



A Corpus-Based Description of Discourse Markers in Arabic Sport Journalistic Texts

Asem Ayed Al-Khawaldeh, Norsimah Mat Awal & Intan Safinaz Zainudin

School of Language Studies and Linguistics, Faculty of Social Sciences and Humanities, UKM

asimabujafar@yahoo.com

Article Info

Received: 04.10.2014

Accepted: 22.11.2014

Published online: 01.12.2014

ISSN: 2231-8968

Abstract

The aim of the present paper is to offer a corpus-based description of Arabic discourse markers in sport news journalistic discourse. To this end, a corpus of 80 articles (around 40,000 words) was compiled from two prominent Arab news websites: Aljazeera.net and Alarabia.net. Both qualitative as well as quantitative methods of research were employed to characterize this phenomenon, based on Fraser's model (2005). Four issues were addressed in the present analysis: identification, classification, frequency, syntactic classes, and position. The study identified a set of 73 DMs and classified them into four classes: elaborative, contrastive, inferential, and temporal. The study indicated that DMs play a crucial role in the organization and interpretation of the Arabic sport text. They are found to relate text units to each other through signaling explicitly the semantic relationships between them. Moreover, the DMs in the study are drawn from various grammatical forms such as conjunction, adverbs, preposition. With respect to their position, the DMs showed strong tendency to occur sentence-initially.

Key words: discourse markers, text, Arabic sport genre, corpus-based description.

Introduction

Since the 1970s, there has been an increasing interest in the study of DMs, resulting in a huge number of studies undertaken on this subject (Schiffrin 1987, Blakemore 1987, Fraser 1999, Aijmer 2002). Research on DMs has attempted to address different issues of discourse markers, particularly, definition, terminology, functions, classification, and features in different languages and genres from different perspectives and approaches (Schourup 1999, Schiffrin 2001, Muller 2005, Fraser 2005). Indeed, it is generally agreed that DMs play a crucial role in the organization and interpretation of text by binding its parts through indicating the relationships existing between them such as elaboration, contrast, and temporality.

Schiffrin (1987: 31) defines them as “sequentially dependent elements which bracket units of talk”. For Fraser (1999: 950), they are “a class of lexical expressions drawn primarily from the syntactic classes of conjunctions, adverbs, and prepositional phrases. With certain exceptions, they signal a relationship between the interpretations of the segment they introduce, S2, and the prior segment, S1”. S2 and S1 refer to the current segment of discourse and the previous one, respectively.

In the literature on DMs, a set of features that are commonly attributed to DMs are extensively discussed. Three of them, Schourup (1999) points out, are the most central:

1-Connectivity: It refers to the function of linking two textual units to each other by marking the relationships between them.

2-Optionality: a DM is syntactically and semantically optional. That is to say, their absence or presence does not affect the grammaticality and semantic content of their host sentence.

3-Non-truth conditionality: DMs contribute nothing to the truth-condition of the proposition expressed by an utterance in which a DM occurs.

DMs derive their significance from two assumed roles. The first one is to connect text units to each other by indicating the relations between them and, thereby, contribute to discourse coherence (Halliday and Hasan 1976, De Beaugrande and Dressler. 1981, Schiffrin 1987 Redeker 1990). The second one is to guide the hearer or reader by constraining the number of the possible interpretations and saving effort and time for the addressee in the task of understanding what is intended by an utterance in context (Blakemore 2002, Müller 2005, Andersen 2001). Thus, DMs are crucial linguistic elements in text that require a particular attention from researchers to provide a clearer and more comprehensive picture in the light of modern linguistic approaches and methodologies. Therefore, the present study intends to shed some light on DMs in one Arabic journalistic discourse, namely, sport new.

Arabic discourse markers -

In the Arabic grammar, the class of DMs is referred to as *huruuf al-atf* "conjunctive particles". It contains a limited, rigid number of elements not more than 10 (these are *wa* "and", *fa* "then", *thumma* "then", *bal* "but", *lakin* "but", *aw* "or", *am* "or", *laa* "not", *amma* "or", *hatta* "even") (Omar et al 1994). Of all Arabic DMs, *wa* "and" is considered as the most frequently used one. It is mainly used to link words, phrases, sentences, and paragraphs (Al-Batal 1985, Khalil 1999, Ryding 2005). Therefore, It has received the major focus of studies on Arabic DMs at the expense of other DMs.

Most of the studies on Arabic DMs are heavily based on the classical Arabic (the language of the Holy Quran and poetry) and ancient Arabic grammarians' views meanwhile very few studies, described as modern treatment, are based on today's written Arabic (known as Modern Standard Arabic, MSA henceforth) and the western scholars' views on this phenomenon. MSA is considered as a modern variety of the classical Arabic developed over the last two centuries due to the

influence of the western culture on Arabic, it has its own features that sets it apart from the classical Arabic albeit the existence of some shared similarities.

