

# Patterns of Textual Coherence in Students' Written Discourse: A Study of Sudanese English Majors

Hamid Abdalla Arabi<sup>1</sup> & Nauman Al Amin Ali<sup>2</sup>

<sup>1</sup> Al Neelain University, Sudan, on secondment to Shaqra University, Kingdom of Saudi Arabia

<sup>2</sup> University of Khartoum, Sudan, on secondment to Shaqra University, Kingdom of Saudi Arabia

Correspondence: Nauman Al Amin Ali, College of Science & Humanities, Dawadami, P. O. Box 18, Saudi Arabia 11911. Tel: 966-50-411-3460. E-mail: [alnauman888@gmail.com](mailto:alnauman888@gmail.com)

Received: November 5, 2014 Accepted: December 1, 2014 Online Published: January 27, 2015

doi:10.5539/ijel.v5n1p91 URL: <http://dx.doi.org/10.5539/ijel.v5n1p91>

## Abstract

Coherence is an inherent element of all effective written communication and, hence, this study purports to investigate the manifestation of textual coherence types in a corpus of fifty answer sheets produced by Sudanese English majors at a large public university. The methodological underpinnings are based on Halliday's (1994) conception of the clause as comprising an information structure and a theme—rheme organization representing the topical constituents, while at the macro- textual level, Fahnestock's (1983) elaboration of continuative clauses was utilized. The analysis of the clausal and textual coherence was undertaken on the syntactic, lexical and discoursal errors embodied in students' texts, accompanied by a detailed study highlighting instances of coherence breaks. It was revealed that the overall coherence of the corpus was negatively affected by the innumerable errors committed at the syntactic and lexical levels. Intra- sentential clausal coherence, for example, was compromised by the incapacity to unambiguously distinguish between given and new information, due to the subjects' countless errors in using definite, indefinite articles and pronominalization processes. Equally, the information structure was often disrupted by faulty use of the passivization function. Regarding inter- sentential textual coherence, it was generally vitiated by such factors as students' poor handling of conjunctions, as reflected in confusion of the basic rhetorical typologies such as cause- result and conditional- consequence, a situation aggravated by the students' verbose writing style and their inattention to textual boundaries and macro- structure. The prevalence of the two major error categories has ultimately resulted in low interpretability, semantic ambiguity and textual opacity.

**Keywords:** coherence, theme and rheme, information structure, circumlocution, language transfer

## 1. Introduction

It can be cogently argued that writing is the most accomplished, yet challenging, language skill, particularly for EFL students. Equally valid is its crucial part in personal, professional and academic success, since it is the only means of communication that affords the opportunity to articulate ideas and synthesize perspectives in a persuasive manner that is independent of time and space constraints (Crowhurst, 1990). This entails an awareness on the part of students and teachers that writing is a complex and multi- faceted entity where discoursal, cognitive, socio-cultural and psychological factors both interplay and compete. On the linguistic side, a text is not a random set of words, clauses and sentences; rather it is an ordered sequence with conventional constraints on the possible orderings if it is to be a meaningful and unified whole (Van Dijk, 1985, p. 108). One of the seven inherent elements of *textuality* (De Beaugrande & Dressler, 1981) is *coherence*. Together with its kindred term *cohesion*, coherence is one of the linguistic system's chief means for text construction. In fact, coherence embodies the presence of overt textual cues rendering it possible for writers to highlight the inner textual propositional logic congruent with pragmatic notions and instinctive human quest for meaning and order.

## 2. Theoretical Framework

### 2.1 The Concept of Coherence

Throughout its history, coherence has proved too elusive for precise definitions, but there are two salient perspectives on which coherence operates: (a) text-based coherence and (b) discourse process-based coherence (Grabe & Kaplan, 1996, p. 67). The proponents of the first orientation consider *coherence* as text-property, and it is realized by certain linguistic signals used by the writer/speaker. In other words, in *reader- responsible prose*

(Hinds, 1987), it is the writer who has to properly use certain linguistic cues that link text patterns and, in turn, guide the reader to arrive at the intended message easily. In this perspective, coherence is commonly equated with cohesion. That is, cohesive relations can help the reader establish the interpretation of the message intended by the writer. The leaders of this position are Halliday and Hasan (1976) who claim that two sentences are coherent as far as their meaning is concerned. More clearly, whether discourse sentences are coherent or not depends on how much they actually inter-relate in meaning. They systematize this semantic notion into five lexico-grammatical categories: reference, substitution, ellipsis, conjunction, and lexical cohesion. These categories, they say, provide continuity from one sentence to another and make practical the analysis of texts. The other coherence elements, Halliday and Hasan (ibid) recognize, are three and they are only secondary to cohesion in building *texture*. Two of these features are textual: the first is thematic patterns and the second is information structure; the linear organization of information which comprises two subdivisions—the Topic/Comment (T/C) and the Old/New (O/N) distributions of information. The third component is macro-structure that categorizes it as one kind of discourse. Yet, it is to be noted that the first two concepts are no less central than cohesion to the overall framework of Hallidayan Systemic Functional Linguistics and, by implication, coherence theory and it is this position that will be espoused in the present study

In contrast, the advocates of discourse-based approach to coherence have been unanimous in criticizing the proponents of text-based definitions of coherence for divorcing the text from its communicative environment assuming that the role of textual features (including cohesive devices) must be examined within the communicative environment in which the text occurs. In other words, the extent to which cohesive devices constitute textuality is not determined by their formal (linguistic) correctness, but rather by virtue of the communicative functions they bear and play between one segment and another. Widdowson (1978, p. 26), for example, argues for the notion that a text can be coherent without overt, linguistically- signaled cohesion. To illustrate this idea the author (ibid, p. 29) provides the following example:

A. *That's the telephone*

B. *I'm in the bath*

A: *Okay.*

According to Widdowson, a hearer of this dialogue utterance can consider them coherent though they do not display any cohesive makers. The first speaker's (A) remark about the telephone is interpreted as a request, and (B)'s response as an apology for not being able to answer the telephone. The A's second remark (OK) is understood as an acceptance of B's apology.

