CHAPTER TWO

THEORETICAL FRAMEWORK
II THEORETICAL FRAMEWORK

2.1 General Remarks

2.1.1 Applied Linguistics

Applied linguistics is a problem based combination of theory and practice, as presented by many linguists such as Trask (2004), Corder (1973), Widdowson (1986) and Aronoff and Rees-Miller (2003). It is, usually, linked with language education and its practical activities as presented by Corder (1973). Unlike Corder (1973), linguists like, Varshney (2005), Candlin (2001) and Aronoff and Rees-Miller (2003) relate applied linguistics to many fields such as philosophy, psychology, sociology, anthropology, literacy, language acquisition, stylistics, translation and others beside education. They make the range of applied linguistic considerably wide; hence, all that influences language use is worth investigation.

While most of the linguists treat applied linguistics as a consumer of linguistic theories, Widdowson (1984, cited in Cook and Seidlhofer, 2001), gives it a dynamic interaction providing useful insights from practical activities. He makes the learners’ experience, as a language user, a conditional goal for researches in applied linguistics. Davies (2007) and Rees-Miller (2003) generalize the role of applied
linguistics into providing the best explanations for certain problems which make generalizations available and applicable. Thus, it is not supposed to solve problems in any field, it only helps diagnosing them.

As applied linguistics raises language-based problems, mostly occurring while acquiring a second or a foreign language, it is essential that applied linguists have certain amount of experience in the field of teaching/learning. It is, equally important that those who are teachers or teacher trainees have enough background of the basics of linguistics, education, and psychology (Rees-Miller, 2003).

2.1.2 Error Analysis

Error analysis is a discipline which was developed from applied linguistics by linguists studying foreign language learners’ errors, like Corder, Crystal, Davies, Wilkins, Richards and others. They investigated the learners’ “approximative system” (Nemser, 1984) or “interlanguage” as Selinker (1984) calls it, especially errors in their learning system. They make a distinction between mistakes (performance rule-breaking) and errors (competence rule-breaking) and regard the latter worth investigation. Jain (1984) justifies the need to investigate errors stating that they would help to determine the areas which need more attention and the type
of that attention. By analyzing those errors a linguist can give accurate diagnosis and help to bring the deficient language area from where it is to where it ought to be (Crystal cited in Davies 2007).

Traditionally, the linguists used to focus on the improvement of language learning. Then they shifted their interest to answer the question ‘What is it that is not being improved?’ The actual situation where learning takes place and learners’ errors can be detected needs observation 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).

Attempting such a work, the analyst must bear in mind how convenient is his work for those who are interested in remedial work. His comments and suggestions are expected to be clear and practical. Consistency in identifying the nature of errors is emphasized by Friermuth (1998). Such results can help in the practical correction of errors if they fulfill the condition stated by Jain (1984), which is to tackle the common erroneous restricted generalizations, not merely correcting their rooted errors.
There are some commonly reported sources of learners’ errors. Richards and Sampson (1984) detected the following sources: (i) language transfer, (ii) intralingual interference, (iii) sociolingual situation, (iv) modality, (v) age, (vi) succession of approximative systems, and (vii) universal hierarchy difficulty. Richards (1984) added (viii) ignorance of rule restrictions, (ix) incomplete application of rules, and (x) false concept hypothesized. Selinker (1984) added an important source that is ‘fossilization’ which includes five central processes; (i) language transfer, (ii) transfer of training, (iii) strategies of second language learning, (iv) strategies of second language communication, and (v) overgeneralization.

Despite all what has been said, by many researchers, about the need and significance of error analysis, Schachter (1974, cited in Ellis, 1999) regards it insufficient to solve the learners’ problems. He believes it is incapable of specifying certain deficient language rules/items in the learners’ interlanguage when the learners’ apply the strategy of avoidance.

2.1.3 Corpus Linguistics

Corpus linguistics is a relatively new branch of linguistics which flourished during
the 1980s and early 1990s (Flowerdew, L 2002). Tognini-Bonelli (2004) has explored some definitions of a corpus (given by Leech 1997, Francis 1982, Sinclair 1991 and Aarts 1991 cited in Tognini-Bonelli, 2004) and has come up with a similar definition to the one given by Mcenrye (2003) as ‘a large body of linguistic evidence typically composed of attested language use…collected together… to be used for a range of research purposes’. According to Tognini-Bonelli (2004), the aim of corpus linguistics is to present analysis and description of language use as realized in text(s). Dehcheshmeh (2007) observes that the introduction of corpus linguistics does not include presenting new information about language, but giving new perspectives to linguistics researches. These researches can be seen as a shift from the formal towards a more practical approach (Tognini-Bonelli, 2004).

