Textual Metadiscourse Resources (MTRs) in Original and Simplified Versions of *Oliver Twist*

Sonour Esmailli  
Islamic Azad University, Tabriz Branch, Iran

Karim Sadeghi  
Urmia University, Iran  
Email: k.sadeghi@urmia.ac.ir

Abstract—The present study was carried out to compare the use of metadiscourse textual resources (MTRs) in terms of frequency and categorical distribution in an original novel and its simplified counterpart (i.e., *Oliver Twist*). The researchers were interested to understand how writers of original and simplified novels made use of metadiscourse elements and whether original and simplified novels differed in the use of these resources. The corpus used in this study was analyzed based on Hyland and Tse's (2004) model. The frequency and categorical distribution of MTRs was calculated per 1,000 words and the difference in their distribution across original novels and their simplified versions was checked using the statistical analysis technique of Chi-Square. The findings indicated that there was not a significant difference in the frequency and categorical distribution of MTRs between the original and simplified versions of *Oliver Twist*, implying that the writers of both original and simplified novels strived to provide a coherent text.

Index Terms—metatexual resources, original novel, simplified novel, text analysis

I. INTRODUCTION

Written language has been regarded as a primary means of communication for several centuries, and literature has been viewed as an example of linguistic excellence, mediated mainly through written language. The elucidation and teaching of the rules of language, accordingly, was restricted to written texts (Faghih & Rahimpour, 2009). Narrative genre of writing in general and stories in particular are well-represented in some form in the language learners' schemata, and this may be a valuable touchstone in introducing interesting, authentic reading material in a form with which the reader is familiar (Harper, 1990). The debate over the appropriateness of using literature in second language classrooms as a relevant source in developing reading, writing and other skills is well-documented (Widdowson, 1984).

Critics of the use of literature for second language instruction cite the potential difficulties caused by the uncontrolled vocabulary of much native language literature. Unfamiliar or inappropriate content may serve as obstacles to comprehension rather than means toward understanding a writer's message. In addition, the stylistic nature of literature often involves the use of more complex syntactical structures, as well as figurative use of language, thus increasing the complexity of reading (Harper, 1990). In order to solve some of these problems associated with the use of original novels in EFL classes, their simplified versions are usually used which of course are not without their problems. For example persuasion is achieved differently in simplified and original novels. Persuasion, as a part of the rhetorical structure, is partly achieved by employing metadiscourse in all kinds of writing. Metadiscourse is the linguistic expressions which refer to the evolving text and to the writer and imagined readers of that text. Although the presence and function of metadiscourse markers has been examined in a number of different contexts, including textbooks (Hyland, 1999, 2000), student writing (Crismore, 1983), science popularizations (Crismore & Farnsworth, 1990) and research articles (Mauranen, 1993; Hyland, 1998, 1999, 2000), surprisingly little attention has been given to the genre of literary discourse. Thus, this study tries to investigate the use of textual metadiscourse resources in an original novel and its simplified version to find out whether there is any significant difference in the original and simplified version of *Oliver Twist* (OT). In other words, this study was an attempt to find answers to the following questions:

1. Is there any significant difference between the original version of OT and its simplified version in the frequency of use of metadiscourse resources?
2. Is there any significant difference in the type of metadiscourse resources employed in the original and simplified versions of OT?

The questions were tentatively answered in the form of null-hypotheses to be tested at 0.05 level of significance.

II. REVIEW OF THE RELATED LITERATURE

A. What is Metadiscourse?
In the 1970s, the study of texts shifted from formal aspects of writing to the organization and structuring of overall discourse. What followed was a “deeper and narrower” (Swales, 1990, p. 3) approach which focused on specific genres and attempted to investigate not only formal features but also communicative aspects of written texts. By the early 1990s, linguists had begun to react against the strong emphasis on propositional meaning in text analysis. This movement resulted in a range of new perspectives on text, among which studies of metadiscourse have gained prominence. The term metadiscourse, according to VandeKopple (2002), goes back to the work of linguist Zellig Harris in 1959 to offer a way of understanding language in use, representing a writer's or speaker's attempts to guide a receiver's perception of a text. According to Hyland (2005), students are often told that successful writing in English is ‘reader-friendly’. It must fit together logically, be signposted to guide readers, and take their likely responses and processing difficulties into account. But it also needs to work for the writer too, as we communicate for a reason: we use language to persuade, inform, entertain or perhaps just engage an audience, and this means conveying an attitude to it with hedges, boosters, self-mentions, and so on. The concept of metadiscourse derives from the postulate that people use language not only to convey ideational meaning, i.e. meaning based on information about the world, but that this referential meaning is complemented with other dimensions of linguistic meaning (Hempel & Degand, 2008).