Despite the recent spate of cognitive coherence theories, there are new perspectives which suggest that much of the coherence-building information is not simply imposed on the text by the reader, especially in ordinary everyday texts (Grabe & Kaplan, 1996, p. 69). Rather, it is built on the text message itself. Moreover, Grabe and Kaplan (ibid) warn that ignoring the text message in text interpretation and comprehension may lead to an undue over-emphasis on top-down processing, and to text irrelevance altogether. Similarly, some cognitive psychology research findings, such as (Van Dijk & Kintch, 1983; Anderson, 1990; Beck et al., 1991; Britten & Gulgoz, 1991; Singer, 1990), have proved that text structuring influences interpretation and it is a major contributor to text coherence. This lends strength to text-based position and the current study which aims to put it into practice.

## 2.2 *The Criteria for Analyzing Text-Based Coherence*

Based on Halliday (1976, 1985, 1994), coherence originates in the function of successive sentences as information units. At sentence level, coherence is achieved by cohesive relations, which denote the way of information organization, such as: from general to specific, from whole to parts, etc. Within a sentence or at clause level, coherence is realized by the use of definite versus indefinite noun phrases, pronominalisation, and so forth. In short, coherence, at textual level (intersententially) is measured by means of cohesion, while within a sentence (intrasententially), coherence is measured via *Given – new information*. Since the relation between cohesive relations and text coherence has so far been illuminated, it remains to shed some light on coherence at clause level (within a sentence).

### 2.2.1 *Intrasentential/Clausal Coherence*

Coherence at clause level is usually measured by means of what is known in Systematic Functional Grammar (SFG) as *information structure* and/or *thematic structure*. According to Bloor and Bloor (1995, p. 65), there are two parallel and interrelated systems, in SFG, for analyzing the clause structure with regards to organizing the message: (a) *Information structure* and (b) *thematic structure*.

### 2.2.1.1 Information Structure (IS)

This concept was propounded by Halliday (1994) who adopts the Prague School concept of information as comprising two categories: *new information* and *given information*. In English, *New information* is introduced by indefinite expressions and subsequently referred to by definite expressions (Brown & Yule, 1983, p. 169) as in

A/ 1. *Yesterday, I saw **a** little girl get bitten by **a** dog.*

2. *I tried to catch **the** dog, but it ran away.*

The noun phrase *a dog* in the sentence (1) is New information, whereas, *the dog* in the second sentence is *Given information*. Brown and Yule (ibid, p. 171) enumerate the following semantic forms as indicators of *Given information*, they organize them into two groups. The first group includes lexical units (articles, lexical cohesion units); the second group includes reference pronouns.

#### group (A):

(i) items mentioned for the second time as in example (A) above.

(ii) items presented in the semantic field of a previously-mentioned lexical unit as in (B):

B/ 1. *Robert found **an** old car.*

2. ***The steering wheel** had broken off.*

The noun phrase ‘an old car’ determined by indefinite article ‘an’ is ‘New information’. The noun phrase (in bold) pre-modified by the definite article *the – steering wheel* is *Given information*.

#### group (B):

(i) pronouns referring anaphorically to a full lexical form in the preceding sentence as in (c), below.

C/ 1. *What happen to the jewels?*

2. ***They** were stolen by a customer.*

The pronoun ‘they’ (in bold) refers to the noun group ‘the jewels’ in the first sentence. Therefore, it is ‘given information’.

(ii) pronouns referring exophorically (the reference is not in the text, rather somewhere in the context) to the physical context of situation, as in (d)

D/ 1. *Look out.*

2. ***It**’s falling.*

(The above-cited examples are all quoted from Brown and Yule, ibid).

### 2.2.1.2 Thematic Structure (TS)

*Thematic structure* is the second system of analysis involved in the text organization. Like, *information structure*, it operates at clause level. All full clauses have thematic structure (Bloor & Bloor, 1995, p. 71). According to Bloor and Bloor (ibid), *thematic structure* is similar to *information structure* and in many clauses, there is a parallel equivalent between *theme* and *given*, on one hand, and between *rheme* and *new*, on the other. Therefore, some linguists conflate the two concepts. Halliday (1994) explains that the *theme* is the point of departure of the message. Hence, it is carried by the first constituent of a clause. The *rheme* is, then, what follows the starting point in a clause. According to Bloor and Bloor (ibid, p. 72), all clauses in English have what is called *topical theme*—that is the first constituent of a meaningful clause. It is realized by one of the following elements: *subjects*, *predicators*, *complements* or *circumstantial adjunct*; the *theme* is of two types: unmarked and marked. The “theme” is said to be unmarked if it is realized by a subject in a declarative clause, as in:

A. Tom sent the letter.

B. The students of engineering are busy with new technologies.

The underlined subjects *Tom* and ‘*the students of engineering*’ function as unmarked themes for the clauses (A) and (B), respectively. The marked *theme* is, by contrast, the one that is found in a clause position where it is not a subject. It is most frequently realized by circumstantial adjuncts (Bloor and Bloor, 1995, p. 76). For example,

C. For a long time, the man has not shown up.

D. After a few hours, the students finished their homework.

The underlined groups represent the *marked themes* for (C) and (D) clauses.

In fact, the speaker / writer is constrained by some syntactic rules to build grammatical sentences. Yet, it is interestingly available to him to convey one message in different sentence structures. This point can be illustrated in the following sentences:

E. *Tom sent the letter.* (active voice)

F. *The letter was sent ...* (passive voice)

The speaker / writer who chooses sentence (E) wants to focus on the subject *Tom*. The opposite is true for the one who uses the passive form (F), i.e. intends focusing on *the letter*. Thus, having a good command on thematic structure system helps a lot in changing the focus of the clause in significant functional ways.

### 2.3 Questions of the Study

The study purports to answer the following questions:

- 1- How well does this sample of Sudanese English majors perform regarding aspects of clausal and textual coherence in their written discourse?
- 2- How do such coherence errors impinge on the readability of their texts?
- 3- Based on The Contrastive Hypothesis and Interlanguage Theory, what are the underlying causes for coherence errors in the corpus?

