Academic Lexical Literacy: Investigating the Cohesion of Arabic Speakers' Essays in English

Nahla Nola Bacha
Lebanese American University

Martin Cortazzi
Brunel University

Fouad Nakhle
Lebanese University

1. Introduction

One way ideas are linked in texts is through cohesion, the network of surface features that connect texts and signal underlying relationships within and between sentences. Cohesive features include reference, conjunctions, substitution, synonyms, to name only a few. Cohesion has been a major area of concern in text analysis since the 1970s, and a number of advances has been made towards the understanding of texts. Since Halliday and Hasan's (1976) seminal work on cohesion, research has included work on both native and non-native student compositions to see how different cohesive devices in these texts contribute to the production of "good" writing. While cohesion may not be a characteristic of all 'good' writing, it seems to be one significant element that marks much high-rated writing (Grabe & Kaplan, 1996).

De Beaugrande & Dressler (1981) identify cohesion as one of the seven standards without which a text would not be a text. Harnett (1986) argues that "...words that signal relations are important resources for writers ... if "good is expressed partially ... through linguistic cohesion, it seems useful to analyze cohesion in writing..." (p.142). Peters (1986) further argues that the use of cohesive devices becomes even more significant in *academic writing* since it is a form of a monologue and "...thus apart from indications of macro-structure, readers expect ...micro-structure details to supply cohesion and to show the logical connections between one statement and another" (p.170).

Linguists note the importance of a specific category of cohesion, that of the connective function of lexis in texts (Bloor & Bloor, 1995; Halliday & Hassan, 1966; Hoey, 1991a; Sinclair, 1966 among others). Although it was with Hoey's (1991a) work that lexical cohesion as a replicable analytical tool in texts became more explicit, linguists have long commented on its importance and attempted to define it. Halliday (1966) mentions the importance of devising methods to study the lexical patterns in language and comments that "... the cohesive power of lexical relations, are of great potential interest" (p.160). Sinclair (1966) adds that "A study of these tendencies ought to tell us facts about languages that cannot be got by grammatical analysis, ..." (p.411).

2. Aim and Significance of Study

L1 cohesion studies have influenced L2 research but few of the latter have focused on lexical cohesion. The present study provides an in-depth textual analysis of lexical cohesion in academic expository essays written by L2 Arabic learners of English in an attempt to offer further evidence for the differential use of lexical cohesion in high- and low-rated academic texts. Specifically, the study attempts to investigate two aspects of lexical cohesion: the repetition of links occurring in these texts, operationally defined by thirteen types (see Table 1b), and the distance between these links over the organization of the text as operationally measured by the number of intervening sentences between the cohesive links.

The purpose of the foregoing is to gain a better linguistic understanding of lexical cohesive patterns in these non-native speakers' academic texts. The study represents the first cohesive analysis in Lebanon and the first application of Hoey's (1991a) model of lexical cohesion to L2 writing, which methodologically could open new avenues for further textual studies on the topic. We first outline relevant theoretical background of research into cohesion in academic writing and Hoey's (1991a) model. Then, we give details of the method, the results obtained and a discussion of the findings.

3. Cohesion in L1 Expository Texts in a Theory of Language

In Halliday and Hasan's (1976) systemic-functional model, language is made up of three major parts: the *ideational,* (concerned with the

expression of content), the *interpersonal*, (concerned with the social, expressive and *conative* functions of language) and the *textual* (concerned with the structural and non-structural systems that create language). They define text as "…a unit of language in use which could be of any length …a semantic unit - a unit not of form but of meaning" (pp.1-2). This semantic communicative function of a text overrides the grammatical unit of sentence; it is something superordinate to a sentence but is realized by sentences. The text must be a unified whole in relation to its environment; that is, the other sentences. To help unify the text, there are what Halliday and Hasan (1976) refer to as *cohesive ties* between and within sentence, structural representations on the surface of discourse to show both grammatical and semantic underlying relations. These cohesive ties are what gives texture (or coherence) to a piece of discourse; they make a text a text. The greater the number of ties, the tighter the cohesion and the clearer the text is. Cohesive ties either point backward to a referent (anaphoric - most common in English) or point forward (cataphoric). A text may be made up of a system of cohesive chains (more than one set of ties) which may be of varying degrees of density depending upon the sophistication of the writer's language, the purpose of the writer, and the type of writing at hand. (Halliday & Hasan, 1976).

According to Halliday and Hasan's (1976) categories, there are three types of cohesion: *grammatical, lexical* and *conjunctions*. Lexical cohesion is of two types: *reiteration* and *collocation* which form a large portion of cohesive ties and "…are independent of structure and may span long passages of intervening discourse" (Halliday, 1994, p.311). In *reiteration* (paradigmatic cohesion), the repetition of words or phrases may be through the use of synonyms, antonyms, hyponyms (superordinate to part e.g. family to father, mother, children etc.), and meronyms (part to whole, e.g. finger to hand) where the two occurrences have the same referent and are accompanied by "the" or a demonstrative; for example, "this" (Halliday & Hasan, 1976, p.318). In *collocation* (syntagmatic cohesion), there is an occurrence of individual items which are associated by a semantic field; for example, "father" collocates with "mother" and "children" in a semantic field of family relationships which exist as hyponomy. Some researchers note that this lexical category is not very clear in Halliday & Hassan's (1976) taxonomy and seems to be a collection of miscellaneous features (see Stotsky, 1983; Hasan, 1984). A revision of the taxonomy (Halliday & Hasan, 1989) shows the two major components divided into non-structural and structural and the

lexical cohesive category more clearly identified. Halliday and Hasan (1989) also revise their view of the text and the cohesive devices as part of a context (context of situation) which they qualify to include three features of the context. The first feature is the field of discourse which refers to what is happening socially "....the general sense of what it is about". For example, in a love poem "..the field of discourse is love.." (p.24). The second feature is the tenor, concerned with the personal relationships or participants in the text. The third feature is the mode: e.g. speech, writing, and genre.

Hoey (1991a), however, claims that cohesion in text is more of a study of lexical patterns. He argues that since text cannot be predicted, it is, therefore, a non-structural element which, along with lexis, mediates and overlaps with syntax and phonology and are realized in either oral or written texts through the interaction level.   It is clear that they are both on the same level of importance in regard to language. The map shows the triple structure of language: phonology, syntax and interaction and the arrows in the map indicate the relations. (see Figure 1).

Figure 1   Hoey's (1991a) Map of Language
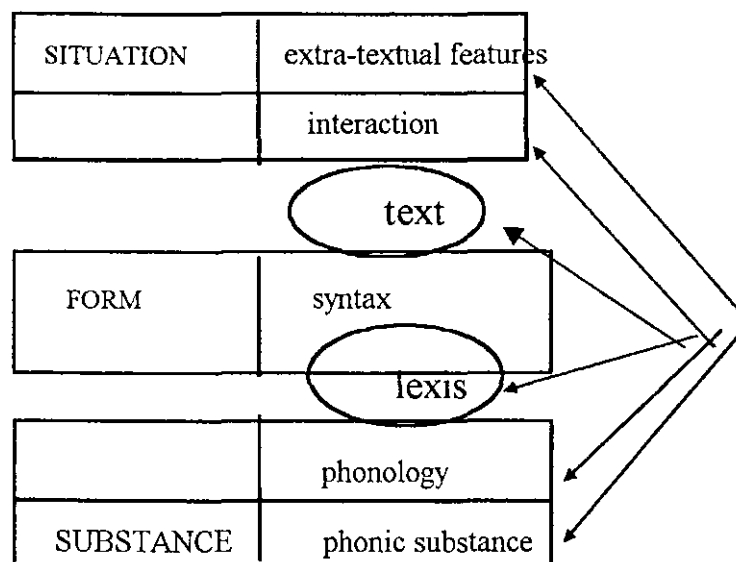
The implication of Hoey's (1991a) theory of language, is therefore, that text and lexis are interrelated and claims that "…it is not only the case that text is lexically signalled, it is also the case that lexis is textually established" (p. 220).   Thus, Hoey points out the importance of investigating "…how lexical cohesion and text organization affect each other" (p.220).