**B. Classification of Metalinguistic Devices**

Within Systemic Functional Linguistics (SFL), language is seen as being organized around, and simultaneously realizing, three broad purposes or 'metafunctions' (Hyland, 2005). Halliday's (1994) macro-functions provide the most appropriate basis for classification of metalinguistic devices. In Halliday’s grammatical theory, the ideational or referential function represents the external world, including not only the representation of physical experiences and internal/mental processes, such as, thoughts and feelings (the experiential subfunction), but also the fundamental logical relations that exist among these experiences and processes (the logical subfunction). The logical relations are expressed linguistically through syntactic devices of co-ordination, indirect speech, etc. The interpersonal function encompasses the relations between the addressee and the addressee in a discourse situation or speech event. Linguistically, the interpersonal function is realized through the use of first/second person pronouns and speech acts, such as, questions and directives. The textual function is concerned with the way language establishes links with itself and the situation to produce text that is linguistically cohesive and semantically coherent.

Generally, metadiscoursal comments are argued to have two main functions: textual and interpersonal. The first kind helps to organize the discourse by pointing out topic shifts, signaling sequences, cross referencing, connecting ideas, previewing material, and so on. The second kind modifies and highlights aspects of the text linguistically through syntactic devices of co-ordination, indirect speech, etc. The interpersonal function realizes the relations between the addressee and the addressee in a discourse situation or speech event. Linguistically, the interpersonal function is realized through the use of first/second person pronouns and speech acts, such as, questions and directives. The textual function is concerned with the way language establishes links with itself and the situation to produce text that is linguistically cohesive and semantically coherent.

1. Transition markers: a set of devices which used to mark additive, contrastive, and consequential steps in the discourse, as opposed to the external world and help readers to interpret pragmatic connections between steps in an argument. They consist of items such as: *in addition, but, thus, and, etc.*
2. Frame markers: they are elements that show schematic text structure and are composed of items used to sequence, to label text stages, to announce discourse goals, and to indicate topic shifts: *my purpose here is to, to conclude, etc.*
3. Endophoric markers: they are expressions that refer to other parts of the text to make additional information available to the reader, such as: *as noted above, see Fig. in section 2, etc.*
4. Evidentials: they are the source of information from others such as: *Z states, According to X, etc.*
5. Code glosses: elements that provide additional information by explaining, rephrasing or elaborating what has been said as: *in other words, e.g., etc.*

**C. Key Principles of Metadiscourse**

Hyland (2005) proposes a functional model of metadiscourse. A model or theory which is based on the assumption that the rhetorical features of metadiscourse can be understood more clearly when they are used or identified in contexts in which they occur. Hence the analyses of metadiscourse have to be conducted as part of that particular context or as part of that particular community practices, values and ideas. The functional theoretical framework of metadiscourse defines writers as the conductors of interaction with the readers. The notion of the writer-reader interactions has underpinnings on the following three key principles of metadiscourse proposed by Hyland and Tse (2004):

- **Metadiscourse** is distinct from propositional aspects of discourse.
- **Metadiscourse** refers to aspects of the text that embody writer-reader interactions.
- **Metadiscourse** refers to only those relations that are internal to the discourse. (p. 159)

The common thread in definitions of metadiscourse is that it concerns meanings other than propositions. VandeKopple’s view (2002, p. 93) of different levels of meaning is as follows: “On one level we expand ideational
material. On the levels of metadiscourse, we do not expand ideational material but help our readers connect, organize, interpret, evaluate, and develop attitudes toward that material. Based on Hyland (2005), the point to be made here is that, VandeKopple and others are simply wrong in stating that metadiscourse is a separate ‘level of meaning’. Texts are communicative acts, not lists of propositions. The meaning of a text depends on the integration of its component elements, both propositional and metadiscoursal, and these do not work independently of each other. Metadiscourse is an essential part of any text and contributes to the ways it is understood and acted upon; it is not a separate and separable set of stylistic devices that can either be included or not without affecting how a text is presented and read.

The second principle of the model sees metadiscourse as embodying the interaction necessary for successful communication. As such it rejects the strict duality of textual and interpersonal functions found in much of the metadiscoursal literature. Instead all metadiscourse is interpersonal in that it takes account of the reader’s knowledge, textual experience, and processing needs and that it provides writers with an armory of rhetorical appeals to achieve this (Hyland & Tse, 2004). Textual elements thus, have an enabling role (Halliday, 1994), facilitating the creation of discourse by allowing writers to generate texts which make sense within their contexts. Their role is crucial for expressing propositional and interpersonal functions, not something they do independently of them. If it is accepted that textual items can realize either interpersonal or propositional functions depending on their context, then there is a need of a means of distinguishing their primary function in the discourse. According to Hyland and Tse (2004), an internal relation thus connects the situations described by the propositions (which is solely communicative), while an external relation refers to those situations themselves.

