Chapter Five
Conclusion

Contents

5.1. Introduction
5.2. Overview of the chapter
5.3. Research objectives and hypotheses
  5.3.1. The first hypothesis
    5.3.1.1. Anaphor reference
    5.3.1.2. Sentence connectives
  5.3.2. The second hypothesis: Lexical and coreferential ties
  5.3.3. The third hypothesis: mental model dimensions
  5.3.4. The fourth Hypothesis: Topical structure analysis
5.4. Limitations of the study
5.5. Pedagogical implications and recommendations
5.6. Suggestions for further research
5.7. Summary of the study
CHAPTER 5
CONCLUSION

5.1. Introduction

In the previous chapter, the statistical results of the study were objectively reported. Two types of statistics regarding the study variables were presented: descriptive and inferential. Descriptive statistics of the variables served to present the data in a clear, logical and comprehensible manner. These statistics were meant to observe and record the patterns resulting from an examination of the means and standard deviation of the variables involved. This procedure proved helpful, especially when the compared variables were not statistically significant but showed underlying trends. The data were also presented in tables and illustrated by graphs. Inferential statistics were used to test the relevant hypotheses and decide whether to accept or reject these hypotheses and hence make inferences about the data. The current chapter will primarily focus on the conclusion, implications, recommendations, and summary of the entire study.

5.2. Overview of the chapter

In chapter 5, the researcher tries to interpret, examine, and qualify the results of the investigation and draw inferences from these results. It includes the following sections. First the hypotheses following from the research objectives will be presented and the relevant statistical procedure used to test each hypothesis along with its results will be restated followed by a discussion based on the analysis performed and results achieved. Second, the implications of the study will be presented. In this section, the researcher will discuss the kind of impact the results can have on the field of study and its contributions to the ESL and EFL writing classes. Limitations of the study will follow too. Third, the pedagogical implications are discussed and recommendations for classroom teaching are put forward. Some suggestions for further research will be made. And finally, a summary of the entire work will be presented in such a way that it can serve as a stand-alone document.
5.3. Research objectives and hypotheses

In this section of the chapter the findings of the research are discussed and possible conclusions and inferences are drawn from these findings.

5.3.1. The first hypothesis

In pursuit of the first research objective to study the possible differences between the Iranian and Indian students’ use of explicit referential and conjunctive cohesive ties, a null hypothesis was formulated as follows:

There are not any statistically significant differences between the Iranian EFL and Indian ESL students as regards their use of the following explicit grammatical and conjunctive cohesive ties as measured by Coh-Metrix.

a) Anaphor references (at adjacent and distant sentence levels)

b) Sentence connectives (at adjacent and distant sentence levels)

5.3.1.1. Anaphor reference

In order to investigate the first group of components of this hypothesis, a t-test was used after the distributions proved to be normal. The results revealed that there were significant differences between the means of the two groups of the students which suggested that the first two components of the null hypothesis were rejected in favor of the Iranian students, i.e., the Iranian EFL students used far more anaphor references between adjacent sentences in their essays than the Indian ESL students. However, the results failed to show any significant difference between the two groups regarding their use of anaphor references among distant sentences. Supplementary analyses of the data revealed that of the three types of anaphoric expressions, personal references were the most frequently used by both Iranian and Indian students. And among the personal references, the Iranian students used specifically more second person you, first person plural we, and third person singular she. Conversely, the Indian ESL students used demonstrative the more frequently than their Iranian peers. The results of the Univariate analysis also indicated that the students’ gender did not play any significant role in the difference of the two means; in other words, gender effect size was too small to have a significant interaction with the linguistic background of the students which was the basis of the comparison.
Discussion

The density of personal references in the Iranian students’ essays can be attributed to different reasons. It appears that the students’ low English proficiency coupled with mother tongue interferences have led to the overuse, misuse, and, at times, underuse of certain referential elements. Learning these types of references requires exposure to natural language. In other words, students learn the true meanings of these cohesive devices when they actually learn and use them in real situations. Students in an EFL setting normally lack natural exposure to the English language, and they do not have the opportunity to use the language; therefore, they develop a poor communicative competence. Such students tend to transfer the equivalent structure of their L1 into their L2.

Persian is a null-subject and verb final language (Taghavipour, 2004); there are no case forms and no gender distinctions in Persian (Megerdoomian n.d.). There are different pronominal enclitics\(^1\) in Persian. These enclitics include -\(\text{æm}\), -\(\text{æt}\), -\(\text{æs}\), -\(\text{emn}\), -\(\text{eten}\), and -\(\text{esan}\) (Rasekh Mahaand, 2007). Depending on the word they are added to, they take on various functions. For instance, if they are added to a noun phrase, it is interpreted as possessive, Ketæb-æt [book + 2sg \(^2\)] (your book). If added to a transitive verb and preposition, the clitic will function as the accusative form of a personal pronoun like dicl-æm-æt [see (past) + 1sg inf.\(^3\) + 2sg] (I saw you). The clitic might also be added to an adverbial, numerical expression and interrogative elements with a partitive meaning like vææætææ [middle + 3sg \(^4\)] (in the middle of it) (Megerdoomian, n.d.). And finally if the enclitic is added to an intransitive verb, it is regarded as the subject pronoun like raft-æm [went (past) + 1sg] (I went). Thus to write a simple sentence like I went in Persian, one has to use two pronominal elements:

\[
(\text{maen}) \text{ raft-æm}.
\]
\[
(\text{I}) \text{ went-1sg}
\]

\(^1\) An enclitic refers to a form of word (or clitic) that cannot stand alone as an utterance, but rather it depends (or leans) on a preceding word in a construction such as the contracted forms of be, as in I’m and he’s. A proclitic, on the other hand, is a form of word (or clitic) that depends on the following word such as English and French articles (Crystal, 2008).
\(^2\) Second person singular
\(^3\) First person singular inflected
\(^4\) Third person singular
In the above sentence, the subject (man) can be omitted in colloquial Persian because of the presence of the first person singular enclitic (-aem). This omission rarely occurs in formal written Persian.