## 3. Methodology

### 3.1 Participants

The subjects of the present study are fifty fourth year English majors who were about to graduate with a B.A. degree in English from The Department of English, Faculty of Arts, Al Neelain University, Sudan. The main rationale for choosing this sample is their homogeneity, since they are all Sudanese nationals who have been learning English as a foreign language under the same syllabus for eleven years: seven at Basic and Higher Secondary Schools and four as a major university subject.

### 3.2 Instrument

The instrument of this study is 50 answer sheets randomly chosen from the Final Examinations of the academic year 2012-2013. The choice of the researchers fell on the examination written production for two reasons, the first of which is that under such condition the attendance is complete and so ensuring a wide and representative research sample. Second, in examinations students perform at their best in the knowledge that the assessment of their answers will be incorporated in their final grades. Consequently, the type of written samples they produce will embody such textual aspects as coherence. One such an ideal environment is a course entitled *Advanced Essay Writing* (one of nine courses prerequisite for fourth year students). In the final examination of this course, the students were required to develop a topic sentence namely, *It is interesting to have a job that involves travelling* into a coherent paragraph. In a second question, the students were provided with the option of developing one of the following thesis statements into a coherent and well- organized essay:

- a) *Banning smoking in public places protects people's health.*
- b) *Learning a foreign language is important for every student*

The body of writing produced in answer to the above questions furnished the setting for the coherence elements analyzed in this study.

### 3.3 The Analysis of Propositional Coherence

Propositional / textual coherence is approached from two perspectives: coherence at clause and textual levels, as further explained below.

#### 3.3.1 The Analysis of Coherence at Clause Level

Coherence errors at clause level are analyzed on the basis of the two analytical systems: information structure and thematic structure. The function of the two systems is organizing the clause message. The first involves *Given* and *New* constituents, while the second involves *Theme* and *Rheme* constituents (Bloor & Bloor, 1995, p. 65).

*Given* constituent usually carries the mutual / shared information between the writer / and reader. It is commonly marked by the definite article *the* or a reference pronoun such as: *he, it, they*. The new constituent is the rest of the clause or what is said about the mutual information.

Concerning the present study, the errors relating to definite and indefinite articles, reference personal pronouns and demonstratives are analyzed in relationship to their functional role in information structure. For example, the

syntax error of *a/the* replacement may functionally influence the information structure. That signifies that if the indefinite article *a* is used in the syntactic context where *the* is obligatory, it may change the *Given* information into *New*, or vice versa, which in turn, impedes the information structure system. For instance, the following two clauses (F-units) are considered incoherent because their information structure is compromised through incorrect use of articles:

*When we entered we faced by the (a) huge dog.*

Although the noun phrase *huge dog* appears for the first time (*New information*), the writer modifies it by the definite article *the*. In so doing, he/she turns *new-information* into *given-information* (mutual / shared one), which breaks the coherence relation holding between the two clauses.

Thematic structure system is also applied for assessing clausal coherence. That means, it is used as an analytic framework for errors associated with deviation from the topical structure (theme) of the sentence / clause. For example, the writer begins the text with certain constituents (Theme-Rheme). And, instead of beginning the second sentence / clause with the same 'theme' or 'rheme' of the first sentence, he/she makes a shift to another new 'theme', which is either indirectly relevant to the first theme or irrelevant altogether. Such type of writing may disrupt the clause message organization, and break the coherence relation between one clause and another.

### 3.3.2 The Analysis of Textual Coherence

The errors relating to textual coherence are analyzed with reference to relational coherence being derived from the logical relationships between clauses (F-units). In this respect, the semantic relations (established by conjunctive words) between one clause and another play a great role in realizing coherence at textual level. To analyze the textual coherence breaks in the subjects' written texts, the taxonomy of continuative and discontinuative relations between clauses, set forth by Fahnestock (1983), is adopted.

Fahnestock (ibid, p. 406) believes that coherence is achieved when readers are able to move from one clause to another without losing meaning, as well as when larger text segments manifest integration among themselves. She makes a distinction between continuative and discontinuative relations. The continuative relation exists when the second clause bears a content expected by the reader or hearer with regard to the first clause. For example, in

*He was very sick, so he did not come to the meeting.*

the second clause of this sentence provides an expected result for the reason mentioned in the first. Hence, the continuative relation known as *reason-result* is established between the two clauses. The whole sentence is, then, coherent by virtue of 'reason-result' continuative relation.

Discontinuative relation, by contrast, apply when the second clause carries a meaning unexpected by the receiver in relation to the first clause. In the following sentence, for instance,

*He was an outstanding student, but he left school.*

leaving school (specified in the second clause), is unexpected attitude or decision by an outstanding student (provided in the first clause). That is why the two clauses are joined together by the adversative conjunction *but*.

In the present study, most textual coherence breaks are caused by incorrect use of subordinating conjunctions such as: *although*, *because*, *if*, etc. Moreover, the most frequent types of such coherence impediments are the ones that involve *cause-effect*, *reason-result*, *condition-consequence* and *contrast* relations. Each of these errors is illustrated, analyzed and discussed in turn.

## 4. Discussion and Analysis

As a grounding for the present study, below are the frequencies, both as percentages and relative ranking, of aggregate syntactic, lexical and textual errors in the corpus.

