-----------|------------|-----------------------|-----------|------------|
| 1     | Concrete<br>Everyday  | 68        | 75%        | Concrete<br>Everyday  | 72        | 79%        |
| 2     | Abstract<br>Generic   | 13        | 15%        | Abstract<br>Generic   | 12        | 13%        |
| 3     | Metaphoric<br>Process | 8         | 9%         | Metaphoric<br>Process | 7         | 8%         |
| 4     | Metaphoric<br>Quality | 2         | 1%         | Metaphoric<br>Quality | 0         | 0%         |
| Total |                       | 91        |            | Total                 | 91        |            |

From the calculation in Table VI above, there is a total of 91 nouns classified into four types of entities that contribute to building the

field of experience in the ST corpus. These 91 nouns are further divided into 68 nouns classified as concrete everyday entities, 13 nouns classified

as abstract generic entities, eight nouns classified as metaphoric process entities, and two nouns classified as metaphoric quality entities. From all the types of entities found, it can be seen that the concrete everyday entities significantly dominate in terms of frequency. Additionally, there happens to be an equal number of data found in the ST and TT corpus. Despite having the same number of data, there are still differences in the frequency and types of entities that appear between the ST and TT corpus. There are three types of entities in the 91 nouns found in the TT corpus, namely 72 nouns classified as concrete everyday entities, 12 nouns classified as abstract generic entities, and seven nouns classified as metaphoric process entities. These data reveal the same finding which show that the concrete everyday entities are still significantly dominating in terms of frequency.

Although there are differences in the types and frequencies of entities that appear between the two corpus, Table VI above also shows that both corpora still maintain the same proportion. This is indicated by the relatively similar percentages of each type of entity. The consistency in patterns and proportions is shown through the percentage of concrete everyday entities dominating in both corpora, followed by abstract generic entities, and lastly by metaphoric process entities, which are found least in both corpus. This means that the translator can be

considered successful in maintaining the substance of the field of experience contained in the original work by preserving the proportions of each entity. The four types of entities found also indicate that the participants involved in constructing the field in this text are common lexicons that are easily understood, and making this novel as a light reading type. This is further supported by the analysis results showing the absence of concrete specialized, abstract technical, abstract institutional, and abstract semiotic entities, indicating the lack of technical elements and making it as a familiar reading material.

The difference in the types of entities found between the ST and TT corpus, specifically the absence of metaphoric quality entities (nominalizations formed from adjectives) in the TT corpus, also indicates how the shift in the translated text occurs. The shift in the frequency of nouns found between the ST and TT corpus further demonstrates how the translator translates and maintains the translation. This can be seen from the quantity of several examples of nouns that have decreased in the TT corpus. For example, the noun 'orang' (people), 'buku' (book), and 'air' (water) that undergo a significant decrease. The examples of the changes that occurred in the source data are displayed in the following table:

**Table VII.** Examples of Entity Changes Found in Some Sentences

| No | ST   | TT   |
|----|--|--|
| 1  | Dialah <u>orang</u> yang paling menunggu-nunggu Kugy selesai berkemas supaya bisa langsung cabut ke Bandung. | <u>Noni</u> couldn't wait for Kugy to finish packing so she could join her in Bandung. |
| 2  | Suatu hari ia bukan hanya seorang kolektor <u>buku</u> dongeng.  | (tidak diterjemahkan/dihapus)  |
| 3  | " <u>Air sungai</u> bakal sampai ke laut?"   | "All <u>rivers</u> flow to the sea?"   |

From Table VII above, we can see that the changes and shifts in frequency and forms of nouns are caused by several factors. These include the fact that each language has features and language styles that characterize its respective grammatical system, the use of

possessive pronouns, explicitation elements, the use of translation procedures (shifting meanings), deletions, and any other factors that result in the differences in the frequency and types of entities found between the two corpus. These findings also illustrate how the field of experience

constructed between the two novels differs. In the ST novel, the field of experience is described with very casual and less diverse language, resulting in many noun repetitions. Meanwhile, in the TT novel, the field of experience has shifted. It is depicted through the reduction of casualness in language use and more varied use of nouns.

This difference is undoubtedly due to the different producers creating the text. As stated by Sholikhati and Mardikantoro (2017), a text is considered as a realm of representation of the understanding and experience of the text maker. Between the two text producers, there are certainly differences in the field of experience as they are different individuals with different native languages and cultures. The differences and shifts also indicate that no language transfer process that will produce a translation identical to the original work. Through the discussion of the analysis results showing similarities and differences in the frequency and types of entities found between the ST and TT corpus, we can also see how the translator's style and approach in transferring and maintaining the translation. The comparison of entity movements between the ST and TT corpus concisely demonstrates how language shifts occur between the two texts.