Another property of the Persian grammar is that it allows personal pronouns to behave resumptively. That is, the language user can use a pronoun whenever he/she expects a gap. For instance,

\[
\text{Mārd-i ke bā } u \text{ sohbat kārdim}
\]

The man with whom we talked

Because of this property of their mother tongue, the Persian speaking Iranian students often tend to write an erroneous English relative clause as follows:

The man who(m) we talked with him

In such sentences some Persian writers use a relative pronoun like who in the above sentence which is equivalent to complementizer “ke” in Persian and a personal pronoun like \(u\) or a pronominal enclitic.

\[
\text{Ketæb-i ke dærbāre-wā bæhns kārd-im}
\]

The book which about it-3sg we talked-1pl

Therefore, in the above sentence, the writer uses pronoun “\(it\)” where it is not grammatically correct. Using such pronouns not only poses difficulty for a Persian English writer but it can improperly add to the density of pronouns in a text as well.

It was also found in this study that there was a significant overuse of the personal pronouns \(you\) and \(we\) by the Iranian students. On the other hand, the Indian students used pronoun \(I\) more frequently than the Iranian students, though this difference was not significant. This could be attributed to the fact there is a tendency among Persian speakers not to use pronoun \(I\) in a formal setting where respect is required. They prefer to use other alternatives including we possibly for reasons of cultural background and social relations such as self-lowering and other-raising (Beeman, 2001); however, the social relations

\[5 \text{ Restrictive relative clause}
\[6 \text{ In generative syntax, a term used to refer to subordinating conjunctions which mark an embedded sentence of a complement type, e.g. which in I said that he was leaving.}
\[7 \text{ First person plural} \]
become rougher as the intimacy between the parties increases. The Indian students, on the other hand, preferred to talk about their personal experiences in their compositions.

Another finding of the research concerned the use of demonstrative “the”, which serves as an anaphoric reference. The supplementary analysis revealed that the Iranian students and the Indian students differed significantly in their use of the English definite article.

The article system of English can be a source of great difficulty and frustration for foreign students of English and, consequently, for ESL and EFL teachers, especially those whose mother tongues do not have specific words for articles or use the articles in different ways (Faghih, 1997; Maslamani, 2008). Persian and Hindi are among the languages that do not have special words for definite articles. In these languages, it is the context of use which determines the definiteness of a noun.

Persian does not have a definite article corresponding to “the” in English. A noun is either definite or indefinite, depending on the context in which it is used. For instance,

**Ketāb**  the book/ a book

**Kolah**  the hat/ a hat

In sentence (a), Ketab (book) is used as an indefinite noun and in sentence (b) it is definite:

a) **Ketāb**  behtærin doost-e ensân xest.
   A book the best friend-of man is.

A book is Man’s best friend.

b) (mān) *ketāb rā be u dādem.*

I gave the book to him.

In order to make the noun clearly indefinite (a, an, etc.), suffix –i can be added to the noun for example,

**Ketāb-i**  a book

**Kolah-i**  a hat

Or sometimes Persian speakers use *yek* (one) before a noun to make it indefinite:

**yek ketāb**  a book

**yek kolah**  a hat
Or even both yek and —i are used simultaneously.

*yek Ketāb-i*

*yek Kolāh-i*

There is no definite article in Persian; however, Persian grammarians put the definite nouns into six categories. These nouns are always definite:

1) Proper nouns: nouns that refer to people (John, Ali, Tom, etc.), places (London, Tehran, Shiraz, etc.), and the nouns that refer to unique objects such as *zaemin*, the earth, *xorshēd*, the sun, *donyā*, the world, etc.

2) Nouns modified by demonstratives (en Ketāb this book)

3) Generic nouns: as in *aesh* heivāne najibi āest. (*The horse* is a noble animal.)

4) Direct objects preceded by the object marker *rā* as in *dar rā* ghofl Kārdām. (*I locked the door.*)

5) Nouns defined by the context of use as in *cherāgh* roušān āest. (*The light is on.*)

Because there is no definite article in Persian, the Iranian students often find it hard to learn the English definite article, *the*. Some tend to use it where it is not at all required, “we can use the internet in the every field...”, some do not use it where it is needed, e.g., “These are some of problems of internet” (*the* is dropped before problem and internet), and finally some wrongly replace it with an indefinite article or the other way around, “I try to use it in a best way”.