3.1 Lexis and Organization of L1 Academic Texts

Linguists have noted the importance of lexis in indicating the macro-structure of texts (Carter & McCarthy, 1988; Dudley-Evans, 1994; Francis, 1994; Halliday & Hasan, 1976; Hoey, 1991a, 1994; Hunston, 1994; Jordan, 1984; Lewis, 1993; McCarthy, 1991; McCarthy & Carter, 1994; Stotsky, 1983). McCarthy (1991) significantly notes three major types of lexis. The first two are grammatical and lexical words (closed or function words as opposed to open or content words). Examples of the former are the articles, verbs, demonstratives, prepositions and the like; examples of the latter are *monkey, noise, toenail* etc. which, according to him, show larger patterns of text (p.74). The third class, the discourse-organizing one (also referred to as lexical signalling) such as *basis, case, cause, these questions, the issues* either refer back or forward in a text and are comparable to Hoey's (1994) and McCarthy's (1994) discourse organizers. Carter and McCarthy (1988) comment that "...lexis is not a boundless chaos; organizational principles are available and simply wait to be more fully exploited" (p.38). Hoey (1991a) argues that "Lexical cohesion is the only type of cohesion that regularly forms multiple relationships...becomes the dominant mode of creating texture ... cohesion is the study of lexis, and the study of cohesion in text is to a considerable degree the study of patterns of lexis in text" (p.10).

3.2 Lexis and Repetition in L1 Academic Texts

Studies have shown that lexical repetition has both a connective and organizing function in developing ideas over text (Halliday & Hasan, 1976; Hoey, 1991a; McCarthy, 1991; Weissberg, 1984; Winter, 1979; Youman, 1991). Winter's (1979) work, particularly relevant to this function, emphasizes the importance of repetition as "replacement" in contributing to the context of adjoining sentences and notes the loss of meaning when sentences are quoted out of context, as Quirk (1952 in Winter, 1979) finds true of words. Winter finds it necessary to define "systematic repetition" "...as the significant repeating of one or more of the constituent features of clause of a first member within the structure of a second member, where it becomes a new sentence or part of a new sentence "(p.102). Thus, he used "repetition" in a broader sense than it had been viewed "... to stand for all its connective functions between clauses: deletion (or ellipsis), substitution and lexical repetition" (p.102).

Hoey (1991a) also notes a different view of lexical repetition in cohesion as " ...the occurrence of one or more items in a sentence that by themselves tell the reader ... nothing new but reinstates some element(s) from earlier sentences so that something new can be said about them" (p.268). He also discusses the difference between text-forming repetition (used in his analysis) and chance repetition in which two criteria distinguish these two types. *Text repetition* is "a pair of lexical items [that] both refer to the same 'object' (real or imaginary)..." whereas *chance repetition* "...is taken to be repetition where the only common ground is the choice of the same lexical item" (p.56) and not contextually determined in the text. He further explains that "When two lexical items are connected, they are said to be linked..." (p.51), defining a link as "... a connection by repetition between items in a text ... (p.266). Hoey (1991a) describes different types of repetition links that connect text (see Table 1a).

4. Taxonomies of Lexical Cohesion in L1 Academic Text Analysis

A number of proposed categories of lexical cohesion has been outlined in taxonomies and used to analyze academic expository writing (Coulthard, 1994; Cruse, 1986; Donnelly, 1994; Ehrlich, 1988; Harnett, 1986; Halliday & Hasan, 1976; Hasan, 1984; Hoey, 1991a; Phillips, 1985; Stotsky, 1983; Winter, 1979). Most of the studies had adopted Halliday & Hasan's (1976) taxonomy.   While the value of the latter was recognized,   its shortcomings in using it to analyze L1 or L2 expository texts were noted by some. Stotsky (1983) points out the unsuitability of using Halliday and Hasan's (1976) taxonomy which is based on an examination of a narrative "conversational literary text" [*Alice in Wonderland*], in studies using academic expository texts. She states the need for a taxonomy that would take into account the different vocabulary and organization found in academic texts.   For example, she finds Halliday and Hasan's (1976) scheme limiting in that it does not account for the 1) use of derivatives and derivational elements (e.g. "nominal", "nominalization"'), 2) use of a superordinate followed by a subordinate (as in expository essay writing it is usual to find instances of a general word preceding an example, e.g. "societies" – "civilians"), 3) different types and preciseness of referential repetition; the identification of lexical repetition need not refer back to a common referent and, in fact, "...the second occurrence may be, as far as reference is concerned identical, inclusive, exclusive and unrelated" (Stotsky, 1983 pp.433-435) and still be counted as reference, and 4) terms that also contrast (e.g.

"employer/worker") are often not picked up if they are part of collocation considered as indicators of high-rated texts (Cooper & Odell, 1977 in Stotsky, 1983).

Stotsky (1983) reformulates Halliday and Hasan's (1976) cohesive categories into a framework more applicable to expository academic essays. She claims that "If lexical rather than grammatical cohesion is the most significant kind of cohesion in academic discourse, future research …may yield clearer … insights about the use of lexical resources … than the original scheme by Halliday and Hassan" (p.440). Ehrlich (1988) in the same spirit argues that "…… it is not the mere presence or frequency of cohesive devices which determines whether a text will be cohesive or not" (p.112) and cites Scarcella's (1984) work with non-native texts wherein it is claimed that "…it is the appropriate fit of these devices to the context and their distribution throughout a text which determines their effectiveness" (p.112).

Hoey (1991a, p.6) observes that most taxonomies of lexical cohesion show ways of repeating, but they do not extend over large stretches of text and do not distinguish between types of texts; his analysis being focused exclusively on those of exposition. He notes that narrative texts connect in different ways to that found in expository ones, and thus the reliability of the findings from studies that rely on Halliday and Hasan's (1976) taxonomy are questionable. Hoey (1991a) illustrates the limitations of the taxonomy, one being in not distinguishing sufficiently between items. For example, he notes how in the sub-class of reiteration relations of exact repetition ("book-book"), synonymy or near synonymy ("book-volume"), superordinate ("spaniel-dog") or general word ("spaniel-dog-creature") are all accepted as referring back to the same item. This makes it difficult to differentiate among them and, therefore, to classify them in any systematized way under reiteration. In another category, collocation, in which lexical items usually occur together, (e.g. "doctor", "nurse", "needle"), Hoey (1991a) further points out that there is a mixture of relations and problems of identification due to the subjective reliance on reader schemata. In proposing a taxonomy and analytic procedure that promises to overcome such limitations, Hoey (1991a) drew upon the work of Hasan, (1984), Philips, (1985) and Winter (1974, 1979).

Hoey (1991a) recognizes the importance of Hasan's (1984) work in the inter-relation of the two types of chains, *identity* (cohesive ties that all

share the same referent) and *similarity* (parallel processes or descriptions) first introduced in the work of Halliday and Hasan (1976). Hoey (1991a, p.14) mentions that Hasan's (1984) analysis is a more integrated than a classificatory view of cohesion; rather than the occurrence of cohesion, it is the combination of cohesion elements that is significant (p.16). He indicates, however, that Hasan's (1984) view does not take into account the relationship of cohesion to the way sentences connect over discourse.

The significance of Winter's (1979) work to Hoey (1991a), is the broader focus on the repetitive function of cohesive devices over text. Hoey (1991a, p.20) comments that "Winter seeks to achieve ... interpretation of pairs of sentences in a text, making use of the way grammar and cohesion interact in the context". Specifically, Hoey (1991a) valued how Winter (1979) used the term *"repetition* to stand for *ellipsis* (which he prefers to refer to as *deletion), substitution* (the label used by Quirk et al. 1972 in Hoey, 1991a), to describe what Halliday and Hasan (1976) term as *reference),* and *lexical repetition* (broadly equivalent to Halliday and Hasan's (1976) *reiteration"* (p.17).   However, Hoey (1991a) found Winter's (1979) work limiting for his purposes as it did not account for the relationship of cohesion and large-scale patterning in text.

