

A Statistical Study of Arabic Discourse Connectors in a Diachronic Bespoke Corpus of the Years 1950 and 2018

Sarah Aijan Alajlan

*School of Languages, Cultures and Societies,
University of Leeds, UK
College of Languages and Translation,
King Saud University, KSA*

salajlan@ksu.edu.sa; mlsala@leeds.ac.uk

Abstract

This paper discusses changes of Arabic discourse connectors in newspaper writing by comparing two distinct time periods: 1950 and 2018. It attempts to provide an answer to the question: How has the usage of Arabic discourse connectors changed, quantitatively and qualitatively, in Arabic newspapers as evidenced in the 1950 and 2018 sub-corpora? A specialized bespoke corpus has been built specifically for this study, Leeds Bespoke Corpus of Arabic Newspaper Writings (LBCANW), that contains rare material from the year 1950 and recent material from the year 2018 (Alajlan, 2019). It is part of an ongoing PhD study that approaches the changes in the lexicon and syntax via 'lexis' (Sinclair, 1991). The research methodology includes: recording the frequency of occurrences; normalizing the frequencies to per million words; and calculating the percentage of change by using %DIFF value (Gabrielatos and Marchi, 2011). Discourse connectors are arranged in descending order according to %DIFF value. Discourse connectors that are found in one sub-corpora are placed in separate tables. The results show noticeable degrees of change in most of the 95 discourse connectors included in this study across the two sub-corpora 1950 and 2018. Finally, detailed linguistic discussion of the changes in five selected discourse connectors is included with concordance lines examples from the corpus using the statistical information obtained from this study.

Keywords: Arabic Newspaper Writing, Language Change, Syntax, Discourse Connectors, Diachronic Corpus, Type of Change, %DIFF Value, Effect Size.

1. INTRODUCTION

This paper discusses and analyses the changes that occur in a specific type of linguistic cohesive device known as discourse connectors. The LBCANW was compiled specifically for this study due to the lack of diachronic Arabic newspaper corpora in spite of the availability of few general Arabic corpora (cf. Almujaivel, 2020). In fact, few recent studies, such as AISenan (2016), have used Arabic newspapers as corpora for their investigations. Discourse connectors act as linking devices between phrases, clauses, and sentences. In this study, 95 Arabic discourse connectors are listed and statistically analyzed. An attempt will be made to link the statistical results that are found across the two sub-corpora of 1950 and 2018 with some literature on language change and corpus-based studies. Firstly, the notion of discourse connectors is defined. A review of some literature about relevant work on Arabic connectors is presented. Then, a brief historical overview of the changes that have happened to Arabic discourse connectors over the past 100 years is provided. After that, an explanation of the methodology that is used to produce the list of discourse connectors, the analytical work that was done on the list of Arabic connectors, and the most prominent results are presented. Finally, some relevant conclusion from the analysis is provided.

2. DISCOURSE CONNECTORS AS A LINGUISTIC AND STYLISTIC DEVICE

The terms 'discourse connectors' and 'discourse relations' are usually used in the field of discourse studies to refer to words or other devices that provide semantic links between clauses, sentences, and paragraphs. These inferential links can explicitly or implicitly indicate a wide number of

relationships such as causality, contrast and temporality. They make a text more coherent by making explicit the relationship between what is mentioned before the discourse connector and what is mentioned after it. Hovy (1993, p 344) defines 'coherence' (and, by extension, 'coherent') as follows: "a discourse is coherent if the hearer knows the communicative role of each portion of it; that is, if the hearer knows how the speaker intends each clause to relate to each clause."

In the past decades, several names have been given to discourse connectors. Recent terms used include 'discourse markers', 'linking words', 'discourse operators', and 'cue phrases'. Prior to this, Hobbs (1985, p. 2) mentions that other names were given to discourse markers such as 'rhetorical predicates' (Grims, 1975; Mann and Thompson, 1986), Fillmore (1974) called them 'sequiturity relations', and 'logical semantic connectives' (Crothers, 1979). Hobbs (1985) calls them 'coherence relations'. Kalajahi et. al. (2012, p. 1669) admit that 'researchers have somewhat failed to explicitly define for language users a complete definition and classification of the term'. In this study, the term 'discourse connectors' will be used to refer to semantic linking words since this is currently the most common name.

There have been many efforts to identify and define the set of discourse relationships. Hobbs (1985, p. 2), for example, proposes a theory of discourse interpretation for 'coherence relations' in his work that depends on 'the knowledge-based theory'. This theory suggests that 'the process of understanding a certain discourse involves using knowledge gained in the past to construct a theory of what is happening in the present' (Hobbs, 1985, p. 2). Knott and Sanders (1998) use two approaches to classify coherence relations, one drawing mainly on psycholinguistic experiments on Dutch-speaking subjects, and the other starting from a study of 'cue phrases' used to signal relations in English texts (Knott and Sanders, 1998, p. 135).

Although he admits that there might be some semantic overlap between them Hovy (1993, pp. 361-362), differentiates three 'top-level' relations: presentational (such as 'well'), interpersonal (such as 'frankly'), and semantic (such as 'consequently'). Hovy (1993, p. 343) explains the function of discourse connectors on the paragraph level. "Discourse structure is the matrix in which clauses are embedded and which, aided by cue words, permits or blocks implicit inferences." Hovy (1993) illustrates several aspects of computational and linguistic theories of discourse structure relations as used by text planners and sentence producers to make a text coherent. He focuses on their nature, number, and extension to associated tasks such as sentence planning and text formatting.

2.1 Relevant Work on Arabic Connectors

The rapid developments in corpus-based studies have paved the way for more studies in language change and attracted more researchers from other disciplines to explore linguistic changes using advanced tools. In recent years, there has been an increasing amount of literature on corpus-based language change studies (Alarfaj and Alsalamn, 2020; Baker, 2011; Bauer, 2002; Buerki, 2013; Gabrielatos, 2007; Hamilton et. al., 2016; Hilpert and Gries, 2009, and Sartori, 2019). However, few studies have focused on Arabic (Abdul Razak, 2011; Al-Hejin, 2012; Mahmoud, 2013). The studies conducted on Arabic language change, and discourse connectors, in the recent years can be described as very rare (cf. Brierley and El-Farahaty, 2019).

Up till now, the notion of discourse connectors in Arabic has not been thoroughly discussed. There have been very few attempts to study Arabic discourse markers. Some researchers studied particular Arabic connectors that are used in Arabic dialects such as Hussein (2009), who focuses on 10 Arabic connectors, five of them used in Modern Standard Arabic and the other half used in Syrian Arabic. Gaddafi (1990) analysis the function of some discourse connectors in spoken Libyan dialect. Mohamed (2000) compares the functions of Arabic and English discourse features including discourse connectors in various text types.

Other researchers have focused on one particular discourse connector such as Khalifa et al. (2011) who address the problem of discourse segmentation in Arabic. They propose a technique to segment Arabic discourse into complete sentences. Their technique is derived from the Arabic Rhetorical system by exploiting one crucial connector *و* as defined by Arabic linguists almost one

thousand years ago such as the linguist Abdul Qaḥer Al-Jarjani عبدالقاهر الجرجاني who addressed the issue of 'faṣl' and 'waṣl' in his book *Dalā'il al-ʿiʿjāz* دلائل الإعجاز (Hemeida, 1997). This approach categorizes the six known rhetorical types of و into two classes, segment and unsegment, known in Arabic as 'faṣl' and 'waṣl' meaning that و can 'separate' segments or 'connect' them. Segmentation places are decided according to the type of connector و. An Arabic discourse corpus was specifically developed for this experiment and collected from newspapers and books. Khalifa et al. (2011) achieved results with an accuracy of 97.95% for discourse segmentation.

Tackling the issue of the absence of a comprehensive list of Arabic discourse connectors, the work of Alsaif (2012) is the first attempt to annotate a corpus with discourse relations for Arabic, and the first corpus study to develop automatic models for the recognition of Arabic discourse relations and connectives. It is thus also the first attempt to draw up an explicit comprehensive list of Arabic discourse connectors and design an annotation tool that can identify them in a reliable way. The final inventory of Alsaif (2012) includes 107 potential Arabic discourse connectors comprising 91 basic connectors and 16 modified forms (Alsaif, 2012, p. 81).

2.2 Changes in Arabic Connectors over the Past 100 Years

Arabic writing style have witnessed many changes in terms of the lexicon and syntax over the past 100 years. The changes are so extensive that many recent western scholars have started differentiating the discourse connector system of older, Classical Arabic from that of the more modern standard version of the language, Modern Standard Arabic, the latter being "the modern descendant of Classical Arabic, unchanged in the essentials of its syntax but very much changed, and still changing in its vocabulary and phraseology" (Holes, 2004, p. 5).

Discourse connectors in particular have witnessed noticeable changes in terms of their function and frequency of occurrence in many forms of Arabic writing. Dickins et. al. addressed the issue of the difference in the system of connectives between Classical Arabic and Modern Standard Arabic. They highlight that "in some types of modern Arabic writing, the system of connectives is rather different from the Classical Arabic norm" (Dickins et. al., 2017, p. 179). Dickins et. al. (2017) divide Arabic discourse connectors into two types; (1) basic connectives: ف, و, and ثم; and (2) secondary connectives such as إذ, حيث, etc. However, in Classical Arabic, there are several ways to express the links between ideas other than explicit discourse connectors. Dickins et. al. (2017, p. 176) illustrate this when they discuss the functions of repetition of words, phrases and even roots in Arabic. They mention two other functions besides rhetorical purposes: "(i) to relate ideas, just like lexical variation in English; (ii) 'text-building' as is done also by connectives.

