A Contrastive Analysis of Lexical Repetition Across the Persuasive Speech and Writing of Arabic and English Majors

Ons Abdi[1] *Faculty of Letters and Humanities, University of Kairouan, Tunisia*

Abstract

The current study aims to verify the claims, in the literature, about the repetitive nature of the Arabic language, its oral origin, as well as its potential transfer into Arabs' EFL discourse. To this end, using a mixed-methods research design, the study investigated the use of lexical repetition (LR) in 40 EFL essays, 40 EFL speeches, 30 Arabic as first language (AL1) essays and 30 AL1 speeches produced by Tunisian university students majoring in English and in Arabic. Through a quantitative analysis of the data, LR was measured by calculating the frequencies, as well as densities of simple and complex lexical repetition (SLR and CLR) and of overall LR. Furthermore, the frequencies and average length of lexical repetition chains (LRCs) were quantified. The statistical tests revealed an absence of a significant difference in the use of LR between AL1 written and oral productions, which implies a possible interplay between speech and writing in Arabic. Similarly, there was no statistically significant difference between AL1 and EFL written essays in LR measures. In contrast, the findings demonstrated a significant difference between the EFL written and oral corpora, with writing unexpectedly displaying higher rates of LR. Overall, the corpora exhibited a prevalence of SLR over CLR. The findings also uncovered longer LRCs and a higher density of CLR in writing in comparison to speech. The qualitative analysis unveiled two major patterns of distribution of LR, namely even dispersions and (densely) clustered distributions. It also accounted for the comparable CLR densities in EFL writing and AL1 productions, despite the highly derivational nature of Arabic. Finally, the study offered some examples of Arabic-specific LR.

Keywords: Lexical repetition, EFL, AL1, speech, writing, argumentative discourse, contrastive rhetoric.

---

[1] ons.abdi@flshk.u-kairouan.tn

Introduction

In the field of contrastive rhetoric, which explores the English discourse of EFL learners in relation to their native language and rhetorical traditions, the writing challenges faced by Arab learners are frequently attributed to the influence of their native language. In this context, the Arabic language is situated in an oral culture with long-standing oral rhetorical traditions and modes of thinking (e.g., Abdi, 2024; Johnstone, 1990; Mohamed & Omer, 2000; Ostler, 1987; Zaharna, 1995). In intercultural studies comparing the discourse and communicative style of Arabs to those of English native speakers, oral-like features, such as indirectness, lack of organization and abundance of coordination are reported, among others, to characterize Arabs' discourse (e.g., Feghali, 1997; Zaharna, 1995). More particularly, the Arabic language and culture are often described as not only tolerant of, but also inclined toward repetitiveness (e.g., Al-Khafaji, 2006; Johnstone, 1990, 1994). Sa'Adeddin (1989) and Hatim and Mason (1997) highlight repetition and redundancy as signs of an orally developed text. While these researchers agree that both oral and visual modes of text development are available to both Arabic and English native speakers, they also argue that Arabic leans more towards oral textual development.

Johnstone (1991), who studied different forms of repetition in Arabic, interprets lexical repetition as being, partly, a manifestation of the "oralness" of Arabic discourse, which is a "hold-over, in writing, of earlier oral norms and requirements" (Johnstone, 1990, p. 226). In the same vein, Van De Wege (2013) links the repetition of

key words in Arabic to the influence of Arabic rhetoric and the *Qur'an* and presents it as a rhetorically favoured means of persuasion.

While many scholars (e.g., Al-Jubouri, 1984; Hatim & Mason, 1997; Károly, 2010) associate the Arabic language and rhetoric with repetitiveness, a review of Arabic rhetorical traditions shows, on the contrary, a preference for brevity, conciseness and the avoidance of repetition. This preference is reflected in most definitions of Arabic rhetoric by Arab rhetoricians. One example of definition is by *Ibn al-Muqaffa'* (d. c. 756/759 CE) who characterized rhetoric as "succinctness" (Abdul-Raof, 2006). Accordingly, he supported views regulating discourse, such as the avoidance of repetition, of initial clichés and of complex propositions. Similarly, in his treatise on rhetoric and brevity, Al-Jāḥiẓ (1991) asserts that all people in the world, Arabs and non-Arabs alike, favour and laud brevity and denigrate excess, lengthiness and repetitiveness.

While I study the influence of Arabic rhetoric on the EFL writing of Tunisian English and Arabic majors through conjunctive relations and the use of connectives in Abdi (2021), I investigate this influence, in the current paper, through the use of lexical repetition, independently of other cohesive devices. This study examines the potential oral origins of lexical repetition by incorporating the dimension of orality

into the analysis. Accordingly, it analyses its use across languages (English and Standard Arabic[2]) and across modes (speech and writing).

The current paper starts with a review of the literature which examines lexical repetition in Arabic and English within, and independently of, cohesion. After identifying the research gap and stating the research questions, the paper outlines the research design, including data collection, transcription and analysis procedures. The model of analysis is also detailed, namely the different measures employed to quantify lexical repetition, and the statistical tests used for the comparative analysis of the corpora. The quantitative findings are subsequently supplemented by a qualitative analysis, which uncovers additional patterns in the use of lexical repetition.

Review of the literature

Repetition in Arabic and English, namely lexical repetition (LR), has been studied in different ways. Some scholars investigated LR quantitatively within the broader framework of cohesion, whereby its occurrence and frequency were compared to other cohesive devices'. In these studies (e.g., Abbas et al., 2016; Al-Shurafa, 1994; Mohamed & Omer, 2000; Mohamed-Sayidina, 2010; Williams, 1984, 1989), the use of LR saliently stood out in comparison to grammatical cohesion, and even within the

---

[2] Standard Arabic is referred to as Tunisian students' L1 in the sense of first language learned in school.

category of lexical cohesion (e.g., Al Huneety et al., 2019; Almuhaysh, 2020; Hellalet, 2013; Khalil, 1989).

Other scholars focused on the study of repetition outside of cohesion as defined by Halliday and Hasan (1976) (e.g., Al-Jubouri, 1984; Johnstone, 1987, 1991; Labidi, 1992; Ouaouicha, 1986). Their qualitative studies described its different manifestations (e.g., repetition of form and structure vs. repetition of content) and its impact on argumentation and the rhetorical organization of discourse. For instance, Al-Jubouri (1984) who investigated parallelism and paraphrase in Arabic newspaper articles concluded that repetitiveness is more than a local linguistic phenomenon, but rather a strategy aimed at the overall effect of persuasion. In the same vein, Labidi (1992), who qualitatively studied repetition in Tunisian learners' EFL argumentative essays, also underscored the strategic use of repetition in Arabic at both the linguistic and rhetorical levels. However, he also noted remarkable levels of repetitiveness in the students' EFL discourse.