According to modern Arabic research (e.g. Al-Batal 1985, Mehamsadji 1988, Badawi et al 2004, Kammansejo (2005) Hussein 2010, Al-Kohlani 2010) which is highly influenced by the western linguistic views on DMs (e.g. Halliday and Hasan 1967), Schiffrin 1987, Blakemore 1987, Fraser 1999), Arabic grammar deals with DMs within the sentence boundary focusing on their semantic and syntactic aspects meanwhile their role and function within a text (above sentence level) are absent.

However, this modern research has revealed a number of expressions (other than the traditional set) that can be accepted as DMs due to the fact that they play a crucial role in organization and interpretation of text by linking its parts through indicating the semantic relations between them. Furthermore, new features of the traditional members (not mentioned in books of Arabic grammar) have been observed when looking at them from the textual perspective. That is, a good number of DMs used in the modern written Arabic, particularly in newspapers, do not exist in the classical Arabic. Some of them entered lately the domain of DMs due to the so-called grammaticalization process (see Brinton 1999) or through borrowing from western languages that influenced Arabic (Holes 1995, Ryding 2005, Al-Kohlani 2010)

MSA has a wide variety of DMs that differs in forms, distribution, function, and classification. It is inarguable that the modern set of DMs is far larger than the traditional class. Following Halliday and Hasan (1976), many contemporary studies (Al-Shurafa 1994, Mehamsadji 1988, Hamed 2014) divide DMs mainly into four classes: additive, causal, temporal, and adversative.

More importantly, as will be discussed in detail below, a number of studies have been undertaken on this phenomenon in Arabic whether in written or spoken Arabic. However, very limited research examines DMs in the journalistic written discourse, particularly, sport news. None of the reviewed studies deal with Arabic DMs in sport genre while this genre has wide readership and popularity in the contemporary society and culture, particularly, in the Arab world. It also has its own language and style that attract the attention of many scholars from various perspectives such as pragmatics, semantics, discourse analysis, and stylistics (Ghadessy, 1988; Romaine 1994; Khalil 2000; Abdulrazaq 2011).

Therefore, we have decided to conduct a study to offer an adequate description of DMs used in the Arabic sport genre. Sport news is deemed as one of the most effective and popular topics that achieve a wide readership all over the world, particularly, in the Arab society. We intend also to relate our findings to the previous studies to find out what this genre shares or differs from other genres in terms of DMs. We will also compare our results with the description of DMs given in MSA grammar books since we expect that there are some DMs or features not referred to in these books.

The significance of our description springs from the fact that it is corpus-driven which helps give an accurate picture of the phenomenon under consideration. Corpus-based studies yield reliable and authentic results as what is common or uncommon in the natural context of language (Cheng 2012). This is not to be claimed that the present study to give a thorough or detailed study on this phenomenon but it is just a stepping stone to focus on underexplored areas of research in Arabic linguistic literature.

Literature on Arabic DMs:

Based on Halliday and Hasan (1976), Al-Batal (1985) examines the cohesive role of DMs within a single modern Arabic expository text. Around 35 DMs are found in his research, indicating various semantic relations between sentences or clauses such as *thumma* "then" indicating chronological relation, *idha* "if" indicating conditional relation, *gayra anna* indicating adversative relation. Of all DMs, *wa* "and", he argues, seems to assume special significance due to its high frequency in the text. He points out that there are some connectives like *kadhalika* "likewise" and *min thamma* "hence, therefore" are not dealt with in books of Arabic grammar. In terms of syntactic types, he classifies these DMs into coordinative conjunctions (e.g. *wa*), subordinative (e.g. *idha*) adverb (e.g. *kadhalika*), prepositional phrase (e.g. *min thamma*)..

Ryding (2005) points out that DMs are “a pervasive feature of MSA”, resulting in a high degree of textual cohesion in Arabic texts (2005: 407). Like Mehamsadji (1988) and Al-Batal (1985), she maintains that most Arabic sentences within a text start with a DM that links each sentence to the previous ones. The DMs listed in her study come from different syntactic forms such as conjunctions, particles, adverbs, and phrases. She lists a wide variety of DMs and their functions, to mention just a few, contrastive DMs (e.g. *bal*), similarity DMs (e.g. *kama*), addition (e.g. *kadhalika*), causal (e.g. *fa*), temporal (e.g. *bainama*), and topic shift (e.g. *amma*).

Employing Fraser’s model, Hamzah (2010) examines the use of DMs in the Quranic texts. The study highlights the role of DMs in creating texture and examines the various meanings of some DMs such as *wa* and *fa*. For *wa*, it is used to open the speech, to introduce topic, to mark topic-change and to link topics together. He concludes that his analysis “lends support to Fraser’s study in matters related to the role played by DMs” in discourse (ibid,2010: 13).