Arabic discourse connectors now to a significant extent mirror those of English. Dickins et.al., (2017, p. 180) commenting on the use of connectives in modern Arabic academic-oriented writing, describe their usage as "much more similar to those used in modern English argumentative writing than are the connectives used in comparable Classical Arabic writing." Furthermore, Dickins (2017) highlights the difference in the usage of discourse connectives between Arabic and English. In many genres, Arabic sentences are typically longer than English sentences, and sentences and clauses in Arabic are typically connected either by one of the three basic connectives, ف, و, and ثم, or by the use of one of the simple secondary connectives, such as إذ, حيث, etc. Dickins argues that Modern Standard Arabic generally makes greater use of subordinating connectors (subordinating conjunctions) than Classical Arabic, probably under the influence of English (and perhaps other European languages) via translation and imitation. (Dickins, 2017, p. 3). Holes (2004) provides a good reason why this might be the case – a shift in Arabic from simple, basic connectors (و and ف) to complex ones (of which و and ف may, in fact, also be a part).

Dickins (2020) states that there is evidence of recent change in theme-rheme 'values' – i.e., the associated connotative meanings – of what he calls phrase-structural para-syntactic features (i.e. meaningful elements of word order which are not analyzable in terms of syntax) in Modern Standard Arabic over the past few decades, and that this affects the way in which Modern Standard Arabic connects sentences, as compared to the way in which Classical Arabic did. Abdelfattah (1990, pp. 76–77) demonstrates, for example, that in the Egyptian newspaper Al-Ahram, nominal

clauses/sentences became increasingly common over the period between 1935 and 1989 and that, correspondingly, verbal clauses/sentences became less common (Dickins, 2020, p. 86). This suggests that the discursual functions (connotative meanings) of initial non-verbal elements, in particular, whether these be nouns or adverbials have changed in this period. However, he admits that “no significant studies seem to have been conducted on changes in the realizational semantic ‘values’ [i.e., discursual functions / connotative meanings] of what are in phrase-structural para-syntactic terms identical structures between Classical Arabic and Modern Standard Arabic” (Dickins, 2020, p. 87). In his explanation of the nature of the changes between Classical Arabic and Modern Standard Arabic, Dickins (2020) illustrates that “there has not therefore been a change in the fundamental phrase-structural para-syntactic categories of Standard Arabic (between Classical Arabic and Modern Standard Arabic), but there has been a change in terms of which elements can fill which positions (‘slots’)” (Dickins 2020, p. 87) – and, as noted a change in the discursual functions / connotative meanings of at least some of the elements in these positions/slots. Thus, for example, preposed adverbials seem to have been extremely ‘marked’ in Classical Arabic – perhaps involving strong contrast with another element expressed adverbially in an adjacent sentence in the text, while in Modern Standard Arabic, a preposed adverbial is not necessarily contrastive (cf. Dickins 2020, p. 87).

Several factors may be responsible for the clear changes in contemporary Arabic writing. Holes (2004, p. 46) holds foreign languages partly responsible for the changes in Modern Standard Arabic, and in newspapers and the media in particular. He argues that “foreign lexical, phraseological, and even syntactic influence has been exerted on Modern Standard Arabic in recent years as a result of loan translation from European languages. More than anything, this is a consequence of the dominance of English, and to a lesser extent French, in the international media. Arabic newspapers in particular are full of rapidly, and often very literally, translated versions of press agency reports with what Joshua Blau calls “Standard European” phrase structure. The ad hoc calques thus formed found their way into everyday use.” On the other hand, Abdelfattah (1990) attributes the changes to the colloquial dialects and the effect of foreign languages more generally on Modern Standard Arabic. However, there are limitations to this influence. Badawi et. al., explain that, for example, “Although there are signs that the morphology of Arabic is moving away from the constraints of the CA [Classical Arabic] syllable structure, loan words are still accommodated into the pattern system as far as possible” (Badawi et. al., 2004. p.10).

Holes (2004) addresses the political style in Modern Standard Arabic and media. He illustrates that prepositional structures have become common in the coining of equivalents for complex noun phrases such as Organization of Petroleum Exporting countries (OPEC) منظمة الدول المصدرة للنفط. But even in simple possessive relationships where there is no adjective and where one might have expected to have a simple construct, there seem to be an increasing tendency to use ـِ, for example, كاتم للصوت. (Holes, 2004, p. 208). Furthermore, he gives examples of Arabic mirroring English structure such as من وإلى المدرسة ‘to and from school’, found in a BBC Arabic Service translation (Holes, 2004, p. 205).

3. DISCOURSE CONNECTORS: A STATISTICAL COMPARISON AND LINGUISTIC DISCUSSION OF THE ARABIC NEWSPAPER CORPORA OF 1950 AND 2018

In the next part of this study, a list of 95 Arabic discourse connectors is statistically analyzed and discussed from a linguistic point of view. The following section explains the derivation of the list of Arabic discourse connectors that are included in this study, the methodology of data collection, and the methodology of statistical analysis. It also presents the lists of the results with particular reference to the changes that have occurred and the possible reasons for these changes.

3.1 The derivation of the list of Arabic connectors in this section and methodology of data collection

The majority of the discourse connectors that are listed in this section are taken from Alsaif (2012), who, as noted, attempts to produce the first comprehensive list of Arabic discourse connectors.

The list of Alsaif (2012) consists of 107 connectors, divided into eight different categories according to their function. In this study, 95 discourse connectors are listed, after excluding some connectors from the analysis due to the inaccuracy in the number of occurrences in the corpus during the process of statistical-based data collection for this study. The discourse connectors included in this study are listed in tables 1, 2, and 3. Some discourse connectors are added to the list of this study by the researcher and marked as such in the list. More details about exclusions and inclusions are given in the following lines.

There were four main reasons for the exclusion of some discourse connectors that were mentioned in the list of Alsaif (2012). The first one is homography (two or more words having the same spelling), and polysemy (one word having more than one meaning). An example of homography, and also in effect, polysemy, is the connector *حال*, an adverbial annexion-head meaning 'at the time of'. This could be confused with the homographic verb *حال*, meaning 'to prevent', or with the noun *حال* meaning 'situation, condition' (although it should be noted that the adverbial, annexion-head *حال* is essentially the same word as the noun *حال*, the former being in the accusative, marking it as adverbial; *حال* 'at the time of' and *حال* 'state' are thus effectively a case of polysemy, with also a case difference and a difference between construct and 'absolute' state). Therefore, the word *حال* was excluded from this study due to the inaccuracy in the number of occurrences in the corpus which homography (and polysemy) would give rise to. However, the connector *في حال* (in the situation/condition) was included in the list for this study since it cannot be confused with the other meanings in the occurrences of the corpus of the study.

The second reason for excluding certain forms is the failure to identify the hamza during the process of searching for some connectors in the corpus. Because the hamza was not recognized by the software of Sketch Engine when searching in the corpus of this study, some confusion between certain words that begin with hamza and ones that do not, or which have a differently positioned hamza, indicating a different following vowel, was noticed (cf. Brierley et. al., 2016). The position of the hamza (whether above or below the alif which 'bears' it) or its absence, can make a difference to the Arabic words. This is illustrated by the connector *إِثْرٌ* (meaning 'after'), with the hamza below the alif. This was excluded because there are other words that have the same spelling, apart from the position of the initial hamza, but a different meaning, such as the noun *أَثْرٌ*, with the hamza above the initial alif, meaning 'effect' or 'mark'. Likewise, some frequent connectors containing hamza were excluded on the same basis, i.e., the program was not able to recognize the hamza in words that begin with it such as *إِن* and *أَنْ*, or in words that have it. Other words were excluded for the same basic reason, that the program cannot recognize hamza in the middle of words, such as *كَأَنَّ* (meaning 'as if') which would be confused with a different word *كَانَ* (meaning 'was'). However, connectors that occurred less often were included since the verification of the position of the hamza was easily made by reading the sentence in which it occurred, as in the cases of *إِلَّا إِنْ* and *أَلَا إِنْ*.

The third reason for excluding some connectors is that they consist of one letter. Thus, *ف* and *ب* were also excluded due to inaccurate counts resulting from wrong identification of the particles by the POS tagger in Sketch Engine. This gave some results in which the initial letter of a main word was misidentified as a connector.

Finally, some connectors are mentioned twice in the list of Alsaif (2012), due to their having more than one function. Thus, *حيث* could be a locative adverbial annexion-head meaning 'where' or a conjunction meaning 'since'. In this study, this connector occurs only once in the list of connectors and all of the occurrences, regardless of their function, are counted together and treated as one connector because the study corpus is not annotated semantically.