Table 1. The distribution of syntactic and textual errors in the corpus

Types of errors	No. of error	Percentage %	Mean	Std. Deviation
1.Tense/aspect	418	15.89656	8.36	4.039
2.Conjunction	392	15.26902	7.84	3.722
3.Reference	241	9.165241	4.82	2.919
4.Lexical/coll.	217	8.252519	4.34	2.536
5.NNSE	178	6.769348	3.56	2.612
6.Morpheme	156	5.932687	3.12	2.173
7.Preposition	147	5.590416	2.94	1.845

8.Article	137	5.210116	2.74	2.530
9.sent.fragments	124	4.715725	2.48	1.940
10.Word class	124	4.715725	2.48	1.843
11.Copula	98	3.726944	1.96	1.937
12.N/V.agreement	88	3.346644	1.76	1.585
13.Circumlocution	85	3.232554	1.70	1.329
14.N/adj.agreement	82	3.118464	1.64	1.758
15.Run-on sent.	74	2.814223	1.48	1.741
16.active/passive	59	2.243773	1.18	1.224
Total	<b>2620</b>	<b>100%</b>	<b>49.82</b>	

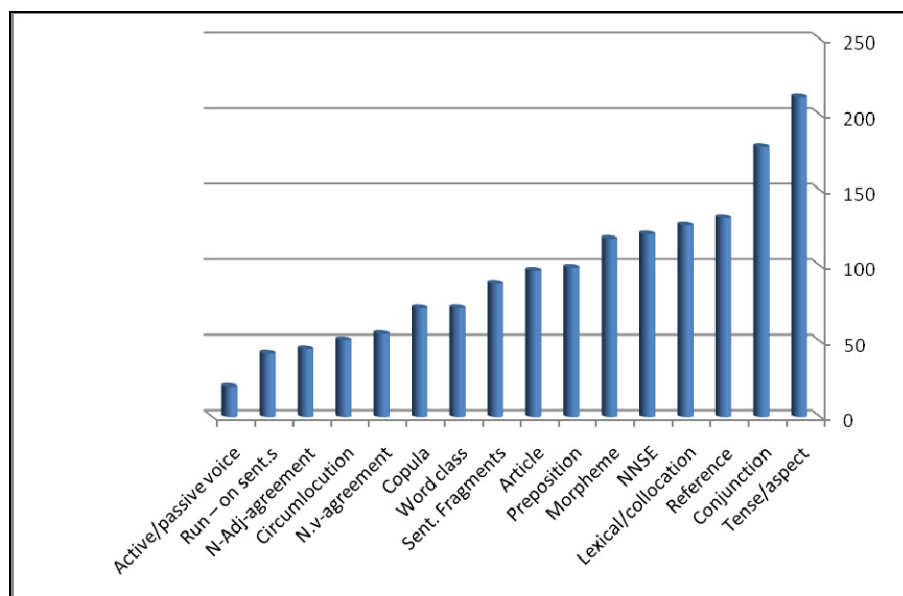


Figure 1. The relationship between error types and no. of errors

Since the objective of this study of coherence breaks, it is worth remembering that text-based coherence is itself of two types: (1) clause / sentence coherence, and (2) text coherence. Following these types, the coherence breaks detected in the corpus are put into two main groups: clausal (intrasentential) coherence breaks, and textual (intersentential) coherence breaks. Each is separately discussed under the following subsections.

#### 4.1 Clausal Coherence Breaks

Coherence at clausal / sentential level is measured through what are known in functional grammar as *information structure* and *thematic structure* (Bloor & Bloor, 1995, p. 65), as detailed below.

##### 4.1.1 Information Structure Breaks

Information structure involves constituents that are labeled *Given – new information*. Any clause (dependent, independent or a combination of both) has *Given – new information*. These constituents are syntactically represented by articles (definite and indefinite) and reference pronouns. Hence, information structure breaks can be described as the ones that are linked to *given – new information* analytical system, including article and pronoun errors. In the following subsections, the contribution of article and pronoun errors in breaking text segments coherence (particularly, at clause level) is illustrated.

##### 4.1.1.1 Article Errors

This category of errors constitutes 5.2%. This result denotes that the coherence relations holding between text segments at clause or sentence level are seriously impeded. This is ascribable to the fact that the definite and indefinite articles are the linguistic markers of *Given-New information* structure. Below are some illustrative examples:

[1] *Only day \*me and \*my friend were invited to birthday of our one of friends we left home with the gift in our hands ....*  
[25/NW]

The constituent *me and my friend* is, technically, Given- information, which syntactically functions as the subject of the clause. The predicate *were invited to birthday of our one of friends* (apart from incorrect word order) is considered New – information. Although the word *birthday* (underlined) is part of New-information, it must be preceded by *the*, because it has become a particular *birthday* after being post-modified by *of one of our friend*. In the second clause, the underlined definite article *the* is incorrectly used in a textual context where it does not apply. The expression *gift* is mentioned in the context for the first time, “new – information”. Therefore, the writer ought to have predicated it with the indefinite article *a* to mark its novelty. He/she, however, does the opposite. i.e. he highlighting the word 'gift' as *given – information*, by pre-modifying it with *the*. In so doing, the writer violates the system of information structure and, consequently, breaks the coherence chain.

Another violation of information structure system can be evidenced in the following written sentence.

[2] *When we entered we faced by the huge dog.* [25/NW]

The same above – cited argument holds true for the second example. The word *dog* in the second clause is made *Given – information* though it is new. It is *New-information* because it is mentioned for the first time, as the context shows. In such a case, the noun "dog" has to be pre-modified by *a* (*a huge dog*), rather than *the*.

In both (1) and (2) examples, coherence breaks involve turning *New-information* into *Give-information*, by dent of indefinite article *a* replacement by definite article *the*. That is to say, the writer uses *the gift* and *the huge dog* instead of *a gift* and *a huge dog*, successively.

One more example to illustrate the point:

[3] *Smoking is Ø expensive habit* (is **an** expensive habit). [28/AW]

The underlined constituent *is expensive habit* is granted to be *New – information* for the reader. Furthermore, the sentence is presented as a thesis statement, which denotes that the pieces of information it comprises are new though technically the word *smoking* is given – information for being the subject of the sentence. Accordingly, the noun phrase *expensive habit* should be marked indefinite and then *New – information* through using the article 'an'. Thus, the deletion of the indefinite article *an* has rendered *New – information expensive habit* unmarked and the whole sentence incoherent as well.

#### 4.1.1.2 Pronominal Errors

An integral element of Given – new information and, hence, clause structure is the presence of pronouns that help weave a fabric of shared knowledge. Three types of pronoun errors are found to be involved in breaking coherence (information structure) at clause level: (1) omission errors. Each group of errors is dealt with below.