In her pioneering work, Al-Kohlani (2010) investigates the functions of DMs in opinion articles such as additive, contrastive, inferential, explanatory, and sequential. To define items as DMs, she relies only on two criteria: initiality (to occur at the initial position of the host sentence and Non-truth-conditionality (to contribute nothing the propositional content of the host sentence). Following Brinton (1996), DMs found in her study are classified into two broad classes: textual DMs (e.g. *wa* "and", *lakin* "but") and interpersonal (e.g. *yabdu* "it seems", *rubbama* "maybe").

In editorial articles, Al-Jarrah (2011) touches on the role of DMs in marking and signaling the organization and relationships between ideas and clauses in text, employing eclectic analytical framework. Similar to Halliday and Hasan (1976) and Fraser (2005), he maintains that DMs are used in his data as cohesive devices to indicate the relationships between sentences or clause. His results reveal that the most frequently used DMs are contrastive.

Khalifa et al (2012) examines the coherence relations in Arabic texts in terms of implicit and explicit relations. Their study addresses the role of DMs in signaling explicitly the relations among parts of discourse. For them, the absence of DMs, makes the process of understanding the text more difficult. Their study identifies around 50 DMs that serve to indicate relations in Arabic text such as *bisabab*, *wa*, *ow*, *lakin*, *raghm*, without classifying them into categories.

Reviewing the relevant studies results in the following comments: i), they have not dealt with DMs in sport genre meanwhile three studies have been conducted on DMs in editorial articles, ii) some studies include items as DMs despite they operate at the sentence level not text-level i.e. they neither connect sentences or clauses nor indicate relations between them, iii) the influence of traditional treatment of DMs that looks at DMs as small, fixed set is still felt in the literature, iv) they do not relate their findings to the previous studies to point out similarities and differences, v) most of the studies base their analysis on a single text or very small data. With these shortcomings and gaps in mind, as discussed earlier, the present study intends to contribute to the literature on Arabic DMs through conducting a corpus-based analysis of DMs in Arabic sport genre and avoiding the shortcomings and gaps felt in the surveyed studies.

Theoretical Framework

The study employs Fraser's model (2005), which is intended to outline and refine Fraser's previous research on DMs (1996 and 1999). It proposes a well-defined and self-explanatory theoretical framework to study DMs in written language (Kurdi 2008, Michailinienè, 2007). That is reflected in his well-established definition and characterization of DMs that offer "a convenient and appealingly rational basis" to identify and describe DMs and to distinguish them from other expressions (Schourup 1999:239), as will be discussed below. Another justification for this choice is that Fraser's classificatory scheme of DMs is "presumably the most comprehensive classification in written discourse" (Dalili and Dastjerdi 2013: 2).

Now we will present in some detail the main issues of Fraser's model on DMs applied in the present analysis in terms of identification, function, classification, grammatical forms, and position.

First of all, we will start with Fraser's definition of DMs that reflects the general headlines of his proposal.

"the class of DMs is defined functionally as those lexical expressions which signal a relationship (which is one of Elaboration; Contrast; Inference; or Temporality) between adjacent messages, all are members of one of five syntactic categories: coordinate conjunction; subordinate conjunction; preposition; prepositional phrase; adverb" (2005:8)

The definition states that the primary function of DMs is to connect two textual segments by indicating the relationship between them. These relationships are one of the following: Elaboration; Contrast; Inference; or Temporality. To identify items as DMs, Fraser argues that a DM has to be connective as well as non-truth conditional.

In terms of classification, Fraser proposes a convenient and comprehensive classification, which is based on four major relations that DMs typically signal in written discourse: Elaborative markers (e.g. moreover, also), Contrastive markers (e.g. but, however), Inferential markers (e.g. therefore, because), and temporal markers (e.g. then, when).

As for syntactic classes of DMs, Fraser maintains that all DMs are drawn from one of the following five syntactic categories: coordinate conjunction; subordinate conjunction; preposition ; prepositional phrase, adverb.

The last issue is the position in which DMs may occur in their host sentence. He argues that DMs tend to occur sentence-initially. However, they may occur in medial and final position.

Methodology

Data Collection

The data needed for the purpose of the present study is based on a small-scale corpus collected from sport reports of two prominent Arabic news websites: Aljazeera and Alarabia. The reason for this choice is that they are the most read news-websites compared with other news websites in the Arab speaking countries. They are mainly the digital edition of the two biggest news satellite channels in the Arab world: Aljazeera channel and Alarabia channel. The present small corpus consists of 80 articles (approximately 40,000 words). The corpus was compiled over a period of two months: April and may 2014.