3.2 Methodology of the Statistical Analysis

The methodology followed in this study is an inductive approach that uses quantitative statistical data and qualitative linguistic discussion of the changes. Corpus data were collected following a sample corpus approach, sometimes called a 'snapshot corpus' approach. The Arabic corpus attempts to represent a particular type of language (newspaper writing) over a specific time span: 1950 and 2018 (McEnery and Hardie, 2012, p. 8). This approach is considered a useful way to

obtain a snapshot of a language when the purpose of a study is to investigate diachronic change, because it enables the researcher to make a wide range of comparisons and contrasts (McEney and Hardie, 2012, p. 9). For each discourse connector mentioned in the list for this study, two main occurrences in both sub-corpora, 1950 and 2018, were verified and recorded. In order to make an accurate comparison between the occurrences in both sub-corpora, the frequency of occurrences was converted to equivalents per million words in each sub-corpus. The conversion is automatically calculated in Sketch Engine when searching for any concordance. Calculating the frequency per million words is done to compare frequencies between corpora of different sizes (Alajlan, 2019). This step was crucial since the word count of the two sub-corpora differs. In order to ensure that the occurrences were accurate, manual scanning was done for all of the discourse connector occurrences to exclude occurrences of quotations from Quran and Hadith, since this category of material falls outside this study. Moreover, occurrences of foreign names that were inaccurately identified as Arabic words were excluded. For example, when entering the word لو ('if'), the two letters ل and و were identified as the connector لو ('if') in several occurrences of foreign names like لودريان 'Lauderian'. After that, the percentages changes in occurrences across the two sub-corpora were calculated using the following equation:

$$\frac{(\text{frequency of 2018} - \text{frequency of 1950})}{(\text{frequency of 1950})} \times 100$$

The above-mentioned equation is a simple and a common way to calculate the percentage of change of any given value. Mathematically, the percentage of increase and decrease are calculated by computing the difference between two values and dividing the result by the initial value (the initial value in this study is the 1950 data). This can tell us how much the initial value has changed. In addition, this equation is what Gabrielatos and Marchi (2011) call the %DIFF value, which is based on Effect size: "The % difference of the frequency of a word in the study corpus when compared to that in the reference corpus." Besides, %DIFF reveals not only differences but also similarities (Gabrielatos and Marchi, 2011).

After the completion of these three steps for every discourse connector (the collection of the frequencies, converting them into per million words, and calculating the percentage of change), the discourse connectors are presented and arranged in Table 1 starting from the highest increase and ending with the highest decrease. This means that those discourse connectors that showed less change are in the middle of the list. As for the discourse connectors that have newly emerged or disappeared, they are mentioned separately in Table 2 and Table 3. The following table, Table 1, lists the discourse connectors, their data before and after the analysis. The abbreviations used in the table are as follows:

RF=raw frequency

PM= per million words

	Arabic Discourse connector	English translation	2018		1950		%DIFF Value
			RF	PM	RF	PM	
1	بالإضافة إلى	in addition to	126	290.58	3	9.86	+2847%
2	في حال	in case of	29	66.88	1	3.29	+1933%
3	رغم أن	although	51	117.62	3	9.86	+1093%
4	على سبيل المثال	for example	17	39.21	2	6.57	+497%
5	خصوصا	specially	51	117.62	6	19.72	+496%
6	بالتالي	consequently	42	96.86	5	16.44	+489%
7	خلال	during	1,310	3,021.14	187	614.73	+391%
8	رغم	though	140	322.87	23	75.61	+327%
9	لذا	for this	32	73.8	6	19.72	+274%
10	فضلا عن	as well as	76	175.27	17	55.88	+214%
11	بينما	while	104	239.85	25	82.18	+192%
12	إلا أن	but	127	292.89	31	101.91	+187%
13	بالفعل	in deed	35	80.72	10	32.87	+146%
14	في أعقاب	after all	6	13.84	2	6.57	+111%
15	في هذه الأثناء	in the meantime	13	29.98	5	16.44	+82.4%
16	حيث	where/since	621	1,432.16	248	815.26	+75.7%
17	لهذا السبب	for this reason	5	11.53	2	6.57	+75.5%
18	جراء	because	25	57.66	10	32.87	+75.4%
19	عندما	when	146	336.71	71	233.4	+44.3%
20	لأجل	for	10	23.06	5	16.44	+40.3%
21	أيضا	also	250	576.55	129	424.06	+36%
22	بسبب	because of	128	295.2	70	230.11	+28.3%
23	على الرغم	although	55	126.84	31	101.91	+24.5%
24	نتيجة لـ	as a result of	26	59.96	15	49.31	+21.6%
25	في ظل	under	69	159.13	45	147.93	+7.57%
26	حين	when	134	309.03	88	289.28	+6.83%
27	خلافًا لـ	unlike	3	6.92	2	6.57	+5.33%
28	لكن	but	414	954.77	283	930.31	+2.63%
29	طالما	as long as	16	36.9	11	36.16	+2.05%
30	بيد	but	0	0	0	0	0%
31	أو	or	998	2,301.6	701	2,304.41	-0.12%
32	و	and	29,377	67,749.67	20,673	67,958.8	-0.31%
33	نظرا لـ	because of	32	73.8	23	75.61	-2.39%
34	منذ	since	280	645.74	205	673.9	-4.18%
35	بعدما	after that	15	34.59	11	36.16	-4.34%
36	كما	as	516	1,190.01	381	1,252.47	-4.99%
37	قبل	before	376	867.14	282	927.02	-6.46%
38	برغم	although	4	9.22	3	9.86	-6.49%
39	في حين	while	33	76.11	25	82.18	-7.39%
40	بل	but	174	401.28	137	450.36	-10.9%
41	عقب	shortly after	66	152.21	53	174.23	-12.6%
42	علاوة على	furthermore	6	13.84	5	16.44	-15.8%
43	نتيجة	a result of	81	186.8	69	226.83	-17.6%
44	حتى	even if	440	1,014.73	375	1,232.75	-17.7%
45	لان	because	214	493.53	186	611.44	-19.3%
46	وقبل	and before	17	39.21	15	49.31	-20.5%
47	لو	if	74	170.63	67	220.25	-22.5%

48	قبل أن	before	42	96.86	39	128.21	-24.5%
49	إلا	but	292	673.41	290	953.32	-29.4%
50	لاسيما	particularly	44	101.47	44	144.64	-29.8%
51	ختاما	finally	1	2.31	1	3.29	-29.8%
52	مع هذا*	although	7	16.14	7	23.01	-29.9%
53	بعد	after	717	1,653.56	736	2,419.47	-31.7%
54	إنما	but/it is indeed	58	133.76	62	203.81	-34.4%
55	بالرغم من	although	13	29.98	15	49.31	-39.2%
56	بفضل	thanks to	18	41.51	21	69.03	-39.9%
57	من ثم	then	21	48.43	26	85.47	-43.3%
58	كي	to	23	53.05	30	98.62	-46.2%
59	وفي الختام	in the end	3	6.92	4	13.15	-47.4%
60	قريب	shortly before	7	16.14	10	32.87	-50.9%
61	لكي	in order to	16	36.9	23	75.61	-51.2%
62	بغية	aiming at	2	4.61	3	9.86	-53.2%
63	مثلا	for example	16	36.9	24	78.9	-53.2%
64	كلما	whenever	11	25.37	17	55.88	-54.6%
65	بيد أن	but	4	9.22	7	23.01	-59.9%
66	أما	as for	131	302.11	232	762.66	-60.4%
67	في حالة*	in case of (feminine)	15	34.59	29	92.05	-62.4%
68	مع ذلك*	although	16	36.9	31	101.91	-63.8%
69	جدير بالذكر	worth mentioning	1	2.31	2	6.57	-64.8%
70	على العكس	by opposite	1	2.31	2	6.57	-64.8%
71	دليلا على	evidence for	4	9.22	8	26.3	-64.9%
72	بعد ذلك	after that	33	76.11	74	243.26	-68.7%
73	لولا	if not	7	16.14	17	55.88	-71.1%
74	ثم	then	201	463.55	525	1,725.84	-73.1%
75	إلا بعد	except after	9	20.76	24	78.9	-73.7%
76	خلاصة	to sum up	3	6.92	9	29.59	-76.6%
77	غير أن	however	16	36.9	53	174.23	-78.8%
78	ذلك أن	that because	19	43.82	64	210.39	-79.2%
79	فعلا	indeed	7	16.14	24	78.9	-79.5%
80	بحجة أن	because of	1	2.31	5	16.44	-85.9%
81	في الواقع	in fact/actually	1	2.31	10	32.87	-93%

*Not in the list of Alsaif (2012).

TABLE 1: List of arranged discourse connectors after the analysis.

Emergence of new items

The following table lists the discourse connectors that have newly emerged in the 2018 sub-corpus compared to the 1950 sub-corpus. The discourse connectors mentioned in this section are well-known in Arabic and are not considered new words or expressions. However, they have newly emerged in this study on the basis of the comparison of the 1950 with the 2018 sub-corpora. The list starts with the most frequent discourse connectors and ends with the least frequent items.

Arabic Discourse connector	English translation	2018		1950	
		RF	PM	RF	PM
1	إضافة إلى	88	202.95	0	0
2	بهدف	80	184.5	0	0
3	حينها	21	48.43	0	0
4	حتى لو	17	39.21	0	0
5	بالإضافة لـ	12	27.67	0	0

6	عموما	generally	10	23.06	0	0
7	في المقابل	in contrast	6	13.84	0	0
8	بمعنى آخر	in other words	3	6.92	0	0
9	باختصار	in sum	3	6.92	0	0
10	بالمقابل	in contrast	2	4.61	0	0
11	لئلا	for not	2	4.61	0	0
12	*على العموم	in general	1	2.31	0	0
13	بالأساس	basically	1	2.31	0	0

*Found in the form على وجه العموم.

TABLE 2: Emergence of new items of discourse connectors.

Disappearance of Old Items

The following table lists the discourse connectors that occurred in the 1950 sub-corpus but disappeared from the 2018 sub-corpus. This does not mean that they are no longer used in Arabic. In fact, given that there is only one such discourse connector, على النقيض, and only one occurrence of it in the 1950 sub-corpus (as compared to none in the 2018 sub-corpus), both the category and the example can be considered marginal.

Discourse connector	English translation	2018		1950	
		RF	PM	RF	PM
1	على النقيض	0	0	1	3.29

TABLE 3: Disappeared old items.