The work of Phillips (1985) offered this patterning over long distances; in fact, over chapters examined in science texts. Hoey (1991a) found that the work was significant in the intercollocations between sentences which linked sentences over chapters of the book. Where Phillips (1985) showed significant long-range lexical connections between chapters, Hoey (1991a) shows these between sentences over texts. Hoey (1991a) comments that "... this vocabulary ("...chapters with shared content will also share vocabulary") is tightly organized in terms of collocation and that in broad terms it allows the identification of topic opening and topic closing and of the text's general pattern of organization. In short, "...he [Phillips] is claiming (or at least it can be claimed for him) that systematic repetition organizes book-length texts through collocation..." (p.24).

Although Hoey's (1991a) analytic model for expository texts claims to be a "way forward" in identifying cohesive patterns over texts, it can be argued that the results he gained are questionable; the analysis being limited to one text   (taken from Michael Foster, 1942, *Masters of Political Thought,* Vol.1, Thomas Nelson & Sons Ltd.) and a few others (Hoey, 1991b). However, the clarity and in-depth explanation of the

model for replicating the analytic procedure, and the well-thought out rationale given for the weaknesses in previous analytic methods for expository text analysis promises new insights into text analysis. Therefore, it was chosen as the tool in the present lexical cohesive study.

5. Studies on Cohesion and Text Quality

Shaughnessy's (1977a,b) studies of skilled and unskilled college students revealed that more proficient texts had "hidden features" of competency. There have been many studies of cohesive devices, on these 'hidden feature of competency', in 'poor' (or low rated) and 'good' (or high rated) writing (Chambers, 1981; Fanning, 1981a,b; Weisberg, 1984; Williams, 1984; Yde and Spoelders, 1985;) mainly based on Halliday and Hasan's (1976) cohesive categories. Some of these studies have been carried out on L2 writers in English (Farghal, 1991; Jafarpur, 1991; Yang, 1989; among others). Research has indicated that better writing may show larger numbers of particular cohesive ties: high-rated texts indicating a higher frequency of lexical variation and "complex" types of cohesion compared to that in low-rated texts by academic writing standards (Ferris, 1994; McCulley,1985; Nunan, 1995; Stotsky, 1983; Witte & Faigley, 1981). Witte and Faigley's (1981) study found that lexical cohesion alone comprised about two thirds of all the cohesive devices in all texts with the high rated texts indicating more lexical density and collocations.

Studies have also shown that quantitative counts are not the only criteria, and perhaps not the most important, in distinguishing between high- and low-rated texts but more how the cohesive devices have been used over discourse. Witte and Faigley, 1981 note that indicators of writing quality are more the writers' invention skills rather than quantitative analyses of cohesion. Harnett (1986) comments that "...simple counts of either types or instances of all cohesive ties cannot be a completely effective index of the quality of prose" (p.151) and notes that "...both good and poor writers may use the same kinds of cohesive ties, but they use them differently"(p.143). Weissberg's (1984) finding that explicit inter-sentential cohesion devices were not used in almost one-quarter of the opportunities presented and asserts that readable texts cannot necessarily be produced "...simply by scattering a certain proportion of repeated words or anaphoric pronouns in the topic portions of their sentences" (p.495). Furthermore, L1 and L2 studies have indicated that high-rated

texts have more cohesive linkage over longer distances of text than that in low rated texts (Ehrlich, 1988; Hoey, 1991a; Granger & Tyson, 1996; Halliday & Hasan, 1989; Neuner, 1987; Norment, 1982; Parsons, 1991).

5.1 Studies on Cohesion and Genre

More recent L1 and L2 studies have found different cohesive devices according to the genre and/or text type being analyzed. (Granger and Tyson, 1996; Harnett, 1986; Jordan, 1984; Norment, 1982; Nunan, 1995; Smith & Frawley, 1983; Stotsky, 1983. This body of research argues that if cohesion text analytic findings are to be valid, they should relate to the specific genre under study since cohesive categories vary with text genre). Although most of this research was on native speakers' texts; a few L2 studies found similar findings.   Norment (1982), for example, found in a study of 180 native and non-native college students that native speakers' different frequencies and cohesive devices in narratives and exposition and across native and non-native student groups with better organization in the former group.   Harnett (1986) comments that researchers should not just adopt cohesion taxonomies that have been outlined for descriptive linguistic purposes (e.g. Halliday & Hasan's, 1976) but should reconsider the ways of analyzing cohesion according to their function in the texts in question.

5.2 Cohesion Studies on L2 Arabic Non-Native speakers' Texts

EFL texts at both the macro-level (over-all organization) and micro-level (sentence and clause levels) indicate different proficiency levels. L2 learners' texts are of relatively lower quality especially in lexical variety and sophistication and lexical cohesion at the discourse level compared to those of L1 learners' texts which affects text coherence (Silva, 1993; Green et.al., 2000).   Arabic non-native speaking students texts have been found to have different cohesive patterns to those in L1 texts   (see for example, Connor, 1996; Connor & Kaplan, 1987; Dudley-Evans & Swales, 1980; Doushaq, 1986; Hamdan, 1988; Kharma & Hajjaj, 1989; Sa'Addedin & Akram,1989, 1991; Shakir, 1991).   These researchers argue that L1 Arabic non-native student texts written in English are characterized by excessive use of coordination, parallelism, repetition (of the same word), and exaggeration with limited academic vocabulary and derivational forms.

These texts also emphasize cohesion at the sentence level rather than that at larger discourse stretches (Al-Abed Al-Haq, 1994; Connor, 1996; Grabe & Kaplan, 1996). It is controversial just how far such texts are influenced by Arabic patterns of cohesion through a process of language transfer. However, it is worth noting that languages differ in the level of lexical repetition that is normally tolerated. Certain repetition of words and phrases is an argumentative strategy in Arabic (Connor 1996), which may make for denser networks of excessive same word repetition in lexical cohesion than is common in English texts (Baker, 1992). Perhaps this is suggestive of how written texts are viewed by writers. A survey (Bacha, 1999) done on another random sample (N=155) of L1 Arabic non-native speakers of English in the same Freshman I English composition course indicated that the respondents considered lexis more significant in listening and reading skills than in their writing essay assignments. This may not be surprising as students feel the need to comprehend university course lectures and required textbook readings than in the production of written texts.

6. Method

6.1 Participants

A random sample of 20 high-rated and 20 low-rated essays (referred to as diagnostic texts) was selected from a corpus of N=261 essays. These were written upon entrance at the beginning of a four month semester as part of the regular program assignment by L1 Arabic non-native speaking students attending the Freshman I EFL Course, Lebanese American, Byblos Campus. It provided a suitable sample in that it was among the students' first courses taken at the university, and thus more representative of their initial proficiency level having limited influencing intervening variables. The subjects represented a stratified random sample of the student population in age (mean = 18), majors in the four schools (30% Arts & Sciences, 25% Business, 30% Engineering & Architecture, and 15% Pharmacy), gender (males 52.5%; females 47.5%), and native language Arabic.

The subjects' high school language background was one in which all had been required to follow Arabic language courses with 50% of the students undertaking their high school studies in the medium of French, 35% in the medium of English and 15% in their native language only, Arabic. English language classes at high school are mainly learned

according to the audio-lingual grammar translation method between 3-5 hours a week. Since results showed no significant differences when the medium of instruction in high school was examined in the cohesive analysis, the high school study-language variable was not included in the present study. Also, since a second comparable set of texts showed similar results as those of the entry texts and due to limitation of space here, the details were also not included.

The Freshman English I course, from which the sample essays were selected, is the first of four EFL composition courses that students are required to take. The main objective of the Freshman I course is to write well-organized and developed paragraphs and essays according to different rhetorical modes of narration, description, illustration, cause-effect, and comparison-contrast. Readings are used to help with the development and support of ideas. Students were instructed to complete an essay on the topic of giving the reasons and results of disagreements between teenagers and parents in sixty minutes and to conform to the traditional essay format of having an introductory paragraph, body paragraphs (two) and a concluding paragraph. The cause-effect mode was considered appropriate for diagnostic purposes, it being relatively easy for students to cope with upon entry and one on which upper level English courses mainly build. Also, it seemed to allow the students from past experience to produce the most vocabulary that their proficiency level would allow. Finally, it was chosen since it complied to an expository genre, the focus of the present study.