Data Analysis

In order to achieve the purpose of the present research, after compiling the needed corpus, the first step is to identify the items used in the sport texts as DMs. Two criteria will be relied on to distinguish DMs from other items. First, a DM has to bind two adjacent textual segments (sentences or clauses) through marking one of the four relations discussed above. Hence, DMs used to introduce paragraphs are out of the scope of the present study. This is because, besides space limitation of our study, Fraser's model is concerned with only local DMs that connect

successive segments while DMs linking paragraphs have not been treated. In addition to paragraph-DMs, items used to function within the sentence limit to bind, for example, nouns or verbs will be neglected. Second, a DM has to contribute nothing to the propositional content of their host sentence. Then, DMs will be classified in respect to the four major relationships discussed above.

Following that, a statistical process will be undertaken to show the frequency of DMs found in the corpus. Then, the syntactic status of the DMs will be described such as conjunctions, adverbs, prepositions, and prepositional phrases. The last step is to find out the position in which DMs tend to occur whether initially, medially, or finally.

Results and Discussion

As argued earlier, the present study aims at giving an adequate description of DMs used in the Arabic sport genre from Fraser's perspective (2005). Four major issues of this model will be addressed in the process of examining DMs found in the corpus under consideration: identification, classification, position, and grammatical forms.

As a challenge we faced in the process of analysis, some DMs are ambiguous and multifunctional that can indicate more than one relation. For example, *fa* (it will be discussed later) has been found in the data to indicate two types of relationships: elaborative and inferential. It is sometimes difficult to identify the intended function. Therefore, in this case, we will rely heavily on the context to arrive at the intended relationship. The most straightforward function will be chosen even if another function is possible in the given context.

As the first step, the identification of DMs will be based on Fraser's definition of DMs. The examination results in 52 different DMs in the corpus. As part of the study, a quantitative analysis is conducted to find out the top ten DMs (see table 1) and the frequency of each class, accompanied with the most frequent DMs within each class (see table 2). They are all put in descending frequency.

Table 1: Top Ten Discourse Markers

Arabic DMs	Transliteration	English translation	Frequency: out of 540 & percentage
و	<i>wa</i>	and	120 22,22
لكن	<i>lakin-lakinna</i>	But	47 8,70
بعد	<i>baad</i>	After	39 7,22
فيما	<i>fima</i>	While	29 5,37

قبل	<i>qabl</i>	Before	24	4,44
ل	<i>li</i>	To, for	20	3,70
كما	<i>kama</i>	Likewise	16	2,96
بسبب	<i>bisabab</i>	Because	15	2,77
بينما	<i>bainama</i>	while	14	2,59
رغم	<i>raghma</i>	Despite, although	14	2,59

Table 2: Classification of DMs and the frequency of the classes

Class	Most frequent DMs	Occurrences : out of 540	percentage
Elaborative	<i>wa, kama, khasatan-khususan, idhan, ay, haithu</i>	170	34 %
Contrastive	<i>fima, bianama, rghma, biada anna, lakin</i>	140	27 %
Inferential	<i>li, bisabab, mimma, liana, idha, ithra, hatta</i>	103	21%
Temporal	<i>baad, qabl ,indama, thumma, iqiba, hinama</i>	93	18%

Evidently, the quantitative analysis of DMs shows that the total frequency of DMs is around 540 times. This confirms that Arabic sport genre abounds with DMs used to bind its parts through signaling one of the four relations discussed earlier. This proves that the use of DMs is a crucial device that contributes to text production and interpretation. Table 1 lists the top DMs found in the corpus, constituting 64% of all DMs identified in the corpus. This is an explicit indication of the substantial reliance of Arabic sport texts on these ten DMs, particularly, *wa*, *lakin*, and *baad*.

In comparison with Arabic editorial genre, Al-Jubouri(1988) provides a relatively different list of the most frequent DMs in his data: *wa, fa, kama, idhan, lakin, li-anna, li, hatta, bal, aw*. Seemingly, *wa, kama, lakin, and li* are frequently used in both sport and editorial genre in the Arabic journalistic discourse.

In the present corpus, the most frequent DM is *wa* “and” with 120 occurrences or 21,40 % of total frequency of DMs. Out of these occurrences, *wa* occurs 17 times combined with other DMs, mostly, contrastive DMs, namely, *lakin* “but” and *raghma* “despite”. This high frequency of *wa* can be ascribed to the fact that the textual function of *wa* is to indicate “that the argument or

discussion is still ongoing with no major breaks” (Al-Batal 1999:246). Indeed, this result is not surprising since many studies on Arabic DMs confirm that *wa* is the most commonly used one in Arabic (Al-Kohlani 2010, Ryding 2005, Cantarino 1975, Al-Batal 1985).