A third type of studies with a mainly quantitative character focused on LR independently of other textual features. Unlike the other descriptive studies, this type relied on systematic quantifications to investigate how repetition and paraphrase built the general lexical structure of discourse. In a first study, Al-Khafaji (2005b) investigated LR in a pair of parallel Arabic and English argumentative texts. His aim was to test the claim made by different scholars, including Johnstone (1990),

that Arabic texts were characterized by more simple lexical repetition than complex lexical repetition[3] compared to English ones. Although his findings showed little difference in the overall frequency of recurrent terms in the pair of Arabic and English texts, their distribution displayed a remarkable difference, as the Arabic text were characterized by more simple lexical repetition than complex lexical repetition at the level of long lexical chains. In a second study, Al-Khafaji (2005a) investigated LR in three Arabic texts of different types (legal, argumentative and fictional narrative), uncovering text-type dependent differences between the corpora. His findings revealed relatively high densities of LR, especially in the legal and argumentative texts. He linked these high densities to the text's high level of informativeness in the legal corpus and to persuasiveness in the argumentative corpus. Unlike Al-Khafaji (2005a, 2005b), who focused on the influence of language and text type on the use of LR, Naser and Almoisheer (2018) explored LR in relation to gender in the EFL writing of 60 Saudi students. Even though their study did not report an extensive use of LR, their findings showed that simple lexical repetition was the most used category.

The aforementioned studies investigated Arabic's tolerance of repetition and its potential transfer to Arabs' EFL writing by analysing LR in their Arabic and EFL written texts. However, a crucial aspect, namely the potential oral origin of LR, remains unexplored. To address this gap, the current study examines LR across

---

[3] These categories are defined on pages 38, 39 and 40.

different modes (speech and writing) in both the EFL and Arabic as first language (AL1) productions of Tunisian students majoring in English and their peers majoring in Arabic.

Objectives of the study

The present study has four major objectives. First, it seeks to verify the uncorroborated claim, in the literature, regarding the repetitiveness of Arabic as a sign of orality manifested in writing. To this end, it compares the use of LR in the AL1 argumentative speech and writing of Tunisian students majoring in Arabic (henceforth, TAM). Second, it examines the influence of Arabic on the students' EFL writing by comparing LR in the argumentative EFL writing of Tunisian Students majoring in English (henceforth, TEM) and the argumentative AL1 writing of TAM. Third, it studies the potential influence of orality on the use of LR in TEM's EFL writing by comparing it to their EFL speech. Last, through a qualitative analysis, the study explores the patterns of LR in the oral and written, AL1 and EFL corpora, gauging similarities and differences in the use of LR across languages and modes.

Research questions

Using multiple levels of comparison, namely between oral and written AL1 productions, written AL1 and EFL essays, as well as oral and written EFL productions, this study aims to answer the following research questions:

1- How do TAM's AL1 oral productions compare to their AL1 written productions in their use of LR? How does this support the claim about Arabic's repetitiveness and its oral origin?

2- How do TAM's AL1 written essays compare to TEM's EFL written essays in their use of LR? What implications do the findings have for the influence of Arabic on EFL writing regarding LR?

3- Does the use of LR in TEM's EFL speech differ from its use in their EFL writing. If so, what does this suggest about the potential influence of orality on EFL writing regarding the use of LR?

4- What patterns of LR emerged in the oral and written, AL1 and EFL corpora, and what are the similarities and/or differences in the use of LR across languages and modes.

Research design

To investigate LR across AL1 and EFL, spoken and written productions, this study adopts a mixed-methods research design, combining quantitative and qualitative analyses of the corpora. In the quantitative analysis, LR measures, as detailed in the model of analysis, are calculated. Statistical tests are then applied to compare the corpora in terms of these LR measures. The qualitative component provides a more nuanced understanding of the findings through the exploration of the similarities and differences in LR use across the corpora. Accordingly, the description of specific

examples of LR use offers deeper insights into the data analysis procedure and the emerging LR patterns in the datasets.

Data collection

The data was collected using convenience sampling. The written and oral, EFL and AL1 samples were produced by final-year students majoring in Arabic and in English at the University of Manouba and the University of Kairouan, Tunisia. To ensure representative datasets, the samples were selected from larger corpora, with overly short productions excluded and within-group variance minimized. This process resulted in a final corpus of 140 productions. These productions are split into four sub-corpora, namely, 40 EFL essays, 40 EFL speeches, 30 AL1 essays and 30 AL1 speeches. The sub-corpora are labelled: WEFL, OEFL, WAL1 and OAL1[4], respectively.

To ensure consistency and comparability across languages and modes, the students completed the same argumentative task using a writing prompt adapted from Connor and Lauer (1985). The prompt was modified to accommodate the speaking task, and translated into Arabic, resulting in four different versions. Each student produced either a written argumentative essay or an argumentative speech, on a problem of their choice, as elicited by the prompt.

---

[4] The labels for the corpora are descriptive of their mode and language of production. The first letter in the label, W or O, stands for the corpora's mode, written or oral. The rest of the letters make up the acronym for the language of production, i.e., either EFL or AL1.

The written essays were invigilated, timed (120 minutes), and produced under controlled conditions and in a consistent manner. The oral productions were recorded individually. First, the students prepared their speeches without taking notes to avoid the potential memorization and recitation of written parts. Once prepared, each student delivered her speech individually in front of the recording researcher.

Transcription

The oral productions were first transcribed verbatim. The Arabic corpora (WAL1 and OAL1) were transliterated using the *ALA-LC Arabic Romanization Table* (2012). To ensure comparability with the written corpora in measures of lexical repetition, the spoken corpora were edited. Specifically, hesitations, false starts and transitional word repetitions were eliminated. Any time a student corrected herself, only the final version was kept.

Data analysis

Given the interpretative element in the identification of LR, a manual computer-assisted annotation of the corpora was opted for. Accordingly, a concordancer, *AntConc 3.5.9* and an additional program, *Repetition Detector*, were used to identify and quantify all instances of LR in the corpora. While these tools were helpful, instances of repetitions and their types (simple or complex) still had to be checked manually. To annotate lexical repetition chains and types of repetitions the software *UAM CorpusTool 3.3v2* (O'Donnell, 2019) was used. The calculations of LR measures

as well as the statistical tests used in the comparison of the four datasets were performed using SAS (Statistical Analysis System).

Model of analysis

To investigate LR across the EFL and AL1 argumentative speech and writing of Tunisian English and Arabic majors, the study focuses on the lexical repetition of form. This type of repetition is not only the most frequently occurring category, as reported in previous research, but also the most quantifiable (Al-Khafaji, 2005a). Its surface-level nature makes it, equally, a particularly suitable form of repetition for examining the potential impact of the students' L1 on their EFL writing and the influence of orality on writing. In contrast, repetition of content may be more indicative of linguistic maturity, and therefore less symptomatic of orality. Like Al-Khafaji (2005a, 2005b), I study formal LR using two categories from Hoey's (1991) model: Simple lexical repetition and complex lexical repetition. They are defined in the following sections.

