The aim of the current study is to give a possible explanation of some common errors committed by Yemeni learners in English Departments in the Faculty of Education.

The actual situations where learning takes place and learners’ errors can be investigated need to be closely observed in order to find out what affects learning. Implementing various second/ foreign language acquisition theories may be a key to understand the activities of teachers and learners which promote or hinder acquisition (Ellis 1999).

The situation which gained interest in this study is the process of teaching/ learning English in Yemen. The common inadequacies in the texts written by the students and the other published forms are being found not naturally sounding English. According to Johan Myles (2002), focusing on written process as a pedagogical tool is appropriate if attention is given to linguistic departments, and if learners are able to get sufficient and efficient feedback with regard to their errors, so the departments of English in the Faculties of Education are selected.
The study attempts a diagnostic process of the students’ errors. It investigates the related errors to the ability of the students to identify and use some lexical cohesion devices. The main linguistic areas included in the study are:

(i) lexical cohesion (collocation)
(ii) coherence
(iii) syntax

The significance of the study lies in implying a new research methodology, Learner Corpus, in analyzing the errors of the Yemeni learners in identifying and using collocation. Another important point is tracing the performance of the same groups throughout their college education period taking their autonomous learning progress into consideration.

The thesis consists of five chapters.

**Chapter one** is a brief introductory chapter. General remarks about English in colleges of education in Yemen have been given in the opening part. The objectives and the rationale behind undertaking the study have been stated. Then, the general organization of the study has been presented. The specific items included in the data have been clearly provided in the scope of the study. As the study investigates errors
in written English, lexical cohesion, syntax and coherence were selected to be covered in the analysis. Moreover, the methods adopted for selecting the population, the test and using the tools in the analysis have been itemized. Then, the core element of the study, hypothesis, have been stated. By the end of the chapter, the significance of the study and the limitations have been mentioned.

Chapter two defines the theoretical concepts establishing the background of the study. The different points of views in Applied Linguistics and Error Analysis have been scanned. Then the relatively new linguistics branches; Corpus Linguistics, Learner Corpus and The Lexical Approach have been brought to light as being the essential platform of collecting and analyzing the data. The importance of referring to these areas of linguistics, when analyzing students’ errors, has been identified. The chapter moves on to the basic language concepts implemented. The traditional and the modern treatment of cohesion especially lexical cohesion have been presented. More importantly, the emergence of collocation as the representative of lexical cohesion has been sufficiently elaborated. Syntax and coherence have been discussed. The interaction of the three language concepts, namely collocation, syntax and coherence, in analyzing errors has been interpreted. The last part of the chapter deals with the contribution of scholars in investigating the status of English in Yemen and the problems faced by Yemeni students. Other studies covering
similar points related to the study in domains other than Yemen have been taken into consideration too.

Chapter three has been classified into three parts. The variables affecting the learning/teaching process; learner autonomy, teachers’ role and teaching materials have been emphasized in the first part. The second part is about the test. The objectives of the successive tests given to students to form the data have been specified. The selected lexical items for the tests have been indicated. Part two ends with clarifying the administration of the test. In the third part the testees educational background has been highlighted.

Chapter four contains the main body of the work. First, the methodology adopted for the analysis of the data has been pointed out. Next, the students’ answers have been thoroughly examined in the result section. The errors spotted in the texts have been classified and listed according to the groups (urban and semi-rural) and levels (second, third and fourth year). Errors in collocation have been identified first, and then the influence of those errors on syntax and coherence is stated. In questions where the students are expected to identify, select or replace certain items, special tables have been prepared. These tables have been interpreted with reference to the statistical figures required.
Chapter five concludes the study and evaluates the students’ performance. The eminent sources of the students’ errors have been summarized. In the second part of the chapter, the remarkable observations of the students’ errors and general progress over the period investigated (second year, third year and fourth year) have been provided.

The arguments of the researcher have been found valid and justified. After that, the pedagogical implications of the findings have been highlighted in order to narrow the wide gap between the present status of the Yemeni learners’ performance and the expected advanced level of the output of the departments of English. Some practical steps to be taken by teachers and teaching material designers have been recommended towards the end of the chapter. The chapter closes with some suggestions of interesting topics as further research work.