The remaining DMs (see the appendix for the full list) achieved lower occurrences in the corpus, approximately less than ten times. For example, *jarra* "as a result" (once), *mundhu* "since" (twice), *thumma* "then" (three times), *idhan* "also" and *ay* "that is" (four times), *min ajli* "in order to" (six times), *idha* "if" (seven times),

Table 2 displays that DMs are classified into four main classes: elaborative, contrastive, inferential, temporal. It also display the frequency of each class in descending order, accompanied with the most frequent DMs for each class. The four classes will be individually discussed as follows:

1-Elaborative DMs (EDMs):

According to table 2, the largest category of DMs is elaborative with 170 occurrences, representing around 33% of the total DMs mass. In comparison with opinion article genre, Al-Jarrah (2011) and Al-Jubouri (1987) find that contrastive DMs constitute the majority of DMs used in their data. In our corpus, it is found that EDMs class contains 15 basic DMs such as *wa* “and”, *idhan* “also”, *bil-idhafati ila* "in addition to" *khasat-an* "particularly", *ay* "that is". Out of 170 occurrences, the predominantly used EDMs are *wa* (100 times) *kama* "likewise" (16 times), *khasatan* (10 times). What seems striking is that *wa* is detected to merely indicate the elaborative relationship in the present corpus despite the fact that *wa* can serve other relationships, particularly, contrastive and temporal (Fareh 1998, Yagi and Ali 2008, Ryding 2005, Al-Batal 1985). However, elaborative *wa* has been observed to co-occur with other DMs such as *wa lakin* "and but" , and *wa baad* "and after". Following Al-Jarrah (2011), we maintain that the elaborative *wa* is superseded by the DMs co-occurring with it. For example, when the elaborative *wa* co-occurs with contrastive *lakkin* "but" , the contrastive sense of *lakkin* is stronger than the elaborative sense of *wa*. Therefore, in this case, the relationship between the successive sentences are contrastive rather than elaborative. It be can be said that elaborative *wa* is neutralized by the function of their co-occurring DM, particularly contrastive markers. The use of *wa* with *lakin/lakinna*, as Kammensjö (2005) suggests, seems to be a matter of style in Arabic writing. As far as the co-occurrence of DMs is concerned, unlike Al-Kohlani (2010) who finds in her study on editorial that *idhan* "also" "never occurs alone in the data" (289), *idhan* in our sport data has been detected to appear four times alone without co-occurring with other DMs.

As a challenge in our analysis, the DM *fa* has been observed to mark two semantic relationships in the present corpus: elaborative and inferential. It has appeared eight times: five as an inferential DM and four as an elaborative DM. Remarkably, the elaborative *fa* is prefixed to nouns, pronouns or the particle *qad*, meanwhile the inferential *fa* is prefixed to verbs in the simple form.

2- Contrastive discourse markers (CDMs):

The elaborative category is followed closely by the contrastive one. In general, a CDM indicates a kind of contrast between the two propositions it connects. This class occurred 140 times, representing around 27% of total DMs occurrence. Identical to editorial genre (Al-Kohlani 2010, Al- Jubouri 1986), *lakin* "but" occupies the top rank within this class. Then, it is followed by *fima* "while" (29 times), *bainama* "meanwhile" (14 times) *raghma* "although" (14 times), and *illa anna* "but" (10 times). The rest of CDMs found in the corpus (such as *bal* "but") *ghayra anna* "however" *fi al-waqt-i nafsi-hi* "at the same time" and *fi hiin* "meanwhile" (4 times) show low frequency. Within this class, two DMs *fi al-waqt-i nafsi-hi* and *fi hiin* constituted a surprising challenge. Although they may suggest a temporal relation of simultaneity between two segments, this temporal sense, however, is overridden by the contrastive meaning of those DMs. This is because our policy in the present study is to choose the predominate function even though other functions are possible in the same context. Despite the fact that it is of wide use in MSA, *bal* "but" exhibits very low frequency (4 times) in the corpus under consideration (Hussein: 2009).

3- Inferential discourse markers (IDMs):

IDMs class represents 20% of the total occurrences of DMs, containing 24 different members. The most frequently used IDMs are *li* "in order to", *bisabab* "because", *mimma* "as a result", *idha* "if". Strikingly, these frequent DMs are not found in editorial texts (Al-Kohlani: 2010). Comparatively speaking, Al-Batal (1984) finds that the top IDMs used in his literary data are *fa* and *li-anna* while they show low occurrence in our corpus. Meanwhile, in Al-Kohlani's study on editorial genre (2010), *li-dhalika* "therefore" is the most widely used DM to mark inferential relation, it occurs only twice in our corpus.

Another worthwhile observation is that the IDM *li* in our corpus is found to prefix verbs, nouns, and pronouns. However, in Al-Kohlani (2010), *li* prefixing only demonstrative pronouns (e.g. *li-hadha*, *li-dha*, *li-dhalik*) is observed.

4- Temporal Discourse Markers (TDMs)

The last category is TDMs, occupying 18% of the total occurrence of DMs in the corpus. Its members are used to indicate a sequential relation between two propositions. The most frequently used TDMs are *baad* "after", with 37 occurrences, filling the highest rank within this category. Then, there are *qabil* "before" *indama* "when" and *thumma* "then". Other TDMs within this class show very low frequency such as *mundhu* "since" (once), *hiina* "when" (once). This wide use of *baad* can be stylistically justified that sport news tend to initiate with the most recent information then move back to narrate the information that precedes it by using the DM *baad*.