Dahl (2004) proposes a taxonomy consisting of two categories of metatextual elements. The first, called locational metatext, comprises linguistic elements which refer to the text itself or to parts of it. Her second category has been termed rhetorical metatext. It includes meta-elements which assist the reader in the processing of the text by making explicit the rhetorical acts performed by the writer in the argumentation process. Rahman (2004) labeled the function of endophoric markers as discourse entities; furthermore, he divided Hyland and Tse’s (2004) frame markers into four parts: discourse acts, discourse labels, illocutionary acts and topic shifting. Dafouz-Milne’s (2008) textual metadiscourse resources are divided into seven categories: logical markers, sequencers, reminders that refer back to previous sections in the text, topicalizers, code glosses, illocutionary markers, and announcements that refer forwards to future sections in the text. Although the above classifications have been applied to the analysis of academic and other genres of writing, the researchers have been unable to locate any particular study in the literature specifically applying the above or any other similar classification to narrative genre in general and novels in particular.

III. Method

A. Materials and Data Selection Criteria

The data for this study came from an original English novel and the simplified version of that novel (i.e. Oliver Twist) written by a native writer (the original version by Charles Dickens and the simplified version by Latif Doss). The text type to be analyzed was, therefore, English novel and its simplified version. The choice of this type of literary writing was due to the importance in developing critical reading skills, as well as cultural knowledge of the second language and writing skills. The choice of the novel was based on a number of criteria: the first criterion was the popularity of this novel such that a person who is familiar with literature knows this novel. The other criterion was the availability of both simplified and original versions of the novel in most English bookstores in our country (Iran). Both versions were analyzed in terms of the frequency and categorical distribution of metadiscourse textual resources (MTRs).

B. Procedures of Data Analysis

The original novel and its simplified version were divided into three sections. Since the length of the original novel was much more than that of its simplified version, and also because what was to be analyzed in both simplified and original texts should have had similar content, based on the chapters of the simplified version, corresponding chapters of original novel was identified. After the original novel was such made into chapters based on the content of its simplified version, the texts were divided into three sections. Then a section was selected randomly in both the original and simplified versions. After these divisions were made, a chapter in the relevant third was selected randomly and that chapter was analyzed for the occurrence of MTRs in both the original and simplified versions. To determine the frequency of MTRs, a list of such items was compiled from Hyland (2005), Dahl (2004), Rahman (2004), Hempel and Degand (2008) and Dafouz-Milne (2008). Then they were classified into six categories of analysis mentioned above.

It should be mentioned again that metadiscourse is closely linked to the context in which it appears, as Hyland (1996, p. 437) asserts: "the choice of a particular device does not always permit a single, unequivocal pragmatic interpretation". As a result, the functions of all items were examined qualitatively based on their actual occurrences in context. After determining the frequency of MTRs in the identified chapters and classifying them into six categories of analysis, the total words used in each chapter were also counted. Counting of items was conducted manually. Since the size of identified section in the original and simplified versions varied, the frequency of MTRs was calculated per 1,000 words. Inferential statistics were used to find out whether the frequency of occurrences and the categorical distribution of MTRs were significantly different in the original and simplified literary texts. The analysis technique employed was Chi-square, with a significance of P=0.05.
IV. RESULTS AND DISCUSSION

In order to find out the differences between the original and simplified versions of OT in the frequency of use and categorical distribution of MTRs, first, we calculated the distribution of MTRs in the identified section of the original and simplified novels, and then the frequency of MTRs was calculated per 1,000 words.

A. Rhetorical Distribution of MTRs

The novel that was analyzed was Oliver Twist in which the first third was selected randomly and in this third, a part was randomly selected in both the original and simplified versions which had the same content. In the first third of the simplified version, chapter five was selected that was equal to chapters five, six and seven of the original novel.

In this study a section for the original novel and its simplified version was randomly selected with the assumption that the identified section in both should have similar content, and then, the frequency of MTRs was calculated per 1,000 words. Table 4.1 presents the total number of words and the frequency of MTRs in the identified section of the original and simplified versions of Oliver Twist. As it is shown in Table 4.1, a total of 467 MTRs were used in the identified section of the original version of Oliver Twist; this represents a frequency rate of 54.96 per 1,000 words. Whereas in the simplified version of Oliver Twist, a total of 86 MTRs used in which the rate is lower, about 45.45 per 1,000 words.