Hindi, like Persian, lacks a definite article. The context in which the noun is used determines its definiteness or indefiniteness. Thus a noun acting as subject or direct object in its sentence is definite rather than indefinite; therefore, depending on its context the word ‘makaan’ (house) may have as its translation equivalent ‘a house’ or ‘the house’ or ‘houses’ or ‘the houses’. Demonstratives may be used in cases where the speaker wishes to stress the identification of a referent. However, most nouns lack determiners, and their information status is generally decidable only by the context (Baldridge, 1996). In spite of this fact, the use of article ‘the’ by Indian speakers of English has been found unaffected by the speakers’ L1 grammar (Sharma, 2005). Sharma found “a nearly even rate of overt and null use of the definite article (p. 552)” in the speech of the Indian speakers of English.
Although both Persian and Hindi lack a definite article, the Indian students tended to use it more frequently and more correctly. The observed significant difference between Iranian and Indian students; therefore, could possibly be explained by their different proficiency levels. Indian students have much more exposure to English than the Iranian students. English is a means of education for many of these students; they have more opportunity to learn and actually use the language and hence develop a higher level of proficiency.

5.3.1.2. Sentence connectives

The second component of the first null hypothesis concerned the density of sentence connectives in general and five different types of connectives in particular. The results of the t-test showed that the Iranian students used significantly more sentence connectors in their essays than the Indian students. Moreover, they used significantly more additive, causal, and temporal connectives. Nonetheless, they didn’t differ in their use of logical and adversative connectives. It should be noted that the gender of the students did not affect the significance of the difference between the means of the two groups.

Discussion

Sentence connectors, like other discourse markers, are vital to the cohesion and coherence of a text. If they are used appropriately, they will not only improve the coherence and cohesion of the text but also enhance comprehension and recall of the text, especially if they signal causal relations between the text segments (Jalilifar 2008; Tapiero, 2007; Sanders and Noordman, 2000.). However, if they are not properly used, these linguistic devices can lead to the false illusion of understanding the text and interfere with the readers’ inference making and generation of elaborations beyond explicit connectives. In other words, overuse of these connectives in the text may increase readers’ passivity. The readers prefer to have recourse to these cohesive signals than using their background knowledge and inference making ability to establish coherence between text segments they may give the reader the false “impression” of understanding the text (Degand et al., 1999).

Martins et al. (2006) give two reasons for the cohesive and facilitative power of connectives. The first reason, as emphasized by Halliday & Hasan (1976), is that
connectives make the text more cohesive and more structured by establishing links between sentences. Furthermore, connectives such as and, because, so, therefore, since, however, but, then, later, and so on clearly tell the reader that the sentences are linked to each other in a precise semantic manner, namely, causal, restrictive, temporal, and so on. Of all the sentence connectives, causal connectives are said to play a more facilitative role. Research has shown that causal connectives may prompt readers to search for knowledge in long-term memory in order to restore local or global text incoherence. For example, Caron et al. (as cited in Martins et al., 2006) showed that when subjects had to recall the following two unrelated sentences:

a) The priest was able to build the new church.

b) The computer had made a serious error.

they recalled them better when they were connected by "because" than by "and". It is possible that “because” prompted the readers to “search for meaning” by finding the reason underlying the semantic connection between the unrelated sentences. This search may enhance integration and memorization. (Millis and Just, 1994) found similar results. Linderholm et al. (2000) found that causal revision of text improved reading comprehension, especially, of difficult texts for more- and less-skilled readers.

The results of the statistical tests in this study revealed that additive connectives and causal connectives were the most frequent types in both the Iranian and the Indian students’ essays. The temporal connectives were the least frequent. This could be due to the nature of the essay. The results of the t-test showed that the Iranian students used significantly more connectives than Indian students, especially additive, causal and temporal ones. It appears that sentence connectives, like pronouns, were the most immediate tool at their disposal to link sentences and achieve coherence, not knowing that the overuse of some connectives does not necessarily increase readability and does not point to argument in particular, rather the redundancy of these connectives would create, as Crew (1990) put it, “hiccup effect” in reading. The findings also go in line with the findings from previous research reporting that non-native speakers of English overuse sentence connectors, yet less appropriately (Asassfeh, 2005).

The results of this study support the findings by research scholars who investigate the role of connectives in discourse and their use by the native and non-native speakers.
They attribute this phenomenon to second and foreign language teaching methodology, specifically writing instruction. Students learn the proper use of such cohesive elements only if they are taught in context. For instance Crewe (1990) believes that rote learning and mechanical exercises are responsible for this problem. He emphasizes the appropriate use of connectives suggesting that it is better to avoid connectives at all if not quite sure about their appropriate use. Byrne (as cited in Tang and Ng, 1995) advises teaching the meaning nuances of connectives to students and providing them with ample exercises so that they can actually comprehend the meaning and practice their use in context.

5.3.2. The second hypothesis: Lexical and coreferential ties

The second null hypothesis comprising five components followed from the second research objective set to investigate the potential differences between the Iranian EFL and Indian ESL students’ use of lexical cohesive devices in their English essays. This null hypothesis comprised five lexical and coreferential indices and was formulated as follows:

There are not any statistically significant differences between the Iranian EFL and Indian ESL students with regard to their use of the following lexical cohesion and coherence devices:

a) Argument overlap (at adjacent and distant sentence levels)
b) Stem overlap (at adjacent and distant sentence levels)
c) Content word overlap
d) Latent semantic analysis (at adjacent and distant sentence levels)
e) Hypernym value of nouns

In order to test these components of this hypothesis a t-test was used. The result of the t-test revealed that there was more overlap of arguments between adjacent sentences in the Iranian students’ essays. However, despite the slight bias of the means of argument overlap between sentences of all distances towards the Iranian students, there was no significant difference between the means of the two groups. The results also indicated that there was significantly more content word overlap in the Iranian students.