While many different TDMs (around 10) are found in the sport corpus, Al-Kohlani (2010) lists only one TDM in her editorial corpus, namely, *thumma* "then". This may indicate that sport genre employs a wider variety of items to mark temporal relationship.

Having identified and classified the DMs found in the present corpus, it is remarkably observed that a set of the DMs identified in the present corpus has not been listed or referred to in the reviewed literature, particularly, grammar books as well as academic theses and papers dealing with DMs in the MSA (Haywood and Nahmad: 1979, Al-Jubouri 1987, Omar et al 1994, Khalil 1999, Alkhuli 1999, Al-Nuqrad: 2003, Badawi et al 2004, Holes: 2004, Al-Makhzumi 2005, Al-Hadid:2005, Ryding 2005, Abu Chakra: 2007, Al-Kohlani 2010,). Among these DMs are *raghma*, *baida anna*, *khassatan*, *ithra*, *jarra*, *bi-fadhl*, *bisabib*, *aqiba*, *nadharan*, *mimma*, *ala-ghirar*, *ala-iitibar*, *fi-hadha al-siyah*, *ala an yakuun*, *fi isharat-in ila*, *bi-shart*.

These DMs should be taken into account in any future attempt to build an exhaustive inventory of DMs in MSA or to examine DMs in other different genres (e.g. academic, scientific, religious discourse). Furthermore, this set of DMs might be useful in teaching MSA due to the fact that it derives its results from representative and authentic corpus of MSA and it shows what is common or in use in language. Hopefully, Our findings contribute to update MSA grammar books and textbooks in respect with DMs.

Grammatical forms of DMs:

The second issue to be addressed here is to describe the grammatical forms of the DMs found in the present corpus. Fraser (2005) argues that all members of DMs come from one of the following five syntactic categories: coordinate conjunction; subordinate conjunction; preposition; prepositional phrase; adverb. The analysis found that Fraser's categories are to large extent fitting to describe the syntactic forms of our DMs. Yet, several DMs are derived from the category of noun, which is not specified in Fraser's account. Therefore, we had to modify Fraser's categories by adding the noun category to become six categories rather than five. Implicitly, this indicates that the English data used in Fraser's (2005) do not contain DMs drawn from the syntactic class of noun. It indicates also that this gap in this model should be taken into account when applying this model on other different Arabic genres.

Table 3 below displays the grammatical forms of DMs found in the sport corpus under examination. It shows a large resemblance to the findings of other studies on Arabic DMs (Alsaif 2012, Alkohlani 2010, Aljubouri 1986,). This confirms by no means the feature of the multi-categoriality that the current literature attribute to DMs (Fraser 1999,2005, Brinton 1996, Schourup 1999, Schiffirin 1987).

Table 3: Syntactic categories of DMs in Arabic sport reports

Syntactic Category	Examples
Coordinating conjunctions	<i>wa</i> "and", <i>thumma</i> "then", <i>lakin-lakinna</i> "but".
Subordinating conjunction	<i>bainma</i> "while", <i>adha</i> "if", <i>lianna</i> "because",
Adverbs	<i>aydhan</i> "also", <i>hiina</i> "when", <i>haithu</i> "where"
Preposition	<i>li</i> "in order to, for"
Prepositional phrase	<i>li-ajl</i> "in order to", <i>bittali</i> "consequently", <i>bi-sabab</i> "because",
Noun	<i>jarraa</i> "because", <i>bughyata</i> "for", <i>raghm</i> "despite", <i>baida</i> "but".

Position of DMs:

The last issue to address here is the position of DMs identified in the present examination. The analysis yields that the DMs have strong tendency to occur at the initial position in their host sentences. The only exception was *idhan* "also", it appeared to occur initially and medially. This goes with what is confirmed as a general property of these expressions in the literature that DMs tend to appear sentence-initially. The reason beyond this tendency is give "wide scope over the whole sentence or paragraph, thereby allow them to influence and guide hearer reader interpretation of everything that follows" (Al-Kohlani (2010:47).

Expectedly, no final-occurrence of DMs is observed in our corpus. This means that Arabic DMs prefer to occur initially but, in very limited cases, they may occur medially (only *idhan* has been found to do so in our data) but never occur finally. This refutes manifestly what Mahmood (2008:18) claims that "Arabic conjunctive particles are restricted to medial position only".

Conclusion

In conclusion, this study aimed at offering an adequate corpus-based description of the DMs used in the Arabic sport genre from the perspective of Fraser (2005). To this end, four issues were

addressed in the present analysis: identification, classification, frequency, syntactic classes, and position.