Of course, as is the case with any corpus-based study, it may well be that the zero occurrence of an item does not mean that this item was not used at all in some texts during the time period that is represented by the corpus; it just did not occur in the corpus of this study. However, zero occurrences give a precise indication of the nature of the language that was used at the time of writing in the newspapers included in corpus of this study. A possible way to corroborate whether a certain item existed or not, and to check the frequency of a certain word in general, is to search for its occurrence in an alternative larger corpus. Additional verification for whether an item was used in the past (before 1950) is used in the discussion of some of the examples that will be analyzed in this section by referring to the KSUCCA corpus, a 42-million corpus of Classical Arabic. The use of this corpus gives this study the advantage of verifying the occurrences of the words of interest at a time when contact with European foreign languages, such as English, was very rare. However, there was extensive contact between Arabic and other languages like Persian and Aramaic – and while Classical Arabic borrowed a certain number of words from these languages, the influence was generally the other way – from Arabic to these languages (Tabasi, 2014). KSUCCA is a compilation of authentic Arabic materials dating back from the seventh until early eleventh century CE (Alrabiah et. al., 2014, p. 29). However, the results of the KSUCCA are not considered in this study; they are used for illustration purposes only.

The verification of zero occurrences in the 2018 sub-corpus and checking frequencies in general is relatively easy because gaining access to current Arabic newspaper corpora is not difficult. For example, Timestamped Arabic corpus 2014-2019 is a 3,096,145,051-word corpus that can be used to verify if a certain item is used or not in current Arabic newspaper writing.

3.3 Analysis of Changes in Discourse Connectors

The investigation of the 95 discourse connectors reveals the following: 29 discourse connectors have increased, 49 have decreased, 2 were almost stable, these being و and أو meaning 'and' and 'or' respectively. The two connectors showed less than 0.3% change. 13 discourse connectors have newly emerged and 1 discourse connector disappeared in 2018 sub-corpus. The following section discusses five of the most noticeable changes in the list and the reasons for selecting them.

3.3.1 Discussion of Changes in Some Connectors

Due to limitations of time and space, the discussion will be limited to the most prominent linguistic changes, even though almost all of the changes that have taken place are worth discussing. Five discourse connectors were selected as main discussion points with some reference to other near-synonymous discourse connectors where needed. The first three discourse connectors in the discussion are the ones that showed the highest increase, while the fourth and the fifth examples showed a noticeable semantic shift in their usage besides the increase in their usage. The change in the discourse connectors in the decrease list showed a relatively smaller change; therefore, priority was given to a separate section for the connectors that showed the greatest change in the increase list. However, eight discourse connectors in the decrease list are mentioned in the discussion of the five selected discourse connectors since they are near-synonyms or semantically related to them. The next section (3.3.1.1) discusses the first discourse connector in the list, which is the one that increased the most.

3.3.1.1 'in addition to' بالإضافة إلى

Starting with the discourse connectors that increased, the discourse connector بالإضافة إلى (meaning 'in addition to') increased by 2,847% in the 2018 sub-corpus as compared to the 1950 sub-corpus. To verify if the expression بالإضافة إلى is a very frequent expression in Arabic in general or not, KSUCCA was referred to. In KSUCCA, 7 occurrences were found, i.e., 0.12 per million words, which is extremely low compared to 9.86 per million words in 1950 and 290.58 in 2018. Figure 1 below illustrates the number of occurrences in each corpus and their per-millions-word equivalents.

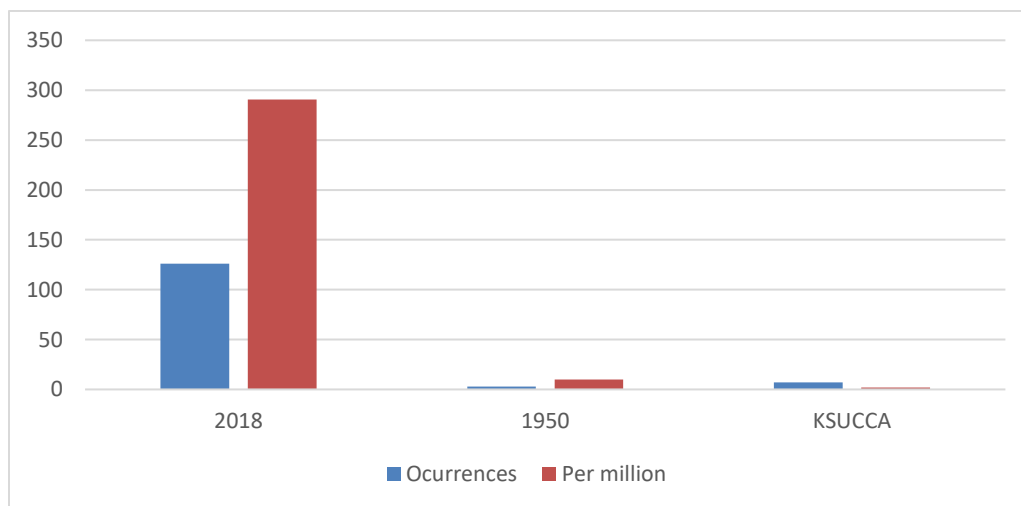


FIGURE 1: Comparison of occurrences of the discourse connector بالإضافة إلى.

As illustrated above, the use of the discourse connector بالإضافة إلى has sharply increased over time. There are other forms which are very similar to بالإضافة إلى, though listed in Alsaif (2012) as separate discourse connectors. In these, the preposition that follows إضافة may be changed, as in بالإضافة لـ or deleted, as in إضافة إلى. These two connectors have newly emerged in the lists of this study; they did not occur in the 1950 sub-corpus but were found in the 2018 one.

There are two other near-synonymous discourse connectors that have increased in use: فضلاً عن and أيضاً. Their occurrences increased by 214% and 36% respectively. For comparison, there are only two near-synonymous discourse connectors that have decreased in use between 1950 and 2018: علاوة على and كما – by 15.8% and 4.99% respectively.

This noteworthy increase in the use of بالإضافة إلى and the related بالإضافة لـ and إضافة إلى can perhaps be attributed to the effect of translation from English into Arabic. English uses many discourse connectors that express the meaning of addition such as 'in addition [to]', 'also', 'moreover',

'furthermore', etc. A possible reason for this increase is given by Badawi et. al. (2016, p. 193), when they the adverbial phrase عن غصباً 'in spite of': "In MWA, however, perhaps in response to the wide range of conjunctions and connectors in the Western languages, this construction has been greatly extended." Thus, the increase in the use of secondary connectors that express the meaning of addition definitely suggests the influence of English on the writing style in Arabic newspapers. The following examples that were taken from the corpus show the use of these connectors across the two sub-corpora for 1950 and 2018.

1950	2018
<ul style="list-style-type: none"> توفى الشيخ عبد الله باحمدين بسكتة قلبية عن عمر نيف على أربعين حولا قضى معظمها مجدا في أعماله التجارية وقد كان يرأس ثلاث شركات وطنية هي شركة مصحف مكة وشركة التوفير والاقتصاد وشركة الثلج فكان يديرها بالإضافة إلى إدارته للمقصف السعودي بجدة <p>(Umm Al-Qurā, national affairs)</p> <ul style="list-style-type: none"> Sheikh Abdullah Bahamdain died of a heart attack at the age of forty. He spent the majority of his life hard working in commercial business. He was heading three national companies, namely the Makkah Holy Qur'an Company, the Thrift and Economy Company, and the Ice Company, which he ran in addition to his management of the Saudi canteen in Jeddah. 	<ul style="list-style-type: none"> وأكد الملا أهمية التعجيل بالمشروعات الجارى تنفيذها في مجال البنية التحتية لنقل المنتجات البترولية مواكبة للزيادة السكانية وزيادة الاستهلاك، بالإضافة إلى مواصلة التوسعات الجارية في الشبكة القومية لنقل المنتجات البترولية عبر خطوط الأنابيب. <p>(Al-Ahram, business)</p> <ul style="list-style-type: none"> Al-Mulla stressed the importance of expediting the projects currently being implemented in the field of infrastructure for transporting petroleum products to keep pace with the population increase and increasing consumption, in addition to continuing the ongoing expansions in the national network for transporting petroleum products via pipelines.
<ul style="list-style-type: none"> وتقول آخر الأنباء أن أكثر من ربع مليون فدان من المراعي الجيدة قد احترقت عن آخرها بالإضافة إلى احتراق كميات كبيرة من العلف المخزون. <p>(Al-Ahram, international news)</p> <ul style="list-style-type: none"> And the latest news says that more than a quarter of a million acres of good pasture have been burnt down, in addition to burning large quantities of stored forage. 	<ul style="list-style-type: none"> وصلت أعداد المتدربين من هم في بعض الأحيان إلى 30 صحفيا ، هذا بالإضافة إلى صحفيين فرنسيين مبتدئين عادوا إلى بلادهم وتقلدوا مناصب صحفية هامة <p>(Al-Ahram, opinion articles)</p> <ul style="list-style-type: none"> The number of trainees reached sometimes 30 journalists, in addition to French novices who have returned to their country and worked in important press positions.
<ul style="list-style-type: none"> وبالإضافة إلى ما تقدم فسيورد بالمصنع الجديد قسم لتجديد محركات السيارات وهي عملية صادفت الكثير من النجاح كما أنها ضرورية جداً في بلد ك مصر نظرا لما تتعرض إليه السيارات من الإنهاك نتيجة لتأثير مناخ البلاد وأرضها. <p>(Umm Al-Qurā, business)</p> <ul style="list-style-type: none"> In addition to the foregoing, the new factory will have a section for renewing car engines, a process that encountered a lot of success and is very necessary in a country like Egypt due to the exhaustion of cars as a result of the impact of the country's climate and land. 	<ul style="list-style-type: none"> وقالت مصادر عسكرية: أن قوات الجيش الوطني صدت هجوما كبيرا للحوثيين على جبل اضيح , وقتلت ما لا يقل عن 40 مسلحا حوثيا بالإضافة إلى جرح العشرات. <p>(Alriyadh, international affairs)</p> <ul style="list-style-type: none"> Military sources said: that the national army forces repelled a major Houthi attack on the mountain of Adeeq, and killed at least 40 Houthi militants in addition to dozens wounded.