6.2 Tasks

Each paper was holistically rated by two experienced TEFL teachers (the class teacher and one from another section) according to the LAU Writing evaluation criteria (see Appendix A). The final percent score assigned to a text was the mean of the two raters' scores or the mean of the two closest scores if a third rating was needed when there was a discrepancy of more than one letter grade between the converted letter grades of A, B, C, D or F. Holistic scoring, giving a percentage grade based on an impressionistic evaluation was found to be more suitable than analytic or more detailed evaluation (Bacha, 2001). Evaluation results indicated high inter-rater reliability coefficients (r= .9, p= .00) using the Person Correlation two-tailed Statistical Test and intra-rater coefficients (r= .86, p= .001) using the Spearman Correlation two-tailed Statistical Test since the latter data was not normally distributed. These

statistical tests examine the strength of the relationship between two sets of independent interval data.

In order to ensure that the text scores were reliable indicators of the text levels, the reliability and validity of the EFL Programs' Writing Evaluation Criteria had been carried out. This involved correlating a different randomly selected sample of 60 Freshman English I essay scores from the N=261 corpus, holistically scored by a comparable group of raters but using the Jacobs et.al (1981) ESL Composition Profile (researched to be valid and reliable for writing evaluation). High coefficients (*p*= .7, p= .001) between the scores obtained on the LAU Evaluation Criteria and those on the Jacobs' ESL Composition Profile, (1981) were obtained confirming the proficiency levels of the texts. Texts were then categorized as *High-rated texts* (HT = 75% and above); *Mid-rated* (MT = 65% - 74.9%); or *Low-rated* (LT =Below 64.9%). The rationale for this division was that at LAU scores of 75% and above are considered "good", texts scoring below 65% "poor", and text scores between the two "fair" to "satisfactory". Mid-rated texts were omitted from the study since the objective of the research was to identify differential patterns of lexical cohesion at high and low proficiency levels.

7. Data Analysis

The lexical cohesion analysis focused on open-set lexical items in context (see Table 1b). The exclusion of repetition between grammatical items such as determiners, prepositions, auxiliaries, negatives, co-ordinators, subordinators, sentence conjunctions (or conjuncts), sub-modifiers or particles, and collocation was only for analytic reasons and not because they may not contribute to the connectedness between sentences (Hoey, 1991a, p.53). Ellipsis and arguable cases constituting 1% of all cases were also omitted from the analysis. It is often not clear in learner's texts whether ellipsis is a flaw or a stylistic technique and arguable cases were difficult to categorize (Hoey, 1991a, pp74-75). Each repetition link needed to conform to the following criteria:

> " a.   Do they have common or related contexts? or
> b.   Do the items share common relationships with neighboring lexical items? or
> c.   Is there whole or partial parallelism between the contexts of the items?

d.   All links are anaphoric; that is, they refer back
     to previous items in the text.
e.   A lexical item can not form more than one
     repetition link with more than one item between
     any two sentences"   (Hoey, 1991a, p.57).

Although the researchers did the analysis manually (electronic analysis non-existent), Hoey (1991a) minimizes any concern with subjectivity since consistency is the overiding criterion in identifying the links. He emphasizes that "...chance repetition forms a small and insignificant proportion of the lexical repetitions identified" (p.57). Another qualification made (Hoey, 1991) is that "Under the heading of repetition we have included a range of lexical relationships, but the controlling requirement has been that the later item must contain the same information as the earlier. Thus, in our terms, a paraphrase has only been considered a true paraphrase if the items under consideration have been interchangeable in the context..." (p.69).   Hoey gives the example in this context of the text "Drug-Crazed Grizzlies" where there is a repetition between *scientists* (sentence 2) and *biologists* (sentence 4) in which it is pointed out "...that we cannot affirm with confidence that they both have identical referents, and so no link is established" (p.70).

Although not explicitly stated, Hoey's (1991a), justification for the priority order in picking up repetition links in a text was based on certainty that no error would be made (Personal communication, Hoey, 1996). That is, it is more certain to identify the simple repetition links (i.e. parents-parents) than it is to identify other links mentioned giving this link priority (see Table 1a).   Since Hoey's (1991a) rank order of lexical categories is genre specific, based on expository native texts, proposing a different rank is justified if clearly outlined and consistently applied (Personal communication, Hoey, 1996). The rationale for the adapted model used in the present study is that lexical discourse links take priority over grammatical ones by EFL essay writing standards. Since lexical links interconnect in quite complex ways, these interconnections will be lost if it is the grammatical links that are given priority and thus recorded in sentences where both grammatical and discourse links occur.   That is, the rank order used in the present study is considered more valuable in researching EFL texts in English as it gives credit for greater use in lexical variety (see Tables 1a and 1b).

Table 1a   Hoey's (1991a) Lexical Cohesion Categories

Code	Hoey's (1991a) Rank Order	Example
Sr	Simple lexical repetition	*Parent... .Parent.. .* (Repetition of the exact same word)
Cr	Complex lexical repetition	(As cr1,cr2,cr3 adapted, but placed all in one category)
Smp	Simple mutual paraphrase	(As smp1,smp2 adapted, but placed all in one category)
Cp	Antonymous complex paraphrase	(As cp3 adapted)
Cp	Other complex paraphrase	(As cp1, cp2 adapted)
spp	Simple partial paraphrase	(Same as adapted)
s	Substitution	(Same as adapted)
co-ref	Co-reference	(Same as adapted)
e	Ellipsis	*of art ... 0 (omitted in the adaptation)*
d	Deixis	Same as adapted

Table 1b   EFL Adapted Lexical Cohesion Categories

Code	Adapted Rank Order	Example
	Discourse Link Types	
cp1	Complex paraphrase 1	*Instruction....teacher (teaching* not mentioned*)*
cp2	Complex paraphrase 2	*1)writings/writer(cr), 2) writer/author(sp), and then* identification of *author/writing (cp2)*
cp3	Complex paraphrase 3	*hot/cold (antonyms with different morphemes)*
spp	Simple partial paraphrase	*Book/volume*
smp1	Simple mutual paraphrase 1	*degree/standard;* (synonym - different morpheme)
smp2	Simple mutual paraphrase 2	disagree/not agree; (antonym – different morpheme)
cr1	Complex repetition 1	*Belief/believe;* (morphological similarity)
cr2	Complex repetition 2	*parent* (n.)/*parent* (adj.); (part of speech change)
cr3	Complex repetition 3	*trust/mistrust;* (morphological change - antonyms)
sr	Simple repetition	*parents* try hard to help their children. *Parents.....*
cor	Co-reference	*Mrs.Thatcher... . The Prime Minister......*
	Grammatical Link Types	
d	Deixis	*The disease has spread. This* will be investigated.
s	Substitution	*Dr. Long* is busy. *He* is lecturing.

A similar procedure (illustrated with a sample text) as that of Hoey's (1991a) was followed in the text cohesion analysis involving 1) identifying the links as outlined in Table 1b according to the adapted

categories, 2) recording the links, and 3) calculating the sentences that have two or more links (referred to as bonded sentences).

7.1 Calculating Links

Each sentence was taken separately and checked against each of the consecutive sentences to identify links. For example, in the sample text, S1 (sentence) was examined for any links with S2, S3, S4 etc.; then S2 with S3, S4, S5, etc. and so forth (see sample text).

Sample Text

---

the teenagers problem.

(1) Since, the creature up till now, All children will attend a certain age to become teenagers, but this period of life will going to be a difficult period and parents will never accept that. (2) As a result. we have many reason for disagreements.   (3) But the question is what are those reasons for disagreements? and the effects of these disagreements on the family.