Simple lexical repetition (SLR)

According to Hoey (1991, p. 55), SLR takes place "when a lexical item that has already occurred in a text is repeated with no greater alternation than is entirely explicable in terms of a closed grammatical paradigm". In other words, lexical items are either repeated in the exact same way, or are changed in a grammatical way, i.e., through inflection. An example of inflectional change is adding the plural morpheme to a noun, or the {-ing} morpheme to the base form of a verb. Examples from the

participants' EFL and AL1 oral and written productions include: *witnesses– witnessed,*
way- ways, affect- affected, people- people, government- government, al-ta'awwud- al-
ta'awwud (verbal noun meaning "the getting used to"), *al-ṭifl- al-aṭfāl- aṭfālihim* (the
child- the children- their children), *atharu– āthāru* (trace- traces), *yusabbibu- sabbabahā*
(causes (v), caused [them]).

As can be seen from the last examples of plural and tense inflections, Arabic
inflectional morphology can result in radical formal changes, such as the addition of
both a prefix and an infix to change the singular noun *ṭifl* (child) *to the plural aṭfāl*
(children). In the same way, the inflectional changes related to tenses paired with
subject-verb agreement can completely transform the structure of verbs in Arabic.
Inflection in English can also lead to morphologically unrelated forms as well, like in
the words "good", "better" and "best".

Complex lexical repetition (CLR)

According to Hoey (1991, p. 55), CLR involves the existence of "two lexical items
[which] share a lexical morpheme, but are not formally identical [...], or when they
are formally identical, but share different grammatical functions". The first part of
the definition suggests that CLR describes items that share a root and relate to each
other through derivational morphology, i.e., through the addition or deletion of one
or more derivational morphemes. The Second part of the definition is a reference to
zero derivation, also called *conversion* or *functional shift* (Brinton & Brinton, 2010),
which results in exactly similar lexical words with different word classes. An

example is *influence* (noun) – *influence* (verb). Other examples of CLR from the students' EFL discourse include *psychology- psychological- psychologically, pollute-pollution, responsible- irresponsible*.

Arabic derivation is different from its English counterpart. For instance, while English derivational morphemes are either prefixes or suffixes, Arabic has derivational prefixes, suffixes and also infixes, like in the following example: *'āṭil* (unemployed) is a noun, more particularly, *ism al-fā'il* (agentive noun) from the verb *'atala* (as in *'atala 'an al-'amal,* or "not in a working state"). The derivation of the noun from the verb involves the addition of an infix. The noun *ta'tīl* (hampering), from the same production, is a CLR of *'āṭil.*

Lexical repetition chains (LRCs)

The study adopts Al-Khafaji's (2005a, 2005b) definition of LRC. Accordingly, a LRC is signalled when a first SLR or CLR of an open-class item is detected. This means that every open-class lexical item in the production is checked against all others. However, unlike Al-Khafaji (2005a, 2005b) who analysed small corpora (a maximum of three texts of unequal length at a time), the current study did not quantify all LRCs in the corpora. Given the study's objective of examining LR to compare languages and modes, and verifying the potential repetitiveness of students' productions, rather than studying its cohesive function in general, the study quantified LRCs that

are at least four words long[5]. In other words, the study set a threshold of three repetitions per chain for a LRC to be computed. The minimum number of repetitions in a long LRC was set to five words. In other words, a long LRC is six or more words long. This choice is motivated by the considerable difference in frequency, in the corpora, between four- and five-word-long LRCs and longer ones. The table below presents the LR measures used in this study, which are selected from the group of measures employed by Al-Khafaji (2005a).

Table 1

General Measures of LR

Measure	Description
Average length of LRCs	1. Total number of lexical words in LRCs divided by number of LRCs in whole text
Average length of long LRCs	2. Total number of lexical words in long LRCs divided by number of long LRCs in whole text
Average density of SLR or CLR in all LRCs	3. Total number of SLR or CLR in all LRCs times 100 divided by total number of LRs: SLR x 100/ (SLR+CLR) or CLR x 100/ (SLR+CLR)
Average density of LR in the text	4. Number of lexical words in LRCs times 100 divided by total number of text (orthographic) words

Examples 1 and 2, below, represent two English and Arabic repetition chains. The first chain is six words long and consists of four complex repetitions and one

---

[5] The length of a LRC is based on the number of repetitions plus the initial word.

simple repetition. The second is five words long and consists of two simple and two complex repetitions.

(1)   Accept (verb)| accepting (noun)| acceptance (noun)| acceptance (noun) | accept (verb) | acceptance (noun) (WEFL)

(2)   al-shabāb/the youth (noun) | shubbān/young men (noun) | shābbāt/young women (noun) | al-shubbān/the young men (noun) | shabāb/youth (noun) (WAL1)

The average measures were first calculated for individual oral and written texts (to account for differences in text sizes) and then the overall average for each corpus was computed. The mean frequencies for SLR, CLR, and the overall LR in both LRCs and long LRCs were also calculated. Then, the appropriate statistical tests to contrast the corpora were applied.

Findings

Lexical repetition chains

The four corpora, WEFL, OEFL, WAL1 and OAL1, were analysed in terms of the LR categories detailed in the model of analysis. Lexical repetition chains (LRCs) were identified, their frequencies quantified, and their length measured. Table 2 presents the absolute frequencies of both LRCs and long LRCs as well as their average lengths across the corpora. According to Table 2, the written corpora are characterized by a higher LRC average length than the oral ones, within the same language (6 vs. 5.64 words in AL1, and 6.35 vs. 5.41 words in EFL). In the EFL corpora, the difference in average length is even more notable as far as long (i.e., six-or-more-words-long) LRCs are concerned (9.47 vs. 7.77 words). The Arabic written corpus also showed a

higher average length value for long LRCs as compared to its oral counterpart, but the difference was less considerable than in the EFL corpora, amounting to roughly a word.

Table 2

Summary of LRCs Frequency and Length

Measures	WEFL	OEFL	WAL1	OAL1
Total number of LRCs	300	156	218	169
Total length of LRCs	1884	832	1351	1067
Average length of LRCs	6.35	5.41	6	5.64
Total number of long LRCs	124	47	95	75
Total length of long LRCs	1108	357	814	658
Average length of long LRCs	9.47	7.77	8.59	7.57

To further investigate LRCs, the percentage of long LRCs within the total number of LRCs[6] was computed. This calculation aimed to determine which corpora had higher percentages of long LRCs compared to others. Notably, despite having the highest average length values for LRCs and long LRCs, the written corpora did not exhibit the highest percentages of long LRCs in comparison to the oral corpora. In fact, OAL1 showed the highest percentage of long LRCs (44.38%). The Arabic and English written corpora, WAL1 and WEFL, had long LRCs representing 43.58% and 41.33% of the total number of LRCs, respectively. Thus, even though OAL1 did not

---

[6] The percentage of long LRCs= number of long LRCs x 100/ total number of LRCs

have the longest average LRC length, it had a higher percentage of long LRCs than the other corpora, with WAL1 ranking second. This suggests that the Arabic corpora, in general, and the Arabic oral corpus, in specific, tend to have a high percentage of long LRCs.