As the results indicate, Arabic sport genre abounds with a wide variety of DMs used to bind its textual parts by signaling explicitly the semantic relationships between them. As the initial step, around 73 DMs were identified, then classified into four major categories: elaborative, contrastive, inferential, and temporal.

As part of the analysis, a statistical small-scale investigation showed that the elaborative class was the most frequently used one, followed by contrastive, inferential, and temporal, respectively. Of all the DMs, the elaborative *wa* occupied the top rank in frequency. A table of the top ten DMs in the corpus was given, accompanied with frequency of each DMs. One of the important contributions is that the study discussed a number of DMs which have not been listed or touched on by the reviewed literature on DMs. This means that our findings are useful for other studies interested in the area of Arabic DMs as well as MSA grammar books and textbooks. In a systematic way, we provide an easy reference to the all DMs identified in the present corpus, classified according to their functions (see the appendix).

The DMs identified in the present study exhibit various grammatical forms: conjunction, adverbs, preposition, prepositional, and noun. The latter category has been added to the syntactic categories that Fraser proposes since "noun" was absent in Fraser (2005). The study also highlights that with the exception of *idhan*, all DMs occurred sentence-initially.

The study recommends that further research is needed to be done on Arabic DMs. For example, to compare and contrast DMs used in the Arabic sport genre with their counterparts in English or other languages or with other genres within MSA. Moreover, examining DMs in other underexplored genres such as literary, religious, academic discourse will be fruitful and useful. This enriches the literature on Arabic DMs with more insights, thereby, a clearer and more thorough picture of DMs in MSA will be arrived at. Hopefully, the findings of our analysis constitute a stepping stone to achieve not only a corpus-based grammar of MSA DMs but other linguistic phenomena in MSA. This is because a corpus-based description is believed to give a real and precise scene of what is common or not in the language in question.

References :

- Abdul Razak, Zainur. (2011). *Modern Media Arabic: A Study of Word Frequency in World Affairs and Sport Sections in Arabic Newspapers*. (Unpublished doctoral dissertation). University of Birmingham, England.
- Abu Chacr, Faruk. (2007). *Arabic an Essential Grammar*. London and new York: Routledge.
- Al-Batal, Mahmoud. (1985). *The Cohesive Role of Connectives in Modern Expository Arabic Text*. (Unpublished doctoral dissertation). the university of Michigan, United States.
- Al-Hadid, Mahmud. (2004). *Mudhakirat Fi Adawat Al-Rrabtwaalwasil Fi Al-lughati Al-Arabiyyah*. Aljamia Alarabiyya Almaftuhah.
- Al-Jarrah, Mohammad. (2011). Arabic Newspaper Discourse: Rhetorical Features, Discourse Markers, Strategies and Organization. *Damascus University Journal*. 27. 3. 150-1772.
- Al-Jubouri, Adnan. (1987). *Computer-Aided Categorization and Quantification of Connectives in English and Arabic*. (Unpublished doctoral dissertation). University of Aston, United Kingdom.
- Al-Judiaa, Yusif. (2007). *Alminhaj Almukhtasir fi Alnahwi wa Alssarf*. Leeds: Muassasit Alrayyan.
- Al-Kohlani, Fatima. (2010). *The Function of Discourse Markers in Arabic Newspaper Opinion Articles*. (Unpublished doctoral dissertation).Georgetown University, United States.
- Al-Khuli, Muhammad. (1999). *Comparative linguistics: English and Arabic*. Amman: Alfalalh house.
- Al-Makhzumi, Mahdi. (2005). *Fi Al-Nahi Al-Arabi: Tawjih wa Naqad*. Baghdad: Da-irit al Shuauun Althaqafyia.
- Al-Nuqrad, Abdullah. (2003). *A-Shamil Fi al-lughati Al-Arabiya*. Benghazi: Dar Alkutub Al-Wataniyah.
- Alsaif, Amal. (2012). *Human and Automatic Annotation of Discourse Relations for Arabic*. (Unpublished doctoral dissertation). The University of Leeds, England.
- Al-Shurafa, Nuha. (1994). Text linguistics and Cohesion in Written Arabic. *Art and humanties*,7,17-33.
- Aijmer, Karin (2002). *English Discourse Particles. Evidence From A Corpus Studies in Corpus Linguistics*. Amsterdam: John Benjamins.
- Andersen, Gisle. (2001). *Pragmatic markers and sociolinguistic variation: A relevance theoretic approach to the language of adolescents*. Amsterdam/Philadelphia: John Benjamin Publishing Company.
- Badawi, Elsaid, Adrian Gully, Mike G. Carter. (2004). *Modern Written Arabic: A Comprehensive Grammar*. New York : Routledge.
- Blakemore, Diane. (2002). *Relevance and Linguistic Meaning. The Semantics and Pragmatics of Discourse Markers*. Cambridge: Cambridge University Press.
- Blakemore, Diane. (1987). *Semantic Constrains on Relevance*. Oxford: Blackwell.