(4) The period of teenagers is one of the most difficult period in our life, because we have many changes in our cycle of life and in our personality. (5) For that we have many reasons for disagreements between Parents and teenagers and these reasons are. (6) First of all, the teenagers fell himself an adulte, he want to do whatever comes into his mind, he want to out after midnight whenever he want, having fun with whoever he want also. (7) Second the school has a big importance in our life and in our way of doing things. (8) Also our environement has a big importance on us. (9) Parents will never accept these things because they still exist in our life and we have to do whatever they want. (10) The disagreement cause a lot of problem between the family and has a big effect. (11) First, the disagreement has a big effect on the teenager's personality, on his action, he becomes, to do bad things like stoling or kiling or robing. (12) Second, he may leave the home and make his own life alone like in the United states of America. (13) When the boy have 18 years old or before he leave the house and never come back to his parents. (14) he do whatever he like.

(15) In conclusion, although of the reasons for desagreement like environment or the way of living, or having the feel to be adult and their effects. (16) our world face these big and dangerous problem because it has many effects in teenagers life. (17) In my opinion we should work to solve this problems in anyway.

The links were identified and then recorded on the repetition matrix, each cell indicating the type of lexical link in abbreviated form and the words forming the link between the earlier sentence and the later (see Figure 2).

Figure 2   Repetition Matrix for Sample Text (for the first 6 sentences)

	1	2	3	4	5
2 **I**					
3		2 sr: r-rs sr: ds-ds			
4 **B** **1**	sr: dif-dif sr: per-per sr: l-l sr: ts-ts		3		
5	sr: ts-ts sr: ps-ps	sr: r-rs sr: ds-ds	sr: rs-rs sr: ds-ds	4 d: s4-that	
6	Sr: ts-ts				5 sr: ts-ts

sr: simple repetition; d: deixis; r(s)=reason(s); dif = difficult
s4= sentence 4; ps= parents; l= life; ts=teenagers; per=period
I=introductory paragraph; B1=first body paragraph

The vertical columns indicate links with earlier sentences and the horizontal columns indicate links with later sentences. If a lexical item has more than one possible type of link between any pair of sentences, then the link recorded is the one that is higher on the priority scale.   No lexical item can enter into more than one link between a pair of sentences. Thus, in the sample text, S2 and S3 (the later sentences) have no links with S1 (the earlier sentence) and the cells are left blank; S4 has four simple repetition (sr) links with S1 which are indicated as sr: difficult-difficult; sr: period-period, sr: life-life, sr: teenagers-teenagers, and so on down the first vertical column.   These links were then converted into raw frequencies onto a cell matrix. Figure 3 indicates 0 links between S1 and S2 and S3; two links between S4 and S5 and so on.

Figure 3   Cell Matrix for Sample Text (for the first six sentences)

	1				
2 I	0				
3	0	**2** 2			
4 B 1	4	0	**3** 0		
5	2	2	2	**4** 1	
6	1	0	0	0	**5** 1

From the cell matrices, the frequencies and percentages of occurrence of types of links for each of the 40 texts were recorded as those in Table 2.

Table 2   Number and Percent Frequency of the Types of Links in Sample Text   (continued next page)

Type	N	%
cp1	0	0

cp2	0	0
cp3	0	0
spp	0	0
smp1	0	0
smp2	0	0
cr1	0	0
cr2	3	4.17
cr3	3	4.17
sr	54	75.00
co-ref	0	0
d	2	2.78
s	10	13.88
Total	**72**	100

Next, the intervening number of sentences (i.e. the *distance*) between each pair of linked sentences was then calculated in percentages for the 40 texts using the cell matrix information. Figure 3 indicates two intervening sentences between the four sr links formed between S4 and S1; however, the deictic (d) link formed between S5 and S4 has no intervening sentences. These distances can be expressed as a percentage of the total number of that type of link in the whole text. Table 3 shows that in the sample text, 10 out of the total 54 (or 18.52%) sr links in the whole text are between adjacent sentences, indicating that most of the sr links in this particular text are adjacent.

Table 3 **Percent Frequency of Intervening Sentences between Types of Links in Sample Text** (continued next page)

Links	0	1	2	3	4	5	6
cp1							
cp2							
cp3							
spp							
smp1							
smp2							
cr1							
cr2					33.33		
cr3			33.3				

			3				
sr	18.52	5.56	12.96	9.26	9.26	9.26	7.41
co-ref							
d	100						
s	20.0			10.0		10.0	10.0
Total	**19.45**	**4.17**	**11.11**	**8.33**	**8.33**	**8.33**	**6.94**

7	8	9	10	11	12	13	14
		33.33	33.33				
33.33						33.33	
9.26	3.70	5.56		3.70	3.70		1.88
10.0	10.0	10.0	10.0	10.0			
9.72	**4.17**	**6.94**	**2.78**	**4.17**	**2.78**	**1.39**	**1.39**

7.2 Calculating Bonded Sentences

Lexical cohesion is operationally defined as the percent frequency of bonds. The data obtained from the links in the texts were used to calculate the bonding. Since bonded sentences are claimed to form coherent texts (Hoey, 1991a), the patterning of the bonded sentences over texts was analyzed as the percentage frequency of bonds in each whole text. First, the cell matrices were used to obtain the number of bonded sentences in each text. The cell matrices show that there is considerable variation in the number of links and, therefore, of bonds that a sentence may have with others.

Table 4 summarizes the information from the cell matrix for the sample text, indicating that out of the 136 cells in the matrix, 81 (or 59.56%) are empty; 41 show one repetition link only; together these give a total of 122 (or 89.71%) sentence pairs which can be taken as not significantly connected by repetition. The *density of linkage* of the rest of the sentences varies between 2 to 3 links (together 10.29%), i.e. only one tenth of the sentence pairs are significantly connected in this low-rated student text. This is a low density linkage text when it is compared to the density of linkage in the native-speaker text analyzed by Hoey, in which 30% of sentence pairs were significantly linked (1991a, p.91).

Table 4   Percent Frequency of Cells with Repetition Links in Sample Text

N of Links	N of Cells	%Cells
0	81	59.56
1	41	30.15
2	11	8.09
3	3	2.20
4	0	0
Total	**136**	**100**

Hoey (1991a) argues that since this lexical repetition appears to serve some text-organizing function, a criterion for repetition is needed that is high enough to distinguish significant pairs from insignificant pairs and low enough to leave something to investigate (in our case to distinguish quality writing among second language writers). After careful analysis of a few texts and examining similar research in the field that note different cut-off points depending on text length and lexical density (Hoey, 1991a,b; Phillips, 1985), Hoey (1991a) regards *three* links as a cut-off point: lexical items form links, and any two sentences sharing three or more links form a bond. Table 4 indicates that in the sample text if three links and above are taken as significant, there will be only 3 bonds (2.20% of bonding) in the text; if two links and above are taken, there will be 14 bonds (10.29% of bonding), which is still lower than that of the text Hoey analyzed (1991a). However, if one link and above is the criterion, then the bonding in Text 1 is 40.44%, probably too high to permit meaningful discrimination between student texts.

After examining tables similar to that of Table 4 for all 40 texts and following the same procedure based on Hoey's (1991a), it was decided to set our criterion as *two* links equal one bond as the cut-off point for these EFL texts. Nevertheless, we still report statistical testing on three links

equal one bond to compare the two and three link criteria for bonding. In order to investigate the distance of the bonding in the texts, the cell matrix data were then converted to a bond matrix to display the frequency and percentage of bonding of each sentence with subsequent (or preceding) sentences.   Table 5 for the sample text was used to convert data from the cell matrix to study the bonding patterns over the text. Bond matrix tables were recorded for each of the 40 texts. The figures in Table 5 indicate the bonding for the sample text for two and three links equal one bond.