The use of computer is a major improvement on the situation which error and interlanguage analysis faced in the 1970s and 1980s (Kaszibiski, 1999). The data should be machine readable as it gained access through sophisticated language processing systems (Mcenrye, 2003). It encouraged the appearance of some softwares designed specially to serve the purpose of looking at the use of language automatically, for instance, Wordsmith Tools developed by Scott in 1996 and MonoConc Pro developed by Barlow in 2002. Concordances, word list, key word, are mainly used in such researches. These tools facilitate analyzing collocation,
colligation, and interrelation between lexis, grammar and semantics. And, for the comparative purpose some general corpora are found useful, for instance, Dehcheshmesh (2007) suggests the 100- million words British National Corpus (BNC) and the 400- million words Bank of English as being among the best known general corpora.

2.1.4 Learner Corpus

Learner corpus is a branch of linguistics which was influenced by the ‘corpus revolution’ and the rise of the Lexical Approach (kaszubski, 1999). The body of the corpus is a collection of texts written by foreign language learners. It focuses on the teaching process, particularly, used for error analysis (Granger 1994 cited in Tognini-Bonelli 2004). Researchers can identify specific patterns in students writing, analyze them and get them compared to native speakers’ corpora (Tognini-Bonelli 2004). The native speakers’ corpora are taken as a norm or at least the target language of learning. The comparison could have the objective to find the difference and similarities in the relative frequencies of words, collocations, idioms, grammatical structures, discourse features etc, and their distribution in text (Kaszubski, 1999). It encourages learner autonomy by developing the skills required to identify, explain and rectify recurrent errors (ibid).
This branch of linguistics is the most relevant area to the objectives and methodology of the current study.

2.1.5 The Lexical Approach

The Lexical Approach is a modern trend which has been presented by some linguists. The first is Lewis who wrote the book ‘The Lexical Approach’ (Lewis, 2001). Like the Communicative and Natural Approaches, it places communication of meaning at the heart of language learning. The emphasis is on acquiring lexis, as being considered the main carrier of meaning, due to repeatedly meeting and noticing them in their natural contexts, then using them naturally. Repetition here is unlike the behaviorists, who emphasized repeated practice as a way of fixing patterns (ibid). The basic argument of this approach according to Lewis (2001) is that language consists of chunks which, when combined, produce continuous coherent texts. The approach explores many language items especially collocation, colligation and semantic prosody.

An important aspect of this approach observed by (Islam and Timmis, 2007) is that lexis and grammar are closely related. Skehan (cited in Lewis, 2001) supports using
corpora in the Lexical Approach saying that it helped realizing that generalizations converted into rules tend to be partial or more restricted to a particular genre- than has usually been believed. In the traditional grammar, accuracy gained more interest for being viewed as the best way to acquire fluency. Whereas, in the lexical Approach, accuracy is the natural outcome of fluency, which is the result of using prefabricated chunks.

Scott Thornbury (cited in Lewis, 2001) criticizes the Lexical Approach expecting ‘fossilization to occur when the learners become dependent on the lexicalized language at the expense of engaging the syntactisized processes. This could be true if the approach ignores syntax. In fact, the approach presents a modified view of syntax and vocabulary interaction as its stable base to produce coherent texts. The fundamental principle of this approach is the insight that language consists of grammaticalized lexis not lexicalized grammar (Lewis, 2001).

Learners’ insight and recognition is very much counted in this approach. Lewis (2001) makes it essential to make learners aware of the lexical nature of language to develop an understanding of the kind of chunk found in the texts they meet.
2.1.6 Cohesion

Cohesion is the relation established in a sentence between it and the sentence preceding and following it by using connectives (Syal and Jindal, 2003). This definition reflects the traditional understanding of cohesion in its dependency on connectors as the basic cohesive devices. Having this in mind many researchers have thoroughly observed four out of the five main cohesive devices identified by Halliday and Hassan (1976). Reference, substitution, ellipsis, conjunction used to be central in cohesion analysis whereas lexical cohesion occupied a minor position. Actually, some linguists regard lexical cohesion as the most difficult type of cohesion. Witte and Faigley (1981) justify this point of view by finding no clear form of connection between collocates as they involve neither repetition, synonymy, super ordination, nor, mention of general items and that they collocate only because they happen to share the same lexical environment.