Lexical repetition densities

Table 3 shows measures of the densities of simple and complex lexical repetition (SLR and CLR) in relation to general LR densities across the corpora. Overall, while SLR densities exceeded by far CLR densities in all the corpora, the EFL and AL1 speeches had the highest SLR average densities, and accordingly their written counterparts had the highest CLR average densities. In this sense, the written mode is linked to more CLR, and the oral mode to more SLR.

Table 3

SLR, CLR and Lexical Density Measures by Corpus

Measures	WEFL	OEFL	WAL1	OAL1
Average density of SLR in all LRCs	79.37	87.21	79.77	81.37
Average density of CLR in all LRCs	20.63	12.79	20.23	18.63
Average density of LR	13.07	7.88	12.99	13.28

EFL essays, AL1 essays and AL1 speeches have relatively comparable average LR densities, which suggests close degrees of lexical repetitiveness between these corpora, except EFL speeches which have a remarkably lower average LR density in comparison.

Frequencies of SLR, CLR and overall LRs

Table 4, below, displays the mean frequencies for the different repetition categories. The normalization of frequencies (1000 words as a basis) allowed for comparisons between datasets and the application of the appropriate tests. The written corpora, WEFL and WAL1, had the highest mean frequencies of LR, not only in general, but also in long LRCs. Interestingly, these values link the written mode to a higher frequency of LR than the oral one, especially in EFL discourse. OAL1 ranks third in the total number of repetitions in general and in long LRCs with mean frequency values that are still comparable to those of WAL1 and WEFL, whereas the OEFL corpus has substantially inferior values (M= 64.06 and M= 30.16).

Table 4

Means and SD for LR Categories

Measures	WEFL	OEFL	WAL1	OAL1
Frequency of SLR	89.32 (41.16)	55.95 (25.88)	87.53 (38.46)	87.02 (49.74)
Frequency of CLR	20.72 (17.19)	8.10 (8.34)	21.08 (15.13)	19.92 (16.35)
Frequency of repetitions in LRC	110.04 (39.48)	64.06 (27.84)	108.61 (45.37)	105.47 (56.51)
Frequency of repetitions in long LRC	68.53 (33.75)	30.16 (27.58)	66.77 (38.11)	57.75 (51.62)

Since SLR is proportional to overall LR, its use in the corpora showed similar trends. Accordingly, WEFL, WAL1 and OAL1 had the same frequency ranking as with overall LR, with comparable SLR mean frequencies, whereas OEFL had a far lower value. The written AL1 corpus exhibited a slightly higher mean frequency of CLR than the oral AL1 corpus. This difference was more notable in the EFL corpora.

SD measures showed a wide dispersion in the use of LR within the corpora, indicating that some productions proved more repetitive than others. Repetition mean frequencies varied the most across Arabic persuasive speeches, while EFL essays showed, mostly, more consistency in the use of repetition, despite the existence of outliers in all four datasets.

Comparing LR measures across the corpora

To compare LR measures detailed in Table 4, the normality of the distribution of the data was first assessed using SAS. Specifically, the Shapiro-Wilk, Kolmogorov-Smirnov, Cramer-von Mises and Anderson-Darling tests were applied to determine the distribution of the data. Based on the results, either a two-tailed t-test or a two-sided Wilcoxon rank-sum test were applied.

Table 5

Statistical Tests for LR's Frequency Measures

		T-test		Wilcoxon Test
OAL1 vs. WAL1				
Feature	parametric	t Value	Pr > |t|	Two-Sided Pr > |Z|
Frequency of SLR	Yes	-0,04	0.9651	
Frequency of CLR	No			0.7558
Frequency of repetitions in LRC	Yes	-0,24	0.8133	
Frequency of repetitions in long LRC	No			0.1446
WAL1 vs. WEFL				
Feature	parametric	t Value	Pr > |t|	Two-Sided Pr > |Z|
Frequency of SLR	Yes	-0,15	0.8538	
Frequency of CLR	No			0.6221
Frequency of repetitions in LRC	Yes	-0,03	0.8888	
Frequency of repetitions in long LRC	Yes	-0,14	0.8396	

| Feature | parametric | T-test | | Wilcoxon Test |
		t Value	Pr > |t|	Two-Sided Pr > |Z|
		OEFL vs. WEFL		
Frequency of SLR	Yes	-4,34	<.0001	
Frequency of CLR	No			<.0001
Frequency of repetitions in LRC	Yes	-6,02	<.0001	
Frequency of repetitions in long LRC	No			<.0001

The statistical tests in Table 5 showed no significant differences in the use of all LR categories between TAM's oral and written productions. Similarly, according to the tests' significance values, there was no significant difference in the way TAM and TEM used LR in their writing. However, unlike TAM whose speech and writing were similar in repetition use, TEM significantly overused repetition categories in their writing in comparison to their speech.

Discussion

To address the first research question relating to the comparison of TAM's persuasive speeches and their persuasive essays in the use of LR, the findings showed the absence of statistically significant differences between the corpora across all LR categories. This suggests a similarity between speech and writing in AL1, with mode exerting no discernible impact on LR use. With speech often reported as being inherently more repetitive than writing (e.g., Crystal, 2005; Kramsch, 1998; Ong, 2012), the comparable repetition rates in Arabic speech and writing suggest a potential interplay between the two modes. Although the overall measures of LR did not differ significantly between the Arabic oral and written corpora, a closer examination of the data uncovered some subtle differences. Notably, the AL1 written

essays showed a slightly higher average length of LRCs in comparison to Arabic speeches, with this difference increasing in long LRCs. Conversely, AL1 speeches displayed a higher percentage of long LRCs than written essays, which aligns with the expected higher repetitiveness of the oral mode. While SLR prevailed in comparison to CLR in both corpora, AL1 writing had a moderately higher density of CLR in comparison to speech. This finding can be attributed to the more sophisticated nature of writing, which allows for the exploitation of more derivational variations of lexical items.