The two groups were also compared for the conceptual coherence or semantic similarity as measured by Latent semantic analysis (LSA) technique. The results of the t-
test showed that there was a significant difference between the Iranian and Indian students’ essays. There was more semantic similarity or conceptual coherence in the Iranian students’ essays than their Indian peers’ both at the levels of adjacent and distant sentences. Thus the results of the statistical tests confirmed three of the components of the second research hypothesis but failed to confirm the other two components, namely stem overlap and noun hypernym value.

**Discussion**

Lexical and coreferential ties are a major source of textual cohesion and an objective of this research. It is “the cohesive effect achieved by the selection of vocabulary” (Halliday and Hasan, 1976, p 234). A lexical item, however, will have a cohesive effect only if it gets into a cohesive relationship with other items. Halliday and Hasan divide lexical cohesion into two main categories: reiteration and collocation. Reiteration, as the name suggests, involves repetition of lexical items. A reiterated item may be a repetition of an earlier item (boy/boy), a synonym (boy/son) or near-synonym (boy/lad), a super-ordinate (vehicle/car). All these reiteration types are cohesive in exactly the same way as the general words, differing only in the general words’ degree of generality (human/male, male/boy).

Referential indices make a major contribution to textual comprehension. The computer application Coh-Metrix employs different forms of lexical coreference, five of which have been selected to measure the lexical cohesion of the students’ essays. This selection was based on the assumption that these indices approximate those in Halliday and Hasan’s taxonomy of lexical cohesion. These indices include argument overlap, stem overlap, content word overlap, Latent Semantic Analysis and hypernym value of nouns.

Argument overlap refers to sentences that share the same noun argument. For instance, in the following two sentences, the same words, *the Internet and uses*, have cohesive functions that bind the two sentences.

*The Internet has many uses and abuses in our daily life. One of the most important uses of the Internet is getting information.*

Another lexical and coreferential index used to assess the coherence of the text is stem overlap. Stem overlap refers to sentences that contain words sharing the same stem. They include not only nouns but also other word forms such as *produce,*
producing, produced, productive, productively, productivity, producer, and production. Like other lexical indices, words sharing the same stem have strong cohesive property that link different sentences in the text.

The computer application measures the argument and stem overlap at two levels: at the level of adjacent sentences and distant-sentence level.

The content word overlap index refers to the overlap of the words that carry most meaning such as verbs, adverbs, adjectives and nouns. And finally, Latent Semantic Analysis employs a statistical technique called singular value decomposition to analyze the semantic relationship between different textual elements such as words, sentences, and paragraphs. This technique allows the textual relations go beyond explicit relationships such as same word overlap, stem overlap or content word overlap into relative semantic similarities such as internet/website, internet/computer, computer/software, chair/table, table/wood, etc. That is, the words that co-occur in the same semantic domain or context are regarded as semantically and conceptually similar.

Findings of this study indicated that the Iranian and Indian students differed in the way they achieved lexical cohesion and coherence in their essays. Iranian students exhibited more coherence through the repetition (or overlap of) noun phrase arguments at adjacent sentence level, overlap of content words (nouns, adjectives, verbs and adverbs), and LSA. The means of other variables were also slightly biased towards the Iranian students. In other words, there was more overlap of noun phrase argument at distant sentence level, stem overlap at both levels of adjacent and distant sentences, and hypernym value of nouns. However, these differences were not statistically significant.

The observed differences could be attributed to language transfer as well as learning effect. These findings apparently comply with the research conducted by Seddigh and Yarmuhammadi (1997). Upon comparing some fifty political journalistic texts in Persian and English, they found out that the Persian texts were denser in lexical repetition than the English texts. Yamini (1999), too, observed that non-native speakers showed more awareness towards factors involved in lexical cohesion than native speakers, possibly due to their learning not to their language intuition.
5.3.3. The third hypothesis: mental model dimensions

The third objective set for this study was to investigate any differences in the types of the situation model dimensions in the English texts written by the Iranian EFL and Indian ESL undergraduate students. This objective led to the following null hypothesis:

There are not any statistically significant differences in the situation model dimensions in the texts written by the Iranian EFL and Indian ESL students as measured by Coh-Metrix. These dimensions are as follows:

a) Causal dimension
b) Intentional dimension
c) Temporal dimension
d) Spatial dimension

Results of the statistical calculations revealed that although causal dimension was the dominant dimension of the students’ essays in both groups, the results of the t-test failed to reject the null hypothesis suggesting that no significant differences existed between the two groups. However, due to the expository nature of the essays, the number of the causal connectives, nouns, verbs, and particles in the essays outnumbered the other types of connectives, nouns, verbs, and particles. This is indicative of the causal cohesion of the essays.

Discussion

One of the responsibilities of the writer towards his prospective readers is to help them to construct a mental representation of the text. In fact, this representation is an integration of the reader’s prior world knowledge with the language input. The presence of certain elements in the text creates the dimensions of situation model or the microworld of what the text is about. The importance of the situation models in the coherence and the comprehensibility of the text made the researcher investigate the degree to which the Indian and Iranian student writers create different mental model dimensions in their English texts. This led to the second research objective and the related research hypothesis:

Although some researchers have emphasized the necessity of processing multiple mental dimensions – temporal, causal, and spatial - in text comprehension (Johnson-Laird, 1983), most scholars consider the causal dimension to be the most important for
comprehension (Noordman & Vonk, 1998, Sanders, 2005) so much so that the greater the number of causal relations that readers identify in a text, the more coherent they perceive the text to be, and the better they remember it (van den Broek, 1988; van den Broek, Tzeng, Risden, Trabasso, & Bashe, 2001).