Table 4.1 displays the observed frequency per 1,000 words. To be able to answer the relevant null hypothesis, that is, there is no significant difference between original and simplified versions of Oliver Twist in the frequency of the use of metadiscourse resources, we need to compare the observed frequency with the expected frequencies. Since there are 100.41 MTRs overall, the expected frequency would be 50.20. According to Hatch and Farhady (1981), when we have a definite hypothesis which we wish to test, we must then, test that hypothesis using inferential statistics. The Chi-Square is one test which allows us to do this; it is a test especially designed for nominal data. Thus, Chi-Square \((x^2)\) is the most appropriate test to find out whether there is any significant difference between the frequency of MTRs in the original and simplified novels. The Chi-Square \((x^2)\) was calculated using the following formula:

\[
x^2 = \sum \frac{(\text{observed} - \text{expected})^2}{E}
\]

According to Hatch and Farhady (1981), the degree of freedom will be based on the number of groups; there are two groups; so the degree of freedom will be 2-1= 1. The critical value of \(x^2\) (Hatch & Farhady, 1981, p. 279) with 1 degree of freedom is 3.84 for the 0.05 level. The observed value of \(x^2\) calculated here is 0.88, which is smaller than the critical value of \(x^2\).

\[
\text{Observed } x^2 = 0.88 < \text{critical } x^2 = 3.84
\]

This observation indicates that the difference between the original and simplified versions of OT in the use of MTRs is not statistically significant. Therefore, the first null hypothesis is accepted.

B. Categorical Distribution of MTRs

In order to find out the distribution of six categories of MTRs in the original and simplified versions of OT, the frequency of MTRs in each category per 1,000 words and also their percentage were calculated.

Table 4.2 shows the distribution of these six categories of MTRs in the original and simplified versions of Oliver Twist. According to the table, in the original version of Oliver Twist, two MTRs with the highest frequency rate are transitions (67.67%) and code glosses (24.41%); in the simplified version, the same applies for transitions with the frequency rate of 88.37% but not for code glosses whose frequency rate is 3.49%. Thus, transitions are the most frequently used category of MTRs in both the original and simplified versions of Oliver Twist. The frequency rate of sequencers is almost identical: in the original version their frequency rate is 5.57% and in the simplified novel it is 5.81%. The use of illocutionary resources and endophorics in the original version is very low while none of these categories exist in the simplified version. Table 4.2 summarizes all this information in one place based on frequencies out of 1000 words as well as their percentage.
### Table 4.2: Distribution of Different Categories of MTRs

<table>
<thead>
<tr>
<th>Categories of MTRs</th>
<th>Original F Per 1,000 W</th>
<th>Original Percent</th>
<th>Original Raw</th>
<th>Simplified F Per 1,000 W</th>
<th>Simplified Percent</th>
<th>Simplified Raw</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transitions</td>
<td>37.19</td>
<td>67.67</td>
<td>316</td>
<td>88.37</td>
<td>76</td>
<td>76</td>
</tr>
<tr>
<td>Code glosses</td>
<td>13.42</td>
<td>24.41</td>
<td>114</td>
<td>34.49</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Sequencers</td>
<td>3.06</td>
<td>5.57</td>
<td>26</td>
<td>5.81</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Topicalizers</td>
<td>0.35</td>
<td>0.64</td>
<td>3</td>
<td>2.33</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Endophorics</td>
<td>0.24</td>
<td>0.43</td>
<td>2</td>
<td>0.00</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Illocutionary Resources</td>
<td>0.71</td>
<td>1.28</td>
<td>6</td>
<td>0.00</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>54.97</td>
<td>100.00</td>
<td>467</td>
<td>100.00</td>
<td>86</td>
<td>86</td>
</tr>
</tbody>
</table>

Note: F= Frequency, MTRs= Metadiscourse Textual Resources, W= Words

The distributions of different categories of MTRs in the original and simplified versions of *Oliver Twist* are shown in Figure 4.1.

In this figure, *transitions* and *code glosses* in the original version and *transitions* in the simplified version of *Oliver Twist* are the most frequently used categories of MTRs.