3.3.1.2 'in the event [of]'

The discourse connector that appears second in the list is في حال. This has increased by 1933% in the 2018 sub-corpus as compared to the 1950 sub-corpus. In its basic (literal) sense, the expression في حال means 'in the condition or situation [of]' and it is commonly used in this sense.

There are 29 occurrences of this discourse connector in the 2018 sub-corpus, 66.88 per million words, and one occurrence, 3.29 per million words, in the 1950 sub-corpus. Interestingly, the feminine form of this connector, which is not found in the list of Alsaif (2012) but is added to the list of this study, *في حالة* has decreased by 62.4% in the 2018 sub-corpus. It occurs 15 times in the 2018 sub-corpus, 34.59 per million words, and 28 times in the 1950 sub-corpus, 92.05 per million words. In Arabic, the feminine form *حالة* and the masculine form *حال* can be used interchangeably. Looking at the percentage increases and decreases in Table 1 above, it is clear that the masculine gender form has become more common than the feminine one in the Arabic newspapers analyzed. One reason for this, we can reasonably speculate, is the effect of the translation of the expression 'in case', which does not have a feminine form in English but is rather neuter in gender (like all words referring to inanimate objects/entities in English). Another possible justification is that English frequently uses a number of synonymous expressions for 'in case of' such as 'in the event that' or 'given the situation that', etc. that are also neuter in gender, a gender which is English can be regarded as neutral (since it is not marked for sex, as are masculine and feminine), and which is in a sense common (rather than specifically sex-marked). As a result, their translations commonly appear in Arabic translations as *في حال* (masculine gender).

Besides changes in the frequency of occurrence, an unusual syntactic change that concerns the use of the discourse connector *في حال* and the words that follow it was noticed. In Arabic grammar, the indefinite noun *حال* (annexion-head) should be followed by another noun (annex) that defines it. However, it is common in the 2018 examples for the noun *حال* to be followed by a verb in the past tense. This awkward grammatical structure, that can be called as 'an analogical grammatical development' of Arabic syntax in Modern Standard Arabic occurs in both newspapers Al-Ahram and Alriyadh in the 2018 sub-corpus. Dickins (2009, pp.137-144) has discussed a very similar phenomenon in the Sudanese Arabic dialect. It is grammatical to say in Arabic *في حالة الفوز* 'in case of winning', but it is not grammatical to say *في حال فاز* 'in case he wins'. It is not grammatical because the noun *حال* (annexion-head) should be followed by another noun (annex) that defines it but not a verb. However, no similar occurrence of a verb following the noun *حال* was found in the 1950 sub-corpus. One justification for this common grammatical mistake or 'change' is that the expression *في حال* is used in a way that its synonymous to the conditional word *إذا* (if) when translating from English without taking into consideration the grammatical constraints of these words in Arabic. The need to translate news texts and publish them in a very short time has probably caused this type of grammatical structure in Arabic writing style. Besides, this could also be the result of the influence of colloquial Arabic on Modern Standard Arabic. Below are examples from the corpus of the occurrences of the feminine and the masculine forms in addition to an illustration of the syntactic change of the word *حال* and the word that follows it.

1950	2018
<ul style="list-style-type: none"> • لقد اقترحت إيطاليا اتخاذ تدابير معينة لصون الأمن في الصومال في حالة الموافقة على أن تتولى السلطات الإيطالية الدفاع عن تلك البلاد (Al-Ahram, international affairs) • Italy has proposed that certain measures be taken to maintain security in Somalia in the event (feminine) that it is agreed (noun/definite) that the Italian authorities take charge of the defense of that country 	<ul style="list-style-type: none"> • ويتزامن ذلك مع مواجهة رئيس الوزراء الحالي ، نجيب عبد الرزاق ، من فضائح فساد واتهامات بتمرير مئات الملايين من الدولارات عبر حساباته المصرفية بطرق غير شرعية وفي حالفازت المعارضة بالانتخابات سيصبح مهاتير محمد أكبر زعماء العالم (Al-Ahram, opinion articles) • This coincides with the confrontation of the current Prime Minister, Najib Abdul Razzaq, with corruption scandals and accusations of passing hundreds of millions of dollars through his bank accounts illegally. In case (masculine/common gender) the opposition won (verb, past) the elections, Mahathir Muhammad will become one of the world's most important leaders.

<ul style="list-style-type: none"> • كل سائق يصيب أحداً بجروح أو رضوض لا يحصل معها كسر وأثبت التحقيق أن الإصابة نشأت من سوء تصرفه أو إهماله يعاقب بالسجن للمدة التي حددها التقرير الطبي لشفاء المصاب وفي حالة ما إذا لم يحدد التقرير العلمي مدة شفاء المصاب يظل المتهم سجيناً وتطالب الجهة المختصة بالتحديد على وجه التحقيق الموصل لتطبيق أحكام هذه المادة . (Umm Al-Qurā, national affairs) • Every driver who injures anyone with wounds or bruises that does not happen to have a fracture and the investigation has proven that the injury arose from his misconduct or neglect is punished with imprisonment for the period specified in the medical report for the recovery of the injured. In the event (feminine) that the scientific report does not specify the period of recovery of the injured, the accused remains imprisoned and demands the competent authority specifically to direct the investigation to implement the provisions of this article. 	<ul style="list-style-type: none"> • حذر سيرجي لافروف وزير الخارجية الروسي من أن الاتفاق النووي الإيراني لا يمكن أن يستمر إذا انسحبت الولايات المتحدة بشكل أحادي منه، وذلك بعد تهديد الرئيس الأمريكي دونالد ترامب بالانسحاب من الاتفاق في حالة عدم إصلاح ما وصفه بالثغرات في الاتفاق الذي توصلت إليه القوى الكبرى مع طهران في 2015. (Al-Ahram, international affairs) • Russian Foreign Minister Sergey Lavrov warned that the Iranian nuclear agreement cannot continue if the United States withdraws unilaterally from it, after US President Donald Trump threatened to withdraw from the agreement in case (feminine) he does not fix what he described as the 'gaps' in the agreement reached by the major powers with Tehran in 2015.
<ul style="list-style-type: none"> • وليس أبلغ في الدلالة على ذلك من أن المعركة قد خلت من أي حادث من حوادث التزوير، هذا ما لم يحدث من قبل الأمن الانتخابي والأمن العام. فسألته رأيه في حال الأمن، بعد ما رآه في هذه الرحلة (Al-Ahram, national affairs) • I am not aware of the evidence for this from the fact that the battle has been free of any incidents of forgery, as this has not happened before electoral security and public security. I asked him how is security (situation) (masculine/common gender), after what he saw on this trip? 	<ul style="list-style-type: none"> • وأكدت الوزارة على أهمية التعاون مع فروع وزارة البيئة والمياه والزراعة عند اعتماد الطرق التي تعبر الغابات أو المتنزهات التابعة لها في حال وجود إزالة للأشجار وفقاً للأنظمة والتعليمات المتبعة . (Alriyadh, national affairs) • The Ministry stressed the importance of cooperation with the branches of the Ministry of Environment, Water and Agriculture when approving roads that cross forests or their parks in the event of (masculine/common gender) deforestation in accordance with the regulations and instructions followed.

3.3.1.3 'although' رغم أن

The third discourse connector in Table 1 is **رغم أن** meaning 'although'. It increased by 1,093% in the 2018 sub-corpus as compared to the 1950 sub-corpus. It occurred 3 times in 1950 and 51 times in 2018, which if converted to their per-million-word equivalents equals 117.62 and 9.83 respectively. To verify the frequency of this discourse connector in Arabic, KSUCCA, the 42 million words corpus, was referred to. Here it occurred 3 times, which if converted to its per-million-word equivalents equals 0.03. In addition, the two other forms of the discourse connector in Table 1, **على الرغم** and **رغم**, have also shown an increase. These two forms, which differ only slightly from **رغم أن**, have increased in 2018, the former, dramatically, by 327% and the latter by 24.5%. However, another form of this discourse connector has clearly decreased in 2018, by 39.2%; this is **بالرغم من**. Finally, the discourse connector **برغم** was almost stable with a decrease of 6.49% in 2018. In general, it can be claimed that in the majority of case **رغم أن** and related forms showed an increase in their use in 2018 compared to 1950. Again, a possible justification for this increase in the use of discourse connectors expressing contrast is that English (also other European languages) that is the main language of source of news material nowadays, use many forms of discourse connectors expressing the ideas of contrast such as 'although', 'even though', 'in spite of', 'despite', etc. As a result, Arabic translations of news articles originally written in English contain many occurrences of

the translations of English discourse connectors such as *رغم أن*. Arabic can express a similar meaning using forms other than *رغم*, such as *مع هذا/مع ذلك* (literally 'with this', 'with that'), which were not mentioned in Alsaif's (2012) list. More traditionally, the coordinators *ولكن* and *ولكن* and even *و* can be used to express contrast in Arabic (cf. Dickins 2017). These two expressions which were added to the list for this study showed a decrease in their frequency of use. The former *مع ذلك* decreased by 63.8% in 2018 as compared to 1950, while the latter *مع هذا* decreased by 29.9%. The following figure compares the frequency of the occurrences of the discourse connector *رغم أن* and their per-million-word equivalents in the two respective corpora. Examples of their uses from the corpora are given afterwards.