Though the Iranian students used significantly more causal connectives, they did not show to have created more causal dimension. This is indicative of the fact that causal connectives per se are not enough for the creation of causal dimension. They must be used in conjunction with other causal elements such as causal verbs and links. As seen in table 4.23, the Iranian students’ mean number of causal verbs, links, and particles is larger than that of the Indian students. However, they must have fallen short of using other major elements of causal dimension and the Iranian students’ bigger mean of causal dimension is because of the significant number of (causal) connectives that these students tend to use in their essays.

Obviously, the use of causal connectives by both groups implies that the students have a fair knowledge of these types of connectives; however, they should be made aware of the fact that these connectives alone cannot signal causal relations and they cannot help their readers to comprehend the text.

5.3.4. The fourth Hypothesis: Topical structure analysis

The researcher also tried to look at the coherence and cohesion of the students’ essays from a different perspective. To this end, the students’ essays were analyzed using Lautamatti’s Topical Structure Analysis (TSA) to see how the Iranian and Indian students developed their topics throughout their paragraphs. Therefore, the fourth null hypothesis was formulated as follows:

There are not any statistically significant differences between the Iranian EFL and Indian ESL undergraduate students with regard to their use of the following topical progressions:

a) Parallel progression
b) Sequential progression
c) Extended parallel progression
To test this hypothesis, a t-test was used. The results of the t-test rejected the null hypothesis suggesting that there were significant differences between these two groups. They developed their topics throughout paragraphs differently. The Iranian students used significantly more sequential progressions. On the other hand, the Indian students used more parallel and extended parallel progressions.

**Discussion**

The present study indicated that both groups used relatively more sequential progressions than the other two types. In the Iranian students’ essays, more than 63% of the t-units appeared in sequential progressions while 18% in parallel progressions and 19% in extended parallel progressions. On the other hand, in the Indian students’ essays, 51% of the t-units were in the form of sequential progressions, almost 24% parallel progressions and 25% extended parallel progressions.

Previous research has established the fact that a proportionately high number of related sequential progressions in a text denote the internal coherence of the text as reflected by the repetition of key words and phrases (Simpson, 2000). Numerous studies have indicated that high-rated and accomplished English essays contained more sequential progressions (Schneider and Connor, 1991; Simpson, 2000; Hoenisch, 2009). These studies recommend a proportion of 65% to 80% of sequential progressions and a good number of extended parallel progressions. This suggests that the authors of high rated essays are better able to both discover content for the topics they introduce and distinguish between crucial and noncrucial topics, thus producing more globally coherent texts. On the other hand, low-rated essays contain more parallel progressions and the topical subjects that often coincide with the grammatical subjects, thus they are more locally coherent.

Considering the criteria set by these researchers, the findings of this study revealed that the Iranian students achieved more topical continuity or internal coherence through the repetition of key words and phrases. This means that they tended to introduce into their essays more new topics and details related to the main topic of the essay. The findings also go in line with the statistical results concerning the lexical cohesion (the second hypothesis), indicating that there was more overlap of content words and the repetition of words and their collocations as measured by LSA technique.
Since topical development is done primarily through the repetition of certain key words and phrases (which serve as topics in the text), the differences between the Iranian and Indian students could be explained by the same facts as those justifying the lexical cohesion differences between these two groups. Firstly, research has shown that Persian texts are lexically denser than English texts (Seddigh and Yarmuhammadi, 1997). This property of the language might have been transferred to their English essays so that their English texts, too, turn out to be lexically denser. Secondly, teaching effect might have played a role in this picture. Language instruction in most Iranian language classes is done mainly through traditional methods. Grammatical rules are taught and students are made conscious of the processes of sentence construction and paragraph development. This could be the reason why Yamini (1999) concluded that Persian speakers showed more awareness towards factors involved in lexical cohesion than native speakers, possibly due to their learning not to their language intuition.

5.4. Limitations of the study

This study was subject to the following limitations:

1) The first limitation concerned the sampling procedure which makes it almost impossible to generalize the findings of this study to populations beyond Chandigarh colleges, Persian Gulf University and Roudehen Azad University. Since it was not practically possible to do a truly random sampling, the researcher followed “area or cluster sampling” (Best, 2003), limiting the data collection to the students studying English in colleges affiliated to Panjab University in Chandigarh in India and the students studying in Persian Gulf University and Roudehen Azad University in Iran.

2) The second limitation concerned the size of the samples. The limitation mentioned above, and the need for variable control like linguistic background, adherence to the essay topic etc. reduced the samples to their present size, i.e., 106 essays from Iran and 73 from India. The unequal size of the Iranian and Indian samples does not affect the ultimate results because, statistically speaking, “when the sample size is 30 or more, the sampling distribution is almost exactly a normal distribution” (Hinton, 2004, p. 55). Therefore, as the number of the essays in both samples exceeded 30, the samples were
assumed to be big enough to secure the normality of the distributions and the validity of the statistical tests.

3) The third limitation of the study was determining the proficiency level of the student writers. It was not practically possible for the researcher to gather so many students in order to administer a proficiency test; therefore, to surmount this problem, he involved only the senior students for the writing task assuming that these students are proficient enough to write expository essays in English.

4) Finally, the last limitation of this study was the exclusion of substitution and ellipsis as two of the cohesive devices. Since the essays were analyzed by a computer application and these two cohesive ties have not been measured by the software, they were excluded from this study. However, upon the inspection of the essays, it was found out that substitution and ellipsis were almost rare in the essays.