Errors involving pronoun disagreement with their referent nouns are considered one of those syntactic errors marring the information structure and impeding coherence as well, especially at clause level. If, for example, a pronoun disagrees with its antecedent, the reader may find it difficult to keep up with the information flow via particular text segments. This can be gleaned in the following example:

[4] *It was in the summer of 1999, when I tried to visit my brother in Kassala ... I reached there .... the street was empty, because... people kept themselves in their houses escaping from the sun ... I phoned my brother to describe th way which lead me to their (his) house, but I failed because his telephone was out of service.* [36/NW]

Drawing on the context, the writer is presumed to be looking for his brother's house, and calls the same brother to help. However, he/she breaks the coherence of these successive situations by referring to the word *brother* as *their*, thus distracting the reader's attention from *my brother* to other endophoric referent nouns, plausibly the plural form *those people who kept themselves in their houses*. As a result, the flow of information related to the above written text may be jeopardized. Yet one more such disagreement is discernible in the following text:

[5] *In order to learn a foreign language, all students should join the schools or institutes which are \*specialist in teaching a foreign language \*so as to meet the need for applying it in our (their) life.* [38/AW]

The flow of information is equally hampered by the incorrect use of reference pronoun *our* whose referent noun / pronoun cannot be approached in the text. It is important to point out here that the above – provided text is the second paragraph of the writer's whole essay having the thesis statement of *learning a foreign language is important for students*. Therefore, the noun phrase *all students* is mentioned for the second time, rendering it *Given- information*. As a consequence, if it is to be mentioned once again in the clause, , it must be replaced by one of its reference pronouns *them*, *their*, *theirs*, etc, according to syntactic norms. However, it can categorically not be referred to by the first person – plural pronoun *our*, as the writer does.

On a level with noun – pronoun agreement, pronoun misuse errors may violate the information structure system and render text-segments incoherent. The following written chunk illustrates the point.

[6] ... *the lorry ran very fast beside the cannal when he (it) reached near the bridge the lorry (it) \*loose his (its) control.* [7/NW]

The reader can perceive the general idea intended by the writer. However, he/she may find it fairly difficult to trace the events being narrated, since the information structure system is violated by the incorrect way the writer marks the reference relationship between the noun phrase *the lorry* (the given information of the first sentence) and its referent *it* (the Given-information of the second sentence). The erroneous way of marking is evident in the noun group *the lorry* disagreement with the pronoun *he*. English syntax does not allow referring to inanimate nouns by such personal pronouns as *he* and *she*, as the writer anaphorically does.

Other coherence breaks are uncovered in the same sentence where the writer tautologically begins the second clause with *the lorry* instead of *it*. What is noteworthy is that, he/she once again, incorrectly refers to *the lorry* as *his*, rather than the correct pronoun *its*. In fact, the text could be split into shorter manageable sentences to replace such an ill – punctuated and uncontrollable one as above. For example

...*The lorry was running very fast along the channel. When it got at the bridge, it lost control.*

Such personification errors as above are attributable to the subjects' mother tongue interference. In Arabic, the reference pronouns equivalent to English *he* and *she* are *هو/huwa/* and *هي/hiya/*. They are not only used to refer to singular masculine and feminine singular nouns as *man* *رجل /rajul/* and *girl* *بنت /bint/* as in English, but also to singular neutral nouns. In Arabic, neutral singular nouns are divided into two types: the ones treated as masculine (equivalent nouns to *book* *كتاب /kitaab/* and *lorry* *لورى /loori/*, etc) and the others, which are treated as feminine (nouns equivalent to *room* *حجرة /hujra/* and *school* *مدرسة /madrasah/*, etc). Each singular neutral nouns from the first group is referred to as the same pronoun 'هو' (equivalent to *he*) which originally refers to masculine singular nouns as *man*, *boy*, etc. A singular neutral noun, from the second group, can be referred to by *she*, in the same way as singular feminine nouns. In short, whereas English reference to in/animate is different, it the same in Arabic. As a result, Arab learners of English, such as the present sample, may refer to inanimate English words like *lorry* and *school* as *he* and *she*, respectively in the same way as in their L1.

Pronoun omission errors come third and they involve deleting a reference pronoun from its syntactic contexts where it functions as a subject. Below are two sentences for the errors.

[7] *After a few minutes Øcalled me (she called me) and said...* [8/NW]

[8] *On the morning Øleaved the hotel (we left hostel) ...* [4/NW]

The pronouns *she* and *we* are missing in these examples. Both pronouns syntactically function as subjects. In the meanwhile, they are supposed to convey the *Given-information*, which is regarded as one of the basic two components of information structure system constituting clause coherence. This signifies that the deletion of *she* and *we* results in unraveling the shared thread, aiding the reader in tracing the propositional contents intended by the writer. Therefore, pronoun deletion errors, like the ones exemplified above, are involved in rendering written texts incoherent.

#### 4.1.2 Thematic Structure Breaks

Concerning the current study, the subjects' written texts have revealed manifold clause coherence problems. These problems are due to the students' unawareness of the function of passivisation, coupled with their syntactic inability to form passive sentences. Another factor for violating thematic structure system is sentence fragment errors discovered in the subjects' surveyed corpus. Therefore, coherence errors of thematic structure are divided into two types: involving active/passive voice, and those involving sentence fragments. Each of the two groups is discussed below.

##### 4.1.2.1 Active / Passive Voice Errors

Although this type of errors is syntactically the least in frequency (2.2%), it is the one that seriously distorts the thematic structure, resulting in clause coherence interference. There are occasions when the reader of the subjects' written texts cannot grasp the intended meanings because of incorrect thematization as shown in the following examples:

[9] *My friends are advanced (went before) and I promise to meet them.* [16/N]

The writer of this sentence does not clearly state what he/she wishes to say about his / her friends. For example, should the reader entirely depend on the manner the writer places the *theme* and *rheme*, he/she may not be able to



reach the intended message. This is a consequence of the writer committing the dual error of putting the first clause in a passive form. First, he/she need not have used the passive form to focus on *my friends*. Second, he/she fails to select the suitable lexical word that transmits the intended meaning. Therefore, the 'theme – rheme structure is disrupted. Moreover, the writer incorrectly uses the simple present to describe successive events that took place in the past.