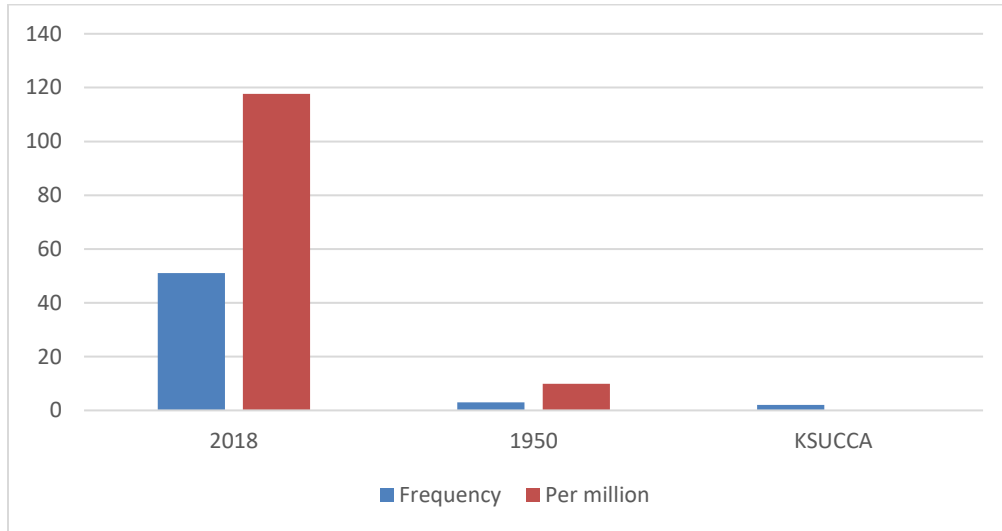


FIGURE 1: Comparison of occurrences of the discourse connector *رغم أن*

1950	2018
<ul style="list-style-type: none"> وهؤلاء الذين تنفقوا من معيها العذب الكثير الزحام قد بكوا يوم دخل الألمان باريس .. إذ تصوروا ، رغم أنهم مشغولون أيضا ببلادهم وما يهددها من اخطار تصوروا مدى ما يحز في نفوس أولئك القوم الشجعان من المرارة والأسى إذ تطأ أقدام المغيرين أرض الشانزلزيه <p>(Al-Ahram, opinion articles)</p> <p>And those who have been educated in it have cried the day the Germans entered Paris. They imagined, although they were also busy with their country and the dangers threatening it, the extent of bitterness and grief for those brave people as the raiders set foot on the land of the Champs Elysees.</p>	<ul style="list-style-type: none"> وكي لأفهم خطأ - رغم أن ذلك لا يشكل هاجسا - لست ضد هذه اللقاءات كفعالية ترفيهية للأخوة المتقاعدين ممن أمضوا سنوات طويلة تحت أسر المنهج <p>(Alriyadh, opinion articles)</p> <ul style="list-style-type: none"> In order not to be misunderstood - although this does not bother me - I am not against these meetings as a recreational activity for retired brothers who have spent long years under the rule of the curriculum
<ul style="list-style-type: none"> وقد طلبت إليها المحكمة بعد ذلك أن تحضر أمامها لتجيب على هذه الدعوى ولكنها لم تحضر على رغم أنها أعلنت مرتين وثلاثا <p>(Al-Ahram, opinion articles)</p> <ul style="list-style-type: none"> The court then asked her to attend the court to answer this lawsuit, but she did not appear, even though she was announced twice and three times. 	<ul style="list-style-type: none"> ورغم أن العقلاء متفقون على خطر هذا السلاح الخبيث - سلاح الشائعات - إلا أنه أيضا يعتبر سلاح الجبناء من أعداء الوطن في الخارج ، والخونة من بعض أبناء الوطن العاقين في الداخل <p>(Alriyadh, opinion articles)</p> <ul style="list-style-type: none"> Although the wise people agree on the danger of this pernicious weapon - the weapon of rumors - it is also considered the weapon of cowards among the enemies abroad, and traitors inside the country.

<ul style="list-style-type: none"> • كما وأن الرفاهية التي وصلت إليها البلاد في عهدنا الحاضر أصبحت تتطلب إلى تأمين الاستهلاك مهما كان رغم أن المواد العمومية قد بلغت (160) موردة بعد أن كانت لا تزيد عن ثلاثين <p>(Al-Ahram, business)</p> <ul style="list-style-type: none"> • Also, the welfare that the country has reached in our present era has become required to secure consumption whatever it was although public materials have, reached 160 supplies after they were not more than thirty. 	<ul style="list-style-type: none"> • ورغم أن الحكومة تنفي هذا الخبر من أكثر من مصدر وأكثر من مسئول فإن هناك أيضا من يؤكد أن الحكومة سوف تعلن مع بداية العام الجديد عن فتح أبواب لفرص عمل جديدة تحتاجها مؤسسات الدولة <p>(Al-Ahram, opinion articles)</p> <ul style="list-style-type: none"> • Although the government denies this news from more than one source and more than one official, there are also people who confirm that the government will announce, with the start of the new year, new job opportunities needed by state institutions.
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

3.3.1.4 'through' خلال

The word خلال can be used as a noun or an adverbial annexion-head. When it is used as a noun, it is the plural of خَلْلٌ, having the basic sense 'a gap between two things' (Ibn ʿabbād, 1994, p. 175; AlfayrozĀbādī, 1994, p. 1285). The recent usage of this word in Modern Standard Arabic is as an adverbial annexion-head, which means 'during' and functions as a discourse connector.

The changes that were noticed in relation to the connector خلال, in Table 1, are interesting in terms of both frequency of occurrence and semantic shift. The use of this connector increased by 391% in 2018 compared to 1950. As for its semantics, the basic sense of this word is, as noted above, 'empty space or gap in something or between two things', and it is used sometimes as an adverbial annexion-head meaning 'throughout' or 'during'. Looking at the occurrences of this discourse connector, all of the occurrences in 2018 sub-corpus are as an adverbial annexion-head meaning 'through/by' or as a noun. None of the occurrences involve the basic physical sense of خلال. However, the basic sense is found in the sub-corpus of 1950 in the expression that is mentioned with the examples below إلا إذا نظرت الورقة في خلال الضوء meaning 'unless you look at the banknote in (through) the openings of light'. It is obvious, from the occurrences in the corpus, that the basic (original) sense of the word has disappeared and it is used solely in the corpus as a temporal adverbial annexion-head (ظرف زمان).

It seems that not only has the basic sense disappeared; the prepositions that proceed the noun خلال have also changed. The proposition في proceeded the word خلال 19 times in 1950 compared to once in 2018. As for the preposition من, its usage increase from just one occurrence in 1950 to 380 occurrences in 2018. However, the use of خلال on its own, as a temporal adverbial annexion-head has not changed very much as it occurs very frequently in this usage in both sub-corpora.

1950	2018
<ul style="list-style-type: none"> • من أبناء أمريكا أن المستر جورج كيران ، رئيس القسم السياسي في وزارة الخارجية هناك ، يعتزم القيام بجولة في القارة الإفريقية يقف خلالها على الأحوال في هذه الأقطار لتحديد موقف مشروع ترومان من إعانة البلاد المتأخرة منها. و سيقوم في خلال جولته هذه بزيارة السودان أيضا للغرض نفسه. <p>(Al-Ahram, national affairs)</p> <ul style="list-style-type: none"> • From the news of America, Mr. George Kiran, head of the political department in the Foreign Ministry there intends to take a tour of the African continent during which he will stand on the conditions in these countries to determine the position of the Truman project for helping developing 	<ul style="list-style-type: none"> • وشهد العام المنصرم ثورة على التعاملات النقدية ، حيث قطعت البنوك شوطا كبيرا في تعزيز مبدأ الشمول المالي ، بهدف دمج جميع المعاملات من خلال النظام المصرفي <p>(Al-Ahram, business)</p> <ul style="list-style-type: none"> • The past year witnessed a revolution in cash transactions, as banks have made great strides in promoting the principle of financial inclusion, with the aim of merging all transactions through the banking system.