5.5. Pedagogical implications and recommendations

In the light of its findings, this study has several useful implications and suggestions for writing instructions and future research. Previous research has shown that non-native speakers of English tend to overuse and less appropriately use discourse markers than native speakers do. For instance, Hinkel (cited in Simpson, 2000; Chen, 2006) reported that non-native speakers of English overused sentence connectors. The easy accessibility of discourse markers, especially, sentence connectors and personal pronouns, makes it possible for the students to insert a connective or personal pronoun into their texts whenever they feel a gap between the discourse segments, often without paying due attention to their appropriate use or reasonable distribution. The findings of the present study are in line with these researches indicating that Iranian learners of English used far more sentence connectors and personal references than the Indian students.

Considering the status of English in Iran as a foreign language and India as a second language, Indian students enjoy much more exposure to the language. English is used in both public and personal domains and its functions, as Kachru (cited in Hohenthal, 2003), points out, “extend far beyond those normally associated with a foreign language, including the instrumental, the regulative, the interpersonal and the innovative, self-expressive function”. It is the main medium of instruction at the postgraduate level, and it
is taught as a second language at every stage of education in all states of India. Now the use of English has been increasing from formal domains like education, government and employment to more informal and personal settings. About 40% of the informants told Hohenthal (2003) that they spoke English with friends, and people get introduced to each other most often in English. Even their conversations in their local languages are interspersed with English words, phrases and sentences. There are a lot of newspapers, magazines, books, and notices and announcements all published in English which offer peripheral learning to Indian students. Indian students have also the advantage of watching TV channels and listening to radio stations broadcast in English. All these provide them a unique chance to learn English within a fairly natural context.

On the other hand, the Iranian students learn English as a foreign language. English is not a means of instruction for them. They normally learn the language within the four walls of a language classroom often with no or very little natural exposure to the language outside the classroom. Teaching is done mainly through their mother tongue with the main focus on teaching grammar, memorizing vocabulary and improving their reading skill. The difference between these two linguistic settings manifested itself in this study: the Indian students made fewer grammatically erroneous T-units. They used appropriately fewer connectives and personal references.

Contextualized instruction of language in general and writing in particular is a must. Teaching grammatical points in isolation will only lead to their mechanical learning. EFL students learn the appropriate meaning of linguistic forms only when they are presented in a context, for it is the context that governs the appropriateness of a cohesive device or whether a stretch of sentences can be regarded as text (Halliday and Hasan, 1976).

In the absence of natural exposure to the foreign or second language, EFL students must profit from communicative exercises. The role of language teachers and textbooks is vital. They can provide the learners with ample contextualized and authentic materials such as technical writings, newspaper articles, research papers, etc so that the learners become conscious of even the slightest semantic variations of coherence and cohesion devices and appreciate the proper use of the discourse markers. Mechanical exercises that encourage rote learning must be substituted by critical reading and hypothesis testing.
exercises. This can help the language learners to acknowledge the appropriateness, articulate the usage, identify the position and understand the meaning of the discourse markers and cohesive devices. As suggested by Hayward and Wilcoxon (cited in Tang and Ng 1995), scrambled sentences and completion exercises should be incorporated into classroom activities.

Activities that aim at raising the students' awareness of the cohesive relations and the elements that contribute to the coherence and unity not only enhance reading comprehension (Williams, 1983; Lubelska, 1991; Liederholm et al., 2000) but also can have good implications for making the student writers conscious of what cohesive devices are, how they are used, and what their distribution is in the text. To this end, the teacher can give the students samples of well-organized and coherent texts and ask them to identify different cohesive ties and specify their referents, focusing on each, one at a time. He can even make cloze passages out of these texts deleting certain cohesive elements like pronouns, connectives, demonstratives etc. and ask the learners to restore the missing items.

Moreover, much of textual cohesion in a well-written text is achieved through lexical ties. Meanwhile, lexis has shown to be responsible for comprehension problems too (Alderson and Cassell as cited in Lubelska, 1991). Helping students understand lexical cohesion and pick up on how successful writers interweave ideas throughout their writings can be a good classroom activity. Students should be taught the ways and means of linking their ideas in the text by the repetition of the same words, the use of synonyms, hyponyms, and collocated words. To do this, the learners should develop a working vocabulary capable of extending the concepts they introduce in their texts. They must be warned against the overuse of any of the lexical devices. Using too many repetitions or too many synonyms can make their writing sound redundant and boring. As to collocations, the students must also beware of false analogies about which lexical items co-occur. Therefore, along with teaching writing, the teacher should try to encourage the students to increase their vocabulary related to the topic of the composition so that he can make the right choice of vocabulary.

Another measure of textual coherence used in this study was the Topical structure analysis (TSA) designed by Lautamatti (1987). Topical structure analysis associates a
coherent text with a text possessing a clear focus. Through topical structure analysis, the relationships between sentence topics and discourse topic can be studied by looking at sequences of sentences and examining how the topics in the sentences work through the text to build meaning progressively (Connor, 1987). Topical structure analysis has been associated with quality of writing. High quality essays contain a good ratio of sequential progression to extended parallel progression. Being conscious of how to develop and progress the topics through the paragraph can help EFL writers achieve topical coherence in their essays and improve their writing ability (Chiu, 2004).