In answer to the second research question relating to the comparison between TAM's and TEM's essays in the use of LR, the statistical tests revealed an absence of significant difference in all LR measures between AL1 and EFL written essays. This finding suggests a possible convergence of LR patterns between AL1 and EFL and consequently implies the potential influence of AL1 on EFL, as Arabic is usually considered lexically more repetitive than English. Measures of average length of LRCs showed LRCs to be slightly longer in EFL essays. Interestingly, this difference became more pronounced in long LRCs. This finding diverges from Al-Khafaji's (2005b) study which reported that long repetition chains had a higher average length in the written Arabic text than written English text. Moreover, contrary to Al-Khafaji's (2005b) finding that Arabic writing showed a higher percentage of CLR than English writing, the present study's findings indicated that both AL1 and EFL written essays had comparable percentages of CLR with EFL writing narrowly surpassing Arabic writing. This result is interesting given the expectation that the

highly derivative nature of Arabic would result in a higher percentage of CLR in TAM's writing. Notwithstanding the comparable percentages in CLR between AL1 and EFL essays, the qualitative analysis shows that CLR was employed differently across these corpora.

In answer to the third research question on how EFL writing compares to EFL speech and the possible influence of orality on writing in EFL, the two corpora showed a significant difference in all measures of LR use. Contrary to the expectations established in the literature about the generally repetitive nature of speech, the findings demonstrated a significantly higher LR density and average length of LRCs, as well as higher amounts of SLR and CLR in EFL writing in comparison to EFL speech.

Qualitative analysis of lexical repetition

Patterns of lexical repetition

To answer the fourth research question, a qualitative analysis of LR across the corpora was carried out, with a focus on examining the immediate context of LR and its patterns to identify potential similarities and/or differences across modes and languages. The qualitative analysis revealed that LR showed two major distribution patterns, namely, even dispersions and clustered distributions.

Even dispersions

The first pattern of LR distribution observed in the corpora was characterized by evenly dispersed recurrences throughout the productions. This type of distribution

was particularly prevalent with the repetition of key terms. It typically characterized repetition chains of medium length, namely four- and five-word-long chains. The uniformity of the distribution decreased with longer repetition chains. This pattern contributed to establishing cohesion by maintaining thematic continuity in discourse, and was, therefore, the least marked distribution across the corpora.

Figure 1

Repetition chains in a WEFL sample

These balanced distributions, which were more prevalent in written texts, typically reflected a well-organized structure that maintained a thematic focus by periodically reminding the reader of key terms and concepts. Figure 1 provides examples of evenly dispersed repetition chains, including those starting with the

words "teenagers" (red), "life" (green) and "internet" (blue). These words represent key terms in this example essay on the negative impact of using the internet.

The repetitions in the lexical chain starting with the term "used" (dark purple) are not as evenly scattered in the text as the repetitions in the afore-mentioned examples of chains. They are rather concentrated in the introduction of the essay which contains five out of the eight total occurrences. I call this pattern of repetition a clustered distribution. Productions across the corpora usually alternated the use of these two patterns, i.e., even dispersions and clustered distributions.

Clustered distributions

Like even lexical dispersions, clustered distributions also served a cohesive function and were an integral part of the regular lexical patterning of discourse. Clustered repetitions often signalled a shift in propositional focus, thus marking a thematic change or the introduction of a new idea or argument. In this way, the emergence of a new cluster of repetitions usually coincided with a propositional reorientation of discourse. For instance, in the example illustrated in Figure 1, the clustered repetition of "use" is closely linked to the description of the problem, namely the negative use of the internet. When the focus shifted to describing the consequences of internet addiction in subsequent parts of the essay, these repetitions ceased to appear.

Figure 2 illustrates the repetition chains in an oral EFL sample. Three out of the four chains are clustered in distinct locations throughout the speech, each marking a significant ideational shift. For example, the chain initiated by the term "think" (red)

is situated in the introduction where the speaker articulates the problematic situation

from her personal perspective, resulting in the repeated use of a verb of opinion.

Figure 2

Repetition chains in an OEFL sample

Similarly, the repetition chain beginning with "problem" (blue) is clustered in

the middle section of discourse, where the speaker focuses on describing the

challenges that the students face at university. Likewise, the chain starting with

"succeed" (purple) is concentrated in the final section, which is devoted to solutions,

as the student posits that "success" is the key to overcoming the previously described

problems. Although the chain starting with "best" (green) is also propositionally

bound to solutions, it shows a less clustered pattern compared to the other chains.

Clustered distributions were also a notable feature of the Arabic written corpus.

Figure 3 displays the repetition chains in a written AL1 sample. In this essay on

unemployment, a significant portion of the repetition chain starting with

"*mashākila*"/problems (red) is concentrated in the opening section, where the student

introduces the problem.

Figure 3

Repetition chains in a WAL1 sample

In contrast, the chain starting with "*ḥall*"/solution (purple) is mainly limited to

the conclusion section. After generally introducing the problem, the student becomes

more specific in the development section by replacing the word "*mushkil*"/problem

by "*al-biṭālah*"/unemployment (blue). This substitution resulted in a new repetition

chain that spread across the remainder of the essay. Occasionally, in an effort to

avoid excessive repetition, TAM tried to use synonyms or other terms to replace

some of the keywords in their writing. However, these substitutions often resulted in the formation of new repetition chains. This is the case of the word "*ẓāhira*"/phenomenon (green) which replaced "*al-biṭālah*"/unemployment in the example WAL1 sample.

In some instances, clustered repetition chains showed a high degree of proximity between recurrences. The chain starting with "*mashākila*"/problems (red) in Figure 3 is a notable illustration of this phenomenon. Likewise, in the chain starting with "*al-madrasah*"/school (yellow), three out of the four total occurrences were concentrated in the same line. This type of dense clustering is rather marked and could, in fact, be attributed to the influence of orality on writing, as it is possible to avoid it through editing and substitution. This thick clustering was more frequent in AL1 essays than their EFL counterparts.

(3) The last obvious negative effects of smoking is having a bad *relationships with other **people**. Smoking is a bad thing for **people**'s life. In fact, it contributed in destroying the different relationships between **people**. Also, it creates the misunderstanding, hostility, minority, violence and inequality between **people**. (WEFL)

(4) *lākin al-mushkil hunā laysa fī hādhā **al-wāqiʻ**, wa lākin al-mushkil yakmun fī ṭarīqat al-taʻāmul maʻa hādhā **al-wāqī'**. fa idhā kāna bi-dākhil al-kull raghbah kāminah fī taghyīr hādhā **al-wāqi'**. hal yakūn al-ḥall fī al-taslīm bi-hādhā **al-wāqi'*** (WAL1)

(but the problem here is not in this **reality**, but the problem consists in the way of dealing with this **reality**. So if there was inside everyone a latent desire to change this **reality**. Does the solution lie in acquiescing to this **reality)**

Examples 3 and 4 provide further illustration of densely clustered repetitions in writing. In both examples, the students appear to have formulated their ideas in a manner similar to speech, focusing on the immediate proposition being generated.

This approach resulted in repetition that served to help the students keep track of their ideas as they wrote. However, the ensuing redundancy also suggests a lack of editing.