Table 5   Bond Matrix for Sample Text

P	S1	2 links = 1 bond %BC	2 links = 1 bond S2	3 links = 1 bond %BC	3 links = 1 bond S2
I	1	23.08	4,5,9	33.3	4
	2	23.08	3,5,15		
	3	15.38	5,10		
B1	4				
	5	23.08	9,11,15	33.3	15
	6				
	7				
	8				
	9				
	10	15.38	15,16	33.3	11
	11				
	12				
	13				
	14				
C	15				
	16				
	17				

P   = Paragraph Type / S1   = Sentence number in the text / I   = Introductory / B1   = Body paragraph 1 etc. /   C   = Concluding / %BC = Percent frequency of bonds the sentence has with consecutive sentences / S2 =   The number of the consecutive / sentences that is bonded to S1

To investigate the bonding over longer distances in the high-rated and low-rated EFL texts, the percent frequency of intervening sentences between bonded pairs of sentences was taken from the bond matrix for each of the 40 texts. When two links equal one bond was considered, the criterion for bonding adopted by the present study, a total of 13 bonds are identified. S1 is bonded to S4,5 and 9, S2 is bonded to S3,5 and 15; S3 is bonded to S5 and 10; S5 is bonded to S9,11, and 15; S10 is bonded to S15 and 16.   The number of intervening sentences between the

bonded sentences and the percent frequency of bonds a particular sentence has with consecutive sentences are indicated.

For example, there are two intervening sentences between the bonded sentences S1 and S4 (excluding S1 and S4), 3 between S1 and S5, 7 between S1 and S9 (i.e., S1 has a 23.08% bonding with consecutive sentences in this particular text), and so forth. . In the sample text, when three links equal one bond, Table 5, shows that S5 and S15 are linked by one bond 9 sentences apart, (excluding S5 and S15) constituting 33% of the bonded sentences; S1 and S4 are linked by one bond with 2 intervening sentences; and S10 and S11 are adjacent (with no intervening sentences) giving a total of three bonds in the text. Tables to show percentages of bonded sentences were drawn up for all 40 texts, both for when two and three links are taken as one bond. Table 6 shows that when three links equal one bond for the sample text, 11 sentences (or 64.71%) are not bonded; these are considered marginal. 6 sentences (or 35.29%) are bonded, considered central.

Table 6   Percent Frequency of Sentences According to Quantity   of Bonds for Sample Text

Quant.of Bonds	Quant.of Sentences	Percent of Sentences
0	11	64.71
1	6	35.29
	17	

7.3 The Bond Distance

These data are further interpreted in tables similar to Table 7 which indicates that 1 bond (or 33.33%) between pairs of sentences had 2 intervening sentences, 1 had 9 and 1 had none.

Table 7   Percent Frequency of Intervening Sentences between Bonded Pairs of Sentences When Three Links Equal One Bond in Sample Text (next page)

Q.of IS	Q.of Bonds	%of Bonds
0	1	33.33
2	1	33.33
9	1	33.33
	3	

Q of IS   =   Quantity of Intervening Sentences
Q of Bonds =   Quantity of Bonds between pairs of sentences
% of Bonds   =   Percent of Bonds between pair of sentences
        calculated on the basis of the total number of bonds in the text

The quantity of intervening sentences between the bonded pairs of sentences is expressed as a proportion out of the total quantity of sentences in the text. To calculate this *bond distance indicator* for each text, the following formula was applied for each of the 40 texts (for two and three links equal one bond):

$$\frac{Quantity\ of\ Bonds\ X\ Distance\ 1\ +\ Quantity\ of\ Bonds\ X\ Distance\ 2,\ etc.}{Quantity\ of\ Sentences\ in\ Text}$$

The higher the proportion, the longer the distance bonds connect over the text. Thus, for the sample text, the formula produces a bond distance indicator of 0.65 when three links equal one bond was considered.   This is low compared to proportions of between 10 to 14 obtained in some of the high-rated texts. It indicates bonding over shorter distances which is expected in a low-rated text.

8.   Summary of Principal Findings

The Mann-Whitney Statistical Test was used to compare the mean ranks (or central locations) of the dependent variables (lexical cohesion categories). The data obtained were recorded on an interval or an ordinal scale and expressed either as raw frequencies or as a percentage of the total observed of the variable discussed (Hatch & Lazaraton 1991). Any two sets of data compared were independent and not normally distributed. The choice of the Mann-Whitney Test was justified as the test compares two groups on the basis of their ranks above and below the median (SPSS, 1997).   Main results indicated the following:

8.1 Types of Lexical Cohesion

- The high-rated texts had a significant higher frequency of more sophisticated types of lexical cohesion: complex paraphrases (cp1) (p=0028), simple mutual paraphrases (smp1) (p=.0022), and complex repetition (cr2) (p= .001) when compared to those of the LT.
- The high-rated texts had higher significant frequencies of simple mutual paraphrase (smp1) (p=.018) and complex repetition (cr2) (p=.25) in adjacent sentences than those in the LT. The LT, however, had a higher significant frequency of simple repetition (sr) in adjacent sentences (p=.014).
- The low-rated texts had a higher significant total frequency of links in adjacent sentences (p=.02).
- In both high- and low-rated texts simple repetition averaged 77% of all lexical cohesion
categories present.

8.2 Distance of Lexical Cohesion

- Tables 8–9 [withheld together with Table 10 due to space limitation but can be obtained from authors] clearly indicate the distances as mean percentages of
intervening sentences in the sample of high and low rated texts.
- The low-rated texts significantly indicated more simple repetition (sr) links in adjacent sentences while the high-rated indicated them over longer distances.
- There were high significant differences in the distance of the bonds (as expressed by the
- bond indicator of the number of intervening sentences between bonds) when both *two* (p=.001) and *three* (p=.001) links equal one bond were considered (see Table 10 for two links equal one bond)

[Table 8 Percentages of Intervening Sentences in Low-Rated Texts (N=20)
Table 9   Percentages of Intervening Sentences in High-Rated Texts (N=20)
Table 10   Distance of Bonding in Pre-Course (Diagnostic1) and Post-Course (Diagnostic 2) Texts When 2 links = 1 bond in Texts (Each group of texts = N=20)]

9. Discussion

The results indicated more lexical variation and the presence of the higher ranking lexical cohesive categories according to the EFL adapted analytic model in the high-rated texts than that in the low-rated.   In the sample text, typical of low-rated texts, there were 54 simple repetition cohesive links (sr) or (75%) of the total links.   However, in high rated texts, although there was a high mean percent of simple repetitions, more sophisticated types of links were present. For example, in one high-rated texts, S17 has two links with S14 and these were recorded in order of importance; the complex repetition type 2 (cr2) being the more important in EFL texts: cr2: teenagers-teenagers', then sr: parents-parents. This confirms previous research that there are different cohesive features depending upon the proficiency level of the text.

We, however, see the need to follow up this study with a comparison of lexical cohesion in two different genres in order to make any claims concerning to what extent the cohesion analysis is genre specific. The fact that derivational lexical forms were present, although minimal, indicated the expository nature of the texts.   It was interesting to note that the high-rated texts indicated these variations in adjacent sentences confirming previous research that "good" writers, and thus their texts, show different invention skills. In this sense, the high-rated texts were of superior quality in that there were significantly fewer simple repetitions in adjacent sentences. Although pedagogical implications are not the focus of this discussion, it can not be avoided to point out that the high-rated texts conformed to academic essay conventions in that repetition of the same word at short distances in texts is not considered a feature of quality texts.   Last, the quality of the high-rated texts was reinforced by the presence of significantly fewer links at short distances again conforming to academic conventions in which global linkage over local is preferred. All in all, both the high and low-rated texts were not quality texts by native standards as the high percentage of simple repetition indicated especially in the low-texts at short distances.

That the high-rated texts showed cohesive bonding over longer stretches of text implies also different patterning at various proficiency levels. Table 3 indicates that for the sample text the maximum number of intervening sentences between any of the links is 14. About 65% of the links have up to 5 intervening sentences between the two lexical items that form the link, typical of low-rated texts. On examining Tables 8 and

9, it is evident that linking is over greater distances in the high-rated texts; almost 50% of the links have a maximum of three intervening sentences and less than 5% have more than 14. Table 9 indicates that 29% of the links have a maximum of three intervening sentences, while approximately 15% have more than 14.

We, however, would like to note that since we have taken 2 links as the criterion for bonded sentences in these learners' texts with interesting results, further research needs to be carried out more samples to confirm the findings. Table 10 clearly shows that in the high-rated pre-course texts (referred in the table as Diagnostic 1 texts used for the present study) 15% of the sentences were bonded according to the criterion of links but less than half that were bonded in the low-rated. Since we also had information on the post-course texts (referred in the table as Diagnostic 2
texts), not included in the present, we still would like to share the results at this point for comparative purposes.