<p>countries. During this tour, he will also visit Sudan for the same purpose.</p>	
<ul style="list-style-type: none"> • تدور المباحثات بين الهيئات النسوية في الأقطار العربية ، كطلاب الجمعية النسوية في سوريا للبحث في اقتراح عقد مؤتمر لنساء الدول العربية خلال الصيف القادم. <p>(Al-Ahram, national affairs)</p> <ul style="list-style-type: none"> • Discussions are taking place between feminist bodies in Arab countries, such as the request of the Women's Association in Syria to discuss a proposal to hold a conference for Arab women (during) next summer. 	<ul style="list-style-type: none"> • إلا أن هذه المعدلات أخذت في التراجع وبدأت تنخفض شهراً بعد آخر لتتراجع إلى 33.26 % في سبتمبر الماضي ، ثم إلى 30.53 % في نهاية أكتوبر الماضي ، ثم هبطت إلى 25.54 % في نهاية نوفمبر ، وسط توقعات بمزيد من التراجع خلال الأشهر المقبلة . <p>(Al-Ahram, business)</p> <ul style="list-style-type: none"> • However, these rates are declining every month reaching 33.26% last September, then declining to 30.53% at the end of last October, and then falling to 25.54% at the end of November, amid expectations of a further decline during the coming months.
<ul style="list-style-type: none"> • هذا وقد توجه سعادة كامل سليم بك أمس إلى وزارة الخارجية حيث اجتمع بحضرات عبدالرحمن حقي بك وكيل الوزارة وعبدالمعص مصطفى بك مندوب مصر لدى لجنة التوفيق الدولية والأستاذ أحمد بهجت قنصل مصر السابق في بنغازي وظلوا مجتمعين وقتاً غير قصير بحثوا من خلاله بعض المسائل المتعلقة بمهمة . <p>(Al-Ahram, international affairs)</p> <ul style="list-style-type: none"> • His Excellency Kamel Salim Bey yesterday went to the Ministry of Foreign Affairs, where he met with the Honorable Abdul Rahman Hakki Bey, Undersecretary of the Ministry, Abdel-Moneim Mostafa Bey, delegate of Egypt to the International Conciliation Commission and Professor Ahmed Bahgat, the former consul of Egypt in Benghazi, and they held a meeting for a short time during which they discussed some issues related to a mission. 	<ul style="list-style-type: none"> • إن تدخل الدولة لتنظيم هذه المهن تنظيمياً قانونياً أصبح ضرورة ملحة من خلال إصدار نظام يحكم كافة جوانب هذه الأعمال والمهن <p>(Alriyadh, opinion articles)</p> <ul style="list-style-type: none"> • The state's interference in regulating these professions legally has become an urgent necessity by establishing a system that governs all aspects of these businesses and occupations
<ul style="list-style-type: none"> • وستحمل واجهة الأوراق الجديدة نفس رسم جلالة الملك الظاهر في ورقة البنكنوت فئة الخمسة جنيهات التي أصدر أخيراً كما تحمل رسماً لمنذنة جامع القاضي يحيى زين الدين وتظهر علامة أبو الهول المائية إلى يسار هلال الشحادة المنذنة ولم يترك فراغ أبيض في هذه الأوراق للعلامة المائية وذلك لا تبدو معالمها إلا إذا نظرت الورقة في خلال الضوء . <p>(Al-Ahram, business)</p> <ul style="list-style-type: none"> • The new banknotes will have the same drawing of His Majesty the King, which appears on the five-pound banknote that was issued lately. It will also have a drawing of the minaret of Al-Qadi Yahya Zain Al-Din Mosque. The sphinx watermark appears to the left of the crescent of the minaret of the Shehadeh, and there is no white space left in these papers for the financial mark, and therefore its features do not 	<ul style="list-style-type: none"> • ثانياً : أن يزيد الدعم الحكومي للبحث في مصر من حوالي - حالياً- أقل من نصف في المائة من الدخل القومي يعني حوالي مليار دولار سنوياً - يصرف معظمها مرتبات - إلي الضعف أي 1 % من الدخل القومي - و هو ما نص عليه الدستور - في خلال الخمس سنوات القادمة <p>(Al-Ahram, opinion articles)</p> <ul style="list-style-type: none"> • Second: The government's support for research in Egypt shall increase from about - currently - less than half a percent of the national income meaning about a billion dollars annually - most of which is paid salaries - to double or 1% of the national income - as stipulated in the constitution - during the next five years

appear unless the paper is looked at in [through] the openings of light.

3.3.1.5 'on the heels of' في أعقاب

The expression **في أعقاب**, literally meaning, 'on the heels of' has increased by 111% in 2018 compared to 1950. In addition to becoming commoner, the expression shows some semantic change. In Arabic, this expression is used in two senses (in addition to its literal sense, which seems very rare), the first as a locative annexion-head 'at the end of', and the second as a temporal annexion-head 'after'. The first sense does not occur in the 2018 sub-corpus, where it is used to express the meaning 'after' in all occurrences. It can be claimed here that the meaning of this expression has undergone 'monosemisation', in which the expression has lost its many senses and only one sense is used at the current time.

The singular form of this expression **في عقب** (literally 'on one heel of', but used to mean 'after') is not in the list of Alsaif (2012), and, it was not found in the 1950 or the 2018 sub-corpora either. It seems that this expression, in addition to disappearing, has lost its literal sense ('at the end of') and the singular form is no longer used.

As for the connector **عقب**, this is a temporal adverbial annexion-head. It is the singular of **أعقاب**, both forms having the root **ع-ق-ب**. Its use decreased by 12.6% between 1950 and 2018, but it seems that its meaning has not changed in the corpus. Below are examples of the connector **في أعقاب**, and also the related verb **أعقب**, meaning 'follow', in addition to some examples from KSUCCA for the clarification of the usage of the two senses of **في أعقاب** and **في عقب** (with the singular form **عقب**).

2018	1950	KSUCCA
<ul style="list-style-type: none"> • ودعت الولايات المتحدة تركيا إلى >> تجنب أي إجراءات من شأنها إثارة نزاع بين القوات التركية والأمريكية << في سوريا، وذلك بحسب بيان للبيت الأبيض في أعقاب اتصال هاتفى بين الرئيس الأمريكى دونالد ترامب و نظيره التركى رجب طيب أردوغان الأربعاء. <p>(Al-Ahram, international affairs)</p> <ul style="list-style-type: none"> • The United States called on Turkey to "avoid any measures that would cause conflict between Turkish and American forces" in Syria, according to a White House statement following a phone call between US President Donald Trump and his Turkish counterpart Recep Tayyip Erdogan on Wednesday. 	<ul style="list-style-type: none"> • من أبرز عناصر حركة البعث القومي التي زادت حدة و قوة في أعقاب الحرب العالمية الأولى أن وجه المصريون اهتمامهم إلى التعليم الجامعي فقامت جامعة فؤاد الاول لتؤدى رسالتها , وأعقبها جامعة فاروق الاول لتسهم في حمل العبء والاضطلاع بالمسئولية <p>(Al-Ahram, opinion articles)</p> <ul style="list-style-type: none"> • One of the most prominent elements of the national Baath movement, which increased in intensity and strength at the end of the First World War, was that the Egyptians directed their attention to university education. Fouad I University was established to fulfil its mission, and Farouk I University followed to contribute to carrying the burden of responsibility. 	<ul style="list-style-type: none"> • قلت : ومن هذا أخذ ما يكتبه الكتاب في أعقاب الكتب (Linguistics) • I said: book writers took from this what they write at the end of the books.
<ul style="list-style-type: none"> • في غضون ذلك ، أفاد المرصد أن القتال الضارى تجدد في منطقة غوطة 	<ul style="list-style-type: none"> • فالمسألة الآن هي : هل نتبع نصيحة الأستاذ عمر و زميله 	<ul style="list-style-type: none"> • وهذا إنما يعرض في الناس في أعقاب الأمراض المجففة.

<p>دمشق الشرقية الأحد بعد فترة من الهدوء النسبي في أعقاب تقارير عن الاتفاق على وقف لإطلاق النار هناك مساء الجمعة .</p> <p>(Alriyadh, international affairs)</p> <ul style="list-style-type: none"> • Meanwhile, the observatory reported that fierce fighting resumed in the eastern Damascus area of Ghouta on Sunday after a period of relative calm following reports of an agreement to a cease-fire there on Friday evening. 	<p>الأستاذ ويلسون أو نتبع النداء الرزين الذي يملا أرجاء الخافقين في أعقاب هذا الركب السائر في طريق الحضارة والعمران؟</p> <p>(Al-Ahram, opinion articles)</p> <ul style="list-style-type: none"> • The issue now is: Do we follow the advice of Professor Omar and his colleague, Professor Wilson, or do we follow the rational call that spread across the two horizons of the East and the West at the end of the movements of the crowds walking in the way of civilization and urbanity? 	<p>(Science)</p> <ul style="list-style-type: none"> • This is what happens to people at the end of desiccative diseases.
<ul style="list-style-type: none"> • كما تتجه المكسيك للتعاون الاقتصادي مع الصين واستقبال الاستثمارات الصينية في العديد من مجالات التصنيع في أعقاب فشل حل الخلافات مع الولايات المتحدة، حيث شاركت المكسيك بصفة مراقب في قمة " بريكس " التي انعقدت بالصين سبتمبر الماضي. <p>(Alriyadh, business)</p> <ul style="list-style-type: none"> • Mexico is also heading for economic cooperation with China and receiving Chinese investment in many areas of industrialization in the wake of the failure to resolve differences with the United States, as Mexico participated as an observer in the BRICS summit held in China last September. 	<ul style="list-style-type: none"> • و أعقبه الدكتور بشر فارس و الأستاذ رمسيس يونان فتحدثا عن فن الأستاذ جورجى وجهوده و إبراز خصائص المصري في النحت والرسم التي سيقف عليها الباريسيون . <p>(Al-Ahram, business)</p> <ul style="list-style-type: none"> • Dr. Bishr Fares and Professor Ramses Younan followed him. They talked about the art of Professor Georgy and his efforts and highlighting the characteristics of the Egyptians in sculpture and painting that the Parisians would see. 	<ul style="list-style-type: none"> • قد تكون استفرافات مثل هذه والجسم صحيح سليم ويكون السبب فيه كثرة ما في الجسم من ذلك الخلط الذي يستفرغ أيضا ويكون في عقب البرء من علل الكبد إذا عادت إليها قوتها وانقضت عنها ما كان مؤذيا مما لا تنهيا إحالته ونضجه <p>(Science)</p> <ul style="list-style-type: none"> • Such vomits may happen if the body is healthy and the reason for it is the increase of the mixing that is being emptied. It Happens at the end of the healing of liver diseases when it returns to its normal functions and the harmful substances disappear.

4. SUMMARY OF THE RESULTS

This paper attempted to provide an answer to the question raised at the beginning of this study: How has the usage of Arabic discourse connectors changed, quantitatively and qualitatively, in Arabic newspapers as evidenced in the 1950 and 2018 sub-corpora? Calculating the percentages of the changes in frequency of occurrence of the 95 discourse connectors made it possible to precisely measure the changes that have taken place. 29 discourse connectors have increased, 49 have decreased, 2 were almost stable, these being و and أو meaning 'and' and 'or' respectively. The two connectors showed less than 0.3% change. 13 discourse connectors have newly emerged and 1 discourse connector disappeared in 2018 sub-corpus.