Figure 4

Repetition chains in an OAL1 sample

Sa-ataḥaddathu lyawm fī mawḍūʻmuthīr wa-huwa maradh al-āyds, ana shakhṣiyyan mathalan arā ʼannū hādhā l-maraḍ, lā ʼaʻrif innahu yamassu l-insān bi-shakl rahīb, yakbut kul *qudurātu l-insān mathalan yuṣbiḥ ladayh ʻāʼiq fi-ttaʻāmul maʻa-nnās fī mujtamaʻih, hādhā mā yuʻaththir salban ʻala *nafsītu, fa-yuṣbiḥ munʻazil muqayyad mukabbal ghayr rāghib fi-lḥayat arā ʼannu lā mushkul fī hādha l-mujtamaʻ aham min hādhā l-mushkul kayfiyyat attaʻāmul maʻa hādhā l-insān ay marīḍ l-āyds yajib an takūn muʻāmla lā aʻrif kayfa aṣifuhā bil-muʻāmala l-jayyida, muʻāmalat attaʼākhī al-silm al-luṭf al-luṭf bidaraja, yaʻnī hādhā l-insān aʻrif *annu hunāka man yakūn khāṭiʼ wa hunāka man yakūn ʼan ghayr qaṣd yaʻnī intaqalat lahu ʻadwā aw mithl dhālik, fa-limādhā nuʻāmil dhālika l-insān yaʻnī bi-iqṣāʼ bi-tahmīsh bi-khawf bi-rahba, wa-naḥnu mudrikīn, wa-naʻrif yaʻnī muthaqqifīn *annu dhālika l-insān lan yanqula lanā l-ʻadwā illā bi-ṭarīqa jinsiyya aw ṭarīqa yaʻnī (le noyant?), lā yakūnu illā bi-dhālika shshakl, fa-limādhā lā nuṣāfiḥahu lā nataḥaddath maʻahu lā nuḥāwiruhu, limādhā yaʻnī fī hādha l-mujtamaʻ, urīd hādhā *l-mushkul an yantahī yantahī fī tūnis mathalan, li-anna l-mushkul yatafāqam kulla yawm al-iḥṣāʼāt kul yawm tazīd, anā mathalan arā fi-ljāmiʻāt istibyānāt kulla yawm yasʼalūn ʼan hādhā l- mushkul limādhā in lam yakun hādhā l- mushkul muthīr, fa-limādhā yasʼalūna ʼanh, li-dhālika arjū an takūna muʻāmalat uṭf, li-annī arā ʻadīd l-ḥālāt, wa-yuʻsifunī yuʼsifunī dhālik, fa-anā arā anna l-insāna yajibu an yakūna laṭīfan maʻahum, fa-limādhā limādhā yuqṣīhum, bil-ʻaks nushajjiʻuhum ʻalā ḥub l-ḥayāt, ʻalā l-istimrār, ʻalā ttaʼākhī, fa-lā maḍarrata lanā fī dhālika.

Heavily clustered repetitions were a normal feature of speech, as exemplified by the oral AL1 sample in Figure 4. With the exception of the chain starting with "*l-insān*"/the human (red), which was evenly dispersed across the speech, the remaining chains consisted of closely proximate repetitions. These recurrences are characteristic of the spontaneous and dynamic nature of the spoken language, where speakers often repeat words in close proximity as they formulate their ideas in real-time.

Lexical repetition chains

When examined in their immediate context, long repetition chains were generally evenly dispersed in long productions. The longest chains in the corpora typically consisted of key terms whose recurrence was necessary for maintaining cohesion. In some cases, the repetition chains were lengthy because the recurrent term was technical or specialized, or because there were no suitable synonyms available to replace it, as seen in examples such as "stress" and "technology". Remarkably, the key terms in long chains were very similar across EFL and AL1 productions, and included words like "smoking", "TV", "children", "*al-tadkhīn*" (the smoking), "*al-aṭfāl*" (the children) and "*al-tilfāz*" (the television), among others.

Generally, four- and five-word-long repetition chains prevailed, quantitatively, over longer repetition chains. Interestingly, many oral and written productions contained no more than one single (very) long repetition chain, while the rest of the chains were considerably shorter. In some cases, productions comprised exclusively four- and five-word-long repetition chains. Against the expectation that long and medium-sized productions (approximately between 300 to 600 words) would include a higher number of repetition chains compared to shorter productions, the length of the productions and the number and length of lexical repetition chains were not clearly proportional across the corpora. Relatedly, the corpora displayed a variation in the use of repetition, independently of production length, with some texts using repetition extensively and others only moderately.

Complex lexical repetition between Arabic and English

A close examination of the oral and written AL1 productions confirmed the highly derivational character of the Arabic language. Yet, this feature was not reflected in the quantitative findings, as the AL1 corpora did not quantitatively stand out in their use of CLR in comparison to EFL writing. The primary reason for that is that English CLR was the outcome of the order in which the repeated items occurred in the productions, rather than a reflection of derivational variation, which is more a characteristic of Arabic. The following examples of repetition chains illustrate this point:

(5) **(a)** society (noun) ǀ social (adjective) ǀ society (noun) ǀ social (adjective) ǀ sociable (adjective) ǀ socially (adverb) ǀ social (adjective) ǀ society (noun) ǀ social (adjective) (WEFL)
(b) creatures (noun) ǀ creatures (noun) ǀ creature (noun) ǀ create (verb) ǀ creatures (noun) ǀ creating (verb) ǀ creatures (noun) (WEFL)

Example 5 (a) presents a nine-word-long repetition chain, where the noun "society" appears three times and the adjective "social" four times. At first glance, the lack of morphological variation in the chain may suggest that simple repetition exceeds complex repetition. Nonetheless, upon closer inspection, it becomes clear that the order of the repeated items, specifically the alternation between the different grammatical classes, results in an exclusively complex repetition chain. In the same way, in example 5 (b), despite the repetition of the noun "creature(s)" five times in a seven-word-long repetition chain, quantification yielded four complex repetitions

against two simple repetitions only. This highlights the importance of considering the order and grammatical context of repeated items when analysing CLR.

The second reason why Arabic did not showcase elevated CLR rates, despite its highly derivational nature, is that many derivationally related words in Arabic were not necessarily semantically related. For the sake of consistency, these words were not counted as cases of CLR, as most derivationally related terms in English were also semantically linked.