Lexical cohesion recently acquired great interest especially with the rapid development of corpus-based studies and computer assisted research work. It is the cohesion that arises from semantic relationships between words (Morris and Hirst, 1991) which makes concordance an essential participant in any analysis of these
devices. The use of concordance in analysis has the merit of locating a particular word in a text wherever it appears and identifying its pattern of usage.

Halliday and Matthiessen (2006, cited in Talib 2007) divided cohesion into paradigmatic and syntagmatic relations, with collocation being the only syntagmatic category. They made collocation a representative of the lexical cohesion devices. Some other linguists supported this view, such as Witte and Faigley (1981) who think that lexical cohesion is achieved through lexical items association – collocation, and that lexical cohesion is the predominant means of connecting sentence meaning (coherence) in discourse.

2.1.7 Collocation

The term collocation was first introduced by Firth (Higuchi, 2004) then it has had many definitions talking about the co-occurrence of words in natural texts with greater than random frequency (Coffin, 2004, Tognini-Bonelli, 2004, Lewis 2001, Firth, cited in Lewis, 2001 and Moon, cited in Williams, 2003). Firth’s statement that ‘You shall know a word by the company it keeps’ is regarded by Walsh (2005) as an ‘opposite view to the ‘atomic’ view that tends to see language as ‘individual words containing inherent meaning which simply fit into structural or grammatical
Islam and Timmis (2007) identify different terms used as alternatives to collocation such as prefabricated phrases, lexical phrases, formulaic language, frozen and semi-frozen phrases and lexical chunks.

Collocation has been neglected for a long time for the difficulty of undertaking a work on it, besides, regarding it as the most difficult aspect in second language learning (Shei and Pain, 2000). It is preferable to teach collocation to advanced learners to help them leave the ‘intermediate plateau’ (Lewis 2001) and continue producing language containing natural sounding elements (ibid). Biber and others (cited in Lewis 2001) drew a great conclusion, that many grammatical generalizations have limited domain of applicability. For Carter (cited in Higuchi, 2004) the patterns of chaining lexis is interpreted as collocation in lexical studies and as structure in grammar. Therefore, there is a reasonable necessity to expose learners to the prefabricated chunks (collocation) together with other language elements to overcome grammatical limitations.

Collocation is of two types; lexical collocation which combines two lexical components and grammatical collocation which combines a lexical word, typically a noun, verb or adjective with a grammatical word (Lewis 2001). McCarthy (1990,
cited in Higuchi, 2004) gives importance to collocation as being the basic organizing principle in teaching content words.

Teaching collocation and training the students to use available collocations in writing enables the students to produce language fluently and accurately. It helps the user of a language to think more quickly and communicate more efficiently (Hill, 2001). In addition, using culturally popular lexical sequences and regular grammatical functions, with adaptation if necessary to contextual fit, give learners a chance to use language coherently, fluently and with less grammatical errors and pragmatical failures (Nyyssonen, 2001)

The first thing to develop among the students’ is the ability to notice the prefabricated chunks. The basic claim in acquiring collocation is first receiving them, then using them again as they were seen at first as one unit. Native speakers, in using collocation, do not analyze them into their constituent parts, they use them as they are.

Teacher should help building learners’ autonomy aiding them with strategies of learning collocation not giving mere rules. Unlike most language items, there are no rules to be given for collocation. Teachers may face problems and difficulties
fulfilling the needs of their students in this area, however, they have to- ignoring it will not help.

Encouraging learners to use collocation does not kill their creativity; they always can say something new (Hoey, M, 2004). It might be worth noticing, while teaching collocation, that each word is expected to have its own collocational field (Williams, 2003), also, that each word has its own grammar (Lewis 2001). Taiwo (2001) spots the importance of keeping the mutual expectation that holds between words unbroken because it is not enough to know the collocates of the words, the context of use is equally important.