Another example of incorrect passivisation is illustrated in the following clause.

[10] ... *they consider (are considered) the languages of technology.* [11/AW]

Contrary to the first, the writer of the second example forms active voice sentence in the context where passive voice is required. He speaks about foreign languages, and the pronoun *they* -at the beginning of the sentence - refers to foreign languages. To make the pronoun *they* the theme of his/her clause, the writer should have turned the whole clause into passive voice as in the following modified version:

*They are considered the languages of technology.*

If this sentence is kept in its active voice form, as the writer does, the focus will not be on the foreign languages, rather on the persons who consider foreign languages as the languages of technology. It can, then, be said that students' ignorance of the function of using the passive voice, besides their failure to syntactically form passive sentences all violate the thematic structure of their written sentences.

#### 4.1.2.2 Sentence Fragments

Sentence fragments are used in this study to include the errors associated with ungrammatical combinations of sentence constituents. These errors may interfere with local coherence (coherence at sentence level) of the written texts. This is evident in Table [1] above, where sentence fragments score 4.7% of all errors in the corpus. In the following written text patterns, the reader can obtain a general idea of the intended meaning, yet it is a demanding task to pinpoint what the writer wishes to state, such as in the following

[11] *My astonishment was very great that was the moment when I realized I was in the wrong place, for the bride was not my intended friend.* [22/NW]

The first problem inherent in the thematic structure is that the writer starts with the unmarked topical structure *my astonishment* in a context where *marked theme* is more appropriate *To my great astonishment*. Technically, the writer is unsuccessful in shifting the main topical structure (unmarked theme) *the bride* from the very beginning of the sentence to the end, hence, piling three subordinating conjunctions, which he/she ultimately fails to command.

Another example of sentential coherence breaks can similarly be noted in the following

[12] *For example, if you introduce for job, there is a condition that requires to be talk more than one language and almost the English language it is very important, because of the how of the foreign investigations and foreign companies.* [25/AW]

It is extremely hard to glean the meaning the writer of the above-cited sentence desires to convey. The reader has to labour after the content of the message in this uncontrollable sentence. The first difficulty lies in the writer stating so many points in only one sentence, and, due to his/her limited vocabulary as well as a poor syntactic knowledge, he/she fails to manage the sentence. As a result, the thematic structure system is marred, if the communication does not break down altogether. Again, the text can make sense and the thematic structure can be retained if it is broken into shorter simple sentences including:

*For example, if one wants to apply for a job, he needs to have more than one language. English, in particular, is the most L communication.*

Some students inexplicably set two unmarked themes for one rheme. To put it another way, they syntactically form two subjects for only one predicate, resulting, in two different topical structures as in:

[13] *This problem, smoker is really effected economically by reducing the production.*

The student is discussing the importance of banning smoking in public. Obviously, there are two subjects for one verb *is* namely, *this problem* and *smoker*. Therefore, the question arises which subject in the topical structure is confusing to the reader. It can be understood from the context that by *this problem*, the writer refers anaphorically to *smoking in public places*. However, it is not clear why he/she inserts the word *smoker* in the same syntactical and thematic position of *smoking in public*.

Another question arises as to how product reduction economically affects *smokers* or *smoking in public*. Plausibly,, what the writer has stated is a complete reversal of the intended meaning. In other words, it is *the economy* that can be negatively affected by smoking in public since the latter may cause serious diseases to people who are supposed

to contribute to economic development. The writer's first failure is, therefore, associated with the unclear topical structure (two themes for one rheme), and, hence, the ambiguity of which part of the sentence is to focus on represents the second problem.

#### 4.2 Textual Coherence Breaks

Information / thematic structure is not confined to establishing coherence at clause / sentence level; rather it extends to realizing coherence relationships at a textual level as well. The students' corpus exhibits errors pertaining to textual coherence impediment. The most frequent errors of this type are *incorrect subordination*, *circumlocution*, and *run-on sentences*, as follows;

##### 4.2.1 Incorrect Subordination

It is shown in Table [1] that conjunctions are the second most frequent errors, amounting to 15.3%, indicating that conjunctions pose difficulties for the subjects. Subordinating conjunctions like *although*, *but*, *however* are no exception. There are three types of subordination errors manifested in students' written texts: (a) *contrast*, (b) *cause – result*, and (c) *conditional – consequence*. The use of these, in textual contexts where their functions do not apply, may hamper the flow of the meaning. The next instance displays the errors of this type:

[14] *Although our \*neighbour had a dog, but I didn't adventure to come close to it.* [26/NW]

It is evident from the context of the above example that the second clause is the result of the first, rather than the opposite. The sentence contradicts our pragmatic notions as the writer intends to state that his/her neighbours had a ferocious dog that is why he/she did not venture to come close to their house. The content of the first clause causes the content of the second. Hence, the subordination relationship that holds between these two clauses is *cause-result* rather than contrast. Conjunctions *because*, *so*, *therefore*, etc have to replace *although* and *but* (note the writer is using both *although* and *but* in the same sentence) for the sentence to make sense. The sentence can be rewritten as:

**Because** *our neighbours had an aggressive dog, I did not venture to come close to their house.*

As an explicit subordinating conjunction, *cause – result* is primarily realized through *because*, *therefore*, *for this reason*, *as a result*, *so*, and is designed to establishing cause – result relation. Yet, the subjects encounter immense difficulties in making such coherence relations as above. This is indicated by the errors they make in subordination, for example:

[15] *... and almost the English language \*it is very important, because of the how of the foreign investigations and foreign companies.* [25/AW]

The writer above fails to signal the reason for English being important, which is supposed to follow the subordinating conjunction *because* at the beginning of the second clause. Therefore, the *cause-result* relation does not hold or it is, at least, incomplete.