Table 6

Examples of Lexical Chains with Derivationally Related Words

Corpus	Root	Lexical chain
WAL1	a) *jīm, mīm, ʻayn*	*mujtamaʻāt* (noun, societies) ǀ *al-ijtimāʻī* (adjective, social) ǀ *jamīʻ* (adjective, all) ǀ *al-jāmiʻāt* (noun, universities) ǀ *al-mujtamaʻ* (noun, society) ǀ *al-jāmiʻiyyah* (adjective, university) ǀ *al-mujtamaʻ* (noun, society)
	b) *ḥāʼ, dāl, dāl*	*ḥadd* (verbal noun, stopping) ǀ *muḥaddad* (adjective, specific) ǀ *al-ḥadd* (verbal noun, the stopping) ǀ *taḥdīd* (verbal noun, specifying)
OAL1	c) *ṭāʼ, lām, bāʼ*	*tataṭallabu* (verb, requires) ǀ *aṭṭalabah* (noun, the students) ǀ *aṭṭalabah* (noun, the students) ǀ *yataṭallabu* (verb, requires)
	d) *ʻayn, wāw, dāl*	*ʻādatan* (adverb, habitually) ǀ *yaʻūd* (verb, goes back) ǀ *yaʻūd* (verb, goes back) ǀ *taʻūdu* (verb, goes back) ǀ *ʻādatan* (adverb, habitually)

The examples in Table 6 illustrate lexical chains that consist of derivationally related words sharing the same root. In example (a), the terms *mujtamaʻāt* (societies), *al-jāmiʻāt* (the universities) and *jamīʻ* (all) do not relate to each other semantically. In example (b), while the terms *muḥaddad* (specific) and *taḥdīd* (specifying) are semantically related, the word *ḥadd* (stopping) does not relate to either. Similarly, in example (c) the verb *yataṭallabu* (requires) and the noun *aṭṭalabah* (the students) are

semantically unrelated, and so are the adverb *ʿādatan* (habitually) and the verb *yaʿūd* (goes back) in example (d).

Arabic-specific lexical repetition

Lexical repetition in the AL1 oral and written samples was at times an inherent feature of the Arabic language itself. In other words, repetition was occasionally embedded in the grammar or formulation of the language, leading to a distinct pattern of repetition that differed from English. This resulted in very close and even back-to-back repetitions that reflected a grammatical/ rhetorical preference specific to Arabic:

(6) **a)** *maththalat al-biṭālah **ʿāʾiqan yaʿūq** al-dhāt al-insāniyyah*

(unemployment represented a **hinderance** that **hindered** the human self)

b) *al-ṭifl **marīḍ** bi-**maraḍ** al-tawaḥḥud*

(the child is ***sick** with the **sickness** of autism)

c) *fī ṣufūf sharīḥat **al-shabāb** min **shubbān** wa **shābbāt***

(in the lines of **young** men and **young** women among the **youth**)

d) *al-amthilah **ʿadīdah** wa **ʿadīdah** min mujtamaʿinā al-yawm*

(the examples are **numerous** and **numerous** from our society today)

e) ***nadhkur** ʿalá sabīl **al-dhikr***

(we **mention** by way of **mentioning**)

f) *rafḍ hādhā al-wāqiʿ lā qabūluh ka-**musallamah** min **musallamāt** ḥayātinā al-bāʾisah* (WAL1)

(refusing this reality not accepting it as a **given** among the **givens** of our miserable life)

These examples from AL1 writing illustrate various forms of repetitions that lack exact English equivalents, as evidenced by their inadequate literal translations. With the exception of example (e), which is part of the formulaic expression "*nadhkur ʿalá sabīl al-dhikr lā alḥaṣr*" (including but not limited to), these structural repetitions

yielded a wordiness that could not be avoided through editing and reformulation. The presence of these repetitions in written essays, suggests the influence of orality on writing. This possibility is corroborated by the occurrence of similar types of structural repetitions in AL1 speech, as illustrated in the following examples:

(7) **a)** *wa ḥawwala hādhā l-ʿunṣur min ʿunṣur ījābī ilā ʿunṣur salbī hunā*

(and transformed this **element** from a positive **element** to a negative **element** here)

b) *wa ḥāttā wa in nuqallidahum nuqallidahum bi-ḥadhar*

(and even if we **imitate** them we **imitate** them carefully)

c) *wa nuṭawwir dhātinā bi-dhātinā*

(and we improve **our self** by **our self**)

d) *hunāka maqūla tuqāl taqūl*

(there is a **saying** that is **said** that **says**)

e) *innamā l-ḥurriyya an tuʿabbira bi-raʾyika, wa an taḥtarima l-ākhar qabla an taḥtarima *li-dhātika, fa iḥtirāmuka *fī l-ākhar hunāka iḥtirāmun fī shakhṣika adhdhātī* (OAL1)

(but freedom is to express your opinion- and to **respect** the other before you **respect** *for yourself- since your **respect** *in the other there is **respect** in your individual person)

These examples establish the oral character of such structural recurrences. Example (d), in particular, highlights their redundancy with two consecutive and superfluous repetitions. Similarly, some of the recurrences in example (e), even though emphatic, could be eliminated without affecting the proposition expressed.

Conclusion

This study explored the use of lexical repetition (LR) in the persuasive discourse of Arabic and English majors in Tunisia. To test the prevalent notion that the Arabic language tends towards oral-like repetitiveness, the study examined LR in both AL1 speech and writing. It also compared the use of LR and lexical repetition chains

(LRCs) in AL1 and EFL persuasive writing to explore the potential impact of the students' L1 on their EFL writing. The findings showed no statistically significant difference between AL1 oral and written productions, nor between EFL and AL1 written productions, which implies a possible interplay between modes and languages. A comparison between EFL speech and writing, however, resulted in a statistically significant difference between the corpora, with writing showing, remarkably, higher LR rates. Overall, the findings uncovered a higher use of SLR in comparison to CLR in all the corpora. They also revealed some differences between the corpora, especially in measure of LRCs average length, with writing having higher values in both AL1 and EFL, especially in long LRCs. There was also variation in the use of LR within the corpora, with some productions displaying a notably higher use of repetition than others.

The qualitative analysis of the data disclosed two main patterns of LR distribution: Even dispersions and clustered distributions. These patterns served different rhetorical and structural purposes, with many LRCs including lengthy ones, serving to maintain cohesion and thematic focus in discourse. While heavily clustered LR was an unmarked feature of speech, its presence in writing could be indicative of oral mode influence. The use of some forms of Arabic-specific repetitions in AL1 writing could also possibly be traced to the influence of the oral mode as similar forms were also used in AL1 speech. The analysis also showed that complex lexical repetition did not necessarily reflect a richer use of derivation in the students' productions.

61

References

Abbas, J. M., Yasin, M. S. M., & Ismail, K. (2016). Arabic language influence on the Iraqi EFL tertiary learners' use of grammatical cohesive devices in their argumentative essays. *European Journal of Social Sciences Education and Research*, *6*(1), 56–64. https://doi.org/10.26417/ejser.v6i1.p56-64

Abdi, O. (2021). The use of discourse connectives in the written academic discourse of students majoring in Arabic and their peers majoring in English. *Arab Journal of Applied Linguistics*, *6*(1), 32–59.