Collocation has often been a broad area of students’ errors (Conzett 2001). Mis-collocations occur due to a problem in the storage of the phrasal mental lexicon of learners. Hill (2001) introduced ‘collocational competence’ and attributed some grammatical errors to its inadequacy in certain occasions. Sonaiya (1988) (cited in Taiwo, 2001) regards lexical errors more serious than other types of errors because inappropriate choice of words affects and hinders communication most. The limited vocabulary students have can be an obstacle they face in acquiring and using new chunks (Kavaliauskienė and Janulevičienė, 2001). Learners after several years of studying a foreign language often know many grammatical frames but have very
little to put into them, as observed by Rudzka et al (cited in Higuchi, 2004). Even advanced learners may suffer from the same problem by making inappropriate or unacceptable collocations (McCarthy, 1990 cited in Higuchi, 2004). They are not able to identify the natural collocates of words to be used. The impact of lexical errors on syntax is emphasized by Shaughnessy (1977) too.

Words in actual use carry additional information beyond their simple meaning, which usually accompanies them when being stored in the mental lexicon and being recalled afterwards. Being similar in meaning does not permit words to be interchangeable. Relying on dictionary meaning without careful association of actual use and the additional information may promote another type of collocational errors. Partial synonym, according to Taiwo (2001), is sometimes responsible for inappropriate choice of lexical items. He tackled another serious problem, which is when learners can not identify the semantic boundaries that separate lexical items and fail to observe the rule of restrictions on the co-occurrences of lexical items; the result is producing incoherent and meaningless texts.

Collocations and expressions are thought to be the most important types of lexical phrases (Kavaliauskienė and Janulevičienė, 2001). These chunks alone, without colligation and semantic prosody, do not account fully for how words are used in
language (Hoey, 2001). These two areas integrate with collocation in studying vocabulary behavior in a discourse.

2.1.8 Coherence

If the text components fit together meaningfully and the sentences flow smoothly, it is said to be a coherent text. Achieving written coherence, in the absence of verbal clues to direct and support the message, requires very carefully planned patterns of coherence, as noted by Kies (2005). He treats coherence as being the product of two factors - paragraph unity and sentence cohesion (ibid).

There are many approaches to study the meaning of a text. One of these approaches is The Language-internal Approach (Taylor, 2002). There are two ways to implement this approach; to focus on pragmatic relation- on the relation between different expressions (synonym, antonym and entailment), and to focus on syntagmatic relations - on relations between items which co-occur within an expression (word collocation) (ibid).

Then, studying collocation is done in the syntagmatic relation which can be part of studying cohesion (Halliday and Matthiessen, 2006, cited in Talib, 2007) and
coherence (Taylor, 2002). As a result, investigating errors in collocation implies reference to cohesion and coherence.

ESL learners’ written texts, exposing the use of simple vocabulary to express both simple and complicated ideas may be a reflection of improper teaching which focuses on the decontextualized lexical items as listed in the dictionaries (Taiwo, 2004). Thereby, these learners lose the insight of word association and are encouraged to use the most general items in their writing.

2.1.9 Syntax

Syntax as Syal and Jindal (2003) defines it as the level at which we study how words combine to form phrases, phrases combine to form clauses and clauses join to make sentences. It also attempts to describe how these elements function in sentences. Combination of words should lead each word to play its assigned function to build accurate structures. Hence, the basic element in syntax is word relationship that constitutes the accurate structures.

Lexis should be handled carefully in texts because if there is a problem in word association it may create a syntactic complication like writing heavily nominalized
sentences when the appropriate verb to be used is not known (Shaughnessy, 1977). Another syntactic problem may arise when students know the bare word but have no sense of the kind of relationship that word is permitted to have with other words in the sentence (ibid). Taiwo (2004) supports the view that lack of collocational competence has serious effects on syntax especially when learners do not know the collocation which expresses precisely their thoughts.

2.2 Literature review

There is lot of work done on corpus linguistics, the lexical approach, collocation and self-access. The researchers and linguists have done a great work and enriched these fields but their work has little to do with the situation in Yemen. They investigated other domains with different objectives and problems.