In order to understand whether the differences between MTR category types in the original and simplified versions of OT were statistically significant, the Chi-square test was utilized. The degree of freedom for χ² analyses is calculated by multiplying the number of categories in each row minus one by the number of categories in each column minus one:

\[ DF = (\text{row}-1) \times (\text{column}-1) = (2-1) \times (6-1) = 5 \]

The observed value of χ² calculated here is 9.97, which is less than the critical value of χ² for 5 degree of freedom at the 0.05 level:

\[ \text{Observed } \chi^2 = 9.97 < \text{critical } \chi^2 = 11.07 \]

Therefore, there is not a significant difference in the type of metadiscourse employed in the original and simplified versions of *Oliver Twist* and the second null hypothesis is also accepted.

V. Discussion and Conclusion

This study examined the frequency distribution of MTRs in the original and simplified versions of OT. Thus, through Chi-Square test the hypotheses were tested to see whether there is a significant difference between an original novel and its simplified version in the frequency of use and categorical distribution of MTRs. The findings suggested that the frequency of MTRs was almost identical in the original novel and its simplified version. This can be explained by the fact that the writers of original and simplified versions seem to be aware of the norms of writing. Generally, we conclude that there is not a significant difference between original and simplified novels in the use of MTRs and the difficulty in comprehending novels is not due to the lack of MTRs. The writers in both versions tend to structure, segment and produce coherent texts. We have been unable to identify any literature related to the analysis of literary genres in terms of MTRs. As far as other genres are concerned however, there is evidence, for example, Crismore and Fransworth (1990), that differences in text type influence the type of metadiscourse. According to Dafouz-Milne (2008), texts with a balanced number of textual and interpersonal metadiscourse are the most persuasive and texts with a low index of metadiscourse markers are less persuasive; and it seems that readers highly value texts that guide and show consideration toward the audience.

In terms of categorical distribution of MTRs, the findings revealed that there was no significant difference in the type of MTRs employed in the original and simplified versions of OT, therefore, the null hypothesis was accepted. The results showed a broad agreement in the use of *transitions* in both versions of OT, demonstrating that the principal
concern of writers was to present information clearly. In line with our findings, Dafouz-Milne (2008) regards transitions as the most frequently used items, the possible explanation could be linked to the way in which different communities view and construct their argumentation. Transitions tended to be more marked in soft fields due to discursive nature of these disciplines (Hyland & Tse, 2004).

**Code glosses** category in the original version had a higher frequency compared to the simplified version, which implies that the writer of the original version of OT presents the text more explicitly than the writer of the simplified version of OT. Dafouz-Milne (2008), regarding the high number of code glosses in the texts, mentions that writers are aware of the broad audience they are addressing and consequently they are also aware of the need to include a number of explicit reading cues as well as more exemplifications, thus, ensuring that the text is read as intended. There were correspondences in the low frequency of sequencers and topicalizers in the original and simplified versions of OT.

The frequency of endophorics and illocutionary resources in the original version was very low while there were no instances in the simplified version. In academic writing, there seems to be more endophorics as compared to novels. Hyland and Tse (2004), in their investigation of students’ postgraduate writing, refer to the high occurrences of endophorics in the engineering discipline due to the fact that engineering’s greater reliance is on multi-modality nature which requires frequent references to tables, figures, examples and equations. This is one of the differences between literary writings and academic writing. Dahl (2004) believes that there is a low frequency of references to the articles themselves in economics and linguistics. In fact, there are far more occurrences of the word study than paper or article. Since medical reporting is so formally organized, the content presented is forced into information categories in a given sequence, and no extra processing effort is needed by the expert reader to orient him or herself within the text. In the case of illocutionary resources, Dahl (2004) believes that in contrast to the medical writers, neither economists nor linguists have a standardized text format to rely on. The consequence is that articles within these two disciplines have a more heterogeneous organization which implies how the writer must signal which discourse act is being performed.

The results of the present study have obvious importance in increasing students’ awareness of the way writers of original and simplified novels organize their writing. Metadiscourse is a valuable tool which provides rhetorical effects in the text such as providing logic and reliance in the text. Writers of both original and simplified novels use MTRs in their novels, so it is impossible to produce coherent texts without such resources. Metadiscoursal analysis can therefore be recommended to be used a useful means by the teachers to help students control their writing practices for effective writing.

### REFERENCES


Sonur Esmaili has an MA in TEFL from Islamic Azad University, Tabriz Branch, Iran. She is interested in researching the relationship between language and literature.

Karim Sadeghi has a PhD in Applied Linguistics from the University of East Anglia, UK and is the Head of English Language Department at Urmia University, Iran. His research focuses mainly on assessment and evaluation as well as on language skills. His work has appeared in TESL Canada Journal, The Modern Language Journal, The Asia-Pacific Education Researcher, and Iranian Journal of Applied Linguistics among others.