Table 10, therefore, shows that on these texts the same ratio of bonding between high and low-rated was found: 14% as compared with a little of 7% for high and low-rated texts respectively. When three links were considered as the criterion for a bond, it is interesting that the results were comparable but with much lower percentages of 4% and 3% of bonded sentences for the high and low-rated respectively on both pre-course and post-course texts. We question what the results would have been if one link had been taken to equal one bond. We also inquire whether in other genres written by the same students in their subject matter classes where the vocabulary might be denser, whether the criterion of linkage would be the same. This is something worth researching further in L2 texts in different genres at various levels. We also acknowledge that investigation into other lexical cohesive categories not included in the present needs to be done which might also be characteristic of quality texts and the need to replicate the present study with a larger sample. Perhaps this will be possible when electronic aid is available.

A further finding that was a by product of the analysis, worth discussion, was the identification of topic sentences as defined in the North American essay. Hoey (1991a) found that when bonded sentences were removed from the whole text and placed together, coherent sub-texts or summaries were possible. There was more than one possible summary of

the whole text, depending upon which bonded sentences were placed together. He showed through extensive examples that topic opening sentences had a high number of bonds with subsequent sentences, as had topic closing sentences with previous sentences. The present study confirms this.   For example, the figures in Table 5 concerning the sample text indicate that S1, 5, 10 have one bond each with one subsequent sentence taking the criterion of three links equal one bond. These are identified as topic opening sentences and their place in the introductory paragraph and beginning of the first body paragraph reflects that topic and thesis statements are placed early in the essay. Similarly, S15 has 1 bond between it and the previous sentence S5 which also reinforces the idea that some sentences later in the text are summative. The bonding between S10 and 11 indicates that topics may be opened in the body.

Table 5 shows that when the criterion was two links equal one bond comparable results were obtained. Most of the bonds appeared between the introductory sentences and those in body and concluding paragraph initial positions.   As we know, many texts may not have explicit topic sentences at the beginning of a paragraph and sometimes might not appear in the text at all.   However, the identification of these topic sentences through the lexical cohesion analysis is indeed a very interesting finding.   We, in fact, find this significant and see the need to follow up this research into non-native speakers of English production of acceptable paragraphs and essays by native reader expectations and conventions especially in light of related recent studies into second sentences and marked themes (Reid, 1996, Allison et.al., 1999).

Previous studies of cohesive devices had indicated little relation between frequency counts of cohesion and quality of texts assessed through holistic ratings (Engber, 1995; Hamdan, 1988; Harnett, 1986; Hasan, 1984; Jafarpur, 1991; Neuner, 1991;Parsons, 1991;Yang, 1989). The contribution of the present study to existing cohesion studies may be viewed in its confirmation that the use of cohesion is not simply a matter of frequency counts but differential use.   That is, the findings reinforce that more proficient texts are characterized by more sophisticated types of lexical cohesive linkage at more global levels of discourse.

10. Conclusion

The aim of this study was to explore the types and patterning of lexical cohesion in L1 Arabic non-native students' academic expository texts. Using an EFL adaptation of Hoey's (1991a) lexical cohesion analytic categories, this first attempt confirmed recent research in the field, shed new light on L2 writing and indicated a reinforcement between lexical categories and the holistic text ratings. Furthermore, if as Hoey (1991a) theory of language claims, a study of text is also a study of lexis, the study showed this relation in one genre, EFL compositions, as a step in characterizing these learners' texts. The map of language that Hoey (1991a) proposes (Figure 1), shows the relation between lexis and text organization and confirms recent research that lexical items are not "a boundless chaos" but are crucial in text patterning.

Although there are numerous other ways to identify devices in quality texts, and the reliance on Hoey's(1991a) system seems to be premature and warrants further researching before any claims can be made, the procedure used here is considered a significant one since it takes into account continuous writing at discourse level. Widdowson (1978) earlier claimed that "...the consideration of use requires us to go beyond the sentence and to look at larger stretches of language. Normal linguistic behavior does not consist in the production of separate sentences but in the use of sentences for the creation of discourse" (p.22). We acknowledge that this is a preliminary exploratory study, but at the same time see the value of the light it has shed on lexical patterning over stretches of texts in L2 academic discourse which seems to promise new frontiers in our efforts in lexical cohesive text analysis.

References

Al-Abed Al Haq, F. & Ahmed, A.S. (1994). Discourse problems in argumentative writing. *World Englishes*, 13:3, 307-323.

Allison, D., Varghese, S., & Mei, W.S. (1999). Local coherence and its limits: A second look at second sentences. *Journal of Second Language Writing*, 8, 77-97.

Bacha, N. N. (1999). *Student perceptions of the Importance of lexis in texts.* Survey work in progress, Lebanese American University.

Bacha, N.N. (2001). "Writing evaluation: What can analytic versus holistic scoring tell us?" *System*, Vol.29, No.3, pp. 371-383)

Baker, M. (1992). *In other words, a coursebook on translation.* London:

Routledge

Bloor, T. & Bloor, M. (1995). *The functional analysis of English: A Hallidayan approach.* London: Arnold Publishers.

Carter, R. & McCarthy, M. (1988). *Vocabulary and language teaching* . London: Longman Group Limited.

Chapman, J. (1983). *Reading development and cohesion.* London: Heinemann.

Chambers, F. (1981), Essay-writing: Cohesion and coherence, *MET,* 9:1, 13-18.

Connor, U. (1996). *Contrastive rhetoric cross-cultural aspects of second – language writing.* Cambridge: Cambridge University Press.

Connor, U. & Kaplan, R. B. (Eds.), (1987). *Writing across languages :Analysis of L2 text.* Reading, Mass. Addison Wesley.

Coulthard, M. (1994). On analysing and evaluating written text. In M . Coulthard (Ed.), *Advances in written text analysis* (pp.1-11*).* London: Routledge

Cruze, D.A. (1986). *Lexical semantics.* Cambridge: CUP.

Donnelly, C. (1994). *Linguistics for writers.* New York: State University of New York Press.

De Beaugrande, R.A. & Dressler, W. U. (1981). *Introduction to text linguistics.* London: Longman Group Limited.

Doushaq, M.H. (1986). An investigation into stylistic errors of Arab students learning English for academic purposes. *English for Specific Purposes,* 5:1, 27-39.

Dudley-Evans T. & Swales, J. (1980). Study modes and students from the Middle East. In *Study modes and academic development of overseas students, ELT Documents, 109.* London: The British Council.

Dudley-Evans T. (1994). Genre analysis: An approach to text analysis for ESP. In M. Coulthard (Ed.), *Advances in written text analysis.* (pp.219-228). London: Routledge.

Ehrlich, S. (1988). Cohesive devices and discourse competence. *World Englishes,* 7:2, 111-118.

Engber, C. (1995). The relationship of lexical proficiency to the quality of ESL Compositions. *Journal of Second Language Writing,* 4:2, 139-155.

Fanning, P. (1981a). Academic essay-writing. *Modern English Teacher,* 9:1, 11-13.

_____ .(1981b). Essay-writing: Cohesion and coherence. *Modern English Teacher,* 9:1,3-17.

Farghal, M. (1992). Naturalness and the notion of cohesion in EFL writing classes. *IRAL ,* 30:1, 45-50.

Ferris, D. R. (1994). Lexical and syntactic features of ESL writing by students at different levels of L2 proficiency. *TESOL Quarterly,* 28:2, 414-420.

Francis, G. (1994). Labelling discourse: An aspect of nominal-group lexical cohesion. In M. Coulthard (Ed.), *Advances in written text analysis* (pp.83-101). London: Routledge.

Grabe, W. & Kaplan, R.B. (1996). *Theory and practice of writing.* London:

Addison Wesley Longman Group Limited.

Granger, S. & Tyson, S. (1996). Connector usage in the English essay writing of native and non-native EFL speakers of English. *World Englishes*, 15:1, 7-27.