Another *cause – result* coherence relation break can be discerned in the following example. *I sat of the last place, so no one on it ....*

The literal meaning of this sentence is that the writer had a seat in the last row, and as a result, there was nobody sitting on it. This type of subordination is conventionally unnecessary, because one does not sit on the seat, which is occupied by somebody else. Yet, even if the writer wishes to subordinate this meaning, he/she must do the opposite. That is, he/she must make the first clause the result of the second:

*There was an empty place in the last row, so I occupied it.*

Condition –consequence relation subordination is also used to signal *modality*, i.e. the probability, possibility, or necessity of events/situations (Hyland, 2005). The subordinating conjunction *if* is usually used to mark a condition under which certain events or situations should be true. The subjects of the present study seem to have problems in forming conditionals. This is apparent in the incorrect use of conjunction *if* revealed in their corpus. The following sentence is an example detected in the corpus.

[16] *And if they are not studying a foreign language and take another studie, they need to learn about foreign language to broaden their knowledge.* [16/AW]

This subordination is incorrect, as the concept enunciated in the *if-subordinate* clause is an illogical condition for the second clause. That is to say 'they are not studying a foreign language, then how do they learn it?'

Such type of subordination does not only impedes the text coherence, but also breaks down the communication. It would have been better for the writer to start the subordination with the last pattern of the sentence *to broaden their knowledge* as an *apriori* of the second clause *they need to learn about a foreign language*.

Coherence impediment is also noticeable in the following written piece:

[17] \**These information generated everyday unless education and sciences are continuing.* [24/AW]

Again, this sentence displays a mismatch between the concepts provided in the two clauses and the reader's axiomatic knowledge of the world. Therefore, coherence relation between the two clauses does not hold. To put it another way, the condition presented in the second clause does not negate the occurrence of the event postulated in the first. On the contrary, the continuity of education entails the continuity in generating information. In fact, the concepts of the two clauses are better subordinated with *as far as*, rather than *if* or *unless* since the relation between them is one of coexistence as shown in:

*This information is daily generated as far as education continues.*

#### 4.2.2 Circumlocution Errors

This type of errors can also weaken the coherence relation holding between text segments. In the corpus, circumlocution is one of the frequent errors accounting for 3.2%. As denoted by the word 'circumlocution', the errors involve indirect statement of ideas by the subjects. They are prone to write so many words to encode simple messages in the way that may render their written texts incoherent. Below are some examples, written out from the subjects' corpus.

[18] *Students who are under graduation or who had already graduate from college need to have another language in addition to their first languages in order to make a high studies for example.* (31 words) [5/AW]

The writer uses abundant words (31) to express a very simple idea, namely, *All students need to learn a foreign language*. In the very beginning, the writer uses two independent clauses (each starts with *who*): *students who are graduation or who had already graduated from college* as post-modifiers to the subject (the topical theme) –*Students*. All these words can be made up for by *graduate/undergraduate students*. In short, the writer fills his/her sentence with redundant, boring expressions in a context where he/she has to be explicit. Consider the revised version

*Students, at whatever level, need to learn a foreign language (English, in particular) to help them carry out their studies.* (15 words) The following is another example for a verbose sentence where meanings are not directly stated.

[19] *Recently, learning a foreign language become necessary for students regardless to the kind of motivation, such kind of learning enable students to contact with foreign countries then allow or make it possible to exchange and interact with those countries, in addition to the ability of reading books that were written in a foreign language.* (55 words) [27/AW]

Each underlined expression is mentioned twice, rendering it difficult for the reader to easily follow the string of the argument. This is due to the repetition and redundancy of words, since the writer rephrases the self- same idea. For instance, he/she employs these two identical pairs of phrases in one sentence (1) *to contact with foreign countries*' and *“to exchange and interact with those countries*; (2) *learning a foreign language*' and *'such kind of learning*. Such tautologous repetitions do not only interfere with the coherence relation holding between written-text patterns, but also causes the intended message to be interrupted. It is unconventional to use more than 55 words to convey such a simple idea as *learning a foreign language is important for students*. Note the revised version.

*Learning a foreign language has recently become necessary for students to communicate with foreigners and access to books written in that language.* (23 words)

Thus, on the whole the subjects' written texts are characterized by indirectness in conveying messages and expressing ideas. This, in turn, renders their writing incoherent and vague.

#### 4.2.3 Run – on Sentences

Comprising 2.8%, *run-on sentences* are one of the salient errors in the corpus. They are closely bound up with the circumlocution errors, discussed above, in the sense that the subjects do not directly identify the ideas they intend to express. Students, being unaware of English syntax rules, do not pay attention to sentence boundaries and so jot down as many words as possible, a practice culminating in the incapacity to control ideas. This situation is aggravated by the subjects' ignorance of punctuation marks. Run – on sentences are major obstacles to text comprehension, as the following sentence illustrates;

[20] *Students are required to take a foreign language //because this new world it required // that the man to be has broader education // as to be easy for him to go beside this technology period //, so that this is he period of internet and computer // and it need a big knowledge.* (50 words) [7/AW]

The sentence (comprising 50 words and 6 clauses) is exceedingly hard to manage, even for English native speakers. The presence of six clauses in only one sentence underscores the fact that the sentence has, at least, three different ideas. Under such circumstances, it would have been better be split into three or more smaller manageable sentences. Another observation is that the writer's limited lexicon compels him to utilize numerous expressions to convey a plain idea.. For example, the following sentence fragment *as to be easy for him to go beside this technology* can be replaced by *to cope with this technology*. Similarly, the wordy clause *because this new world it required that the man to be has broader education*, can be compensated with *because the new technology involves good education*.