There are some researchers who investigated the status of English in Yemen and the problems faced by Yemeni learners though with a focus different from that of the present study. The accessible relevant work (for the researcher) was the work done by Al-Raymi (1999), Al-kadasi (1999), Abdul-Mughni (1999), Al-Magalih (1998) and Mahdi (2003).
Al-Raymi (1999) focused on the textual components which deal with the creation and organization of the text, namely, theme structure, information structure and cohesion. His aim was to assess the teaching of cohesive devices in English in Yemeni colleges of education. The study investigated errors in cohesion. It included grammatical devices (clausal substitution, reference, conjunctions, ellipsis, nominal and verbal substitution) and excluded lexical devices. He summarized the problems related to cohesive devices in overuse, misuse and omission. The study showed that learner deficiency in recognizing and using the cohesive devices may be due to the fact that they are generally ignored in the teaching process and to the mother tongue interference.

Abdul-Mughni (1999) worked on the comparison of adverbials in Arabic and English. His main work was about the grammatical function of adverbs. In his analysis he did not include collocations yet he mentioned it in his findings as an area of interest which teachers and syllabus designers should take into account. He, also, suggested it as an important area for further investigation.

Al-Magalih (1998) studied Yemeni learners’ abilities in translation from Arabic into English and from English into Arabic. Analyzing the texts, he found out that those who are not efficient in writing English compositions lack this ability in their mother
tongue too. The study revealed that when learners compose in English, they first compose the sentences in their mother tongue (in their minds if not in writing) then translate them into English. In this process the selection of words depends on their mother tongue which is not always reliable due to the difference between the two languages.

The aim of Al-Kadasi (1999) was to focus on the most common syntactic errors committed by Yemeni learners. He found fifteen areas of errors; tense, article, preposition, relative clauses, word order, concord, auxiliaries, reported speech, adverbs, connectors, conditionals, passive, copula, modals and adjectives. The sources of errors, as he noticed, were the mother tongue interference, intra-English influence and faulty teaching techniques. He does not include lexis and syntactic interaction in his study.

Mahdi (2003) made a contrastive study of prepositions in Arabic and English. He tested a first year group and compared their performance to a fourth year group. The study focused on using the prepositions correctly in sentences. The company a preposition keeps in various expressions was not investigated. He finds that correct answers were given in familiar subject matters. The main factor affecting the students’ performance was mother tongue interference.
There are scholars who studied some related aspects of students’ errors in other areas and came out with useful findings. One of these scholars is Inani (2000) who adopted a contrastive-rhetoric approach in studying some texts written by Egyptian learners. He found some common errors in the structure of sentences and the composition of the texts. He claimed that some of these problems were also found in essays written in English by native speakers of English. He argued that the errors emanate from the negative effect of spoken English and mother tongue interference.

Biber and Conrad (2004) provided a brief introduction to the use of corpus in research work. They made a comparison among registers. They developed a systematic method to provide empirical evidence of how language users adapt their language for different situations.

Joodi (1998) examined ellipsis in English and Arabic in a comparative study. He concluded that ellipsis is a problematic area mainly because of the negative transfer of mother tongue.

The problem of de-lexicalized words was discussed by Dowling (2004). For him, the students’ avoidance of using de-lexicalized verbs prevents them achieving more naturally sounding language.
Kavaliauskienė and Janulevičienė (2001) presented a good example of implementing the lexical approach in a classroom. They believed that in ESP language learning, high-frequency lexical items should be taught first. In addition, they emphasized that learning ESP in multi-word chunks means a change for the better in the L2 vocabulary acquisition. Learners’ autonomy was considered in the study and some practical suggestions to promote it were provided.

Tono (2002) undertook a research on learner corpus definition. His work interpreted the area and the development in understanding learners’ difficulties. He connected learning with the learners’ external factors, learners’ internal factors and individual learner differences. His basic argument was that data is fed and received reciprocally in-between these areas.

Mahmoud (2008) raised the point that learners’ linguistic and pragmatic errors are intensively studied while research on the ‘strangeness’ of linguistic forms and expressions is lagging behind. He studied some errors made in collocation by Arab learners and found that some collocation types are more problematic than others. He suggests teaching collocation at early stages with the help of special lists. The need for bilingual dictionaries of collocation was emphasized in teaching foreign learners.
Gerson (2006) conducted a study on teacher’s response to collocational errors at the college level. The purpose of his study was to explore connection between instructor’s belief and practice of formulaic language as well as to provide evidence that more emphasis on the area is needed in teacher training courses. He highlighted some areas which need to be introduced to the trainees, especially non-native like selection errors.