## CONCLUSION

This study aims to answer two questions: the dominant types of entities in the ST and TT corpus and the changes of entities between the two corpus. Four types of entities were found in the ST corpus are concrete everyday, abstract generic, metaphoric process, and metaphoric quality. Meanwhile, there is a difference as three types of entities were found in the TSa corpus. They are concrete everyday, abstract generic, and metaphoric process. Although there are changes in the types and frequencies of entities between the two corpora, the percentage patterns for each entity type remain relatively consistent. This is indicated by the percentage of concrete everyday entities dominating in both corpora, followed by abstract generic entities, and lastly by metaphoric process entities, which are found least in both corpora. This means that the translator is

considered successful in maintaining the substance of the field of experience contained in the original work by preserving the proportions of each entity.

A clearer pattern of shifts is actually reflected through the significant differences in the quantity of some nouns between the ST and TT corpus. For example, the nouns like "orang" (people) and "air" (water), which initially appeared more than a dozen times in the ST corpus, but then appeared less than 10 times in the TT corpus after the translation process. The differences and shifts, in both the types and frequencies of entities, including the frequency of noun comparisons between the ST and TT corpus, further reinforce the idea that, no matter how faithful a translator tries to translate, marked by the similarity of entity proportions between the two corpus, a translated work will always experience shifts compared to its source text. Broadly, this is due to various factors. First, there are differences in grammatical structures and language systems in each language. Second, there are differences in the field of experience of each text producer because they are different individuals with distinct native languages and cultures. Third, there are decisions made by translators in transferring messages based on the skopos and the use of translation procedures such as explicitation, deletion, addition, and others.

To this point, this study has only utilized parallel corpora (i.e., STu and STa) from excerpts of a novel, analyzed using the taxonomic relations framework by Martin (1992) and Martin and Rose (2007) to identify shifts in entities. This study does not examine other aspects of discourse and grammar that can indicate pronouns as substitutes for marked entities, so further analysis considering these aspects can be conducted in future research. Further research on the entire contents of the novel and research analyzing comparable corpora is suggested. Additionally, a similar study would be more comprehensive if it is applied to analyze other types of texts, such as academic texts and journalistic texts that generally contain a more diverse range of nouns and a richer variety of entity types.

## REFERENCES

- Budianta, M., Husen, I. D., Budiman, M., & Wahyudi, I. (2006). *Membaca sastra*. Magelang: IndonesiaTera.
- Colina, S. (2015). *Fundamentals of translation*. Cambridge: Cambridge University Press.
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. John Benjamins.
- Grahmayanuri, N. (2017). Ideational taxonomic relations of hobar on parpokatan orja of south tapanuli. *The 2nd Annual International Seminar on Transformative Education and Educational Leadership (AISTEEL) 2017*, 16-17 Oktober 2017, Medan.
- Halliday, M. A. K., & Matthiessen C. M. I. M. (2014). *Halliday's introduction to functional grammar*. Routledge.
- Hoed, B. H. (2006). *Penerjemahan dan kebudayaan*. Jakarta: Pustaka Jaya.
- Iqbal, M., & Grahmayanuri, N. (2019). Taxonomic relations in surah Ar-Rahman. *Anglo-Saxon: Journal of the English Language Education Study Program*, 10(1), 87-94.
- Landers, C. E. (2001). *Literary translation: A practical guide*. Cromwell Press Ltd.
- Lubis, S. I. S. (2020). The text density from the realization of taxonomic relation in english translation of surah Al-Baqarah. *English Education: English Journal for Teaching and Learning*, 8(2), 124-137.
- Lubis, S. I. S., & Khairuna, R. (2019). Taxonomic relation in the quran. *In Sixth of International Conference on English Language and Teaching (ICOELT 2018)* (pp.163-166). Atlantis Press.
- Lubis, S. I. S., Murni, S. M., & Zainuddin. (2019). Ideational taxonomic relation of surah Al-Baqarah in english translation. *In 4th Annual International Seminar on Transformative education and Educational Leadership (AISTEEL 2019)* (pp. 88-92). Atlantis Press.
- Martin, J. R. (1992). *English text: System and structure*. Philadelphia: John Benjamins.
- Martin, J. R., & Rose, D. (2007). *Working with discourse: meaning beyond the clause*. (Second Ed.). London: Continuum
- Newmark, Peter. (1988). *A textbook of translation*. Hertfordshire: Prentice Hall.
- Putri, M., Mardikantoro, H. B., & Baehaqie, I. (2022). Critical discourse analysis of Dua Garis Biru novel (novel adaptation of Lucia Priandarini from Gina S. Noer's scenario). *Seloka: Jurnal Pendidikan Bahasa dan Sastra Indonesia*, 11(3), 283-292.
- Sholikhati, N. I., & Mardikantoro, H. B. (2017). Analisis tekstual dalam konstruksi wacana berita korupsi di Metro TV dan NET dalam perspektif analisis wacana kritis Norman Fairclough. *Seloka: Jurnal Pendidikan Bahasa Dan Sastra Indonesia*, 6(2), 123-129.
- Suryawinata, Z., & Hariyanto, S. (2003). *Translation: Bahasan Teori & Penuntun Praktis Menerjemahkan*. Yogyakarta: Kanisius.