Green, F.C., Christopher, E.R., & Lam K. M. (2000). The incidence and effects on coherence of marked themes in interlanguage texts: A corpus-based enquiry. *English for Specific Purposes*, 19, 99-113.

Halliday, M. (1966). Lexis as a linguistic level. In C.E. Bazell, et.al. *In Memory of J.R. Firth* (pp.148-162). London: Longman Group Limited.

Halliday, M.& Hasan, R. (1976). *Cohesion in English.* London: Longman .

_____. (1989). *Language, context, and text: Aspects of language in a social-semiotic perspective.* Oxford: OUP.

_____.(1994). *An introduction to functional grammar.* (2nd ed.). London: Arnold Publishers.

Hamdan, A.S. (1988), *Coherence and cohesion in texts written in English by Jordanian university students.* Unpublished Ph.D. Thesis. Manchester University, England.

Hartnett, C.G. (1986). Determining how written language is valued: Static and dynamic cohesion: Signals of thinking in writing. In B. Couture (Ed.), (1994) (pp.142-153)..

Hasan, R. (1984). Coherence and Cohesive Harmony. In J. Flood (Ed.),*Understanding reading comprehension* (pp.181-219). Delaware: International Reading Association.

Hatch, E. & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics.* New York: Newbury House Publishers.

Hoey, M. (1983). *On the surface of discourse.* London: George Allen & Unwin.

_____ . (1991a). *Patterns of lexis in text.* Oxford: Oxford University Press.

_____ .(1991b). Another perspective on coherence and cohesive harmony. In E. Ventola (Ed.), *Functional and systemic linguistics, approaches and uses* (pp.385-413). Berlin: Mouton deGruyter.

_____ .(1994). Signalling in discourse: A functional analysis of a common discourse pattern in written and spoken English. In M.Coulthard (Ed.), (1994) ,(pp.26-45)..

Hunston, S. (1994). Evaluation and organization in a sample of writte academic discourse. In M. Coulthard (Ed.), (1994) *Advances in written text analysis* (pp.191-218). London: Routledge.

Jafarpur, A.(1991). Cohesiveness as a basis for evaluating composition *System*, 19:4, 459-465

Jordan, M. P. (1984). *Rhetoric of everyday English texts.* London: George Allen and Unwin.

Khalil, A. (1989). A study of cohesion and coherence in Arab EFL college students' writing". *System*, 17, 3:89, 359-371.

Kharma, N. & Hajjaj, A (1989). *Errors in English among Arabic speakers: Analysis and remedy.* London: Longman.

Lewis, M. (1993). *The lexical approach*. Hove: England, Language Teaching Publications.

McCarthy, M. (1991). *Discourse analysis for language teachers*. Cambridge: Cambridge University Press.

McCarthy, M. & Carter, R. (1994). *Language as discourse: Perspectives for language teaching*. London: Longman Group Limited.

McCulley, G. (1985). Writing quality, coherence, and cohesion. *Research in the teaching of English*,19:3, 269-282.

Nattinger, J. R. & De Carrico, S. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.

Neuner, J.L.(1987). Cohesive ties and chains in good and poor freshman essays. *Research in the Teaching of English*, 21:1, 92-105.

Norment, N. Jr. (1982). Contrastive analyses of organizational structures and cohesive elements in English, Spanish (ESL) and Chines (ESL) students' writing in narrative and expository modes . *Education Resource Information Center (ERIC)*.

Nunan, D.(1995). The write stuff: Achieving coherence in scientific writing. In P. Bruthiaux, et. al. (Eds.), *Explorations in English for Professional Communication* (pp.99-115). Hong Kong: City University of Hong Kong.

Odlin, T. (1989). *Language transfer, cross-linguistic influence in language learning*. Cambridge:CUP.

Parsons, G. (1991). Cohesion coherence: Scientific texts. In E. Ventola (Ed.), *Functional and systemic linguistics, approaches and uses* (pp.415-429). Berlin: Houton de Gruyter.

Peters, P. (1986). Getting the theme across: A study of dominant function in the academic writing of university students. In B. Couture (Ed.), *Functional approaches to writing: Research perspectives* (pp.169-185). London: Frances Pinter Publishers.

Phillips, M. (1985). *Aspects of text structure: An investigation of the lexical organization of text*. Amsterdam: North-Holland.

Reid, J. (1996). U.S. academic readers, ESL writers, and second sentence. *Journal of Second Language Writing*, 5, 129-161.

_____. (2000). Comments on "Local coherence and its limits: A second   look at second sentences" - Another look. *Journal of Second Language Writing*, 9, 77-88.

Sa'Addedin, M. A (1989). Text development and Arabic- English negative interference. *Applied Linguistics*, 10:1, 36-51.

_____ . (1991) Writing across language communities: The structure of   Arabic text. In P. Adams et. al. (Eds.), *Socio-cultural issues in English for academic purposes*, 1:2, London: MacMillian.

Shakir, A. (1991). Coherence in EFL student-written texts: Two perspectives. *Foreign Language Annals*, 24:5, 399-411.

Shaughnessy, M. P. (1977a). *Errors and expectations: A guide for the teacher of basic writing*. New York: Oxford University Press.

_____ . (1977b). Needed research on writing. *College Composition and Communication*, 28, 317-320.

Sinclair, J. (1966). Beginning the study of lexis. In C. Bazell et. al., *In memory of J.R. Firth.* London: Longman Group Limited.

Smith, R. & Frawley, W. (1983). Conjunctive cohesion in four English genres. *Text,* 3-4, 347-373.

SPSS base 7.5 for windows' user's guide. (1997). Chicago: SPSS Inc.

Stotsky, S. (1983). Types of lexical cohesion in expository writing: Implications for developing the vocabulary of academic discourse. *College Composition and Communication*, 34:4, 430-446.

Weissberg, R.C. (1984). Given and new: paragraph development model from scientific English. *TESOL Quarterly,* 18:3, 215-229.

Widdowson, H.G. (1978). *Teaching language as communication.* Oxford: OUP.

Williams, M.P. (1984). A problem of cohesion. In J Swales & H. Mustafa, H. (Eds.), *English for Specific Purposes in the Arab World* (pp.118-128). Birmingham, UK: Language Studies Unit,University of Aston.

Winter, E.O. (1979). Replacement as a fundamental function of the sentence in context. *Forum Linguisticum,* 4:2, 95-133.

Witte, S. & Faigley, L. (1981). Coherence, cohesion and writing quality. *College Composition and Communication,* 32, 189-204.

Yang, A.W. (1989).Cohesive chains and writing quality. *WORD*, 40:1-2, 235-54.

Yde, P. & Spoelders, M. (1985). Text Cohesion: An exploratory study with beginning writers. *Applied Psycholinguistics*, 6, 407-416.

Youman, G. (1991). A new tool for discourse analysis: The vocabulary-management profile. *Language*, 67:4, 763-789.

APPENDIX A

LAU EFL WRITING EVALUATION CRITERIA (LOC)

Scale: A (5) 90-100% ; B (4) 80-89% ;C (3) 70-79% ; D (2) 60-69%; F (1) below 60%

A   5   L Fluent in use of language; syntactic and vocabulary variety
          O Global and local organization excellent
          C Knowledgeable and support extensive and relevant
B   4   L Good command; shows some syntactic and vocabulary variety
          O Global and local organization adequate
          C Wide range of knowledge and coverage; has support but not
             varied nor specific enough
C 3 L Contains some errors but meaning not distorted; very
      little variety
          O Global and logical organization is fair but local incoherent
          sometimes
          C Shows fair knowledge and coverage; needs more specific
          support
D   2   L Errors in sentence structure and usage – no variety
          O Global fair, but logical and local development lacks focus
          C Very little knowledge and coverage of topic; little support;
          redundant
F   1   L Many and serious errors throughout
          O May have global paragraph order but extremely weak in
          logical and local development of ideas
          C Extremely limited; no support or inadequate; little relevance,
          pointless, poor quality

Language = Tone, style, sentence structure, grammar, vocabulary, coherence, mechanics
Organization = Format, logical order of ideas (globally and locally)
 Content = Main and minor ideas.