In line with the conscious raising activities, the teachers in writing classes can use the topical structure analysis as a revision strategy to help EFL student writers at intermediate and advanced levels to revise their own essays. To begin with, the teacher can give sample paragraphs to the students and teach them how to identify the topical subjects of each sentence of the paragraph. The next step is to teach them how to draw the TSA diagram and how to distinguish between different types of topical progression: parallel progression, sequential progression and extended parallel progression. Once he is sure that the students know how to identify the topical subjects of sentences (which are sometimes different from the grammatical subjects) and know how to draw TSA diagrams, the teacher can help the learners to apply their TSA knowledge to revise their own essays.

In the like manner, the teacher can help the learners to do topical structure analysis to revise their own essays. He should assign composition topics to the students. Then he should ask the students to read their essays and try and identify the topical subjects of every sentence in their essays and draw the TSA or coherence diagram of their paragraphs. In the beginning there might be a disproportionate number of parallel progressions, i.e., either too many parallel progressions or too few sequential progressions, which suggests that the students repeat the same topic in every sentence. This leads to a lot of redundancy and repetitions. Again the teacher should direct the students to develop their paragraphs and introduce new subtopics into their essays and discuss every aspect of the discourse topic (sequential progression) while they make occasional references to the main topic of the discourse (extended parallel progression).
5.6. Suggestions for further research

The present work documented that Iranian and Indian students’ essays differed in terms of the cohesive and coherence features. Certainly there is a need for more contrastive studies to decide whether or not these findings are valid in other types of texts, written or oral, before making any generalizations. During the course of the research, many more questions arose and required more investigations. Some of these areas are outlined below.

1) Future research can focus on comparing Iranian EFL and Indian ESL essays with those of the native speakers of English in order to investigate the coherence and cohesion devices in general and the types of devices, the distribution of the ties and the position of these devices in particular. Also, the topical progressions used in the essays written in both languages can be studied.

2) In the present study, all the Indian participants came from only Hindi linguistic background. Therefore, another research work can study the differences in the coherence and cohesion features of the English essays written by the Indian students who come from other linguistic backgrounds such as Panjabi. This can show if a vernacular makes a difference in the use of the discourse markers in the English writings of the students who live in the same country and study in the same educational setting.

3) Supplementary analyses revealed that essays written by the male students in this study were lexically denser than those of the female students. Future research can focus on reasons for this observation.

4) As discussed before, findings of this study showed that Iranian students’ essays were denser in terms of lexical cohesive ties, i.e., there was more content word overlap and more semantic and conceptual coherence. This was assumed to be either due to teaching and learning effect or mother tongue interference, drawing upon the previous researches claiming that Persian texts are richer in lexical cohesive devices and the Iranian students were more conscious of lexical cohesion than native students. A similar study can be conducted analyzing the students’ writings on the same topic in English and in their vernacular and see if they use the same strategy to achieve coherence and cohesion in both languages.
5.7. Summary of the Study

During the last few decades, especially following the publication of Halliday and Hasan’s “Cohesion in English” in 1976, cohesion and coherence have become a topic of interest for research in linguistics and its related fields. A huge amount of research has been carried out to investigate the role that these two concepts play in bringing unity to a text and how they contribute to the comprehensibility of a text. All these research works unanimously agree that cohesion and coherence play a vital role in the understanding of discourse so much so that coherence has equaled understandability and comprehensibility. The research findings also indicate that if a writer wants to compose a well-organized and comprehensible text, he or she must follow the writing conventions of the native speakers and make proper use of cohesive devices such as referential and lexical ties, coreferential features and conjunctions.

In order to assess the coherence and cohesion of a text, researchers have tried to think up models and tools that can account for the textual coherence and help analyze a text and measure the degree to which different segments of a text hang together (for example, Halliday and Hasan, 1976, Lautamatti, 1978, Givon, 1983b, and Hoey, 1991). Halliday and Hasan proposed a taxonomy of coherence ties that can be used to identify cohesive ties between different text segments and thereby assess the cohesion and coherence of that text. According to their model, a text is regarded as a semantic unit. The parts of the text are linked by some explicit cohesive ties. These cohesive devices go beyond “sentence boundaries” and make the sequences of sentences to be understood as a text. These cohesive ties fall into three classes: grammatical, lexical and conjunctive. Each of these classes is divided into subclasses.

Lautamatti’s (1987) topical structure analysis views a text’s coherence as the way in which certain key words or topics are repeated in a text and the way these topics progress across sentences leaving three different progression patterns: parallel, sequential and extended parallel. Though the density of parallel progression will lead to more locally coherent texts, they often result in underdevelopment of the main discourse topic. On the other hand, more sequential progressions will result in more development of the discourse topic but too many sequential progressions will lead to drifting away from the main topic of the discourse unless occasional references are made to the main topic of discourse and
extended parallel progressions are formed. Based on TSA, a coherent text is a text which contains a proportionate number of sequential (65% to 80%) parallel and extended parallel progressions.

For Givon (as cited in Jin, 1998), coherence can be viewed in terms of topic continuity. Distinguishing between micro and macro organizational levels of language, he suggests the study of referential distance (the gap between two occurrences of a referent/topic in the discourse) topic persistence (frequent occurrence of important topics in discourse), and potential interference (the disruptive effect that other referents may have on topic availability and identification). And finally, by the same token, Hoey (as cited in Jin, 1998) sees cohesion as the property of a text and coherence is the facet of readers’ evaluation of a text. For him, cohesion and coherence are interrelated textual and cognitive features, the former facilitating the latter, and lexical repetition is the key cohesive device forming a web of bonds within a text.