Abdi, O. (2024). *The impact of Arabic rhetoric on the argumentative discourse of Tunisian students of English: A cross-linguistic study of speech and writing* [Doctoral Thesis]. University of Manouba.

Abdul-Raof, H. (2006). *Arabic rhetoric: A pragmatic analysis*. Routledge.

Al Huneety, A., Al Mashaqba, B., Al-Omari, M., Zuraiq, W., & Alshboul, S. (2019). Patterns of lexical cohesion in Arabic newspaper editorials. *Jordan Journal of Modern Languages and Literature*, *11*(3), 273–296.

ALA-LC Arabic romanization table. (2012). The Library of Congress. https://www.loc.gov/catdir/cpso/romanization/arabic.pdf

Al-Jāḥiẓ, A. I. B. (1991). Al-balāghah wal ījāz. In A. M. Haroun (Ed.), *Rassā'il Al-Jāḥiẓ* (Vol. 4, pp. 151–154). Maktabah al-Khānijī.

Al-Jubouri, A. J. R. (1984). The role of repetition in Arabic argumentative discourse. *English for Specific Purposes in the Arab World*, 99–117.

Al-Khafaji, R. (2005a). Patterns of lexis in arabic text types. *Current Issues in the Analysis of Semitic Grammar and Lexicon*, *56*(3), 125–148.

Al-Khafaji, R. (2005b). Variation and recurrence in the lexical chains of Arabic and English texts. *Poznan Studies in Contemporary Linguistics*, *40*, 5–25.

Al-Khafaji, R. (2006). In search of translational norms: The case of shifts in lexical repetition in Arabic-English translations. *Babel. Revue Internationale de La Traduction / International Journal of Translation*, *52*(1), 39–65. https://doi.org/10.1075/babel.52.1.03alk

Almuhaysh, M. A. (2020). Reiteration Relations in Arab ESL Student Writing: Effect of First Language Transfer. *English Language Teaching*, *13*(7), 8–18.

Al-Shurafa, N. S. D. (1994). Text linguistics and cohesion in written Arabic. *Arts and Humanities. King Abdulaziz University, Jeddah*, *7*, 17–30. https://doi.org/10.4197/art.7-1.2

Brinton, L. J., & Brinton, D. (2010). *The linguistic structure of modern English*. John Benjamins Publishing.

Connor, U., & Lauer, J. (1985). Understanding persuasive essay writing: Linguistic/rhetorical approach. *Text-Interdisciplinary Journal for the Study of Discourse*, *5*(4), 309–326.

Crystal, D. (2005). Speaking of writing and writing of speaking. *Recuperado de: Http://Www. Pearsonlongman. Com/Dictionaries/Pdfs/Speaking*.

Feghali, E. (1997). Arab cultural communication patterns. *International Journal of Intercultural Relations*, *21*(3), 345–378.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.

Hatim, B., & Mason, I. (1997). *The translator as communicator*. Routledge.

Hellalet, N. (2013). Reiteration Relations in EFL Student Writing: The Case of Moroccan University Students. *English Language Teaching*, *6*(11), 160–166.

Hoey, M. (1991). *Lexical Patterns in Text*. Hong Kong, OUP.

Johnstone, B. (1987). Parataxis in Arabic: Modification as a model for persuasion. *Studies in Language*, *11*(1), 85–98. https://doi.org/10.1075/sl.11.1.04joh

Johnstone, B. (1990). 'Orality' and Discourse Structure in Modern Standard Arabic. In M. Eid (Ed.), *Papers from the first Annual Symposium on Arabic Linguistics: Held April 24—25, 1987 at the University of Utah* (Vol. 63, pp. 215–233). John Benjamins Publishing.

Johnstone, B. (1991). *Repetition in Arabic discourse: Paradigms, syntagms, and the ecology of language* (Vol. 18). John Benjamins Publishing.

Johnstone, B. (Ed.). (1994). *Repetition in discourse: Interdisciplinary perspectives*. Ablex Publishing Corporation.

Károly, K. (2010). Shifts in repetition vs. Shifts in text meaning: A study of the textual role of lexical repetition in non-literary translation. *Target. International Journal of Translation Studies*, 22(1), 40–70.

Khalil, A. (1989). A study of cohesion and coherence in Arab EFL college students' writing. *System*, *17*(3), 359–371. https://doi.org/10.1016/0346-251X(89)90008-0

Kramsch, C. (1998). *Language and culture*. Oxford university press.

Labidi, A. (1992). *Arabic cultural/educational and linguistic background as factors affecting EFL writing performance* [Doctoral dissertation, University of Salford, UK]. http://usir.salford.ac.uk/id/eprint/2188

Mohamed, A. H., & Omer, M. R. (2000). Texture and culture: Cohesion as a marker of rhetorical organisation in Arabic and English narrative texts. *RELC Journal*, *31*(2), 45–75. https://doi.org/10.1177/003368820003100203

Mohamed-Sayidina, A. (2010). Transfer of L1 cohesive devices and transition words into L2 academic texts: The case of Arab students. *RELC Journal*, *41*(3), 253–266. https://doi.org/10.1177/0033688210380569

Naser, M., & Almoisheer, R. (2018). Lexical repetition and written text's unity from gender perspectives: A case of languages and translation students at the

University of Tabuk. *African Journal of Humanities and Social Science Research*, *2*(3), 14--22.

O'Donnell, M. (2019). *UAM CorpusTool* (Version 3.3v2) [Java]. http://www.corpustool.com/

Ong, W. J. (2012). *Orality and literacy: The technologizing of the word* (J. Hartley, Ed.; 30th anniversary ed.; 3rd ed). Routledge.

Ostler, S. E. (1987). *A study of the contrastive rhetoric of Arabic, English, Japanese, and Spanish* [Doctoral Dissertation]. University of Southern California.

Ouaouicha, D. (1986). *Contrastive rhetoric and the structure of learner-produced argumentative texts in Arabic and English* [Doctoral dissertation]. The University of Texas.

Sa'Adeddin, M. A. A. M. (1989). Text development and Arabic-English negative interference. *Applied Linguistics*, *10*(1), 36–51. https://doi.org/10.1093/applin/10.1.36

Van De Wege, M. (2013). *Arabic rhetoric: Main idea, development, parallelism, and word repetition* [MA Thesis]. Eastern Washington University.

Williams, M. P. (1984). A problem of cohesion. In J. Swales & H. Mustafa (Eds.), *English for specific purposes in the Arab world* (pp. 118–128). University of Aston in Birmingham.

Williams, M. P. (1989). *A comparison of the textual structures of Arabic and English written texts: A Study in the comparative orality of Arabic* [Doctoral Thesis]. University of Leeds.

Zaharna, R. S. (1995). Understanding cultural preferences of Arab communication patterns. *Public Relations Review*, *21*(3), 241–255.