- Brinton, L. (1996). *Pragmatic Markers in English. Grammaticalization and Discourse Functions*. Berlin: Mouton.
- Brinton, Laurel J. (1990). The development of discourse markers in English. In J. Fisiak (Ed.), *Historical Linguistics and Philology* (pp. 45–71). Berlin: Mouton de Gruyter.
- Chenge, W. (2012). *Exploring Corpus Linguistics: Language in Action*. New York: Routledge.
- Cantarino, V. (1975). *Syntax of Modern Arabic Prose*. Indiana: Indiana University Press.
- Dalili, Mehdi and Dastjerdi, Vahid. (2013). A Contrastive Corpus-Based Analysis of the Frequency of Discourse Markers in NE and NNE Media Discourse. *Corpus Linguistics and Ling. Theory*, 9, 1, 39-69.
- De Beaugrande, Robert and Wolfgang Dressler. (1981). *Introduction to text linguistics*. London: Longman.
- Fareh, Shehdeh. (1998). The functions of AND and WA in English and Arabic written discourse university. *Papers of studies in contrastive linguistics*, 434, 303-312.
- Fraser, B. (1999). What are Discourse Markers. *Journal of Pragmatics*, 31, 931-952.
- Fraser, B. (2005). Towards a Theory of Discourse Markers. In *Approaches to Discourse Particles* edited by K. Fischer. Elsevier Press.
- Ghadessy, Mohsen. (1988). The Language of Written Sports Commentary: Soccer – a Description, PP.17-51.” In Ghadessy, Mohsen. (ed.) *Registers of Written English. Situational Factors and Linguistic Features*. London: Pinter Publishers.
- Halliday, M. A. K. and R. Hasan. (1976). *Cohesion in English*. London: Longman.
- Haywood, J. A and Nahmad, H.M. (1979). *A new Arabic grammar of the written language*. London: Lund Humphries.
- Hamza, Mohammed. (2010). Discourse Markers in Written Arabic. *Teachers college researchers journal*, 22, 4, 44-60.
- Holes, Clive. (2004). *Modern Arabic: structures, functions, and varieties*. Georgetown university press.
- Hussein, Miri. (2009). *Relevance Theory and Procedural Meaning: the semantics and pragmatics of discourse markers in English and Arabic*. (unpublished doctoral thesis). Newcastle University.
- Kammensjö, Helene. (2005). *Discourse Connectives in Arabic Lecturing Monologue*. Goteborg: Acta universitatis Gothoburgensis Göteborg
- Khalifa, Iraky, Zakareya, Al Feki, and Abdelfatah, Farawila. (2012). A Comprehensive Taxonomy of Arabic Discourse Coherence Relations. ICCIT.
- Khalil, Esam N. (2000). *Grounding in English and Arabic news discourse*. Amsterdam: John Benjamins Publishing Co.
- Khalil, Aziz. (1999). *A contrastive grammar of English and Arabic*. Jordan Book Centre.

- Kurdi, Heeven. (2008). The use of discourse markers by Syrian Arabic learners of English. Unpublished disseration thesis.
- Mahmood, lubna. (2008). A contrastive study of conjuncts in English and Arabic.
- Mehamsadji, Mokhtar. (1988). Cohesion and text development in written Arabic.
- Michailinienė, Viktorija. (2007). The translation of discourse markers from English into Lithuanian. Unpublished M.A thesis. *Šiauliai University*: Lithuania.
- Müller, Simone. (2005). *Discourse markers in native and non-native English discourse*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Omar, ahmadmukhtar, zahran, mistafa, abdlatif mihammed. (1994). al-nahu al-asasi. Dhat al-salasil: Kuwait.
- Redeker, Gisela (1990). "Ideational and pragmatic markers of discourse structure". *Journal of Pragmatics*, 14, 367–381.
- Romaine, Suzanne. (1994). On the creation and expansion of registers: sport reporting in TokPisin. Edited by. Douglas Biber. Edward Finegan, New York Oxford. Oxford university press.
- Ryding, Karin C. (2005). *A Reference grammar of modern standard Arabic*. Cambridge : Cambridge University Press.
- Schiffrin, Deborah. (2001). Discourse markers: language , meaning, and context. In the handbook of discourse analysis. Edit: schiffrin, D, Tannen, D, and Hamilton, H. Blakwell publishers.
- Schiffrin, D. (1987). *Discourse markers*. Cambridge, New York & Melbourne: Cambridge University Press.
- Schourup, Lawrence (1999). "Tutorial overview: Discourse markers". *Lingua*, 107, 227–265.
- Yagi, Sane M. and Ali, Mohamed Yunis (2008). "Arabic conjunction wa: a conflict in pragmatic principles" in *Poznań Studies in Contemporary Linguistics* Vol. 44, No. 4, pg. 617-627.