Recent advances in computer technology and the development of computer applications have made it possible for the linguists, especially, the researchers working in the areas of textual analysis and Natural Language Generation (NLG), to analyze different aspects of text, using huge corpora of texts in a twinkling of an eye and with acceptable accuracy and objectivity. In the light of these developments, once again, the analysis of textual coherence and cohesion has gained a new momentum. In line with these research developments, the present study aimed at analyzing the expository essays written by the Iranian EFL and Indian ESL students in a bid to investigate the possible differences between these two groups of learners considering their use of cohesion and coherence features in their compositions. Due to the potentially large number of texts involved in the present study and the need for speed, accuracy, and objectivity, the researcher decided to use a computer application called Coh-Metrix to analyze the students’ essays. Among the numerous measures of cohesion and coherence that this software offers to assess textual coherence, the researcher picked out only those indices that approximated the cohesive ties in Halliday and Hasan’s model. These indices include explicit grammatical ties (anaphor references and sentence connectives), lexical and coreferential ties (argument overlap, word stem overlap, content word overlap, Latent semantic analysis, and hypernym value of nouns), and mental model dimensions (causality, intentionality,
temporality, and spatiality). The essay were also analyzed for topical and internal coherence through Lautamatti’s Topical Structure Analysis.

Analyzing 106 essays by the Iranian students and 73 by the Indian students, this contrastive study sought to investigate the possible differences between the Iranian and Indian students with regard to their use of cohesion and coherence features in their English expository essays. This research work built on an interdisciplinary approach that included academic writing research, cross-linguistic studies and contrastive rhetoric to achieve its goals.

Four major objectives were set for this study. Each of these objectives led to a null hypothesis comprising other components in order to become more manageable and statistically testable.

The first research objective concerned the students’ use of explicit grammatical cohesive ties such as anaphor reference (including personal reference, demonstratives, and comparatives both at adjacent and distant sentence levels) and connectives. Results of the statistical tests revealed that these two groups of students differed in the use of anaphor reference only at adjacent sentence level. Further detailed analysis of the anaphor reference also demonstrated that the Iranian students used far more personal references you, she, and we. The Iranian students, generally, used more sentence connectives. They particularly used more additive, causal and temporal connectives. On the other hand, the Indian students used definite article the more frequently and more correctly than their Iranian peers did. The observed differences could be explained by the lower proficiency of the Iranian students and mother tongue interference as Persian is said to behave resumptively allowing its speakers to insert a pronoun whenever they expect a gap.

The second research objective was set in order to investigate any differences between the Iranian EFL and Indian ESL undergraduate students as regards their use of lexical and referential features. These features included argument overlap, word stem overlap, content word overlap, and Latent semantic analysis (an index of conceptual coherence or semantic similarity). The results of the statistical analyses revealed that there was more argument (noun-pronoun) overlap difference between these two groups of students only at adjacent sentences. The results also indicated that Iranian students’ essays exhibited more semantic similarity and conceptual coherence (LSA) both at adjacent and
distant sentence levels. By the same token, there was more overlap (reiteration) of content words (words that carry most meaning such as adjectives, nouns, verbs and adverbs). In a word, the Iranian students’ essays were richer in lexical ties. However, there were no differences between the means of word stem overlap and hyponym value of nouns in the essays written by these two groups. These differences could be attributed to the Iranian students learning experience (Yamini, 1999) and mother tongue interference since Persian texts tend to be denser in lexical repetition than English texts (Sedigh and Yarmuhammadi, 1997).

The third research objective aimed at investigating possible differences between the means of situation or mental model dimensions (text’s microworld) created by the essays written by the Iranian and Indian students. In other words, the study sought to see if these two groups of students differed in the way they structured their essays so that they could be linked to the reader’s long term memory and world knowledge and contribute to the construction of mental representation of the text through employing certain connectives, nouns, verbs and particles. Four of these situation model dimensions quantified by the computer application were targeted by this study: causal, intentional, temporal, and spatial. Results of the statistical analyses revealed that despite slight differences, these two groups of students did not differ significantly. However, causal dimension appeared to surpass the other three dimensions in both groups; consequently, they tended to be more causally coherent, presumably due to the expository and cause-effect nature of their essays.

And finally, the last objective of this study was to investigate any difference between the Iranian and Indian students in the way they developed the topics in their essays. Using Lautamatti’s (1987) Topical Structure Analysis model, the researcher aimed at three types of progression in the students’ compositions: parallel, sequential, and extended parallel progressions. A null hypothesis followed from this objective. The statistical analyses rejected the null hypothesis indicating that Iranian and Indian students differed significantly in their use of different topical progressions. The Iranian students used far more sequential progressions while their Indian peers used more parallel and extended parallel progressions. Further supplementary analyses showed that both groups
could not live up to the progression norms set for quality essays (65% to 80% of sequential progression).

Results of the Univariate analyses proved that the gender of the students did not significantly affect the results of any of the tests of significance, although the essays written by the male and female students differed in lexical devices and types of topical progression patterns.

Supplementary analyses of the essays also indicated that although the Iranian EFL students showed a tendency to write more yet shorter paragraphs and Indian ESL students preferred to write significantly fewer but longer paragraphs, these students’ essays demonstrated similarities in the physical characteristics such as the number of words, the number of sentences, the length of their words and sentences, the (Flesch-Kincaid) grade level and the (Flesch) ease score of their compositions.

The End