Another example of run - on sentences is also shown in the following:

[19] *In all the world the ministries of healthes are warning the people about the dangerous of smoking, and in the \*last years we find that the governments is direct their sitizens to let that habit in public places, but unfortunately after all these directions, also there are some perons never take care about this bad habit.* (54 words) [28/AW]

Again, this sentence has more than one idea, which is the primary reason for the writer's failure to cope with it. Apart from some syntactic problems, 54 words in one sentence is fairly difficult to control, especially for EFL students. Furthermore, the writer is most likely influenced by the way Sudanese people chat in colloquial Arabic, as shown in this sentence

"لكن لسوء الحظ بعد كل التوجيهات دي برضو في بعض الناس ما بيعملوا حسابهم أبداً للعادة الضارة دي" / *laakin lisuu ai- hadh ba3d kul al- tawjiihaat di bardu fi ba3d al- naas maa bikhutu baalun lil al- 3aada al- daara di* /.

.Spoken Arabic intrusion in Arab students' EFL- writing has recently been proved by research in contrastive rhetoric (Connor, 1996). Ostler (1987), cited in Grabe & Kaplan, 1996, p. 193) notes that: *Arabic ESL writers are heavily influenced by classical Arabic, a language which is more reflective of oral traditions in language use*. Likewise, Sa'Adeddin (1989) reinterprets the many textually oral features in Arabic texts as representing one rhetorical option in written Arabic, though one which is typically preferred – an “aural” mode of writing, (quoted from Garbe & Kaplan, *ibid*).As a consequence, he/she does not abide either by English syntax system or English rhetorical conventions. This is illustrated by the last clause, where the writer uses four conjunctive words (three of which are in succession):

.... *but unfortunately after all these directions, also there some persons never take care about this bad habit.*

To turn the text into proper English style, some modifications are required. First, it must be split into smaller manageable sentences. Some redundant items have to be deleted. The revised version can be thus:

*The Ministries of health, all over the world, keep warning people on the danger of smoking. Recently, governments have set regulations that ban smoking in public. However, some smokers never take care about them.*

Thus, run – on sentences are one of the linguistic problems being encountered by the subjects. Run – on – sentence errors can be attributed to the subjects' ignorance of writing conventions and English syntax system, on one hand, and mother tongue (Arabic) interference, on the other, as illustrated above.

## 5. Conclusion

Though the present study is based on a moderate sample of fifty texts, it was proved that the students *do* commit numerous syntactic, lexical and semantic errors and, more importantly, that some of the above syntactical and cohesive errors directly result in text coherence breaks, particularly at clause levels. It was demonstrated through innumerable citations from the corpus that incorrect use of definite/indefinite articles, for example, along with wrong subordination and misused central and demonstrative pronouns combined may violate the *Given-New* information and *thematic* structure systems set forth by Systemic Functional Linguistics. It is well- attested that both information structure and thematic structure are responsible for establishing relational (propositional) coherence at clause/sentence and discourse levels alike Violating these two systems for one reason or another, as in the present study, will, by definition, block the written text coherence. Four structural errors were found to be directly responsible for coherence impediment namely, *sentence fragments*, *circumlocution*, *run-on sentences* and *active/passive replacement*. To answer the research question of *To what extent the errors in the present study correlate with the subjects' overall writing*, the results have indisputably proved that coherence errors indeed correlate with the overall writing quality to such an extent that the corpus is hardly interpretable.

## References

- Anderson, J. (1990). *Cognitive Psychology and its Implications*. Oxford & New York: Freeman.
- Beaugrande, De & Dressler, W. (1981). *Introduction to TextLinguistics*. New York: Longman
- Beck, I., Mckeowin, W., Sinatra, G., & Loxterman, J. (1991). Revisin studies text from a text-processing perspective: evidence of improved comprehensibility. *Reading Research Quarterly*, 26, 251-276. <http://dx.doi.org/10.2307/747763>
- Bloor, T., & Bloor, M. (1995). *The Functional Analysis of English: A Hallidayan Approach*. Oxford: Oxford University Press.
- Britten, B., & Black, J. (1985). *Understanding Expository Text*. Hove, Sussex and Hillsdale, NJ: L. Erlbaum.
- Britten, B., & Gulgoz, S. (1991). Using Kintch's computational model to improve instructional text: effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psychology*, 83, 329-345. <http://dx.doi.org/10.1037/0022-0663.83.3.329>
- Brown, G., & Yule, G. (1983). *Discourse Analysis*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511805226>
- Connor, U. (1996). *Contrastive Rhetoric: Cross-Cultural Aspects of Second Language Writing*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139524599>
- Crowhurst, M. (1990). Reading/writing relationships: An intervention study. *Canadian Journal of Education*, 15, 155-172. <http://dx.doi.org/10.2307/1495373>
- Fahnestok, J. (1983). Semantic & lexical coherence. *College Composition and Communication*, 34(4), 400-416. <http://dx.doi.org/10.2307/357897>
- Grabe, W., & Kaplan, R. B. (1996). *Theory and practice of writing: An applied linguistics perspective*. London: Longman.
- Halliday, M. (1994). *An Introduction to Functional Grammar* (2nd ed.). London: Arnold.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Hinds, J. (1987). Reader vs. writer responsibility: a new typology. In U. Connor & R.B. Kaplan (Eds.), *Writing Across Languages: Analtsis of L2 texts* (pp.141-152). Wokingham, England and Reading, MA: Addison-Wesley.
- Hyland, K. (2005). *Metadiscourse*. London: Continuum.
- Ostler, S. (1987). English in parallel: a comparison of English & Arabic prose. In U. Connor & R. Kaplan, R. (Eds.), *Writing across languages: Analysis of L2 text* (pp. 169-185). Reading, M.A: Addison. Welsey Publishing Company.
- Sa'Adeddin, M. A. (1989). Text development and Arabic-English negative interference. *Applied Linguistics*, 10(1), 36-51. <http://dx.doi.org/10.1093/applin/10.1.36>
- Singer, M. (1990). *Psychology of Language, an introduction to sentence and discourse processes*. Hove, Sussex and Hillsdale, NJ: L. Erlbaum.
- Van Dijk, T. (1985). *Handbook of Discourse Analysis*. London & New York: Academic Press.
- Van Dijk, T., & Kintsh, W. (1983). *Strategies of Discourse Comprehension*. London and New York: Academic Press.
- Widdowson, H. (1978). *Teaching Language as Communication*. London: Oxford University Press.

## Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).