As for semantic changes, this study introduces the notion of 'monosemisation' into historical linguistics to deal with the observation of lexical items which originally had a range of different meanings, but for which all but one of the different meanings are no longer used. Monosemisation (i.e. loss of all except for one sense of the discourse connector) was noticed in some discourse connectors. For example, it was found that some connectors that have multiple meanings like *في أعقاب*, are used in one sense only.

The study also recorded some changes in the forms of some discourse connectors. The tendency to use the plural form of some discourse connectors was noticed in the study corpus. For example, the discourse connector *في أعقاب* occurs in the plural form but not the singular, as in *في عقب*, even though they both considered correct in Arabic. Moreover, changes were noticed in the form of connectors that can be used in either the feminine form (by adding the suffix *ة*), or the masculine form. The analysis of these connectors reveals a decrease in the use of the feminine form compared to the masculine form in the study corpus.

It is hoped that this study will offer useful insights for researchers of Arabic, discourse connectors, and linguistics. The study is also intended to help raise awareness of Arabic diachronic linguistics in general and highlight statistical methods in this field for which it would be beneficial to apply in other studies. Linguists can investigate the changes found in this study by analysing them phonologically, morphologically, and syntactically. Statisticians can work with linguists to find solutions to other statistical methods in linguistic studies. Finally, further studies can be conducted on the reasons that may have contributed to changes in Arabic discourse connectors, or the lexicon and syntax of Arabic in general, with a particular focus on the role of foreign languages, mainly English, from a linguistic and a statistical perspective.

5. REFERENCES

Abdelfattah, N. M. S. (1990). *Linguistic changes in journalistic language in Egypt, 1935-1989: A quantitative and comparative analysis*. PhD thesis, University of Texas at Austin.

Abdul Razak, Z. R. (2011). *Modern media Arabic: A study of word frequency in world affairs and sports sections in Arabic newspapers*. PhD thesis, University of Birmingham.

Alajlan, S. (2019). Compiling a diachronic corpus of Arabic newspaper texts: methodology and challenges. International Corpus Linguistics Conference (CL2019), 22-26 July 2019, University of Cardiff, UK.

Alarfaj, A. & Alsalamn, A. (2020). An efficient semantic relation extraction method for Arabic texts based on similarity measures. *International Journal of Computational Linguistics (IJCL)*, 9(1), pp. 22–38.

AlfayrozĀbādī, M. (1994). *القاموس المحيط*. 4th ed. Heritage Investigating Office at Alresalah Publishing, Mohammed Naeem Alerqososi (Eds). Alresalah Publishing. Beirut, Lebanon. Published in Arabic.

Al-Hejin, B. (2012). *Covering Muslim women: A corpus-based critical discourse analysis of the BBC and Arab News*. PhD thesis, University of Lancaster.

Almujaiwel, S. (2020). A comparative evaluation of POS Tagging and N-Gram measures in Arabic corpus resources and tools. *International Journal of Computational Linguistics (IJCL)*, 11(1), pp. 1–17.

Alrabiah, M., Al-Salman, A., Atwell, E., and Alhelewh, N. (2014). KSUCCA: a key to exploring Arabic historical linguistics. *International Journal of Computational Linguistics (IJCL)*, 5(2), pp. 27–36.

Alsaif, A. (2012). *Human and automatic annotation of discourse relations for Arabic*. PhD thesis, University of Leeds.

AlSenan, H. (2016). The change in vocabularies of freedoms and rights in Egyptian political writings from al-Ṭaḥṭāwī until 1952: A diachronic approach to lexical semantics. PhD thesis, University of Exeter.

Badawi, E., Carter, M. G., and Gully, A. (2016). *Modern Written Arabic: A comprehensive grammar*. London and New York: Routledge.

Baker, P. (2011). Times may change, but we will always have money: Diachronic variation in recent British English. *Journal of English Linguistics*. 39(1), pp. 65–88.

Bauer, L. (2002). Inferring variation and change from public corpora. In: Chambers, J.K., Trude-gill, P., and Schilling-Estes, N. eds. *The handbook of language change and variations*. Malden Massachusetts, USA: Blackwell Publishers Inc, pp. 95–114.

Brierley, C. and El-Farahaty, H. (2019). An interdisciplinary corpus-based analysis of the translation of كرامة (karāma, 'dignity') and its collocates in Arabic-English constitutions. *The Journal of Specialised Translation*. (32), pp. 121–145.

Brierley, C., Sawalha, M., Heselwood, B. and Atwell, E. (2016). A verified Arabic-IPA mapping for Arabic transcription technology, informed by Quranic recitation, traditional Arabic linguistics, and modern phonetics. *Journal of Semitic Studies*. 61(1), pp. 157–186.

Buerki, A. (2013). Automatically identifying instances of change in diachronic corpus data. *Corpus Linguistics 2013*, 22–26 July 2013, Lancaster University, UK.

Dickins, J. (2009). Clausal annexes and related phenomena in Sudanese Arabic. In Watson, J.C.E. and Retsö, J. (Eds), *Genitive structures in Semitic* (special publication of the *Journal of Semitic Studies*). Pp. 135-151.

Dickins, J. (2017). The pervasiveness of coordination in Arabic, with reference to Arabic>English translation. *Languages in Contrast*. 17(2), pp. 229–254. ISSN 1387-6759

Dickins, J. Hervey, S. and Higgins, I. (2017). *Thinking Arabic translation: a course in translation method, Arabic to English* (2ndedn.). London: Routledge.

Dickins, J. (2020). *Thematic structure and para-syntax: Arabic as a case Study*. London and New York: Routledge.

Gabrielatos, C. (2007). If-conditionals as modal colligations: A corpus-based investigation. In: Davies, M., Rayson, P., Hunston, S., and P. Danielsson eds. *Proceedings of the Corpus Linguistics Conference: Corpus Linguistics 2007 Birmingham*: University of Birmingham, pp. 1–20.

Gabrielatos, C. and Marchi, A. (2011). Keyness: Matching metrics to definitions. *Corpus Linguistics in the South: Theoretical-methodological challenges in corpus approaches to discourse studies – and some ways of addressing them*. University of Portsmouth, 5 November 2011, pp. 1–28.

Gaddafi, A. (1990). *Study of discourse markers in Libyan spoken Arabic*. PhD thesis, University of London.

Hamilton, L. W., Leskovec, J., Jurafsky, D. (2016). Cultural shift or linguistic drift? Comparing two computational measures of semantic change. In: Su, J., Duh, K., Carreras, X. eds. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, Austin, TX. Association for Computational Linguistics, pp. 2116–2121.

Hemeida, M. (1997). نظام الارتباط والربط في تركيب الجملة العربية. The Egyptian Int. Company for Pub, Longman, Lebanon Library, Beirut, Lebanon. Published in Arabic.

Hilpert, M. and Gries., S. Th. (2009). Assessing frequency changes in multistage diachronic corpora: Applications for historical corpus linguistics and the study of language acquisition. *Literary and Linguistic Computing*. 24(4), pp. 385–401.

Hobbs, J. R. (1985). On the coherence and structure of discourse. Tech. Report No. CSLI-85-37, *Center for the Study of Language and Information (CSLI)*, Stanford University.

Holes, C. (2004). *Modern Arabic: structures, functions and varieties*. London and New York: Longman.

Hovy, E. H. (1993). Automated discourse generation using discourse structure relations. *Artificial Intelligence (Special Issue on Natural Language Processing)*. Elsevier, 63 (1–2), pp. 341–385.

Hussein, M. (2009). *Relevance theory and procedural meaning: The semantics and pragmatics of discourse markers in English and Arabic*. PhD thesis, Newcastle University.

Ibn ʿabbād, A. (1994a). المحيط في اللغة. Shaikh Mohammed Hassan Al-Yaseen (Ed). Vol. 4, AlamAlkotub, Beirut, Lebanon. Published in Arabic.

Kalajahi, S. A. R., Abdullah, A. N., Mukundan, J., and Tannacito, D. J. (2012). Discourse connectors: An overview of the history, definition and classification of the term. *World Applied Sciences Journal*. 19(11): pp. 1659–1673.

Khalifa, I., Al Feki, Z., and Farawila, A. (2011). Arabic discourse segmentation based on rhetorical methods. *International Journal of Electric & Computer Sciences*. 11(1): pp. 10–15.

Knott, A. and Sanders, T. (1998). The classification of coherence relations and their linguistic markers: An exploration of two languages. *Journal of Pragmatics*. 30(2), pp. 135–175.

Mahmoud. A., (2013). A linguistic perspective of the effect of English on MSA: Manifestations and ramifications. *Journal of King Saud University – Languages and Translation*. 25(1), pp. 35–43.

Mann, W. and Thompson, S. (1986). Relational Propositions in Discourse. *Discourse Processes*. 9(1), pp. 57–90.

McEnery, T. and Hardie, A. (2012). *Corpus linguistics: method, theory and practice*. Cambridge: Cambridge University Press.

Mohamed, A. (2000). *A contrastive study of discourse features in English and Arabic writing in connection with two text types: research article abstracts and newspaper letters to the editor*. PhD thesis, University of Exeter.

Sartori, M. (2019). A relational approach to modern literary Arabic conditional clauses. In: McEnery, T., Hardie, A., and Younis, N. eds. *Arabic corpus linguistics*. Edinburgh: Edinburgh University Press, pp. 143–169.

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Tabasi, A. (2014). Effect of Arabic Language and Grammar on Persian language and Grammar. *Research Journal of Recent Sciences*. 3(8): pp. 125–129.